Genome **Medicine**

DNA methylation signatures for breast cancer classification and prognosis

Moshe Szyf*

Abstract

Changes in gene expression that reset a cell program from a normal to a diseased state involve multiple genetic circuitries, creating a characteristic signature of gene expression that defines the cell's unique identity. Such signatures have been demonstrated to classify subtypes of breast cancers. Because DNA methylation is critical in programming gene expression, a change in methylation from a normal to diseased state should be similarly reflected in a signature of DNA methylation that involves multiple gene pathways. Whole-genome approaches have recently been used with different levels of success to delineate breast-cancer-specific DNA methylation signatures, and to test whether they can classify breast cancer and whether they could be associated with specific clinical outcomes. Recent work suggests that DNA methylation signatures will extend our ability to classify breast cancer and predict outcome beyond what is currently possible. DNA methylation is a robust biomarker, vastly more stable than RNA or proteins, and is therefore a promising target for the development of new approaches for diagnosis and prognosis of breast cancer and other diseases. Here, I review the scientific basis for using DNA methylation signatures in breast cancer classification and prognosis. I discuss the role of DNA methylation in normal gene regulation, the aberrations in DNA methylation in cancer, and candidate-gene and whole-genome approaches to classify breast cancer subtypes using DNA methylation markers.

Keywords DNA methylation, epigenetics, hypermethylation, hypomethylation, breast cancer metastasis, breast cancer stages, whole-genome methods

*Correspondence: moshe.szyf@mcgill.ca
Department of Pharmacology and Therapeutics, Sackler Program in Epigenetics and Psychobiology, McGill University, 3,655 Sir William Osler Promenade, Montreal H3G1Y6, Canada

BioMed Central

Breast cancer classification, staging and prognosis: the importance of DNA methylation signatures

Breast cancer is a heterogeneous disease with very different therapeutic responses and outcomes. It has traditionally been staged by histopathological criteria that are based on size, level of invasiveness and lymph node infiltration, and by immunochemical characterization of cell surface receptors, including estrogen receptor (ER), the progesterone receptor (PR) and the human epidermal growth factor receptor 2 (HER2). However, in many instances staging breast cancer fails to predict prognosis or therapeutic response because of the heterogeneity of the disease. More recently, molecular approaches focusing on gene expression profiles have been used. Classifications based on gene expression profiles have expanded the detailed classification of breast cancer by revealing cellular identity profiles, with particular emphasis on presence of stem cells and the nature of the immune response to the tumor. These new molecular-based classifications are termed 'intrinsic subtypes of breast cancer' because they reveal the molecular identity of the breast cancer cell in the tumor rather than its stage. Several distinct tumor types and normal breast-like intrinsic classes of breast cancer were previously described [1] (see list in Table 1). These different subtypes are found in all stages of breast cancer, even in the early stages, and therefore serve as early prognostic and therapeutic predictors. They have contributed prognostic value in breast cancer management as they are now guiding prediction of patient relapse, survival and response to chemotherapy. However, there are still major challenges in accurate early prediction of breast cancer, prognosis, and prediction of therapeutic outcome. There is significant room for improving our predictive and prognostic tools, particularly in guiding therapeutic choices.

DNA methylation is a molecular modification of DNA that is tightly associated with gene function and cell-type-specific gene function and therefore provides an exquisite identity descriptor of a cell. In the recent past, DNA methylation analysis was targeted at a few candidate genes using either methylation-sensitive restriction enzymes or gene-specific DNA methylation mapping by

Table 1 Intrinsic classification of breast cancer by gene expression profiles and cell surface hormonal expression

ER-positive	ER-low	ER-negative, PR-negative, HER2-negative	Normal breast-like
Luminal A	**HER2-enriched**	**Claudin-low**	**High stromal content**
Luminal B		**Basal-like subtype**	**High lymphocyte infiltration**
			True normal epithelial cell contamination

Column headings indicate primary immunohistochemical criteria only (ER-positive and so on; in italics). The subtypes are indicated below each heading. Subtypes in bold exhibit worse prognosis (relapse and mortality). There are four most commonly referred subtypes in the literature: luminal A, luminal B, HER2-enriched and triple-negative basal-like. Additional subtypes that have been proposed are claudin-low in the triple-negative group and normal breast-like, which are subdivided into additional subtypes. Subtypes in bold show significantly poorer outcome than luminal A subtypes. ER, estrogen receptor; HER2, human epidermal growth factor receptor 2; PR, progesterone receptor.

sequencing bisulfite-converted DNA. These studies provided the initial proof of principle that DNA methylation patterns are different between tissue types and between tumors and normal surrounding tissue. However, because these studies focused on a small number of genes, they provided a narrow, low-content portrait of the DNA methylation pattern. The utility of low-content DNA methylation profiles in classification of different subtypes of cancer in general and breast cancer in particular is inherently limited, as it is evident that genes do not act on their own and that gene networks and modules define cellular identities [2]. Therefore, DNA methylation in cancer would be predicted to influence multiple gene networks rather than single genes. Recent technological advances in DNA methylation mapping, including high-density oligonucleotide arrays, Illumina bead arrays and next-generation high-throughput sequencing, together with advances in bioinformatics, have allowed examination of broad regions of the genome and delineation of high-content profiles of DNA methylation for the first time. Moreover, it is possible to study genome function at several levels, including analysis of microRNA levels, DNA copy number, DNA methylation and histone modifications, and integrate these into combined genomic pathways. Several such studies have examined associations between whole-genome DNA methylation analyses and breast cancer classification and prognosis, and will be reviewed below. For example, Flanagan *et al.* [3] have delineated DNA methylation signatures that are associated with *BRCA* mutation state, but these were not predictive of subtypes defined by gene expression profiling. Kristensen *et al.* [4] used an integrated molecular approach that examined genome-wide transcription, DNA copy number, microRNA and DNA methylation profiles. This study did not reveal improvement to prognostic value by adding DNA methylation and microRNA to the analysis. However, Dedeurwaerder *et al.* [5] have described DNA methylation profiles in a relatively large study that reveal and classify the existence of new breast cancer groups that are not classified by current expression subtypes. The study points to the prospect that DNA methylation signatures will extend our ability to classify breast cancer and predict outcome beyond what is currently possible. It is anticipated that such integrated methods will reveal the genomic basis for heterogeneity of breast cancer.

Hypothesis and critical questions

Given that DNA methylation changes are plausibly critical components of the molecular mechanisms involved in breast cancer, breast cancers as a group and specific subtypes of breast cancer might be expected to show distinct DNA methylation states. These DNA methylation states could serve as diagnostic tools in breast cancer care. The important questions are: 1) are the changes in DNA methylation in breast cancer limited to a narrow set of candidate genes? 2) could DNA methylation states serve as early predictors of breast cancer? 3) could DNA methylation states provide information regarding the stage of breast cancer? and 4) could DNA methylation states provide tools for prognosis and stratification for different therapeutic approaches?

Overall, it is crucial to delineate how specific changes in DNA methylation patterns of subsets of genes relate to the molecular pathologies involved in breast cancer initiation, progression and metastasis. It is also important to determine whether a limited set of specific gene methylation events would be sufficient in breast cancer diagnostics, or whether this would require more complex 'signatures' that involve coordinated changes in groups of genes. Here, I discuss the state of knowledge in this emerging field as well as future directions and prospects. First, I provide a short introduction to DNA methylation and its role in regulating gene function, and the changes in DNA methylation that occur in cancer. I then review candidate-gene and whole-genome DNA methylation mapping approaches aimed at associating DNA methylation profiles with breast cancer subtypes and prognosis.

DNA methylation

Vertebrate DNA is covalently modified by addition of methyl residues at the 5' position of cytosines residing mostly in CG (also known as CpG) dinucleotides [6]. Not all CGs are methylated in vertebrate genomes, and the

distribution of methylated and unmethylated CGs in the genome is tissue-specific, resulting in a cell-specific pattern of DNA methylation [7]. The idea that different cell types have different patterns of methylation was introduced three decades ago [7] and was recently confirmed by whole-genome DNA methylation mapping of differentiating human embryonic stem cells [8].

The DNA methylation reaction is catalyzed by DNA methyltransferases (DNMTs) [9,10]. DNA methylation is unique among all the factors that are involved in programming gene expression because the methyl moiety is a component of the chemical structure of the genome. Thus, a DNA molecule contains, in addition to the ancestral genetic information encoded by the four bases comprising the DNA sequence, a coating of methyl moieties that contains epigenetic information. The genetic information is inherited and copied by the DNA replication enzymatic complex, while the DNA methylation pattern is established during embryonic development by an independent enzymatic process that includes DNMTs and as-yet unknown demethylating enzymes and proteins that target DNMTs to specific positions in the genome [6].

Three distinct DNMTs have been identified in mammals. DNMT1 shows preference for hemimethylated DNA *in vitro*, which is consistent with its role as a maintenance DNMT [11,12], an enzyme that copies the DNA methylation pattern from the methylated parental strand to the unmethylated daughter strand during cell division. DNMT3a and DNMT3b are *de novo* DNMTs, as they methylate unmethylated and methylated DNA at an equal rate [13].

Several proteins have been shown to target DNMTs to specific positions in the genome. For example, EZH2, a member of the multi-protein Polycomb complex PRC2/3, which methylates histone H3 at lysine 27, is believed to target DNMTs to specific locations in the genome [14-17]. This relationship between EZH2 and DNMT is thought to be important in the methylation of tumor suppressor genes in cancer [14-17]. URHF1 targets the maintenance DNMT1 to hemimethylated DNA generated during DNA replication and is required for the copying of the DNA methylation pattern from the template to the daughter DNA strand [18,19]. The binding of transcription factors to specific DNA sequences is also important in targeting or preventing DNA methylation during development, as has been suggested previously [20,21] and confirmed recently [22]. Early candidate-gene approaches and recent genome-wide approaches for measuring DNA methylation are described in Box 1.

DNA methylation and its role in programming gene expression

DNA methylation patterns in vertebrates are distinguished by their correlation with chromatin structure.

Active regions of the chromatin, which enable gene expression, are associated with hypomethylated DNA, whereas hypermethylated DNA is packaged in inactive chromatin [23]. It has been known for more than three decades that DNA methylation in regulatory regions such as promoters and enhancers can silence gene expression and that there is an inverse correlation between gene expression and DNA methylation in promoters [7]. Recent whole-genome approaches have also revealed that promoters of vertebrate genes are generally devoid of DNA methylation and that overall there is an inverse correlation between promoter DNA methylation and gene expression [24].

Two important mechanisms for inhibition of gene expression by promoter DNA methylation are well established. First, methylcytosine residues in the recognition elements of transcription factors block their binding, resulting in reduced transcriptional activity [25,26]. A second mechanism involves recruitment of methylated DNA binding domain (MBD) proteins to methylated cytosines in promoters [27]. MBDs recruit histone-modifying complexes containing histone deacetylases (HDACs), such as the NurD complex, and histone methyltransferases (HMTases) to promoters, resulting in an inactive chromatin configuration around the genes [28]. It is also emerging that gene bodies of actively transcribed genes are more methylated than gene bodies of silent genes [24,29,30]. The regulatory role of gene-body methylation is unclear but if indeed it has a role in gene regulation, gene bodies should also be of interest for DNA-methylation-based diagnostics. Gene bodies have attracted almost no attention in the mapping of cancer methylomes and this might need to change.

DNA methylation and human disease

It is likely that all common human diseases involve changes in gene expression. Genes act to shape normal physiology through interacting networks and functional circuitries [31]. If indeed DNA methylation is involved in stable regulation of gene expression, it then makes sense that changes in DNA methylation would be detected in human disease. Aberrations in DNA methylation have been reported in schizophrenia [32-34], lupus [35-37] and type II diabetes [38-42] and have been proposed to be involved in cardiovascular disease [38,43,44].

However, changes in DNA methylation associated with human disease are just associations, and it is unclear whether these changes are causal or not. It is extremely difficult to demonstrate a causal relationship between differential DNA methylation and the pathobiology. Nevertheless, the plausibility of a causal relationship is increased in particular examples by additional lines of evidence. In systemic lupus erythematosus (lupus) patients, T-cell DNA is hypomethylated relative to

---

Box 1 Methods for measuring DNA methylation: from candidate-gene approaches to whole-genome methods

Many of the first principles of DNA methylation and its involvement in cancer were derived from the analysis of the DNA methylation state of a limited number of genes. The first method that allowed studies of DNA methylation used methylation-sensitive bacterial restriction enzymes and their methylation-insensitive isoschizomers (enzymes that cleave the same sequence). The most commonly used enzymatic pair is *MspI* and its isoschizomer *HpaII* [103]. Southern blotting and hybridization with gene-specific probes first enabled studies of the state of methylation of specific regions in the genome. Later, bisulfite treatment was used to convert unmethylated cytosine residues to uracil residues, while methylated cytosines were protected from bisulfite conversion. The bisulfite treatment therefore creates a sequence difference between methylated and unmethylated cytosines. Specific regions in the genome are then amplified using gene-specific primers and PCR and the fragments are cloned and sequenced [104,105].

Several technological advances have enabled the extension of these initial studies of several genes to genome-wide mapping. High-density oligonucleotide arrays combined with different methods for enrichment of methylated DNA have enabled studies of DNA methylation states of broad regions of the genome. Several methods for enriching methylated DNA have been developed, including immunoprecipitation with methylated cytosine antibodies [106] or capturing methylated DNA with methylated DNA binding domain proteins [83]. Illumina has introduced the bead array platforms, which have allowed interrogation of the state of methylation of thousands of CGs concurrently [107]. Current arrays can examine up to 450,000 CG sites [108]. The current Illumina 450K technology combines two methods of differentiating the methylated from unmethylated alleles, the original Infinium I assay used with the 27K arrays and Infinium II assays. While the Infinum I assay differentiates between the methylated and unmethylated alleles by differential hybridization to methylated (C) or unmethylated (U) versions of the beads followed by fluorescent single base extension, the Infinum II assay uses a common oligo on bead primer for both the methylated and unmethylated alleles followed by differential fluorescent nucleotide base extension across a methylated (C) or unmethylated (T) CG site in the sample template.

High-throughput genome sequencing of bisulfite-treated DNA has enabled the mapping of cancer methylomes genome-wide [8]. This method is still costly and unfeasible for high-throughput DNA methylation profiling. However, a limited number of studies have provided important insights into the organization of the cancer epigenome. For example, a recent study used shotgun bisulfite genome sequencing for three colorectal cancers and matched normal colonic mucosa and described large hypomethylated blocks of DNA in these cancers [59].

There are strengths and weaknesses to each of these genome-wide methods. Antibody-enrichment-based methods such as methylated DNA immunoprecipitation (MeDIP) are not biased towards CG sequences and will immunoprecipitate DNA that contains methylated CGs and methylated cytosines in other dinucleotide sequences. In addition, this method does not require bisulfite conversion of DNA and therefore avoids biases in amplification of bisulfite-converted DNA that could dramatically affect results for samples that usually contain a mixture of methylated and unmethylated sequences [109]. MeDIP methods are extremely effective when a small fraction of the population of cells is methylated, because the method focuses on the methylated DNA population and measures the change in this population. This is particularly important in cancer given that DNA methylation patterns in tumors and other tissues are heterogeneous [59]. Other methods based on bisulfite conversion measure both methylated and unmethylated DNA and lose sensitivity when the percentage of cells whose DNA is methylated in a certain region is lower than the noise-to-signal ratio of the method. For example, a change in methylation from 1% to 2% will theoretically generate a duplication of the signal size by MeDIP approaches and be easily detected, but a methylation change of 1% in a range from 0% to 100% will be within the noise range of pyrosequencing or bisulfite mapping. The main disadvantage of MeDIP is that it provides an overall average of the state of methylation but does not provide information at single base resolution.

Illumina bead arrays and high-throughput sequencing provide information at single base resolution but are potentially confounded by bias in amplification of mixtures of methylated and unmethylated bisulfite-converted DNA. The extent of the bias varies from region to region, confounding interpretation of the data. Although altering the temperature of amplification could reduce the bias in specific regions [109], this is not currently feasible in a whole-genome approach as the optimal temperature of amplification differs from region to region. Illumina bead arrays are also limited in representation of CGs and biased towards CG dinucleotides.

---

normal controls; demethylating drugs such as hydralazine and procainamide reduce T-cell DNA methylation and also induce autoreactivity in culture and lupus symptoms in humans [45], which is consistent with a causal link between demethylation and the lupus phenotype. In addition, genes that are differentially methylated in lupus are known or suspected to be involved in pathobiology of the disease from other lines of study. Genes encoding interleukin-4 and interleukin-6 are demethylated in T cells from lupus patients and are also activated and demethylated in normal T cells following treatment with the demethylating drug 5-azacytidine [46]. A recent whole-genome approach was used to map differential DNA methylation profiles in pancreatic islets from type II diabetes patients and non-diabetic donors. Differentially methylated regions were uncovered in 254 genes in diabetic islets. A fraction of these genes also showed concordant transcriptional changes, suggesting a function for these DNA methylation differences. A biocomputational analysis of the functional pathways involving these genes revealed pathways implicated in β-cell survival and function, supporting a role for these DNA methylation changes in the disease [42].

DNA methylation serves as a mechanism that provides different functions to identical sequences. A clear example is parental imprinting of genes, whereby an allele of a gene that was paternally inherited has a different DNA methylation state and expression from the allele that was maternally inherited [47-49]. Several human disease states, such as Prader-Willi syndrome (deletion or absence of a paternal contribution to chromosome 15q11-q13), Angelman syndrome (absence of a normal maternal copy of the same region) and Beckwith-Wiedemann syndrome (deregulated parental imprinting by DNA methylation of chromosome 11p15) [50] involve disruption of parental imprinting, providing perhaps the strongest evidence for a causal link between aberrations in DNA methylation and human disease. Loss of imprinting (LOI) in chromosome 11p15 is associated with several cancers, including Wilms tumor [51]. Furthermore, LOI of specific genes in this region has been associated with different cancers; LOI of *H19* (which encodes a long non-coding RNA) is associated with lung cancer [52] and hepatoblastoma [53], and LOI of *IGF2* (which encodes insulin-like growth factor 2) and *H19* is associated with cervical cancer [54].

DNA methylation and cancer

Cancer was the first group of diseases to be associated with DNA methylation and to be considered for DNA-methylation-targeted therapeutics, and it serves as a prototype for determining the role of DNA methylation and DNA-methylation-targeted therapeutics in other diseases [55]. Several types of aberration in DNA methylation and in the proteins involved in DNA methylation occur in cancer: hypermethylation of tumor suppressor genes, aberrant expression of DNMT1 and other DNMTs, and hypomethylation of unique genes and repetitive sequences [56-58]. Silencing of tumor suppressor genes by DNA methylation provides a powerful molecular mechanism by which DNA methylation can trigger cancer, and also provides a rationale for therapeutics aimed at inhibition of DNA methylation and re-expression of silenced tumor suppressor genes. Recent genome-wide studies suggest that not only high-density CG islands but also regions of lower CG density near islands, termed shores, are differentially methylated in several cancers, and that the same regions are differentially methylated between tissues, suggesting a role for these regions in defining tissue specificity of gene function [59].

DNA methylation of tumor suppressor genes has been the focus of numerous studies that have aimed to identify DNA methylation biomarkers of cancer. However, it is becoming clear that hypomethylation is equally important, because critical genes for cancer growth and metastasis are hypomethylated in cancer [60-63]. DNA demethylation has an important role in cancer by turning on the expression of pro-metastatic genes, such as the heparanase gene [60], *MMP2* (which encodes matrix metalloproteinase-2) [61] and *uPA* (which encodes urokinase plasminogen activator) [62]. We have recently delineated the DNA hypomethylation landscape of liver cancer. My colleagues and I [63] showed that there is an equal number of genes that are demethylated and hypermethylated in hepatocellular carcinoma in comparison with surrounding normal liver tissue. The hypomethylated genes are clustered in broad genomic regions, suggesting a high level of organization of demethylation in liver cancer. Functional biocomputational analysis of the hypomethylated genes suggests that they are involved in functions relating to cell growth, invasion and metastasis [63]. A causal role for demethylation in cancer metastasis is supported by the fact that treatment of non-metastatic breast cancer cells with demethylating agents increases their invasiveness [64,65], and that treatment of invasive breast cancer and liver cancer cell lines with agents that reverse demethylation results in inhibition of invasiveness and metastasis [62,63]. A recent genome-wide approach involving bisulfite mapping of several colorectal cancer samples revealed blocks of hypomethylation encompassing half the genome relative to normal colon tissue [59]. Therefore, an interesting question that has important diagnostic implications is whether the hypomethylation state of certain genes is characteristic of a more advanced and metastatic stage of breast cancer and could be of use in breast cancer staging.

DNA methylation in breast cancer

DNA methylation of candidate genes in breast cancer

The original concept driving investigation of changes in DNA methylation in diseased states was that limited sets of candidate genes were critical for disease initiation and progression. However, unbiased approaches could potentially reveal new genes and new functional gene networks that are associated with a disease, whereas candidate approaches essentially allow validation of genes that are already known to be involved. Early studies attempting to take advantage of the emerging role of methylation of promoters of tumor suppressor genes in cancer examined whether methylation of specific CGs in tumor suppressor genes correlates with different breast cancer clinical states [66]. Methylation of the p16 tumor suppressor gene was proposed to be an early biomarker for detection of breast cancer [67]. Methylation-specific PCR of six known tumor suppressor genes was used to generate a hypermethylation profile of primary breast tumors, and the methylation states of different genes were found to be significantly associated with several known prognostic factors [68]. However, our current understanding of the functional pathways of gene expression in physiological

and pathological processes suggests that it is highly unlikely that analysis of a few specific CG sites will be sufficient to stage and provide prognostic information on breast cancer with high accuracy and specificity.

DNA methylation signatures and early whole-genome approaches

One of the lessons learnt from gene expression analyses in breast cancer is that the transcription profiles that distinguish breast cancer stages involve coordinated changes in expression of many genes generating a 'signature' that characterizes a stage of disease (Table 1). Transcription signatures have been used in classifying breast cancer, and in differentiating molecular signatures in the primary tumor of breast cancers that metastasize to bone marrow or lymphatic nodes [69]. Expression signatures differentiate tumors with *BRCA1* and *BRCA2* mutations [70], supporting the idea of unique molecular signatures for subtypes of breast cancer. Therefore, it is likely that, similar to transcription profiles, DNA methylation signatures involve multiple coordinate changes in several genes and that specific patterns of DNA methylation across a broader spectrum of genes will be able to differentiate subtypes of breast cancer and their prognosis with high accuracy.

Early methods were sufficiently developed to examine a limited set of genes with bisulfite sequencing, methylation-sensitive PCR or methylation-sensitive restriction enzyme analysis. However, to delineate DNA methylation signatures without bias, whole-genome methods for mapping DNA methylation were required.

Over a decade ago, more comprehensive methods were developed to interrogate a large number of CG islands in either cell lines or tumor samples using differential methylation hybridization. This method used methylation-sensitive restriction enzymes to enrich for methylated DNA fragments, followed by hybridization to CG island arrays (containing 1,000 CG islands). By focusing on CG islands the bias for hypermethylated CG islands was preserved, and the basic assumption remained that the informative DNA methylation event in cancer is hypermethylation of CG islands. A pioneering study by the Huang group [71] used this approach to identify DNA methylation signatures by comparing 28 paired primary breast tumor and normal samples, and to determine whether patterns of specific CG hypermethylation correlate with pathological parameters in the patients analyzed. The study found that the number of CG hypermethylated islands increased with decreased differentiation of the tumors [71]. This was an early demonstration of the potential of broad DNA methylation signatures for differentiating and staging breast cancer. The main caveat of this approach is its bias towards hypermethylation of CG islands.

The use of breast cancer cell lines for identification of DNA methylation signatures specific to breast cancer stages

As an alternative to genome-wide delineation of differentially methylated genes in breast cancer, a pharmacological transcriptome-wide approach has been used to identify novel genes that are putatively hypermethylated in cancer cells. This approach is methodologically biased towards promoters of genes that are hypermethylated in cancer and ignores the hypomethylated genes that are emerging as important players in the advanced metastatic stages [60-62,72-74]. Breast cancer cell lines were treated with 5-aza-cytidine, a demethylating drug, to reveal genes whose expression was induced in response to the drug treatment using expression arrays. It was considered that this broad collection of genes whose expression was induced by a DNA demethylating drug represented a group of genes that are hypermethylated and silenced in breast cancer, and that could be useful for diagnostic and prognostic purposes clinically [75]. This approach, which makes necessary the use of breast cancer cell lines rather than primary tumors for genome-wide discovery, has several caveats. Firstly, demethylating agents could affect gene expression indirectly and by DNA-demethylation-independent mechanisms [76]. Secondly, cells in culture show DNA methylation changes that are different in many instances from the situation *in vivo* [77,78]. For example, in a study by Stefanska *et al.* [63] primary liver cells in culture showed a signature of DNA methylation that was categorically different from that of normal liver, and Fu *et al.* [79] showed distinct expression and DNA methylation patterns of Hedgehog ligands between primary colorectal tumors and colorectal cancer cell lines. Thirdly, although this approach could reveal genes that are potentially broadly methylated in breast cancer, fine demarcation and exquisite phenotyping of breast cancer cell lines is required for a cell-culture-based approach to deliver revealing DNA methylation signatures that differentiate breast cancer subtypes or signatures that have prognostic value.

The use of distinctly phenotyped but highly related breast cancer cell lines for delineating DNA methylation signatures that differentiate and classify breast cancers has recently been reported [80]. A multi-dimensional comprehensive analysis of DNA methylation was performed, using enrichment of differentially methylated regions with methylation-specific restriction sites followed by microarray hybridization and gene expression, and this was repeated more recently by including copy number analysis. The study compared two isogenic MDA-MB-231 breast cancer cell lines, MDA-MB-468GFP and MDA-MB-468GFP-LN (the latter derived from a lymphatic metastasis). This study [81] revealed broad changes in DNA methylation that included both hypomethylation and hypermethylation, and measured their association with gene expression and copy number

variation. The correspondence between some of the hypomethylation and hypermethylation events with gain and loss of copy number suggested a linkage of these two events that needs to be further explored. The changes in DNA methylation were highly organized structurally and functionally; specific networks and functional pathways were affected. These results support the hypothesis that broad signatures exquisitely define variations between closely related breast cancer cells that take a different metastatic course. However, the main caveat, again, of this approach is the use of breast cancer cell lines. It is unclear what fraction of the DNA methylation signature identified *in vitro* will be relevant in primary breast cancer tumors. If this is true *in vivo* as well, such broad signatures could be instrumental in prognosis and have a profound impact on breast cancer care.

A different study [82] examined whether broad DNA methylation, expression and copy number signatures differentiate ER+ and ER- breast cancer cell lines. The study was able to identify a cluster of differentially methylated genes that differentiate ER+ and ER- cells. The relevance of this signature to primary breast cancer was highlighted by the finding in primary tumors of 84 genes that are components of the methylation signature identified in ER+ cell lines [82]. However, the clinical advantage of these DNA methylation biomarkers over current immunochemical and histopathological methods remains to be tested in larger studies.

Differential DNA methylation of a panel of 15 CG islands was used [83] to define the cellular origin of breast cancer cells, particularly focusing on stem cells. Park *et al.* [83] first determined cancer stem cell phenotype by CD44/CD24 and ALDH1 immunohistochemistry in 36 luminal A, 33 luminal B, 30 luminal-HER2, 40 HER2-enriched, and 40 basal-like subtypes of breast cancer (Table 1) [83]. They reported that the number of CG island regions that were methylated was different between the subtypes. The basal-like subtype was enriched with the CD44+/CD24- and aldehyde dehydrogenase 1-positive (ALDH1+) putative stem cell population. For example, methylation of promoter CG islands was significantly lower in CD44+/CD24-cell+ tumors than in CD44+/CD24-cell- tumors, even within the basal-like subtype, suggesting that DNA methylation could detect the 'stemness' of breast cancers. However, it is unclear whether these differences created a true DNA methylation profile that would increase the accuracy of prognosis or identification of cell of origin beyond the classification achieved by traditional immunohistochemistry [83].

The use of current genome-wide methods to delineate stage-specific DNA methylation signatures for staging primary breast cancer

In recent years several methods have been developed to provide a genome-wide picture of the state of DNA methylation, including: next-generation genome-wide sequencing of bisulfite-converted DNA [8]; methylated DNA immunoprecipitation (MeDIP) followed by either hybridization to high-density oligonucleotide arrays [84] or next-generation sequencing [85]; and dedicated Illumina 27K and 450K arrays [86] that measure the state of methylation of well characterized CG sites distributed in the genome. Although genome-wide sequencing is still prohibitively costly for large population studies, array approaches are being frequently used to delineate DNA methylation signatures of disease states in primary clinical material rather than cell lines. Several studies have used this approach to differentiate breast cancer subtypes and their prognosis. Li *et al.* [87] used 27K arrays in a small sample of ER/PR+ and ER/PR- breast cancer samples, and identified and validated four genes whose DNA methylation was affected by ER/PR status. A similar approach was recently used to identify a group of genes that showed an association with relapse-free survival [88].

Fang *et al.* [80] also used the 27K array to delineate DNA methylation signatures that would differentiate breast cancers based on their metastatic potential. The study first discovered a 'methylator' phenotype, a co-ordinated methylation of a large group of CG islands in groups of tumors, which they termed 'breast CpG island methylator phenotype' (B-CIMP), and which resembles the previously characterized methylator phenotype in colorectal cancer [89]. The methylator phenotype was associated with low risk of breast cancer metastasis and improved rates of survival independently of other known breast cancer prognostic markers, such as ER+ status. This provides strong evidence for the potential of DNA methylation signatures to prognostically differentiate breast cancers beyond current classifications.

These results have important implications for further development of DNA methylation signature biomarkers and epigenetic cancer therapeutics and highlight the importance of genome-wide and unbiased approaches for DNA methylation mapping in breast cancer. Firstly, this study by Fang *et al*. [80] in primary tumors provides strong support to studies from my laboratory [62,90] that proposed that DNA hypomethylation is a driving force in breast cancer metastasis. It further highlights the importance of DNA hypomethylation markers in molecular diagnosis of aggressive breast cancers. The study also illustrates how DNA methylation signatures could have important therapeutic implications in guiding the use of epigenetic drugs in anticancer therapy [91]. It supports the conclusion that the use of hypomethylating drugs exclusively might exacerbate rather than cure cancer by unleashing the expression of hypermethylated pro-metastatic genes and converting nonaggressive breast cancers to highly metastatic aggressive tumors with low survival

prognosis [64,92]. These data are consistent with our data in primary liver cancer that showed extensive hypomethylation in advanced liver cancer [74]. The genes whose promoters were demethylated in liver cancer were mainly involved in cell growth, cell adhesion and communication, signal transduction, mobility and invasion, functions that are essential for cancer progression and metastasis [74].

The use of genome-wide approaches and a larger number of breast cancer samples and controls in the past 2 years has enabled further investigation of the classification and prognostic value of DNA methylation profiles in breast cancer. More interestingly, recent studies suggest that DNA methylation profiling might provide information on the cellular origin of cancer cells in a breast tumor, as well as the microenvironment, particularly the immune cell types, that are present in the tumor [93].

Related to this, a distinct profile of T cell subtype gene expression could be detected in mixed populations of tumors and stroma [4]. Kristensen *et al.* [4] used an integrated approach termed 'Pathway Recognition Algorithm using Data Integration on Genomic Models' (PARADIGM), integrating DNA methylation and microRNA profiling with mRNA expression and DNA copy number. The analysis was conducted on approximately 110 breast carcinomas and then the PARADIGM clusters derived from the discovery sample set were tested in two other breast cancer cohorts [4]. The authors identified key tumor and stromal signatures in the mixed tumor-stroma samples, suggesting that it is possible to obtain informative stromal molecular signatures without dissecting the stromal cells. This would simplify the diagnostic protocol. In addition to molecular signatures that classify breast cancer cell subtypes, they found a chronic inflammatory signature in all breast cancers. The strongest predictor of good outcome was a high T-helper 1 (Th1)/cytotoxic T-lymphocyte signature, in contrast to a Th2 signature. The PARADIGM clustering seems to expand classification beyond traditional immunohistochemistry, as a distinction was found between two clusters within luminal A breast cancer (called the PDGM3 and PDGM4 clusters) and luminal B breast cancer (PDGM4) clusters. However, although the differential DNA methylation profiles were mapped onto functional pathways that were identified by gene expression analysis, the authors did not demonstrate that adding DNA methylation profiling improved the prognostic value of the PARADIGM clusters over using mRNA expression and copy number variation [4].

Flanagan *et al.* [3] used a MeDIP approach to determine whether genome-wide DNA methylation profiles would predict tumor mutation status and intrinsic subtypes. Although DNA methylation profiles predicted tumor subtypes with some estimated error rates, they did not accurately predict the intrinsic subtypes defined by gene expression [3]. A distinct subgroup of BRCAx tumors defined by methylation profiles was identified, supporting the hypothesis that DNA methylation profiling might expand subtype classification beyond mutation analysis.

DNA methylation profiling will become important for breast cancer diagnosis and prognosis only if it provides additional classification value to other currently used methods, namely immunohistochemistry and mRNA expression analysis. A recent detailed whole-genome DNA methylation analysis by Dedeurwaerder *et al.* using the Illumina 27K arrays [93,94] suggests that DNA methylation profiling might expand current classifications of breast cancer subtypes. The analysis of 248 breast cancer tumor samples, comprising a 'main set' of 123 samples (4 normal and 119 infiltrating ductal carcinomas (IDCs)), and a 'validation set' of 125 samples (8 normal and 117 IDCs), revealed an immune 'signature' in a mixed tumor stromal population, as also reported by Kristensen *et al.* [4]. DNA methylation profiles revealed six classes, three of which defined new groups that were not classified by expression subtypes, and these might reflect different cells of origin [94]. However, the sample size of the main set was too small to allow investigation of the prognostic value of these methylation classes.

Clinical testing

DNA methylation is cell-type-specific [7]. Given that it is considered that molecular changes in cancer occur in the cancer cell, it is anticipated that changes in DNA methylation that characterize the cancer stage will be limited to the cancerous cell. This hypothesis, if true, would in practice limit DNA methylation diagnostics to biopsies rather than fluid samples. Carefully designed clinical studies will be needed to determine whether DNA methylation signatures can form the basis of more accurate and specific diagnostic and prognostic tests than currently available histopathological tools and immuno-chemical tests. Another important area in which DNA methylation signatures in biopsies might be of value is in stratifying patients for therapy and predicting therapeutic outcomes.

Blood-based tests

If indeed DNA methylation signatures are informative only in tumor sample biopsies, this limits the utility of such markers for use in routine follow-up and population-wide early screening. Biopsies are invasive, so it is highly unlikely that they will become part of a routine screening procedure. Moreover, even in breast cancer patients, biopsies are not applicable for routine follow-up following surgery and particularly when there is no visible tumor growth. Noninvasive methods are essential for early prediction and follow-up of therapeutic response following surgery.

Nevertheless, it is possible that circulating tumor cells display DNA methylation signatures that are reflective of the state of methylation in the tumor mass. Informative breast cancer DNA methylation signatures in tumor cells found in blood would be extremely important in early screening, diagnosis, staging and follow-up of treatment, and there is increased interest in their potential clinical importance. An important area of research is DNA methylation mapping of circulating tumor cells to identify DNA methylation signatures of breast cancer in these circulating cells. The initial focus has been on hypermethylated genes that are characteristic of many cancers.

Jing *et al.* [95] recently tested a CIMP in serum from 50 sporadic breast cancer (SBC) patients and paired controls, by examining the state of methylation of CG sites in 10 genes known to be methylated in cancer using methylation-specific PCR. CIMP was found to be more prevalent in serum from SBC patients than controls, the methylation rate was 92% (46/50) at least in one gene in SBC, and serum from only four patients showed no methylation of any of the ten genes. This study demonstrated that it is possible to identify changes in DNA methylation in serum from breast cancer patients. It also showed that a combination of methylated genes provided high specificity and sensitivity markers for breast cancer as well as prognostic value, as CIMP+ status in serum was associated with a relative risk of recurrence of 8.6.

In another study, Radpour *et al.* [96] focused on ten candidate genes and investigated two cohorts: a first cohort with 36 plasma samples from breast cancer patients and 30 plasma samples from healthy controls, and a second cohort of 60 triple matched samples (cancerous tissue, and matched normal tissue and serum samples) from 20 patients with non-familial breast cancer. Seven of the genes showed concordant methylation in serum and tumor tissue from the same patient. This supports the hypothesis that serum DNA is derived from and accurately reflects the DNA methylation profile of the primary tumor. A panel of eight genes out of the ten studied was proposed as a highly specific and sensitive test for breast cancer [96]. Furthermore, methylation of particular genes was associated with particular clinical parameters. It is unclear whether this specific set of DNA methylation biomarkers provides any early predictive or prognostic value, and further and more extensive studies are required. Nevertheless, these recent studies suggest that there is potential for using blood samples to detect DNA methylation markers in breast cancer.

Conclusions and future directions
DNA methylation states are involved in long-term gene expression programming of cell-type identity and are therefore exquisite descriptors of the functional state of a cell. The differences in DNA methylation between cell types involve multiple genes that are components of functional gene networks. Therefore, it is reasonable to consider that cancer cells might have a unique profile of DNA methylation that reflects not only their identity as tumor cells, but also will differentiate between tumor stages and predict clinical outcomes and response to therapy. Early attempts to discover breast cancer DNA methylation markers focused on a shortlist of candidate genes and were highly biased towards DNA hypermethylation of CG islands in tumor suppressor genes. The advent of genome-wide methods for DNA methylation mapping should allow us to delineate comprehensive and unbiased high-definition DNA methylation signatures that could provide accurate classification of breast cancers. Such maps might be used in prognosis, prediction of therapeutic outcomes and stratification for different treatment strategies. Several studies have supported the prospect that DNA methylation signatures can be effective diagnostic markers. However, the data so far are very limited and the predictive value of the small number of DNA methylation signatures that have been identified is unclear. Several studies were limited to different breast cancer cell line manipulations and only few studies have looked at a sizeable number of genes in primary tumors. The critical challenge is to derive high-quality DNA methylation signatures that are confirmed in prospective studies as specific and sensitive predictors of clinical outcome and therapeutic responses. An additional question is to determine whether DNA methylation signatures would provide advantages over current histopathological and immunochemical methods.

It will be particularly important to develop noninvasive molecular markers for breast cancer. Preliminary studies suggest that circulating tumor DNA in plasma samples bear tumor-specific DNA methylation markers that provide potential molecular markers for breast cancer. It is so far unknown whether blood-based DNA methylation markers have prognostic or early predictive value, specifically in follow-up of response to therapy. Furthermore, the DNA methylation signatures of the earliest transition into a transformed state are unknown; if such DNA methylation signatures exist, they could provide early molecular markers of breast cancer, especially if found in serum DNA.

RNA transcription profiles have been used in breast cancer molecular diagnosis [70,97-102]. DNA methylation markers potentially have several advantages over transcription profiles as diagnostic tools in breast cancer. Firstly, DNA is a robust clinical material and could be preserved under harsh conditions, including incubation in serum, whereas RNA is a highly labile material. Secondly, DNA methylation profiles represent a stable long-term programming of the genome, whereas

transcription assays provide a snapshot of the transcription activity at a specific time point and in response in part to transient signals. It is therefore anticipated that the noise-to-signal ratio should be significantly lower for DNA methylation signatures, which constitute a stable definition of the molecular state of a cell. The limited data that are available, the advent of genome-wide methods for DNA methylation mapping and the emerging understanding of the cardinal role of DNA methylation in controlling cell-type-specific genome function support continuing studies in this emerging area and provide reasons for optimism that DNA methylation markers could serve as exquisite molecular markers for prediction, prognosis and follow-up of breast cancer therapy.

Abbreviations

ALDH, aldehyde dehydrogenase; BRCA, breast cancer gene; CIMP, CpG island methylator phenotype; DNMT, DNA methyl transferase; ER, estrogen receptor; EZH2, Enhancer of Zeste 2; HER2, human epidermal growth factor receptor 2; IDC, infiltrating ductal carcinomas; MeDIP, methylated DNA immunoprecipitation; MBD, methylated DNA binding domain; PARADIGM, Pathway Recognition Algorithm using Data Integration on Genomic Models; PR, progesterone receptor; PRC1, Polycomb group repressive complex; SBC, sporadic breast cancer.

References

1. Perou CM, Børresen-Dale AL: **Systems biology and genomics of breast cancer.** *Cold Spring Harb Perspect Biol* 2011, **3:**a003293.
2. Costanzo M, Baryshnikova A, Myers CL, Andrews B, Boone C: **Charting the genetic interaction map of a cell.** *Curr Opin Biotechnol* 2011, **22:**66-74.
3. Flanagan JM, Cocciardi S, Waddell N, Johnstone CN, Marsh A, Henderson S, Simpson P, da Silva L, Khanna K, Lakhani S, Boshoff C, Chenevix-Trench G: **DNA methylome of familial breast cancer identifies distinct profiles defined by mutation status.** *Am J Hum Genet* 2010, **86:**420-433.
4. Kristensen VN, Vaske CJ, Ursini-Siegel J, Van Loo P, Nordgard SH, Sachidanandam R, Sørlie T, Wärnberg F, Haakensen VD, Helland A, Naume B, Perou CM, Haussler D, Troyanskaya OG, Børresen-Dale AL: **Integrated molecular profiles of invasive breast tumors and ductal carcinoma in situ (DCIS) reveal differential vascular and interleukin signaling.** *Proc Natl Acad Sci USA* 2012, **109:**2802-2807.
5. Dedeurwaerder S, Desmedt C, Calonne E: **Largest ever DNA methylation dataset for breast cancer completed.** *Expert Rev Mol Diagn* 2011, **11:**470.
6. Razin A, Riggs AD: **DNA methylation and gene function.** *Science* 1980, **210:**604-610.
7. Razin A, Szyf M: **DNA methylation patterns. Formation and function.** *Biochim Biophys Acta* 1984, **782:**331-342.
8. Lister R, Pelizzola M, Dowen RH, Hawkins RD, Hon G, Tonti-Filippini J, Nery JR, Lee L, Ye Z, Ngo QM, Edsall L, Antosiewicz-Bourget J, Stewart R, Ruotti V, Millar AH, Thomson JA, Ren B, Ecker JR: **Human DNA methylomes at base resolution show widespread epigenomic differences.** *Nature* 2009, **462:**315-322.
9. Adams RI, Turnbull J, Smillie Ej, Burdon Rh: **DNA methylation in nuclei and studies using a purified DNA methylase from ascites cells.** In *Post-synthetic Modification of Macromolecules.* Edited by Antoni F, Farago A. Amsterdam: North-Holland; 1975: 39-48.
10. Jeltsch A: **Molecular enzymology of mammalian DNA methyltransferases.** *Curr Top Microbiol Immunol* 2006, **301:**203-225.
11. Gruenbaum Y, Cedar H, Razin A: **Substrate and sequence specificity of a eukaryotic DNA methylase.** *Nature* 1982, **295:**620-622.
12. Stein R, Gruenbaum Y, Pollack Y, Razin A, Cedar H: **Clonal inheritance of the pattern of DNA methylation in mouse cells.** *Proc Natl Acad Sci U S A* 1982, **79:**61-65.
13. Okano M, Xie S, Li E: **Cloning and characterization of a family of novel mammalian DNA (cytosine-5) methyltransferases.** *Nat Genet* 1998, **19:**219-220.
14. Fuks F, Burgers WA, Brehm A, Hughes-Davies L, Kouzarides T: **DNA methyltransferase Dnmt1 associates with histone deacetylase activity.** *Nat Genet* 2000, **24:**88-91.
15. Rountree MR, Bachman KE, Baylin SB: **DNMT1 binds HDAC2 and a new co-repressor, DMAP1, to form a complex at replication foci.** *Nat Genet* 2000, **25:**269-277.
16. Fuks F, Hurd PJ, Wolf D, Nan X, Bird AP, Kouzarides T: **The methyl-CpG-binding protein MeCP2 links DNA methylation to histone methylation.** *J Biol Chem* 2003, **278:**4035-4040.
17. Viré E, Brenner C, Deplus R, Blanchon L, Fraga M, Didelot C, Morey L, Van Eynde A, Bernard D, Vanderwinden JM, Bollen M, Esteller M, Di Croce L, de Launoit Y, Fuks F: **The Polycomb group protein EZH2 directly controls DNA methylation.** *Nature* 2006, **439:**871-874.
18. Bostick M, Kim JK, Estève PO, Clark A, Pradhan S, Jacobsen SE: **UHRF1 plays a role in maintaining DNA methylation in mammalian cells.** *Science* 2007, **317:**1760-1764.
19. Sharif J, Muto M, Takebayashi S, Suetake I, Iwamatsu A, Endo TA, Shinga J, Mizutani-Koseki Y, Toyoda T, Okamura K, Tajima S, Mitsuya K, Okano M, Koseki H: **The SRA protein Np95 mediates epigenetic inheritance by recruiting Dnmt1 to methylated DNA.** *Nature* 2007, **450:**908-912.
20. Szyf M, Tanigawa G, McCarthy PL: **A DNA signal from the Thy-1 gene defines de novo methylation patterns in embryonic stem cells.** *Mol Cell Biol* 1990, **10:**4396-4400.
21. Brandeis M, Frank D, Keshet I, Siegfried Z, Mendelsohn M, Nemes A, Temper V, Razin A, Cedar H: **Sp1 elements protect a CpG island from de novo methylation.** *Nature* 1994, **371:**435-438.
22. Stadler MB, Murr R, Burger L, Ivanek R, Lienert F, Schöler A, Wirbelauer C, Oakeley EJ, Gaidatzis D, Tiwari VK, Schübeler D: **DNA-binding factors shape the mouse methylome at distal regulatory regions.** *Nature* 2011, **480:**490-495.
23. Razin A, Cedar H: **Distribution of 5-methylcytosine in chromatin.** *Proc Natl Acad Sci U S A* 1977, **74:**2725-2728.
24. Rauch TA, Wu X, Zhong X, Riggs AD, Pfeifer GP: **A human B cell methylome at 100-base pair resolution.** *Proc Natl Acad Sci U S A* 2009, **106:**671-678.
25. Comb M, Goodman HM: **CpG methylation inhibits proenkephalin gene expression and binding of the transcription factor AP-2.** *Nucleic Acids Res* 1990, **18:**3975-3982.
26. Inamdar NM, Ehrlich KC, Ehrlich M: **CpG methylation inhibits binding of several sequence-specific DNA-binding proteins from pea, wheat, soybean and cauliflower.** *Plant Mol Biol* 1991, **17:**111-123.
27. Nan X, Campoy FJ, Bird A: **MeCP2 is a transcriptional repressor with abundant binding sites in genomic chromatin.** *Cell* 1997, **88:**471-481.
28. Zhang Y, Ng HH, Erdjument-Bromage H, Tempst P, Bird A, Reinberg D: **Analysis of the NuRD subunits reveals a histone deacetylase core complex and a connection with DNA methylation.** *Genes Dev* 1999, **13:**1924-1935.
29. Hellman A, Chess A: **Gene body-specific methylation on the active X chromosome.** *Science* 2007, **315:**1141-1143.
30. Suzuki M, Oda M, Ramos MP, Pascual M, Lau K, Stasiek E, Agyiri F, Thompson RF, Glass JL, Jing Q, Sandstrom R, Fazzari MJ, Hansen RS, Stamatoyannopoulos JA, McLellan AS, Greally JM: **Late-replicating heterochromatin is characterized by decreased cytosine methylation in the human genome.** *Genome Res* 2011, **21:**1833-1840.
31. Siegal ML, Promislow DE, Bergman A: **Functional and evolutionary inference in gene networks: does topology matter?** *Genetica* 2007, **129:**83-103.
32. Grayson DR, Jia X, Chen Y, Sharma RP, Mitchell CP, Guidotti A, Costa E: **Reelin promoter hypermethylation in schizophrenia.** *Proc Natl Acad Sci U S A* 2005, **102:**9341-9346.
33. Petronis A, Gottesman II, Kan P, Kennedy JL, Basile VS, Paterson AD, Popendikyte V: **Monozygotic twins exhibit numerous epigenetic differences: clues to twin discordance?** *Schizophr Bull* 2003, **29:**169-178.
34. Veldic M, Guidotti A, Maloku E, Davis JM, Costa E: **In psychosis, cortical**

interneurons overexpress DNA-methyltransferase 1. *Proc Natl Acad Sci U S A* 2005, **102**:2152-2157.

35. Balada E, Ordi-Ros J, Serrano-Acedo S, Martinez-Lostao L, Rosa-Leyva M, Vilardell-Tarrés M: **Transcript levels of DNA methyltransferases DNMT1, DNMT3A and DNMT3B in CD4+ T cells from patients with systemic lupus erythematosus.** *Immunology* 2008, **124**:339-347.

36. Balada E, Ordi-Ros J, Vilardell-Tarrés M: **DNA methylation and systemic lupus erythematosus.** *Ann N Y Acad Sci* 2007, **1108**:127-136.

37. Yung RL, Richardson BC: **Role of T cell DNA methylation in lupus syndromes.** *Lupus* 1994, **3**:487-491.

38. Gluckman PD, Hanson MA, Bateson P, Beedle AS, Law CM, Bhutta ZA, Anokhin KV, Bougnères P, Chandak GR, Dasgupta P, Smith GD, Ellison PT, Forrester TE, Gilbert SF, Jablonka E, Kaplan H, Prentice AM, Simpson SJ, Uauy R, West-Eberhard MJ: **Towards a new developmental synthesis: adaptive developmental plasticity and human disease.** *Lancet* 2009, **373**:1654-1657.

39. Stenvinkel P, Karimi M, Johansson S, Axelsson J, Suliman M, Lindholm B, Heimbürger O, Barany P, Alvestrand A, Nordfors L, Qureshi AR, Ekström TJ, Schalling M: **Impact of inflammation on epigenetic DNA methylation - a novel risk factor for cardiovascular disease?** *J Intern Med* 2007, **261**:488-499.

40. Ling C, Del Guerra S, Lupi R, Rönn T, Granhall C, Luthman H, Masiello P, Marchetti P, Groop L, Del Prato S: **Epigenetic regulation of PPARGC1A in human type 2 diabetic islets and effect on insulin secretion.** *Diabetologia* 2008, **51**:615-622.

41. Maier S, Olek A: **Diabetes: a candidate disease for efficient DNA methylation profiling.** *J Nutr* 2002, **132**:2440S-2443S.

42. Volkmar M, Dedeurwaerder S, Cunha DA, Ndlovu MN, Defrance M, Deplus R, Calonne E, Volkmar U, Igoillo-Esteve M, Naamane N, Del Guerra S, Masini M, Bugliani M, Marchetti P, Cnop M, Eizirik DL, Fuks F: **DNA methylation profiling identifies epigenetic dysregulation in pancreatic islets from type 2 diabetic patients.** *EMBO J* 2012, **31**:1405-1426. doi: 10.1038/emboj.2011.503.

43. Corwin EJ: **The concept of epigenetics and its role in the development of cardiovascular disease: commentary on "new and emerging theories of cardiovascular disease".** *Biol Res Nurs* 2004, **6**:11-6; discussion 21-3.

44. Hanson MA, Gluckman PD: **Developmental origins of health and disease: new insights.** *Basic Clin Pharmacol Toxicol* 2008, **102**:90-93.

45. Cornacchia E, Golbus J, Maybaum J, Strahler J, Hanash S, Richardson B: **Hydralazine and procainamide inhibit T cell DNA methylation and induce autoreactivity.** *J Immunol* 1988, **140**:2197-2200.

46. Mi XB, Zeng FQ: **Hypomethylation of interleukin-4 and -6 promoters in T cells from systemic lupus erythematosus patients.** *Acta Pharmacol Sin* 2008, **29**:105-112.

47. Sapienza C: **Parental imprinting of genes.** *Sci Am* 1990, **263**:52-60.

48. Polychronakos C: **Parental imprinting of the genes for IGF-II and its receptor.** *Adv Exp Med Biol* 1993, **343**:189-203.

49. Tilghman SM, Bartolomei MS, Webber AL, Brunkow ME, Saam J, Leighton PA, Pfeifer K, Zemel S: **Parental imprinting of the H19 and Igf2 genes in the mouse.** *Cold Spring Harb Symp Quant Biol* 1993, **58**:287-295.

50. Mannens M, Hoovers JM, Redeker E, Verjaal M, Feinberg AP, Little P, Boavida M, Coad N, Steenman M, Bliek J: **Parental imprinting of human chromosome region 11p15.3-pter in the Beckwith-Wiedemann syndrome and various human neoplasia.** *Eur J Hum Genet* 1994, **2**:3-23.

51. Lalande M: **Parental imprinting and human disease.** *Annu Rev Genet* 1996, **30**:173-195.

52. Kondo M, Suzuki H, Ueda R, Osada H, Takagi K, Takahashi T, Takahashi T: **Frequent loss of imprinting of the H19 gene is often associated with its overexpression in human lung cancers.** *Oncogene* 1995, **10**:1193-1198.

53. Fukuzawa R, Umezawa A, Ochi K, Urano F, Ikeda H, Hata J: **High frequency of inactivation of the imprinted H19 gene in "sporadic" hepatoblastoma.** *Int J Cancer* 1999, **82**:490-497.

54. Douc-Rasy S, Barrois M, Fogel S, Ahomadegbe JC, Stéhelin D, Coll J, Riou G: **High incidence of loss of heterozygosity and abnormal imprinting of H19 and IGF2 genes in invasive cervical carcinomas. Uncoupling of H19 and IGF2 expression and biallelic hypomethylation of H19.** *Oncogene* 1996, **12**:423-430.

55. Szyf M: **DNA methylation properties: consequences for pharmacology.** *Trends Pharmacol Sci* 1994, **15**:233-238.

56. Baylin SB, Esteller M, Rountree MR, Bachman KE, Schuebel K, Herman JG: **Aberrant patterns of DNA methylation, chromatin formation and gene expression in cancer.** *Hum Mol Genet* 2001, **10**:687-692.

57. Issa JP, Vertino PM, Wu J, Sazawal S, Celano P, Nelkin BD, Hamilton SR, Baylin SB: **Increased cytosine DNA-methyltransferase activity during colon

cancer progression.** *J Natl Cancer Inst* 1993, **85**:1235-1240.

58. Ehrlich M: **DNA methylation in cancer: too much, but also too little.** *Oncogene* 2002, **21**:5400-5413.

59. Hansen KD, Timp W, Bravo HC, Sabunciyan S, Langmead B, McDonald OG, Wen B, Wu H, Liu Y, Diep D, Briem E, Zhang K, Irizarry RA, Feinberg AP: **Increased methylation variation in epigenetic domains across cancer types.** *Nat Genet* 2011, **43**:768-775.

60. Shteper PJ, Zcharia E, Ashhab Y, Peretz T, Vlodavsky I, Ben-Yehuda D: **Role of promoter methylation in regulation of the mammalian heparanase gene.** *Oncogene* 2003, **22**:7737-7749.

61. Shukeir N, Pakneshan P, Chen G, Szyf M, Rabbani SA: **Alteration of the methylation status of tumor-promoting genes decreases prostate cancer cell invasiveness and tumorigenesis in vitro and in vivo.** *Cancer Res* 2006, **66**:9202-9210.

62. Pakneshan P, Szyf M, Farias-Eisner R, Rabbani SA: **Reversal of the hypomethylation status of urokinase (uPA) promoter blocks breast cancer growth and metastasis.** *J Biol Chem* 2004, **279**:31735-31744.

63. Stefanska B, Huang J, Bhattacharyya B, Suderman M, Hallett M, Han ZG, Szyf M: **Definition of the landscape of promoter DNA hypomethylation in liver cancer.** *Cancer Res* 2011, **71**:5891-5903.

64. Ateeq B, Unterberger A, Szyf M, Rabbani SA: **Pharmacological inhibition of DNA methylation induces proinvasive and prometastatic genes in vitro and in vivo.** *Neoplasia* 2008, **10**:266-278.

65. Chik F, Szyf M: **Effects of specific DNMT gene depletion on cancer cell transformation and breast cancer cell invasion; toward selective DNMT inhibitors.** *Carcinogenesis* 2011, **32**:224-232.

66. Dickinson RE, Dalloc A, Bieche I, Krex D, Morton D, Maher ER, Latif F: **Epigenetic inactivation of SLIT3 and SLIT1 genes in human cancers.** *Br J Cancer* 2004, **91**:2071-2078.

67. Berman H, Zhang J, Crawford YG, Gauthier ML, Fordyce CA, McDermott KM, Sigaroudinia M, Kozakiewicz K, Tlsty TD: **Genetic and epigenetic changes in mammary epithelial cells identify a subpopulation of cells involved in early carcinogenesis.** *Cold Spring Harb Symp Quant Biol* 2005, **70**:317-327.

68. Shinozaki M, Hoon DS, Giuliano AE, Hansen NM, Wang HJ, Turner R, Taback B: **Distinct hypermethylation profile of primary breast cancer is associated with sentinel lymph node metastasis.** *Clin Cancer Res* 2005, **11**:2156-2162.

69. Woelfle U, Cloos J, Sauter G, Riethdorf L, Jänicke F, van Diest P, Brakenhoff R, Pantel K: **Molecular signature associated with bone marrow micrometastasis in human breast cancer.** *Cancer Res* 2003, **63**:5679-5684.

70. Hedenfalk I, Duggan D, Chen Y, Radmacher M, Bittner M, Simon R, Meltzer P, Gusterson B, Esteller M, Kallioniemi OP, Wilfond B, Borg A, Trent J, Raffeld M, Yakhini Z, Ben-Dor A, Dougherty E, Kononen J, Bubendorf L, Fehrle W, Pittaluga S, Gruvberger S, Loman N, Johannsson O, Olsson H, Sauter G: **Gene-expression profiles in hereditary breast cancer.** *N Engl J Med* 2001, **344**:539-548.

71. Yan PS, Perry MR, Laux DE, Asare AL, Caldwell CW, Huang TH: **CpG island arrays: an application toward deciphering epigenetic signatures of breast cancer.** *Clin Cancer Res* 2000, **6**:1432-1438.

72. Pakneshan P, Szyf M, Rabbani SA: **Hypomethylation of urokinase (uPA) promoter in breast and prostate cancer: prognostic and therapeutic implications.** *Curr Cancer Drug Targets* 2005, **5**:471-488.

73. Szyf M, Pakneshan P, Rabbani SA: **DNA methylation and breast cancer.** *Biochem Pharmacol* 2004, **68**:1187-1197.

74. Stefanska B, Huang J, Bhattacharyya B, Suderman M, Hallett M, Han ZG, Szyf M: **Definition of the landscape of promoter DNA hypomethylation in liver cancer.** *Cancer Res* 2011, **71**:5891-5903.

75. Shi H, Yan PS, Chen CM, Rahmatpanah F, Lofton-Day C, Caldwell CW, Huang TH: **Expressed CpG island sequence tag microarray for dual screening of DNA hypermethylation and gene silencing in cancer cells.** *Cancer Res* 2002, **62**:3214-3220.

76. Jiemjit A, Fandy TE, Carraway H, Bailey KA, Baylin S, Herman JG, Gore SD: **p21(WAF1/CIP1) induction by 5-azacytosine nucleosides requires DNA damage.** *Oncogene* 2008, **27**:3615-3623.

77. Bork S, Pfister S, Witt H, Horn P, Korn B, Ho AD, Wagner W: **DNA methylation pattern changes upon long-term culture and aging of human mesenchymal stromal cells.** *Aging Cell* 2010, **9**:54-63.

78. Grafodatskaya D, Choufani S, Ferreira JC, Butcher DT, Lou Y, Zhao C, Scherer SW, Weksberg R: **EBV transformation and cell culturing destabilizes DNA methylation in human lymphoblastoid cell lines.** *Genomics* 2010, **95**:73-83.

79. Fu X, Deng H, Zhao L, Li J, Zhou Y, Zhang Y: **Distinct expression patterns of hedgehog ligands between cultured and primary colorectal cancers are

associated with aberrant methylation of their promoters. *Mol Cell Biochem* 2010, **337**:185-192.

80. Fang F, Turcan S, Rimner A, Kaufman A, Giri D, Morris LG, Shen R, Seshan V, Mo Q, Heguy A, Baylin SB, Ahuja N, Viale A, Massague J, Norton L, Vahdat LT, Moynahan ME, Chan TA: **Breast cancer methylomes establish an epigenomic foundation for metastasis.** *Sci Transl Med* 2011, **3**:75ra25.

81. Andrews J, Kennette W, Pilon J, Hodgson A, Tuck AB, Chambers AF, Rodenhiser DI: **Multi-platform whole-genome microarray analyses refine the epigenetic signature of breast cancer metastasis with gene expression and copy number.** *PLoS One* 2010, **5**:e8665.

82. Sun Z, Asmann YW, Kalari KR, Bot B, Eckel-Passow JE, Baker TR, Carr JM, Khrebtukova I, Luo S, Zhang L, Schroth GP, Perez EA, Thompson EA: **Integrated analysis of gene expression, CpG island methylation, and gene copy number in breast cancer cells by deep sequencing.** *PLoS One* 2011, **6**:e17490.

83. Park SY, Kwon HJ, Choi Y, Lee HE, Kim SW, Kim JH, Kim IA, Jung N, Cho NY, Kang GH: **Distinct patterns of promoter CpG island methylation of breast cancer subtypes are associated with stem cell phenotypes.** *Mod Pathol* 2012, **25**:185-196.

84. Keshet I, Schlesinger Y, Farkash S, Rand E, Hecht M, Segal E, Pikarski E, Young RA, Niveleau A, Cedar H, Simon I: **Evidence for an instructive mechanism of de novo methylation in cancer cells.** *Nat Genet* 2006, **38**:149-153.

85. Ruike Y, Imanaka Y, Sato F, Shimizu K, Tsujimoto G: **Genome-wide analysis of aberrant methylation in human breast cancer cells using methyl-DNA immunoprecipitation combined with high-throughput sequencing.** *BMC Genomics* 2010, **11**:137.

86. Bibikova M, Barnes B, Tsan C, Ho V, Klotzle B, Le JM, Delano D, Zhang L, Schroth GP, Gunderson KL, Fan JB, Shen R: **High density DNA methylation array with single CpG site resolution.** *Genomics* 2011, **98**:288-295.

87. Li L, Lee KM, Han W, Choi JY, Lee JY, Kang GH, Park SK, Noh DY, Yoo KY, Kang D: **Estrogen and progesterone receptor status affect genome-wide DNA methylation profile in breast cancer.** *Hum Mol Genet* 2010, **19**:4273-4277.

88. Hill VK, Ricketts C, Bieche I, Vacher S, Gentle D, Lewis C, Maher ER, Latif F: **Genome-wide DNA methylation profiling of CpG islands in breast cancer identifies novel genes associated with tumorigenicity.** *Cancer Res* 2011, **71**:2988-2999.

89. Toyota M, Ahuja N, Ohe-Toyota M, Herman JG, Baylin SB, Issa JP: **CpG island methylator phenotype in colorectal cancer.** *Proc Natl Acad Sci U S A* 1999, **96**:8681-8686.

90. Campbell PM, Bovenzi V, Szyf M: **Methylated DNA-binding protein 2 antisense inhibitors suppress tumourigenesis of human cancer cell lines in vitro and in vivo.** *Carcinogenesis* 2004, **25**:499-507.

91. Szyf M: **Epigenetics, DNA methylation, and chromatin modifying drugs.** *Annu Rev Pharmacol Toxicol* 2009, **49**:243-263.

92. Chik F, Szyf M: **Effects of specific DNMT gene depletion on cancer cell transformation and breast cancer cell invasion; toward selective DNMT inhibitors.** *Carcinogenesis* 2011, **32**:224-232.

93. Dedeurwaerder S, Desmedt C, Calonne E, Singhal SK, Haibe-Kains B, Defrance M, Michiels S, Volkmar M, Deplus R, Luciani J, Lallemand F, Larsimont D, Toussaint J, Haussy S, Rothé F, Rouas G, Metzger O, Majjaj S, Saini K, Putmans P, Hames G, van Baren N, Coulie PG, Piccart M, Sotiriou C, Fuks F: **DNA methylation profiling reveals a predominant immune component in breast cancers.** *EMBO Mol Med* 2011, **3**:726-741.

94. Dedeurwaerder S, Fumagalli D, Fuks F: **Unravelling the epigenomic dimension of breast cancers.** *Curr Opin Oncol* 2011, **23**:559-565.

95. Jing F, Yuping W, Yong C, Jie L, Jun L, Xuanbing T, Lihua H: **CpG island methylator phenotype of multigene in serum of sporadic breast carcinoma.** *Tumour Biol* 2010, **31**:321-331.

96. Radpour R, Barekati Z, Kohler C, Lv Q, Bürki N, Diesch C, Bitzer J, Zheng H, Schmid S, Zhong XY: **Hypermethylation of tumor suppressor genes involved in critical regulatory pathways for developing a blood-based test in breast cancer.** *PLoS One* 2011, **6**:e16080.

97. Prat A, Perou CM: **Deconstructing the molecular portraits of breast cancer.** *Mol Oncol* 2011, **5**:5-23.

98. Finak G, Bertos N, Pepin F, Sadekova S, Souleimanova M, Zhao H, Chen H, Omeroglu G, Meterissian S, Omeroglu A, Hallett M, Park M: **Stromal gene expression predicts clinical outcome in breast cancer.** *Nat Med* 2008, **14**:518-527.

99. Hu Z, Fan C, Oh DS, Marron JS, He X, Qaqish BF, Livasy C, Carey LA, Reynolds E, Dressler L, Nobel A, Parker J, Ewend MG, Sawyer LR, Wu J, Liu Y, Nanda R, Tretiakova M, Ruiz Orrico A, Dreher D, Palazzo JP, Perreard L, Nelson E, Mone M, Hansen H, Mullins M, Quackenbush JF, Ellis MJ, Olopade OI, Bernard PS, *et al.*: **The molecular portraits of breast tumors are conserved across microarray platforms.** *BMC Genomics* 2006, **7**:96.

100. Weigelt B, Hu Z, He X, Livasy C, Carey LA, Ewend MG, Glas AM, Perou CM, Van't Veer LJ: **Molecular portraits and 70-gene prognosis signature are preserved throughout the metastatic process of breast cancer.** *Cancer Res* 2005, **65**:9155-9158.

101. Sørlie T, Perou CM, Tibshirani R, Aas T, Geisler S, Johnsen H, Hastie T, Eisen MB, van de Rijn M, Jeffrey SS, Thorsen T, Quist H, Matese JC, Brown PO, Botstein D, Eystein Lønning P, Børresen-Dale AL: **Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications.** *Proc Natl Acad Sci U S A* 2001, **98**:10869-10874.

102. Perou CM, Sørlie T, Eisen MB, van de Rijn M, Jeffrey SS, Rees CA, Pollack JR, Ross DT, Johnsen H, Akslen LA, Fluge O, Pergamenschikov A, Williams C, Zhu SX, Lønning PE, Børresen-Dale AL, Brown PO, Botstein D: **Molecular portraits of human breast tumours.** *Nature* 2000, **406**:747-752.

103. Cedar H, Solage A, Glaser G, Razin A: **Direct detection of methylated cytosine in DNA by use of the restriction enzyme MspI.** *Nucleic Acids Res* 1979, **6**:2125-2132.

104. Frommer M, McDonald LE, Millar DS, Collis CM, Watt F, Grigg GW, Molloy PL, Paul CL: **A genomic sequencing protocol that yields a positive display of 5-methylcytosine residues in individual DNA strands.** *Proc Natl Acad Sci U S A* 1992, **89**:1827-1831.

105. Clark SJ, Harrison J, Paul CL, Frommer M: **High sensitivity mapping of methylated cytosines.** *Nucleic Acids Res* 1994, **22**:2990-2997.

106. Brinkman AB, Simmer F, Ma K, Kaan A, Zhu J, Stunnenberg HG: **Whole-genome DNA methylation profiling using MethylCap-seq.** *Methods* 2010, **52**:232-236.

107. Bibikova M, Lin Z, Zhou L, Chudin E, Garcia EW, Wu B, Doucet D, Thomas NJ, Wang Y, Vollmer E, Goldmann T, Seifart C, Jiang W, Barker DL, Chee MS, Floros J, Fan JB: **High-throughput DNA methylation profiling using universal bead arrays.** *Genome Res* 2006, **16**:383-393.

108. Dedeurwaerder S, Defrance M, Calonne E, Denis H, Sotiriou C, Fuks F: **Evaluation of the Infinium Methylation 450K technology.** *Epigenomics* 2011, **3**:771-784.

109. Shen L, Guo Y, Chen X, Ahmed S, Issa JP: **Optimizing annealing temperature overcomes bias in bisulfite PCR methylation analysis.** *Biotechniques* 2007, **42**:48, 50, 52 passim.