Methylation status of potential genes in breast cancer patients and correlate them with gene expression

Naiyar Azmi 1 and Ayesha shahid 2,*

1 Vedica College of B. Pharmacy, Bhopal, M. P, India.
2 Department of Pharmaceutical Sciences and Technology, BIT, Mesra, India.

World Journal of Advanced Research and Reviews, 2023, 17(01), 815–824

Abstract
Breast cancer is one of the world’s most prevalent cancer among women. Globally approximately 2.3 million women are diagnosed with breast cancer in 2020. It arises due to epigenetic modification which outlooks aberrant methylation as its major cause. The methylation of DNA involves a covalent chemical modification which is one of the major drawback, promoting development of breast cancer. In addition, evidence and investigation suggest that these methylations at promoter region of CpG islands leads to silencing of gene associated with tumor suppressor which is involved in crucial gene expression leading to establishment of breast cancer. Among variety of hyper methylated gene involved in different function linked with breast cancer, the hyper methylation of BRCA1 gene is identified as the targeted gene in precipitating disorder. A lot of susceptible genes are identified and based on their penetrance ability are being classified.

In order to analyze DNA methylation status Bisulfite sequencing is employed which is based on the mechanism of converting every un-methylated cytosine into uracil. With the advancement in gene sequencing the concept of whole genome bisulfite sequencing is also on the way to determine the DNA methylation occurring at single base. Another robust and sensitive approach to determine the methylated status is RT-PCR, a diagnostic tool applied in order to detect the expression of single gene.

Keywords: Breast cancer; DNA methylation; Genes; CpG islands

1. Introduction
Initially the development of breast begins in fetal or embryonic life [1]. The epidermal cells that originated from ectoderm are found to project onto the mesenchyme thereby forming mammary ridges and ultimately forms lactiferous ducts. The breast is made of two types of tissues i.e., glandular tissues as well as stromal (supporting) tissue respectively[2]. The glandular tissues are the milk-producing glands called as lobules and the milk passages are called as ducts. The stromal tissues constitute fatty and fibrous connective tissues of the breast. It is also known fact that the breast is also made up of lymphatic tissue-immune system tissue that removes cellular fluids and waste. The stem cells obtained from mammary stem cells form and give rise to inner luminal epithelial as well as outer basal myoepithelial cells that are found in the labulo ductal area. Majority of tumor that develop are reported to have originated from epithelial cells. The development of breast cancer can occur in varied areas such as lobules, ducts and even in between the tissue [1].

In general, the stages of breast cancer include [3]
1.1. **Stage 0**
Stage 0 is used to describe non-invasive breast cancers, such as DCIS (ductal carcinoma in situ). In stage 0, there is no evidence of cancer cells or non-cancerous abnormal cells breaking out of the part of the breast in which they started, or getting through to or invading neighbouring normal tissue.

1.2. **Stage I**
Stage I describes invasive breast cancer (cancer cells are breaking through to or invading normal surrounding breast tissue). Stage I is divided into subcategories known as IA and IB.

In general, **stage IA** describes invasive breast cancer in which

- The tumour measures up to 2 centimetres (cm) and
- The cancer has not spread outside the breast; no lymph nodes are involved

In general, **stage IB** describes invasive breast cancer in which

- There is no tumour in the breast; instead, small groups of cancer cells — larger than 0.2 millimetre (mm) but not larger than 2 mm — are found in the lymph nodes or
- There is a tumour in the breast that is no larger than 2 cm, and there are small groups of cancer cells — larger than 0.2 mm but not larger than 2 mm — in the lymph nodes

1.3. **Stage II**
Stage II is divided into subcategories known as IIA and IIB.

In general, **stage IIA** describes invasive breast cancer in which

- No tumour can be found in the breast, but cancer (larger than 2 millimetres [mm]) is found in 1 to 3 axillary lymph nodes (the lymph nodes under the arm) or in the lymph nodes near the breast bone (found during a sentinel node biopsy) or
- The tumour measures 2 centimetres (cm) or smaller and has spread to the axillary lymph nodes or
- The tumour is larger than 2 cm but not larger than 5 cm and has not spread to the axillary lymph nodes

In general, **stage IIB** describes invasive breast cancer in which:

- The tumor is larger than 2 cm but no larger than 5 centimeters; small groups of breast cancer cells — larger than 0.2 millimetre [mm] but not larger than 2 mm — are found in the lymph nodes or
- The tumor is larger than 2 cm but no larger than 5 cm; cancer has spread to 1 to 3 axillary lymph nodes or to lymph nodes near the breastbone (found during a sentinel node biopsy) or the tumor is larger than 5 cm but has not spread to the axillary lymph nodes

1.4. **Stage III**
Stage III is divided into subcategories known as IIIA, IIIB, and IIIC.

In general, **stage IIIA** describes invasive breast cancer in which either

- No tumour is found in the breast or the tumour may be any size; cancer is found in 4 to 9 axillary lymph nodes or in the lymph nodes near the breastbone (found during imaging tests or a physical exam) or
- The tumour is larger than 5 centimetres (cm); small groups of breast cancer cells (larger than 0.2 millimetre [mm] but not larger than 2 mm) are found in the lymph nodes or
- The tumour is larger than 5 cm; cancer has spread to 1 to 3 axillary lymph nodes or to the lymph nodes near the breastbone (found during a sentinel lymph node biopsy)

In general, **stage IIIB** describes invasive breast cancer in which

- The tumour may be any size and has spread to the chest wall and/or skin of the breast and caused swelling or an ulcer and
May have spread to up to 9 axillary lymph nodes or
May have spread to lymph nodes near the breastbone

Inflammatory breast cancer is considered at least stage IIIB. Typical features of inflammatory breast cancer include

- Reddening of a large portion of the breast skin
- The breast feels warm and may be swollen
- Cancer cells have spread to the lymph nodes and may be found in the skin

In general, stage IIIC describes invasive breast cancer in which

- There may be no sign of cancer in the breast or, if there is a tumour, it may be any size and may have spread to the chest wall and/or the skin of the breast and
- The cancer has spread to 10 or more axillary lymph nodes or
- The cancer has spread to lymph nodes above or below the collarbone
- The cancer has spread to axillary lymph nodes or to lymph nodes near the breastbone

1.5. Stage IV

Stage IV describes invasive breast cancer that has spread beyond the breast and nearby lymph nodes to other organs of the body, such as the lungs, distant lymph nodes, skin, bones, liver, or brain. Cancer may be stage IV at first diagnosis, called “de novo” by doctors, or it can be a recurrence of a previous breast cancer that has spread to other parts of the body.

The underlying major cause in precipitating the disorder highlights to two aberrant methylation i.e., global hypo-methylation taking place in the genome and hyper-methylation taking place at cytosine residue that precede a guanine residue being located at the promoter regions, commonly known as CpG islands. The global hypo-methylation often corresponds to enhanced instability of chromosome as well as involves in activation of transcription oncogenes, and protein encoding ultimately leading to malignancy.

2. Genes involved in breast cancer

In the history of breast cancer a huge range of major and minor susceptible genes that has been identified.

2.1. Highly penetrating genes

Are the ones which have the susceptibility of being very rare in the population with a minor allele frequency of <0.005.

2.2. BRCA1 and BRCA2 genes

These are tumor suppressor genes located respectively on chromosome 17 and 11 that encodes protein often associated with the stability of genes. These are basically involved in the DNA repair process via homologous recombination pathway. It is observed and reported in various research that BRCA1 and BRCA 2 mutation is found to be inherited in autosomal dominant manner but due to the presence of tumour suppressor, it acts recessively to repair DNA break. It is investigated and found that these genes are inherited from parents and are involved in protecting individuals from certain types of cancer [4] But its mutation at molecular level taking place by either frameshift or missense mutation inhibits its efficient working and thus are prone to get breast, ovarian and other cancer [5].

2.3. TP53 gene

This tumour suppressor gene is found to be highly penetrating and precipitates a variety of tumor. A mutation in TP53 is found to cause breast cancer at an early age [6].

2.4. PTEN gene

Basically involved in major functions like cell differentiation, proliferation, adhesion, apoptosis, migration and also in genome stability and cell cycle. Due to abnormal degeneration or H1 genes. These genes belong to family of cadherin. Being glycoprotein in nature, it is cell to cell adhesion protein present in chromosome number 16. Mutation and loss of function of this gene results in cancer and metastasis particularly lobular breast cancer and hereditary diffuse gastric cancer (HDGC). Clinical studies and investigations have shown that its difficult to differentiate the two cancer caused by
CDH1 mutation due to occurrence of signet ring cells and diffuse pattern of growth being prominent in both the cases. [7]

2.5. Moderate penetrant genes

These genes have the susceptibility of being rare, with a minor allele frequency of 0.005–0.01.

2.5.1. CHEK2 gene

Checkpoint kinase 2 is the gene basically involved in the repair mechanism of DNA. Being located on the chromosome 22, it encodes a protein that acts as a tumor suppressor gene. Furthermore mutation of CHEK2 gene is found to cause a varied variety of cancer since on mutation the cells are potentially able to divide rapidly [8,9]. Certain germline mutation of this gene is associated to increase risk of breast cancer in females two fold times and also male breast cancer.

2.5.2. ATM

The ataxia telangiectasia mutated gene (ATM) is involved in the break repair pathway of DNA strands. ATM mutation is linked with a neurodegenerative gene called Ataxia Telangiectasia syndrome and findings suggest that it often causes breast cancer [10]. The risk is severe since it interacts with C terminal of BRCA1 gene involved in DNA repair damage.

2.5.3. PALB2

This moderate penetrant gene also confers breast and ovarian cancer. Besides BRCA1 and BRCA2 this gene carries high risk of breast cancer in terms of risk factor. Studies and investigations reveal that PALB2 has major role in functions linked with BRCA2 and therefore it was earlier called as BRCA2 interacting protein and also its association with BRCA1 is also being found [11].

2.5.4. BARD1

Along with BRCA1 gene, BARD1 is found in DNA repair damage, regulation of cell cycle and RNA processing [12]. The structure of both the protein structure are alike and so form a heterodimer due to the presence of N-terminal RING finger domains that form a stable complex. The susceptibility of this particular gene is also found to cause triple negative breast cancer as well. There are two cancerous isoforms of this gene i.e, BARD1β and BARD1δ, furthermore recent reports throw light on the findings that suggest that families with negative BRCA1 and BRCA2 genes are found to have been affected by BARD1 in hereditary breast cancer [13].

2.5.5. BRIP1

BRIP1 (Breast Cancer 1 Interacting Helicase 1) is a tumor suppressor gene. Being located on chromosome number 21, it shows interaction with BRCA1 and has vital function in preserving the genetic stability by repairing DNA damage though have significant associations with the onset of breast cancer (BC) if mutated or overexpressed. The overexpression of BRIP1 has been found to be associated with different clinical features such as breast tumor subtypes, promoter methylation status, and survival of BC patients that often pose a risk of ensuing malignant transformation. While, lower expression of BRIP1 hinder Breast Cancer prognosis [14].

2.5.6. MRN complex

The MRE11–RAD50–NBS1 complex (MRN), is a hexameric protein complex involved in maintaining genomic stability and also in oncogene-induced replication stress. The protein complex consisting of MRE11, RAD50 and NBS1 are involved in cell regulation, recombination of DNA, and in double stranded DNA break repair. Being involved in cell homeostasis, a small disruption of MRN complex or its mutation leads to carcinogenesis [15].

2.5.7. XRCC2

X-ray repair cross complementing gene 2 is one of the breast cancer susceptibility gene. This gene encodes a member of the Rec A/Rad51-related protein family that participates in homologous recombination to maintain chromosome stability and repair DNA damage [16]. The risk of precipitating breast cancer due to XRCC2 is associated with family history, missense mutation, base pair mutation and protein truncating variant [17].
2.5.8. CDK1
Cyclin dependent kinase 1 is an important gene involved in DNA repair. Cyclin-dependent kinase 1 (CDK1) is necessary not only for BRCA1-mediated S phase checkpoint activation, but also for Homologous Recombination, because it phosphorylates BRCA1 for the efficient formation of BRCA1 foci [18].

2.5.9. ATR
Ataxia-telangiectasia- and Rad3-related is required for maintenance of genomic integrity. ATR, along with ATR-interacting protein, acts in parallel with ATM (ataxia-telangiectasia mutated), which is defective in the neurodegenerative disorder ataxia-telangiectasia and is also associated with breast cancer susceptibility [19]. ATR protein within the DNA Damage Response machinery is found to exhibit its major role as a sensor of replication stress (RS), that found to be elevated in cancer due to activation of oncogenes and impairment of G1 checkpoint control.

2.6. Low penetrant gene
These genes have the susceptibility associated with a small to moderate increased relative risk for breast cancer contributing to minor allele frequency >0.05.

2.6.1. ABRAXAS
Abraxas is part of a BRCA1-centered tumor suppressor gene network. The Abraxas protein links BRCA1 to a protein complex involved in BRCA1-dependent DNA damage responses. BRCA1 is needed for Abraxas' function in DNA repair. The BRCA1 protein C terminus (BRCT) domains interact with multiple proteins and are required for BRCA1's tumor-suppressor function [20].

2.6.2. MDM2
This gene (mouse double minute 2 also known as HDM2) encodes a nuclear-localized E3 ubiquitin ligase. It promotes tumor formation by targeting tumor suppressor proteins, such as p53, for proteasomal degradation. This gene is itself transcriptionally-regulated by p53. Overexpression or amplification of this locus is detected in a variety of different cancers including breast cancer [21]. Moreover, numerous studies have shown that MDM2 overexpression is associated with tumors that have a higher degree of invasiveness, later stages, greater metastatic potentiality and resistance to chemotherapeutic agents and radiation.

2.6.3. ARHLTSI1.
This gene ADP ribosylation factor like tumour suppressor gene encodes GTP binding protein that are found to be exhibiting major cell regulatory functions. Mutation in this gene results in truncated protein that does not get bind to GTP [22].

Moreover a lot of low-susceptibility loci have been identified through genome wide association studies (GWAS). Various low penetrant genes have been reported so far that are found to be causing risk of breast cancer such as NF1, CASP8, VEGF, LSP1, MAP3K1, TNRC9, TGFB1, TOX3, ESR1, and FGFR2.

3. DNA methylation and its underlying mechanism
DNA methylation is the addition of methyl group at the 5' position of cytosine ring [23]. This donor methyl is obtained from S-adenosyl methionine and it's a covalent modification. The whole alteration take place on cytosine that precedes a guanine in the sequence of DNA nucleotide located at the promoter site of gene known as CpG islands [25]. The process of methylation is catalysed by DNA (cytosine-5) methyl-transferases (DNMTs. It is predominantly seen that in mammals the DNMTs exist in the form a highly conversed family protein constituting DNMT1, DNMT2, DNMT3A, DNMT3B and DNMT3L [24]. The DNMT1 helps to methylate newly-synthesized in-order to copy paternal strands to daughter strands and the establishment of methylation is maintained via this. It takes place before chromatin packaging during DNA replication and is subjected to the replication foci occurring during the S-phase of cell cycle. DNMT3L functions as a stimulatory factor for de novo methylation and is also involved in methylation of DNMT3A and DNMT3B. It is investigated and found that the expression of DNMT3B is more enhanced in different human cancer types, whereas its suppression corresponds in apoptosis of tumour cells. It is evident from studies that DNMT3B plays a vital role and is much more overexpressed over DNMT3A and DNMT1 when seen in breast cancer malignancy. [27]
The hyper-methylation at Cytosine guanine islands (CGIs) situated at tumour suppressor genes often result in silencing of transcriptional genes that occurs via a number of mechanisms such as hyper-methylation of DNA [26] directly affects the RNA polymerase II whereby inhibiting the binding of DNA with transcriptional factors on certain specific sequences and also hyper-methylated DNA binds with methyl-CpG binding proteins. Furthermore Tumour suppressor genes (TSGs) normally functions to suppress or negatively regulate proliferation of the cells by the use of proteins that are encoded which necessarily block or inhibit the action of growth-promoting proteins. It is postulated that cancer involves the loss of function of TSGs that marks the silencing of genetic information. The altered 5-methylcytosine located in the promoter region of CpG island induces mutations and translocations, leading to disruption of the possibly controlled cell proliferation and thereby ultimately leading to malignancy. The hyper-methylation occurring in CGIs promoter of BRCA1 gene is now evidently recognized as the causative mutation precipitating breast cancer. The gene involved encodes a nuclear phosphoprotein which helps in maintaining stability of genome. The encoded protein often combines with other tumour suppressor, DNA damage sensors, and signal transducers to form a large multi-subunit protein complex that is called as BRCA1-associated genome surveillance complex (BASC) [27]. Furthermore, the BRCA1 protein is associated in multifunction involving repairing of damaged DNA of double stranded break, regulation of the transcription, recombination and controlling of the cell cycle check points. The hyper-methylation taking place at the BRCA1 promoter region is regarded as an inactivating mechanism for the expression of BRCA1 gene, leading to breast tumorigenesis. Evidences suggest that there is an important association between the inactivation or low expression of BRCA1 protein expression with the aberrant methylation status of CGIs in the BRCA1 promoter in breast cancer.

3.1. Diagnosis of DNA methylation

Studies confirm that a wide range of methylated genes are found in the DNA of tumour tissue of patients with breast cancer. Studies have also shown that aberrant methylation events occur early on in breast cancer development, and are detectable in tissue from in situ carcinomas (both lobular and ductal) [28] and early stage breast cancer (stages 0 and 1)[29] [30]. The changes that occur at methylation status of promoter site are natural phenomenon and are easily detected by various methods. An array of genes related to cancer have been found to be methylated in breast cancer. Furthermore adding them as markers, exhibit a potential for distinguishing between malignant disease and benign disease or normal tissue, and they also serve as detection of lobular carcinomas. It is observed that the combination of this minimally invasive methods along with mammography have shown an improvement in the sensitivity of tumor detection in women with high breast density, a characteristic that is associated with an increased breast cancer risk [31] and reduced sensitivity towards mammography[32]. Breast cancer screening methods are there to evaluate the status of methylation of DNA . The future of methylation analysis will likely involve a combination of isolation of the methylated fraction of DNA either using MBD proteins (Methylated-CpG Island Recovery Assay (MIRA)) or immunocapture (Methylated DNA Immunoprecipitation on Chips (MeDIP-chip)) methods and next generation microarray or sequencing technologies [33].

Analysis of methylation status is done in combination with real-time methylation-specific PCR and specific primers to amplify CpG islands. A genome-wide approach with high-throughput sequencing can also be performed using bisulfite-converted DNA from different non-invasive biological sources such as whole blood, serum and plasma [34]. In breast cancer, [35] using methyl DNA immunoprecipitation (IP) followed by a promoter array, found that hypermethylation was more frequent than hypomethylation in differentially methylated regions (DMRs) between tumor vs normal. However, methyl DNA IP is biased for hypermethylated DNA. Similarly,[36] found that hypermethylation is more frequent than hypomethylation in breast cancer compared with non-malignant tissue.

Another method to access the methylation status in breast, tumor stromal and epithelial cells by[37] using methylation-specific digital karyotyping (MSDK) was done to to differentiate between hypomethylated condition in epithelial and normal stromal cells. This technique was found to be biased for CpG islands, but the regions that were analyzed were distributed all across the genome inspite of being clustered around transcription start sites.

Findings suggest that genome wide methylation status can be assessed by different methodologies but there is biasness observed due to methods employed and their considerations. The use of Genome wide DNA methylation has been extensively applied to discriminate between tumor and normal or histologically non-malignant breast tissue. A wide range of loci have been identified that are found to be with aberrant DNA methylation state as observed between normal and tumor cells. Likewise several studies also confirmed on the profile of DNA status as well as their subtypes [38] used cluster analysis to show that ER+PR+ tumors had high methylation, whereas triple-negative breast cancers had low methylation. Genome wide DNA methylation studies in breast cancer cell lines have also shown clustering according to hormone receptor status based on DNA methylation levels [39].
In coming future a genome wide DNA methylation status along with its gene pathway and histone modifications needs to be integrated in order to analyse the gene expression and to identify the underlying causes that occur in its underway. Meta-analysis of all the currently undergoing studies along with its raw data that could correlate overlapping genes/gene families subject to aberrant DNA methylation in breast and other cancer as well.

4. Proposed methodology

To analyse and determine the methylation status of genes the most widespread technique is Bisulphite Sequencing which enables and prepares the genomic DNA in-order to identify specific gene methylation. This method is based in the use of Sodium bisulphite that helps in conversion of cytosine to uracil. In the meantime, the altered cytosine i.e. methyl cytosine does not get changed and thus remains intact. Thus with the help of this method methylation taking place at CpG islands are determined since a 5-methylcytosine appears as a matching C occurring in the bisulphite sequence, and the transition from C to T implies un-methylated cytosine’s. Still this method of bisulphite sequencing is considered inappropriate. The reagent bisulphite converts only single-stranded DNA but not double-stranded DNA and therefore this further leads to incomplete denaturation or reannealing resulting in incomplete conversion [24,25].

Therefore, this creates a complicated data, as majority of the time it is impossible to determine whether the unconverted cytosine is an indication of bona fide methylation or whether it is an experimental artefact. It also has demerit concerned with depuration, strand breakage and DNA degradation which take place due to harsh reaction conditions.

The gene specific way to determine DNA methylation is Real-timePCR. This method is employed due to its high sensitivity as well as its ability of being sequence specific. This technique has dynamic accuracy and conclusion. Meanwhile it is a method that uses minimal amount of DNA, RNA or cDNA to quantify and is based on fluorescence that bind to double stranded DNA or Sequence specific probes. This is monitored using a reporter fluorescent molecule that emits fluorescence in manifolds as the PCR products gets accumulated after amplification cycle.

Further by analysis of DNA methylation and the identification of corresponding genomic sequence for gene expression can be potentially utilized for determining the causative aberrant methylation. Both of these results can be correlated to determine the inappropriate gene silencing leading to breast cancer.

5. Conclusion

Methylation arrays thus are a potential target in identifying aberrant DNA. In order to assess and identify the methylation status changes occurring, PCR based strategies have been an option to detect the molecular changes in methylated alleles. Methylation marker can be further explored to predict the cancer initiation via nipple duct lavage fluid or by use of fine needle aspirates of breast. The circulating tumour cells is found to be major marker in detection and identification of metastatic diseases. Various newer approaches like isolation of the methylated fraction of DNA either using MBD proteins (Methylated-CpG Island Recovery Assay (MIRA)) or immunocapture (Methylated DNA Immunoprecipitation on Chips (MeDIP-chip)) methods and next generation microarray or sequencing technologies have been on the way to develop in order to identify and evaluate the status of methylation. Drug combination therapy along with more nanomedicine of targeted drug delivery and nanoparticle are some of the promising strategies towards cure of this disease.

Compliance with ethical standards

Acknowledgments

The completion of this review article could not have possible without the support of my co-author whose insightful knowledge and data analysis made this article a success.

Disclosure of conflict of interest

The authors declare that they have no conflict of interest.

References


