Chapter-4

Results

Overview

The results of this study are presented in four parts.

In chapter-4.1 entitled **"Survey, Cytology and Immunohistochemistry"**, data collected from the survey of the cancer registries of Cachar Cancer Hospital and Research Centre, Silchar and Silchar Medical College and Hospital, Silchar, are respectively presented. In the removal of cancer sample through surgery from breast cancer patients, the tissue subjected to tissue processing for further analysis. The histological slides are examined and evaluated under a microscope by a pathologist and he indicated whether the surgical margin is cleared or was involved. In the middle, through immunohistochemistry technique, we observed specific Anti-BRCA1 (Ab-1423) primary antibody staining in specific tumor of breast cancer patients. With the positive and negative results of cytology, we went for experimenting our second confirmation test through Molecular Biology technique. Chapter-4.2 entitled **"DNA isolation, Primer design and PCR"**, deals with the isolation of DNA from various samples with specialized protocol. After isolation we quantify the DNA and stored it into a freezer (-20^0C) for further study. With the help specific designed primer all the preserved samples were amplified through the PCR technique. Chapter-4.3 entitled **"Hotspot mutation detection and COLD PCR"**, deals with mutation analysis through bidirectional Sanger sequencing method. From the study some hotspot mutations in specific regions (Exon 2 and 11d of *BRCA1*) were identified. With the help of new Tc technique of COLD PCR, we detected the mutation without bidirectional Sanger sequencing of all samples in concordance to referenced data (Prior Sequencing data). Finally in chapter-4.4 **"Low penetrance candidate genes"**, the polymorphism of low penetrance candidate genes were determined in the studied samples of breast cancer patients and a specific correlation of mutation and polymorphism was analyzed.

4.1 Survey, Cytology and Immunohistochemistry

4.1.1 Survey of breast cancer patients in hospitals

A survey of the breast cancer patients at the Cachar Cancer Hospital and Research Center and Silchar Medical College was conducted for the first three years of the study and statistical analysis of breast cancer prevalence was performed. Cancerous tissue was collected with the help of medical oncologist or surgeon and examined by a pathologist. The collected samples were then subjected to different laboratory experiments starting from the DNA isolation to the sequencing of the amplicon products of each individual. Blood was also collected from healthy volunteers to compare the genetic changes to the mutated reast cancer patients. Fresh cancerous tissue, Tissue specimens from primary tumors, matched normal (adjacent non-neoplastic) tissues, Paraffin embedded tissue as well as the blood were collected from 59 breast cancer patients (with consent letter) treated at Cachar Cancer Hospital according to the Institutional Review Board approved protocol (Table 4.1).

Sample ID	S/A/R	Year	Sample ID	S/A/R	Year
AUBC1	F/35/H	2009	**AUBC31**	F/57/H	2010
AUBC2	F/46/H	2009	**AUBC32**	F/67/H	2010
AUBC3	F/70/H	2009	**AUBC33**	F/40/H	2011
AUBC4	F/56/H	2009	**AUBC34**	F/53/H	2011
AUBC5	F/67/H	2009	**AUBC35**	F/60/H	2011
AUBC6	F/70/H	2009	**AUBC36**	F/55/H	2011
AUBC7	F/76/H	2009	**AUBC37**	F/60/H	2011
AUBC8	F/65/H	2009	**AUBC38**	F/40/H	2011
AUBC9	F/45/H	2009	**AUBC39**	F/47/H	2011
AUBC10	F/49/H	2009	**AUBC40**	F/42/H	2011
AUBC11	F/56/H	2009	**AUBC41**	F/55/H	2011
AUBC12	F/64/M	2009	**AUBC42**	F/49/H	2011
AUBC13	F/70/M	2009	**AUBC43**	F/42/H	2011

AUBC14	F/56/H	2009	**AUBC44**	F/48/H	2011
AUBC15	F/66/H	2009	**AUBC45**	F/26/H	2011
AUBC16	F/65/M	2010	**AUBC46**	F/35/H	2011
AUBC17	F/63/H	2010	**AUBC47**	F/31/H	2011
AUBC18	F/35/H	2010	**AUBC48**	F/57/H	2012
AUBC19	F/39/H	2010	**AUBC49**	F/50/H	2012
AUBC20	F/42/M	2010	**AUBC50**	F/43/H	2012
AUBC21	F/68/H	2010	**AUBC51**	F/76/H	2012
AUBC22	F/43/H	2010	**AUBC52**	F/52/H	2012
AUBC23	F/53/H	2010	**AUBC53**	F/46/H	2012
AUBC24	F/44/H	2010	**AUBC54**	F/38/H	2012
AUBC25	F/34/H	2010	**AUBC55**	F/59/H	2012
AUBC26	F/36/H	2010	**AUBC56**	F/35/H	2012
AUBC27	F/45/M	2010	**AUBC57**	F/55/H	2012
AUBC28	F/80/H	2010	**AUBC58**	F/45/H	2012
AUBC29	F/69/H	2010	**AUBC59**	F/60/H	2012
AUBC30	F/56/H	2010			

Table 4.1 Collection of samples from Cachar Cancer Hospitals and Research Centre, Southern Assam.

Cancer Incidences in Southern Assam:

The survey of the hospital based data from ICMR suggested that the cancer incidences in northeast is high as comparision to rest part of India. In southern Assam, higher incidences of cancer was observed in Cachar district as compared to the other districts of Southern Assam in our surveyed data. From the data it was found that the percentage of occurrence of cancer in 2010 was more than the ratio observed in 2008 and 2009 (Figure 4.1).

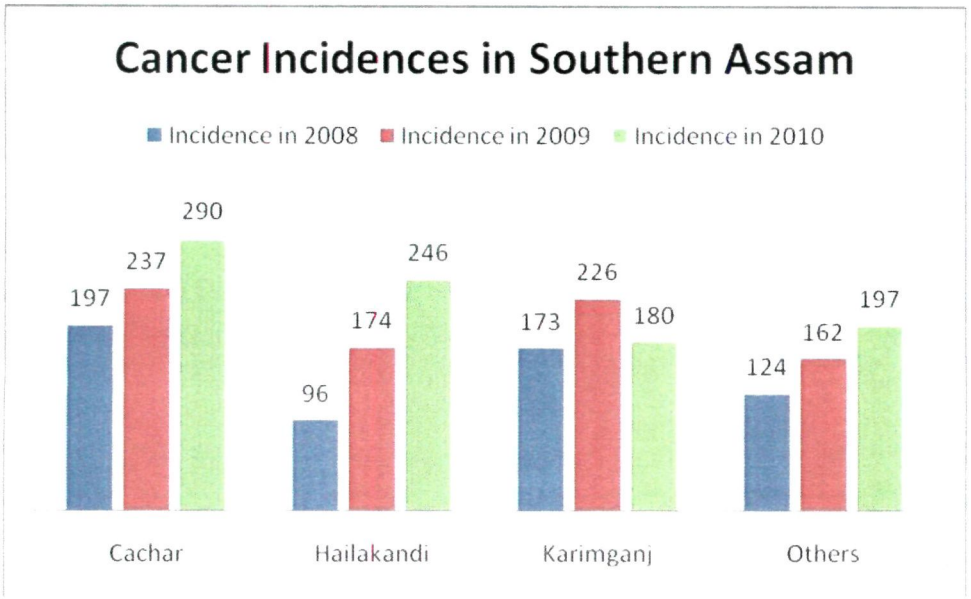


Figure 4.1 Two hospital survey (CCHRC and SMC) for cancer incidences in Southern Assam. Incidences of cancer in different parts of the Southern Assam for three years 2008, 2009 and 2010.

Prevelence of cancer in Southern Assam:

The data given below showed that the prevalence of different types of cancer found in Southern Assam in the Year of 2008, 2009 and 2010. From this data, breast cancer occupies the third highest threat in the year of 2008 and 2009 and second highest position in 2010. The head and neck cancer was severely reported from this area and the lung cancer and breast cancer incidences were higher at Southern Assam (Figure 4.2). There is gallbladder cancer, stomach cancer as well as cervical cancer found in the cancer prone area of Northeast India.

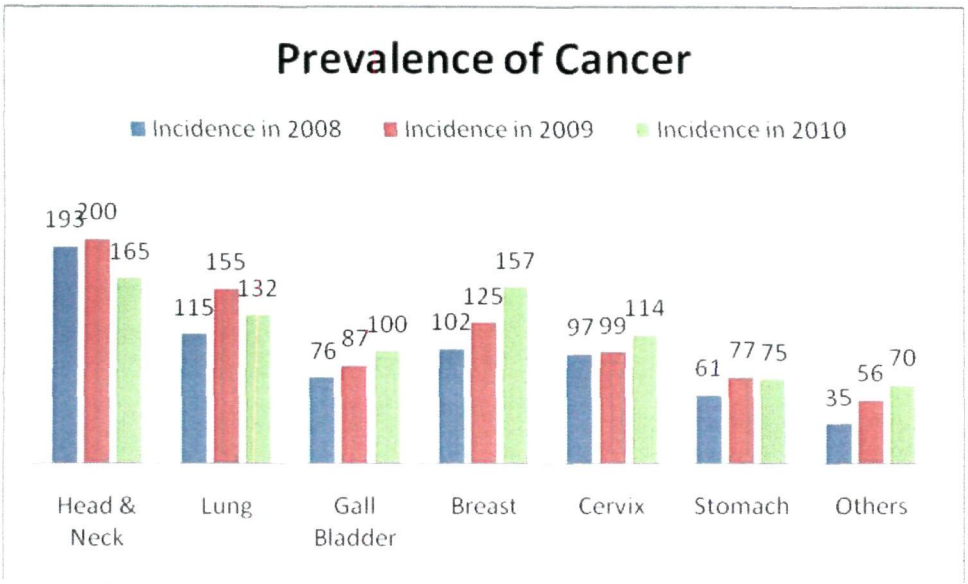


Figure 4.2 Prevelence of cancer in Southern Assam. The incidences of Head & Neck cancer was severely reported from this area and the breast cancer was after the incidence of head and neck cancer.

Breast cancer by age group in Southern Assam:

Breast cancer is highly prone to older age people. And in Southern Assam the result was alike to previously reported breast cancer of different place in India. The breast cancer by different age group has higher incidences in average 40-79 ages of people in southern Assam, in the year 2008 the highest incidence was found in 40-49 age group, and in the year of 2009 the highest incidences are at 40-59 ages, but from the year 2010 the incidences by age group is higher in the range of 50-59 years of age (Figure 4.3). Unfortunately some lower age female also affected from the breast cancer and it might be the inheritance of affected genes from parents to offspring.

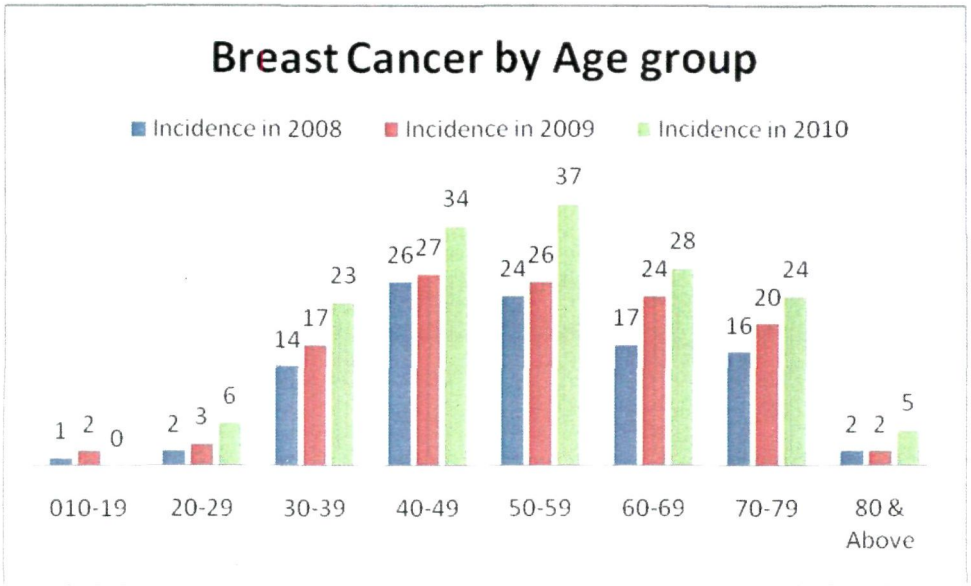


Figure 4.3 Breast Cancer incidences by different age group at Cachar Cancer Hospital and Research Centre. The incidences of breast cancer are higher in older people.

Basis of diagnosis of Breast cancer:

Basis of diagnosis of breast cancer in Cachar Cancer hospital and research center were H&E, FNAC, clinical, post operative, radiological. In the year 2008, 2009 and 2010, the state of diagnosis of H&E and FNAC was most operative diagnosed in Southern Assam, this is the basic diagnosis for all patients coming to hospital. Radiological diagnosis was also a basis of diagnosis but very less patients opt for this diagnosis. The state of diagnosis is more by HPE and FNAC (Figure 4.4).

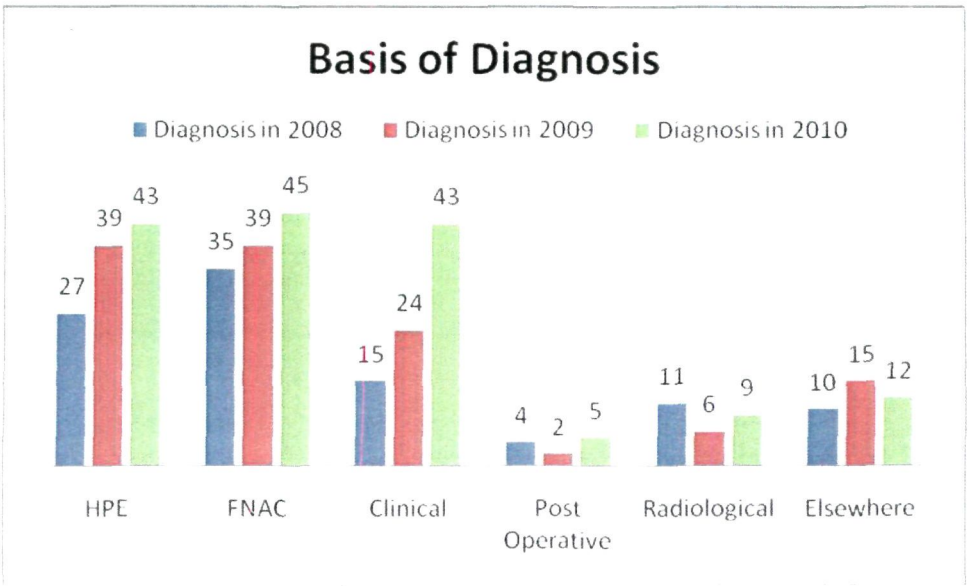


Figure 4.4 Different types of breast cancer diagnosis from the patients of Cachar Cancer Hospital and Research Centre and Silchar Medical College.

4.1.2 Histopathology of breast cancer cells

Haematoxylin and Eosin (H & E) is one of the most widely used histological staining methods for all and is a primary contrast method in the medical diagnosis of biopsy specimens. In our work it was routinely adopted to check the cell changes at the Cachar cancer hospital and Research Centre. The H & E stain yielded a surprising amount of useful information to our data. Though H & E stains were first implemented at least a century ago, the generalized method is still considered essential for recognition of tissue types and for recognizing morphological indicators diagnostic of cancer pathology. H & E staining used in detection of cancer routinely, the approach has evolved to include a variety of specialized, though related, protocol for different tissues and contrast emphasis. H & E staining is used for its remarkably robust and works well with a variety of fixatives and is used to discriminate between a broad range of cytoplasmic, nuclear and extracellular matrix features. The H & E staining method involves application of haematoxylin, a basic dye, to yield a blue-purple contrast on basophilic structures. An acidic eosin counter

stains eosinophilic structures bright pink. Various hues can also be present in the sample, including yellow and brown due to intrinsic pigments such as melanin.

The H & E staining first involved in the application of the basic dye haematoxylin. Haematoxylin stained basophilic structures, primarily those containing nucleic acid moieties such as chromatin, ribosomes and cytoplasmic regions rich in RNA. Nuclei showed varying cell and cancer specific patterns of heterochromatin condensation that are of noteworthy importance in pathological diagnosis.

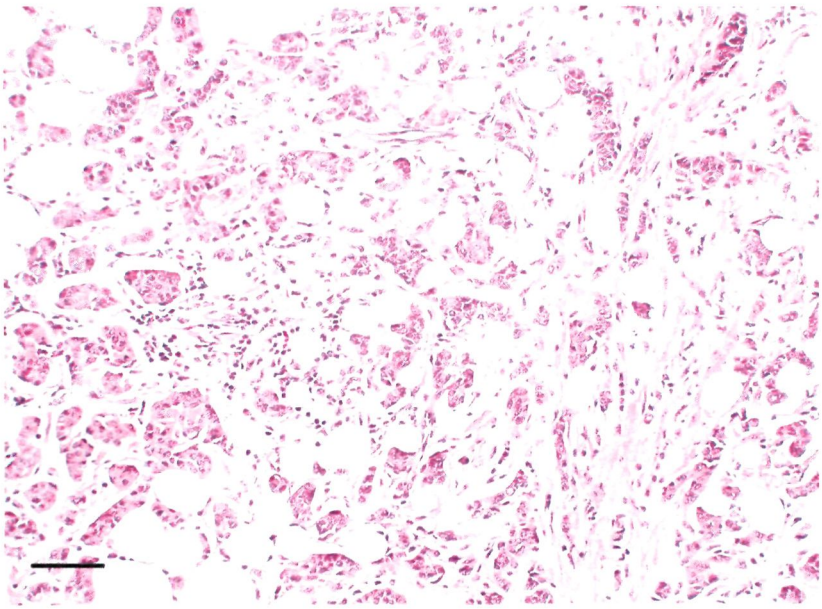


Figure 4.5 H&E stained section of a poorly differentiated, infiltrating breast cancer at magnification X 40 (bar = 20 μm). The infiltration of the fat by small nests of tumor cells. Granules appear in the cytoplasm and the cells become converted to hyaline structure masses of keratin which stain brightly with eosin and are identical with the horny material on the surface of the skin.

The staining strategy for application of haematoxylin used in this study was the progressive staining method. A progressive staining approach referred to a process of bathing the sample for an amount of time in the stain; the length of this step determines the staining characteristics and fewer steps are required

overall. There was a regressive staining strategy, where the specimen was effectively over stained then put through a series of acid ethanol washing steps to remove unwanted staining to the desired depth of color and staining characteristics.

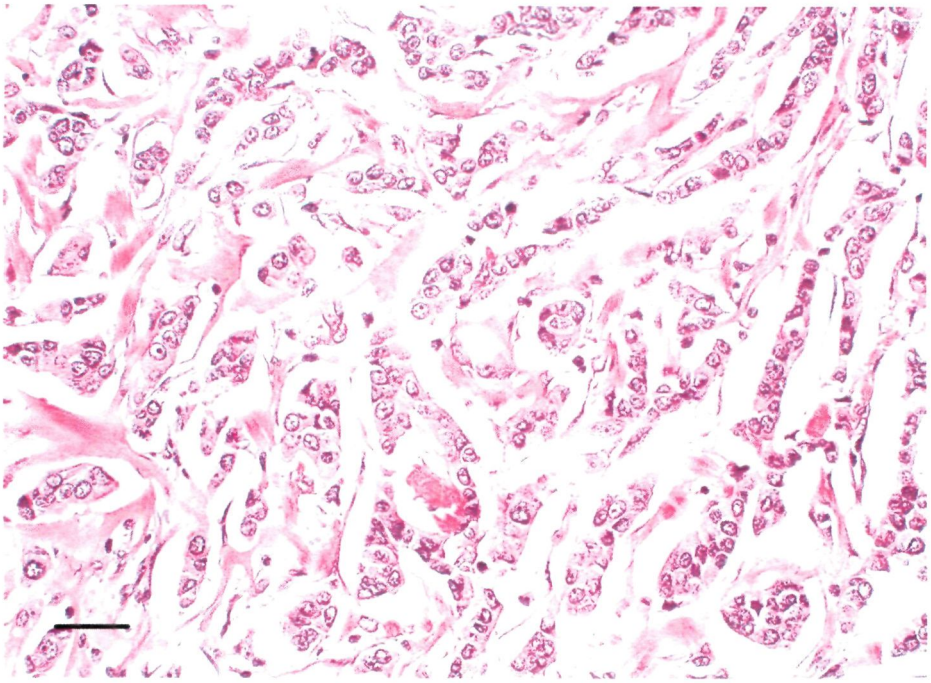


Figure 4.6 H&E stained section of a breast cancer tissue at magnification X 400 (bar = 20 μm). The cornified masses are known as cell nests or epithelial pearls. The outer cells of the pearls are arranged in a concentric manner as shown in figure. Note the infiltration by small nests of tumor cells surrounded by a dense connective tissue. This was graded as a SBR score = 8 and a Grade III.

This washing step was referred to as "differentiation." Haematin exhibits pH indicator-like properties; haematin is red and more soluble under alcoholic acidic conditions and becomes blue and less soluble under aqueous-alkaline conditions. When applied in breast tissue specimens, it is administered from an acid medium and the initial staining is brick red. In this phase the haematin was soluble and must be converted to the blue insoluble

form or it will leach into the mounting medium under the coverslip. The conversion is done through altering the pH of solutions bathing the slide. It was customary to place the sample in an alkaline solution to perform the "blueing" step then to wash with water to neutralize the pH of the tissue section. In Figure 4.5 the eosinophilic structures were covered the cytoplasmic region, which was generally composed of protein. Nucleoli stained with eosin. There were a variety of synthetic eosins, varying in hue. There are three commonly used forms of eosin: eosin yellowish (eosin Y), eosin bluish (eosin B), and eosin alcohol soluble (ethyl eosin). These dyes are all similar in staining characteristics and can be used both progressively and regressively. We used eosin Y, the most popular and also displayed a yellow-green fluorescence. Muscle tissue is somewhat paler but still pink, and collagen generally exhibits a pale pink color (Figure 4.6).

Ductal carcinoma in-situ (DCIS) is the presence of abnormal cells inside a milk duct in the breast. DCIS is considered the earliest form of breast cancer. DCIS is noninvasive, meaning it hasn't spread out of the milk duct to invade other parts of the breast (Figure 4.7). H & E stained tissue carries a remarkable amount of information to the trained eye, and an accurate and reliable reproduction of the many hues of high resolution in a manner consistent with the image as perceived through the eyepieces of a microscope a challenging application of camera technology.

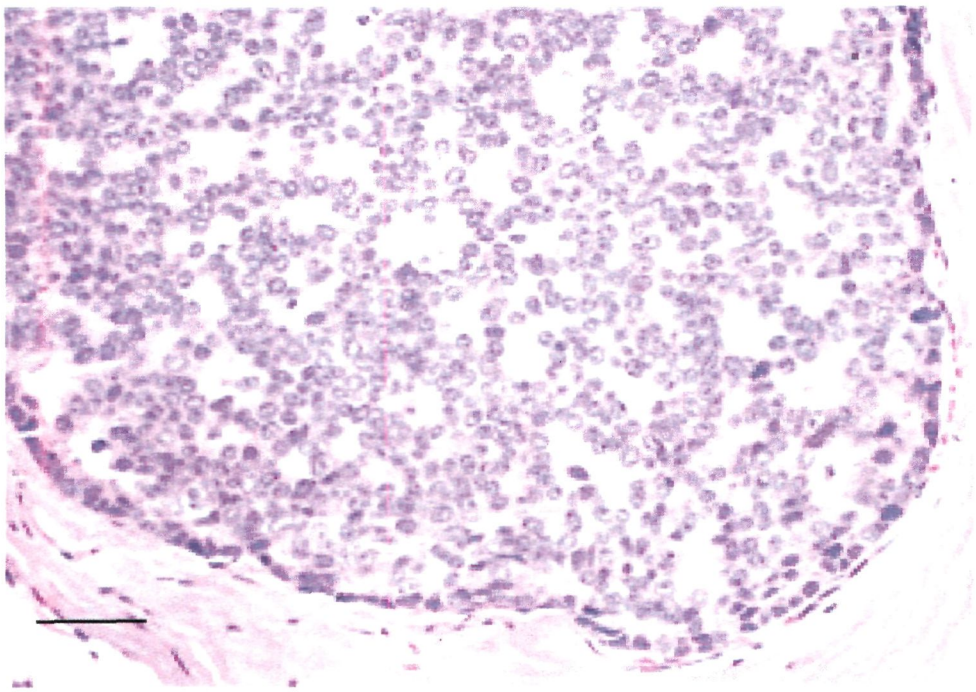


Figure 4.7 Low grade mammary biopsies with Ductal carcinoma in-situ at magnification X 400 (bar = 20 μm)

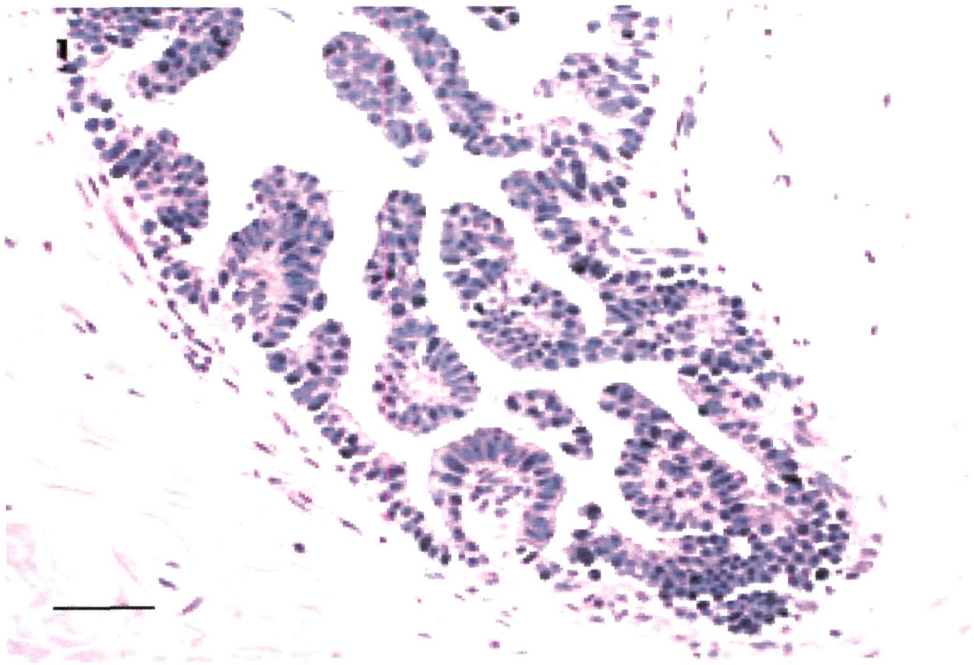


Figure 4.8 Intermediate grade biospies with Ductal Carcinoma in-situ at magnification X 400 (bar = 20 μm).

Furthermore, in Figure 4.8 the intermediate grade biopsies with ductal carcinoma in-situ were intermediate grade, which was nuclear grade 2 level where the mitotic rate is intermediate. After this stage the patient was at high risk for breast cancer because it is more likely to come back (recur) after complete removal through surgery. Patients with higher grade DCIS needed additional treatment for breast cancer. Most of the patients coming into Cachar Cancer Hospital and Research Centre were at high grade stage with ductal carcinoma in-situ.

4.1.3 Immunohistochemistry expression using Anti-BRCA1 (Ab-1423) antibody

Immunohistochemistry expression is most important nowadays because it express the specific protein in cancer tissues, it is the technique that microscopically detects cellular constituents via specific antibodies, has revolutionized the field of surgical pathology by tremendously empowering the hematoxylin and eosin (H&E) slide. It has almost perfected the pathologist's ability to make specific diagnoses such as those of mesothelioma, melanoma, and neuroendocrine tumors. In the investigation of the unknown primary, it enables detecting the origin of the tumor by use of tissue- or organ-specific antibodies as well as antibodies associated with minimal histogenetic differentiation. These diagnoses should be reached by interpreting the immunohistochemical results in the context of a detailed morphological analysis and differential diagnosis based on the H&E.

We used specific Anti-BRCA1 (Ab-1423) primary antibody and monoclonal Anti-Rabbit IgG (γ-chain specific) alkaline phosphatase antibody produced in mouse for detection of specific tumor for breast cancer. Immunohistochemistry can provide valuable information regarding tumor behavior and therapy. The interpretation of immunohistochemistry expression was made in a qualitative and subjective manner. The result depends on antibody staining to particular molecule for cellular response. Majority of

studied of slides were in positive staining, the result known as "positive" refers simply to the presence of nuclear staining (peroxidase) in any part of the studied tissue. The interpretation of immunostaining depends on the quantity of antigen present in the tissue and the determination of cut off values between what must be considered as positive and negative results. Among all these reasons, certainly the inter-laboratory reproducibility of the results of immunohistochemistry reactions is one of the most difficult challenges faced. In the normal tissue the immunoexpression in cellular compartments or in extra-cellular matrix components of the same marker was less as compare to cancerous one. As all we know the functions and the biological phenomena of BRCA1 protein with the cancer cells the correct interpretation of an immunohistochemistry expression can be through the figure that the antibody over express in the cancerous cells.

The validation of results in cases of ambiguity can be solved using antibodies against different epitopes of the same molecule, or by the detection of correlated antibodies. However, in our study we didn't go for this as because this antibody is very specific to BRCA protein, when the nature of an antigen or its function has not yet been totally elucidated, determining if the positivity of a given immunohistochemistry expression is relevant could prove hard. The possibility exists of a false-positive result, but also that the cell in question plays a distinct biological roles depending on the cellular compartment where the immunostaining is present.

Quantification of Immunohistochemistry Expression

Many studies have demonstrated that there is a correlation between the results obtained from the immunohistochemistry quantification and the tissue concentration of the antigen in question, efforts were made in order to try quantifying protein expression using immunohistochemistry. The biological colorations (which include those performed with aniline, hematoxylin and/or eosin for example) are usually difficult to control in terms of staining

intensity. This makes the comparison of cell to cell difficult as well as from slide to slide (between different tissues and between slides prepared on different days). This difficulty tends to decrease with the introduction of automatic techniques of coloration. The reagents employed in the immunohistochemistry technique present the potential to give true quantitative results. Most researchers, however, do not consider this possibility because they often do not observe the fact that this technique is no more than an immunological test carried out in situ or in histological slices.

As the need for an accurate immunostaining measurement is rising, quantitative biochemical methods of tissue detection are being progressively substituted by immunohistochemistry. Some question whether this precision is in fact achievable, or even necessary. However, advances in molecular biology and the emergence of new treatments for cancer will certainly increase the demand for precise results of a series of new molecules or target-genes, as a patient selection method for a given treatment. Therefore, studies will be developed in a progressively higher number of tissues, because the immunohistochemistry will likely be the chosen tool in the detection of these molecules.

Semi-quantitative analysis

As seen previously, the tissue expression of biomarkers employed in the immunohistochemistry technique can occur in different cellular compartments and even in extra-cellular matrix constituents. The evaluation of this reactivity may vary from essentially positive or negative to immunostaining intensity and/or extension, which constitutes an attempt towards immunohistochemistry technique quantification, frequently denoted in the literature as a "semiquantitative method".

When the intensity is the evaluation focus, the inclusion of reaction controls containing different levels of staining are required for comparing

criteria. Subjective scores, such as those that categorize the reaction in groups of null, weak, moderate and intense immunostaining, depend very much on the researcher's experience and are therefore unsuitable. To estimate the extension of immunostaining can also vary from a meticulous counting to a "glance" over the slides. The evaluation of the percentage of labeled cells trough the categorization of scores of the obtained percentage are less precise. Scoring systems were introduced into clinical practice in an attempt to overcome variances, particularly for markers that aim to select patients for specific treatments. It is important to emphasize that all scores, including those mainly used in daily practice, have demonstrated statistically significant relevance with regards to clinical variables when used by experienced researchers in the area, although they are laborious.

4.2 Molecular diagnosis of breast cancer patients

4.2.1 Qualitative and quantitative characteristics of sample DNA

For the maximal tumor content we cut the block in 5 μm thick sections. And then isolation of DNA was carried out by the standard phenol: chloroform: isoamylalcohol protocol as well Control DNA from blood and matched normal tissue of each individual was also extracted by ProteinaseK digestion followed by phenol/chloroform extraction. From most of the samples the genomic DNA extracted were of high molecular weight (greater than 10 kb). The extractions were in the range of 50 – 250 ng/μl and revealed purity in the range of 1.35 – 1.60 in a term ratio of the absorbance at 260/280 nm. The blood samples showed relatively ease in terms of digestion and cell lysis in comparison to tissue samples, which were further easier than the preserved samples. A representative image of DNA isolation from blood was shown in Figure 4.9. Furthermore, the genomic DNA extracted from blood samples in most cases were intact and therefore were less degraded. In case of tissue samples and mostly in case of preserved samples, the extracted genomic DNA showed smearing in few cases and may be mainly due to hydrodynamic stress and long term handling of samples involved during the extraction procedure (Figure 4.10). All the DNA samples were of high molecular weight and there were no degradation observed during the isolation period.

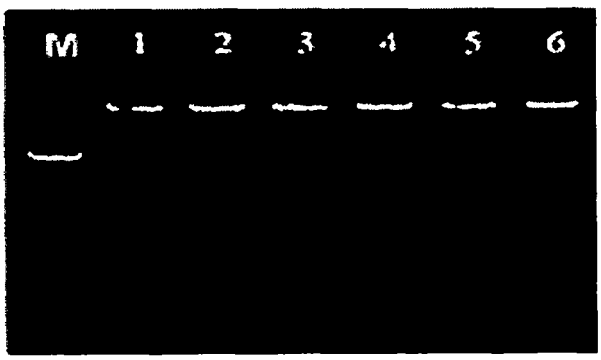


Figure 4.9 Isolated DNA from blood samples of breast cancer patients. M stands for DNA ladder

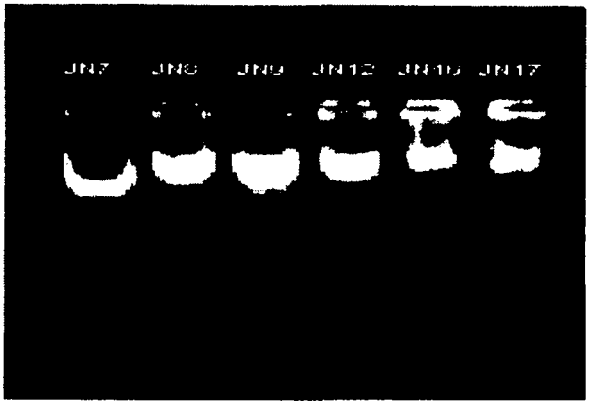


Figure 4.10 Isolated DNA, from breast cancer tissue samples in 1% agarose gel. JN7 to JN17 stand for breast cancer sample id.

4.2.2 Designing and preparation of primers

Optimal primer sequence and appropriate primer concentration are essential for maximal specificity and efficiency of PCR. A poorly designed primer can result in little or no product due to non-specific amplification and/or primer-dimer formation, which can become competitive enough to suppress product formation. There are several online tools devoted to serving molecular biologist design effective PCR primers. Primers were designed using the different software listed in Table 4.2. Details of the primers used for amplification of all target regions are listed in Table 4.4. Primers supplied in lyophilized form were resuspended in the required volume of sterilized Millipore water to obtain a final concentration of 100 pmol μl^{-1} and were stored refrigerated at -20^0C.

For the design of the primer different Bioinformatics tool was used considering the Tm value. Specificity and the temperature of annealing are dependent on primer length. It is generally accepted that the optimal length of PCR primers is 18-22bp. This length is long enough for adequate specificity, and short enough for primers to bind easily to the template at the annealing temperature. Researchers agreed early on that the design of PCR primers was difficult and unreliable. Computer programs were devised to take all of the design criteria into account. One of the first programs written for primer design was

Olga which made use of the implementation of Digital Research GEM (Graphics Environment Manager) on the Atari ST. Olga was specifically suited to the polymerase chain reaction (PCR) allowing simultaneous analysis of two primer sequences. The advantage of Olga was that it provided in one program analyses for direct repeats, secondary structures and primer dimerization as well as several useful 'finishing' tools for workers engaged in PCR optimization and oligonucleotide syntheses. The Primer3 program at the Whitehead Institute is now thought to be the most reliable and versatile tool currently available.

1. Primer Length: It is generally accepted that the optimal length of PCR primers is 18-22 bp. This length is long enough for adequate specificity and short enough for primers bound easily to the template at the annealing temperature (Table 4.2).

Primer	Length	GC content	Mol. Weight	A	T	G	C
BRCA1-185delAG-F	20 mer	35%	6247.2	10	3	6	1
BRCA1-185delAG-R	22 mer	36%	6724.5	6	8	3	5
BRCA1-1014delGT-F	18 mer	50%	5550.7	7	1	4	6
BRCA1-1014delGT-R	21 mer	52%	6464.3	6	3	6	6
BRCA1-3889delAG-F	21 mer	48%	6381.2	5	6	4	6
BRCA1-3889delAG-R	21 mer	48%	6332.2	4	7	3	7
p53-E8-F	20 mer	45%	5975.9	1	10	2	7
p53-E8-R	18 mer	50%	5416.6	2	7	3	6

Table 4.2 Details of primer properties length, GC Content, Molecular weight, ATGC composition of each primer.

2. Primer Melting Temperature: Primer Melting Temperature (Tm) by definition is the temperature at which one half of the DNA duplex will dissociate to become single stranded and indicates the duplex stability. Primers with melting temperatures in the range of 52-58^{0}C generally produce the best results. Primers with melting temperatures above 65^{0}C have a tendency for secondary annealing. The GC content of the sequence gives a fair indication of the primer Tm (Table 4.3).

3. Primer annealing temperature: The primer melting temperature was the estimate of the DNA-DNA hybrid stability and critical in determining the annealing temperature. Too high Ta produced insufficient primer-template hybridization resulting in low PCR product yield. Too low Ta lead to non-specific products caused by a high number of base pair mismatches (Table 4.3). So to overcome from these hurdles we used a formula:

Ta = 0.3 x Tm (primer) + 0.7 Tm (product) – 14.9

Where,

Tm (primer) = Melting Temperature of the primers

Tm (product) = Melting temperature of the product

4. GC Content: The GC content of primer designed in such a way that the percentage of GC was in between 40-60% (Table 4.3).

5. GC Clamp: The presence of G or C bases within the last five bases from the 3' end of primers (GC clamp) helps promote specific binding at the 3' end due to the strong bonding of G and C bases. More than 3 G's or C's should be avoided in the last 5 bases at the 3' end of the primer. In BRCA1-1014delGT-F, BRCA1-3889delAG-F, BRCA1-3889delAG-R, p53-E8-F and p53-E8-R primers contained

3 G's or C's in the 3' end of each sequence and rest primers contained less (Table 4.3).

Primer Name	Sequence (5'→3')	Tm	Ta
***BRCA1*-185delAG-F**	ATT GGA ACA GAA AGA AAT GG	52.3 ^{0}C	50^{0}C
***BRCA1*-185delAG-R**	CCT AGT ATG TAA GGT CAA TTC T	56.4 ^{0}C	
***BRCA1*-1014delGT-F**	ACA GCA TGA GAA CAG CAG	53.8 ^{0}C	54^{0}C
***BRCA1*-1014delGT-R**	CAC AGG GGA TCA GCA TTC AGA	61.2 ^{0}C	
***BRCA1*-3889delAG-F**	TCT ACT AGG CAT AGC ACC GTT	59.5 ^{0}C	53^{0}C
***BRCA1*-3889delAG-R**	CTT CCA ATT CAC TGC ACT GTG	59.5 ^{0}C	
***p53*-E8-F**	GCT TCT CTT TTC CTA TCC TG	56.4 ^{0}C	50^{0}C
***p53*-E8-R**	CTT ACC TCG CTT AGT GCT	53.8 ^{0}C	

Table 4.3 Primer sequence of different portion from the *BRCA1* and *p53* gene.

3'-end Sequence: It is well established that the 3' terminal position in PCR primers is essential for the control of miss priming. We have already explored the problem of primer homologies occurring in these regions. Another variable to look at is the inclusion of a G or C residue at the 3' end of the primers. This "GC Clamp" helps to ensure correct binding at the 3' end due to the strong hydrogen bonding of G/C residues. It also helps to improve the efficiency of the reaction by minimizing any "breathing" that might occur.

6. Primer Secondary Structures: Presence of the primer secondary structures produced by intermolecular or intra molecular interactions can lead to poor or no yield of the product. They adversely affect primer template annealing and thus the amplification. They greatly reduce the availability of the primers to the reaction.

i) Hairpins: It is formed by the intramolecular interaction within the primer and should be avoided. Optimally a 3' end hairpin with a ΔG of -2 kcal/mol and an internal hairpin with a ΔG of -3 kcal/mol are tolerated generally.

ΔG definition: the Gibbs Free Energy G is the measure of the amount of work that can be extracted from a process operating at a constant pressure. It is the measure of the spontaneity of the reaction. The stability of hairpin is commonly represented by its ΔG value, the energy required to break the secondary structure. A larger negative value for ΔG indicates stable, undesirable hairpins. Presence of hairpins at the 3' end most adversely affects the reaction.

$\Delta G = \Delta H - T\Delta S$

ii) Self Dimer: A primer self-dimer is formed by intermolecular interactions between the two (same sense) primers, where the primer is homologous to it. Generally a large amount of primers is used in PCR compared to the amount of the target gene. When primers form intermolecular dimers much more readily than hybridizes to target DNA, they reduce the product yield. Optimally a 3' end self-dimer with a ΔG of -5 kcal/mol and an internal self-dimer with a ΔG of -6 kcal /mol are tolerated generally.

iii) Cross Dimer: Primer cross dimers are formed by intermolecular interaction between sense and antisense primers, where they are homologous. Optimally a 3' end cross dimer with a ΔG of -5 kcal /mol and an internal cross dimer with a ΔG of -6 kcal /mol is tolerated generally.

7. Repeats: A repeat is a di-nucleotide occurring many times consecutively and should be avoided because they can misprime. For example: ATATATAT. A maximum number of di-nucleotide repeats acceptable in an oligo are 4 di-

nucleotides. And in our designed primer, not a single primer contains these type of repeats (Table 4.3).

8. Avoid Template secondary structure: A single stranded Nucleic acid sequences is highly unstable and fold into conformations (secondary structures). The *stability of these template secondary structures depends largely on their free* energy and melting temperatures (Tm). Our primers are designed for secondary structures which are stable even above the annealing temperatures, the primers are unable to bind to the template and the yield of PCR product is significantly affected. Our products determine the secondary structures of the template and design primers avoiding them.

9. Avoid Cross homology: To improve the specificity of the primers it is necessary to avoid regions of homology. Primers designed for a sequence must not amplify other genes in the mixture. Commonly, primers are designed and then BLASTed to test the specificity. Our products offer a better alternative. For better result we avoided the regions of cross homology. We checked our primer pairs with the BLAST and templates against the appropriate non-redundant database and the software had given us a good result. Software can identify regions with significant cross homologies in each template and avoided them during primer design.

4.2.3 Amplification of targeted mutation

Fresh cancerous and matched normal (adjacent non-neoplastic) tissue's specimen as well as peripheral blood samples were collected into EDTA vials from fifty nine randomly selected patients from Cachar Cancer Hospital and Research Centre (CCHRC). The extraction of DNA from breast cancer tissue and blood were done by phenol/chloroform method (Ghosh et al., 2011). The supernatant containing total genomic DNA was aliquot and stored at -20°C. The isolated

DNA was checked by the spectrophotometer and gel electrophoresis for its purity and quantity. Subsequently the DNA was processed for the PCR amplification through the exon 11 specific primer pair. The quantity of DNA used for PCR reaction was about 50-100ng. The reaction mixture contained total volume of 30µl, which was prepared by the GeneAmp® High Fidelity PCR system (Part Number 4328212, Applied Biosystem) and the final concentrations of reagents as follows: 1X buffer, 200µM each dNTP, 20pmol forward and reverse primers, 5U/µL High Fidelity enzyme, 3mM $MgCl_2$ and DNA. It was performed and monitored in Veriti PCR system (Applied Biosystems Inc., CA, USA) for 30 cycles. Then, PCR products were eluted from the agarose gel by the help of QIA kit® (Cat. No. 28704, USA) and processed for bidirectional Sanger sequencing at GCC Biotech, India Pvt. Ltd., Kolkata, India.

The primer pairs mentioned in the methodology for amplification of partial length of specific exons of *BRCA1* and p53 genes led successful PCR amplification in all the cases. A single uniform band was amplified for each exon of the respective genes each primer pair that carried traces of primers that migrated further by passing the PCR products. After purification, none of the PCR amplicons got degraded and therefore no smearing observed. Also there were no traces of primers on the purified PCR products that are essential for getting good sequencing results. The amplified PCR products were shown in Figure panel 4.11, 4.12, 4.13 and 4.14.

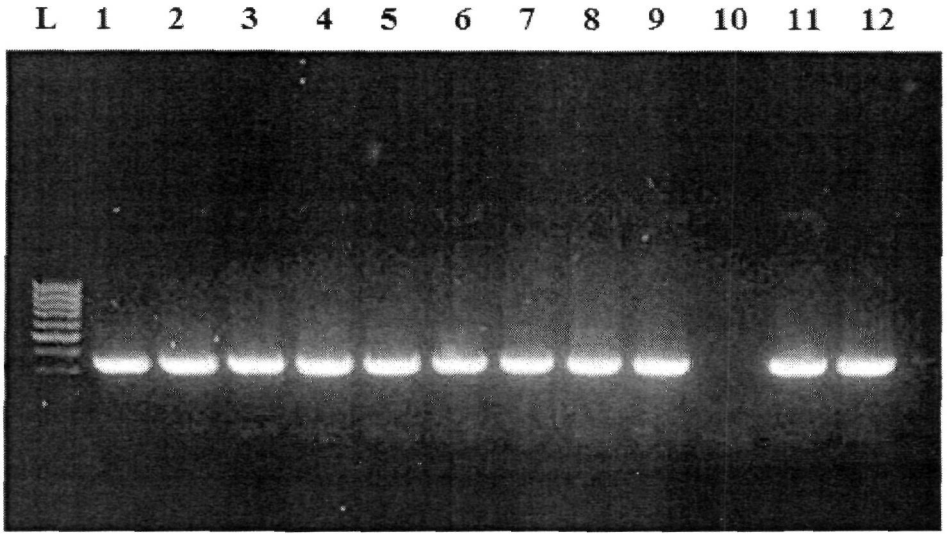


Figure 4.11 PCR amplification through *BRCA1*-1014DelGT primer pairs from breast cancer tissue DNA in 2% agarose gel. Where L was 100bp ladder, amplicon product size was approx. 195 bp and all numerical were sample id. Sample number 10 was not amplified due to less concentration of DNA.

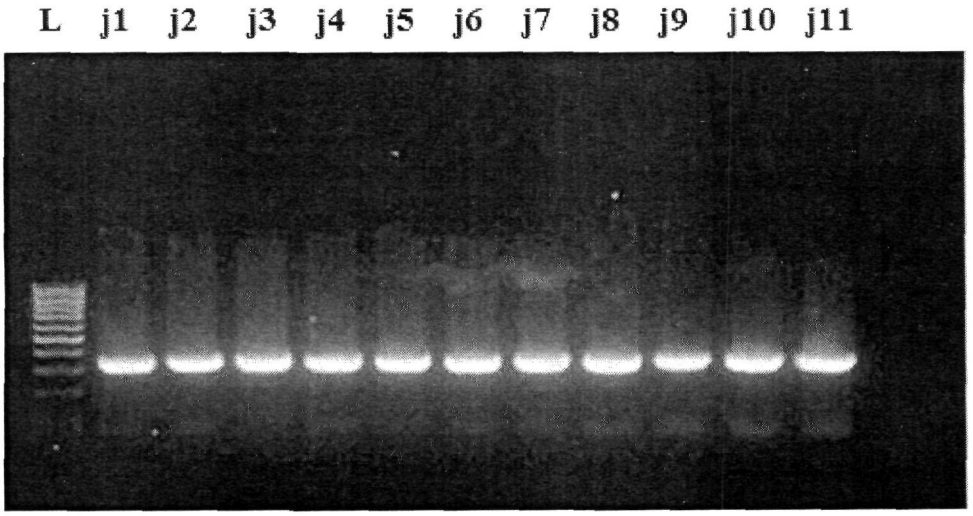


Figure 4.12 PCR amplification of *BRCA1*-3889DelAG primer pairs from breast cancer DNA in 2% agarose gel. Where j1, j2, j3 to j11were sample id of respective isolated DNA. The amplicon size was approx. 192bp.

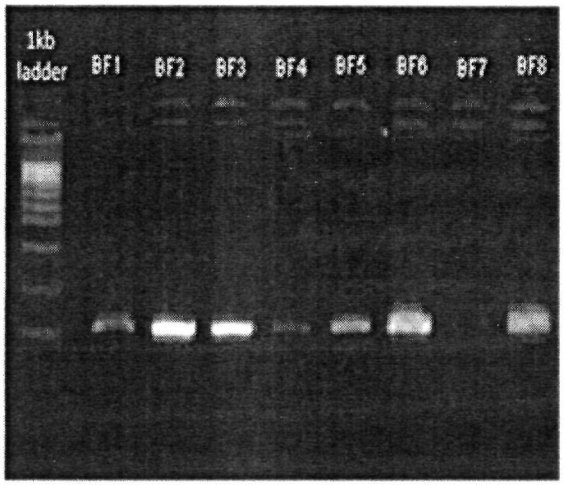


Figure 4.13 PCR amplification of *BRCA1*-185DelAG primers pair from breast cancer DNA in 2% agarose gel. Where 1kb ladder was used and the amplicon size was approx. 180 bp.

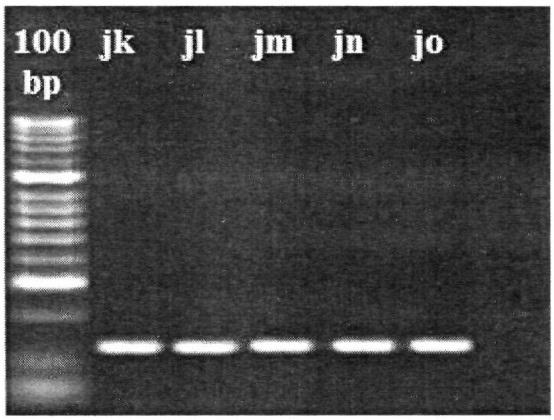


Figure 4.14 PCR amplification of p53 E8 primers pair from breast cancer DNA in 2% agarose gel. The amplicon size was approx. 167 bp.

4.3 Enhancement of molecular detection

From the diagnosis of thirty two women for breast cancer during 2009-2011, 90.62% women were belongs to the age group 46-57 years and 9.37% belongs to the age group less than 40 years. In comparison to the relation of age and breast cancer, the percentage of breast cancer in the younger women is quite high in this region. 6.25% patients have the personal history of breast cancer and they found with the breast cancer disease. Reminiscent of this, the personal history of a patient with breast cancer in one breast or related cancer in other parts of the body has a 3 to 4 fold increased risk of developing a new cancer in other breast or in another part of the same breast. From the present exploration of mutational research study, it has been revealed that 12.5 % of the patients have breast cancer in first-degree relative. The details of which can be seen in Table-4.4.

4.3.1 Hotspot mutation detection

By the help of three sets of primer, we amplified the particular regions of the *BRCA1* gene and screened for alterations of the particular position of 185DelAG, 1014DelGT and 3889DelAG of exon 2, 11*a* and 11*d* respectively. To attain the objective of the study, we used the PCR technique and DNA sequencing technology for screening of all patients for the particular region of mutation with this population. The mutation rate is high, more than 40% were found to be positive with these three variations of the *BRCA1* after screening. And from that, most of the mutation occurrences were from the exon 11, out of the 32 patients 12 have the mutation which was around 37.5% and in exon 2, the incidences of mutations were quite low (around 3.12%) to the other part of world ethnicity. In commencing exon 11, we have screened two parts of the partial sequence; those were exon 11*a* and 11*d* of 192 and 194 bp product of amplicon respectively. The mutation rate of 1014DelAG and 3889DelAG was 09.37% and 28.12% respectively (Table-4.5).

Raw nucleotide sequences from the sequencing results were processed through Sequence Scanner v1.0 (Applied Biosystem). The exported sequences were analyzed by using BLASTN software on NCBI site. According to the highest similarities, we depict the peak through the Chromas 2 software, and show the variations among the sequences in figure panel (Figure-4.15, 4.16, 4.17).

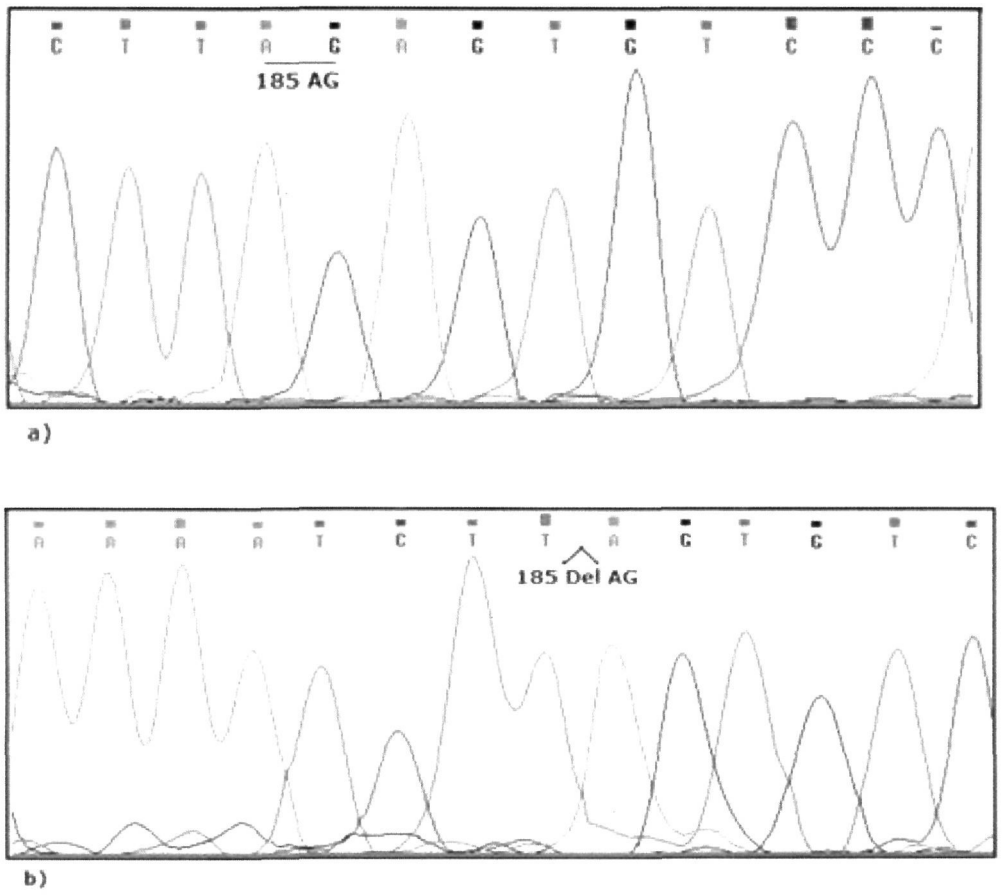
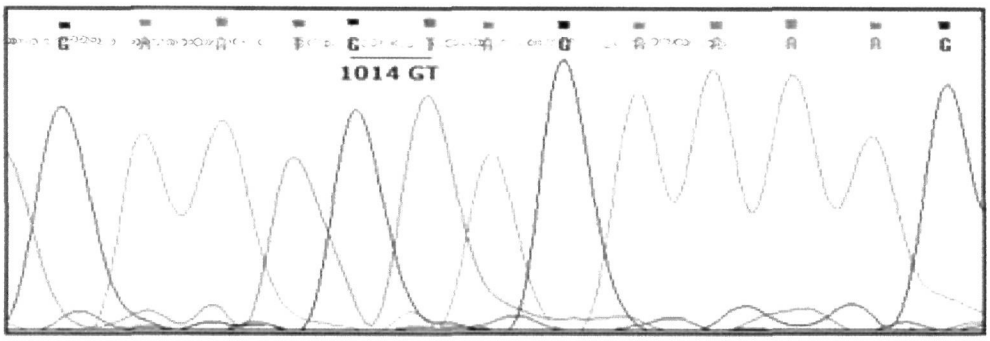
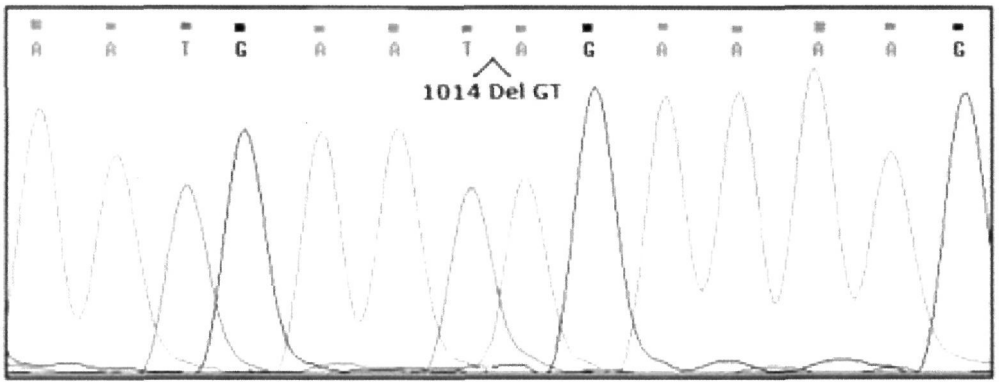


Figure 4.15 Variation in the sequence chromatograms of *BRCA1* gene amplified and screened for alterations of the specific position at 185 Del AG. A) Normal sequence, no deletion at 185 AG position from exon 2. b) Mutated sequences, deletion at 185 AG position, from exon 2.

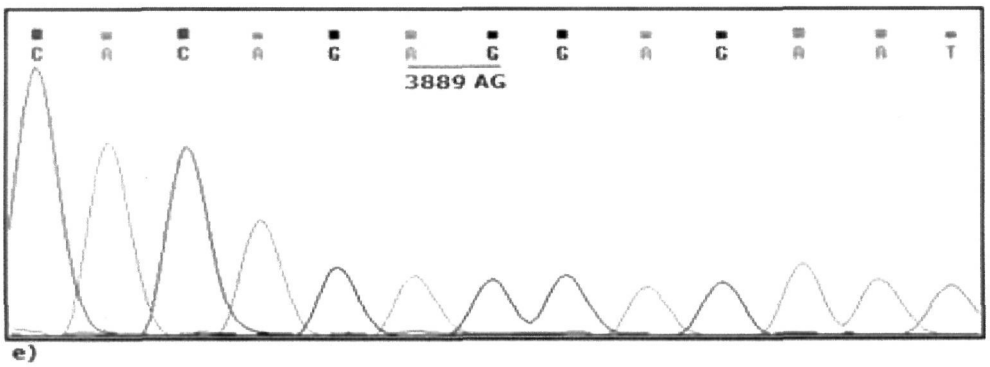


c)

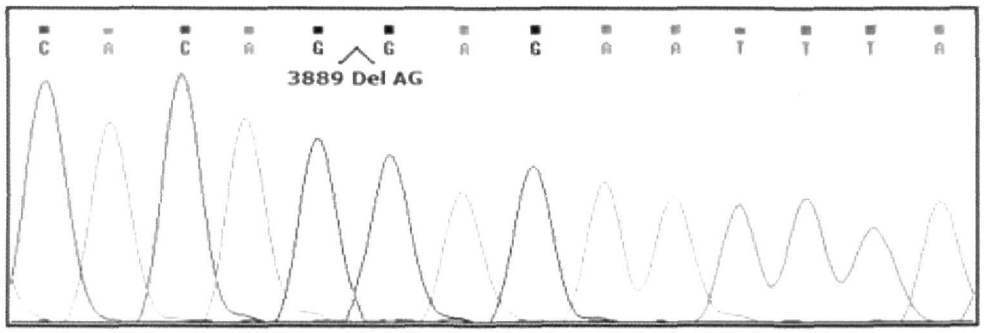


d)

Figure 4.16 Variation in the sequence chromatograms of *BRCA1* gene amplified and screened for alterations of the specific position at 1014DelGT. c) Normal sequence, no deletion at 1014 GT position from exon 11a. d) Mutated sequences, deletion at 1014 GT position, from exon 11a.



e)



f)

Figure 4.17 Variation in the sequence chromatograms of *BRCA1* gene amplified and screened for alterations of the specific position at 3889DelAG. c) Normal sequence, no deletion at 3889 AG position from exon 11d. d) Mutated sequences, deletion at 3889 AG position, from exon 11d.

4.3.2 Enrichment of sequence through COLD-PCR technique

A prominent concern confronting the clinical and diagnostic applications of mutation detection in breast cancerous tissue is the precise ability to detect mutations in Breast cancer patients. Mutation detection result can have major clinical significance for planning & screening of the population in high risk & low risk groups for treatment. To brazen out the mutation detection Jin Li et al. discover Co-amplification of target gene at lower denaturation temperature-PCR (COLD-PCR) technique in 2009. This approach was based on determination of critical denaturation temperature (Tc) for a given DNA sequence, that the percent of denaturation becomes sensitive to the exact DNA sequence, such that even point mutations make a substantial difference.

Co-amplification at lower denaturation temperature-PCR (COLD-PCR) is a new technology discovered by Jin Li et al., (2008b, 2009). It is a single-step method that results in the enhancement of both known and unknown minority alleles during PCR, irrespective of mutation type and position. This approach is based on critical denaturation temperature (Tc) for a given DNA sequence, that the percent of denaturation becomes sensitive to the

exact DNA sequence, such that even point mutations make a substantial difference. It can be carried out in 2 formats, full COLD-PCR and fast COLD-PCR, depending on whether it is necessary to detect all mutations comprehensively or to detect specific 'Tm' reducing mutations in a rapid and highly selective fashion (Mancinni et al., 2010).

After a number of regular PCR cycles that enable an initial build-up of target amplicon(s), the PCR program was shown in Table-3. After denaturation at 94^0C, the PCR amplicon(s) are allowed to cross-hybridize at an intermediate temperature (for example, 70^0C for 2–8 min). Because mutant alleles are in the minority, most mutant alleles end up in a mismatch-containing structure (heteroduplex) that has a lower melting temperature than the fully matched structure (homoduplex). Next, the PCR temperature is raised to the Tc to denature the mismatch containing sequences preferentially over the fully matched sequences. Finally, the temperature is reduced to 55^0C to allow the primers to bind and prime replication of the preferentially denatured sequences. Because this crucial denaturation is performed at every PCR cycle, the differential enrichment of mutation-containing alleles is compounded exponentially, and results in a large difference in overall amplification efficiency between mutant and wild-type alleles at the end of the cycling (Li, and Makrigiorgos (2009), Nature medicine).

Construction of data library

From the sequencing results it has been found that among three studying mutations type, 3889delAG was found in more percentage (28.12% from thirty two patients), which results in protein truncation of *BRCA1* protein by forming stop codons individually at 1265 position of the amino acid. So we constructed three library having three mutation type details and there Ta and Tm value. From constructed library of *BRCA1* mutation, we selected 3889delAG mutation for further experiment of COLD PCR. To provide mutational screening of 3889delAG without sequencing, we used Tc concept from COLD-PCR to detect the wild and mutant type among the breast cancer

population which is effective and inexpensive. In order to verify the accurate mutation identification through Tc concept and to overcome from the expensive mutational screening, we compared the new experiment result data with the constructed library data and it is totally concord with the referenced data. Hence, we propose a strategic use of Tc, that we prefer to call Clever-Tc, can be applied to detect mutations in *BRCA1* cutting down the overall molecular diagnostic cost.

"The Clever-Tc"

We divided the final constructed library in two categories i.e. wild and mutant type of BRCA1. Those samples with no detectable mutations came under the category of wild type and samples with mutations resided in mutant category. To identify the optimal critical denaturation temperature experimentally from both wild and mutant types, a PCR reaction for the targeted sample was performed using conventional gradient PCR skipping initial denaturation temperature (Figure- 4.18). For both amplicon there is a specific Tc, below which PCR efficiency drops abruptly. The denaturation temperature that reproducibly produces robust PCR products combined with substantial enrichment of the amplicon was selected as the Tc and the temperature difference of Tc between each wild and mutant type amplicon was found typically in the range of 0.2 to 1.5^{0}C (Figure- 4.20). Using these two important critical denaturation temperature we detected the BRCA1 genetic variation among wild and mutant types.

In a PCR reaction, we set the following graded denaturation temperatures:

(a) A conventional temperature, 98^{0}C;

(b) Temperature equal to conventional temperature –0.5^{0}C;

(c) Temperature equal to conventional temperature –1.0^{0}C;

(d) Temperature equal to conventional temperature –1.5^{0}C;

(e) Till the Critical denaturation formation.

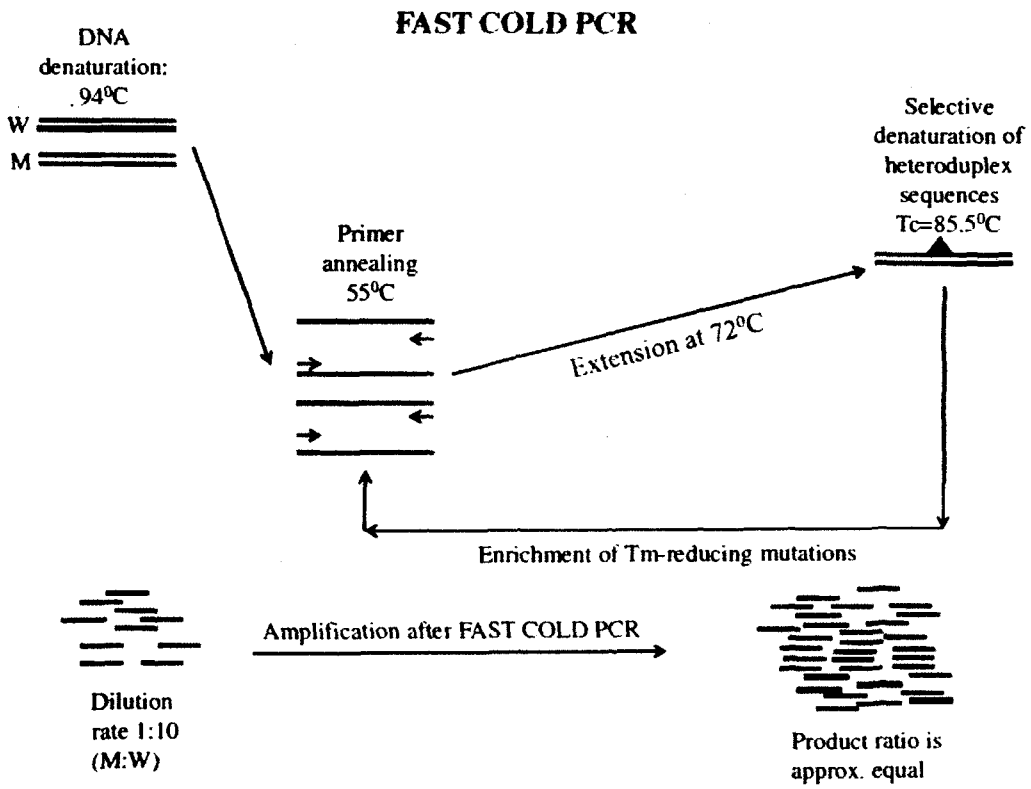


Figure 4.18 *fast* COLD-PCR for enrichment of Tm-reducing mutations, it also contain number of regular PCR cycles that enable an initial build-up of target amplicon(s), the *fast* COLD-PCR has following steps; 1) the amplicon DNA was denatured at a high temperature 94 °C. 2) Followed by primer annealing step for amplification of the both type of DNA amplicon 55 °C. 3) Then extension of the amplicon in preferred temperature 72°C, it helps to form the heteroduplex sequences without any requirement of intermediate temperature. 4) By the help of critical denaturation temperature 85.5 °C, mismatch heteroduplex were selectively denatured, excluding the homoduplex DNA which is form due to normal sequence present in the reaction. 5) Then thirty steps of cycles with each different step was carried out for the amplification of the mutant amplicon by the help of critical denaturation which enrich the Tm reducing sequence, and hence a larger proportion of minor variant DNA was amplified by subsequent rounds of PCR.

Tc of wild type was determined at 78^0C and its mutant type at 77.5^0C. Incorporating Tc of mutant type in both cases during PCR reaction, mutation at 3889delAG position along the BRCA1 sequence was amplified, which was absent in wild type. DNA amplicons differing by a single nucleotide reproducibly resulted in different amplification efficiencies relative to wild-type when the PCR denaturation temperature is set at the Tc. In deletion of specific mutation the Tc of wild type is more than the mutant type and in insertion it is reversible (Figure- 4.19).

In subsequent experiments, we followed the Clever-Tc concept to other samples with reference to the constructed library. The DNA of each individual was amplified using conventional PCR for determination of BRCA1 wild & mutant type and consequently follow another set of PCR reaction with the predetermined Tc to analyze the data through agarose gel under the UV visualization (Figure- 4.20). And it is found total agreement to the sequenced mutation/ wild data.

Total Agreement

By the help of sets of primer, we amplified the particular regions of the *BRCA1* gene and screened for alterations of exon 11d using BLASTN and chormas 2 software tool. We used the PCR and DNA sequencing technology for screening of all patients for the particular region of 3889delAG variation within these fifty nine patients (Figure 4.20), nine out of 32 breast cancer patients (28.12%) were found to have 3889delAG mutation in BRCA1. The main purpose of this study was to evaluate the potential of Tc to detect the genetic mutational changes in the breast cancer samples. To validate the new Tc concept, we considered sequenced data as the reference data and divided accordingly in two categories. Our result suggested successful application of Tc concepts for mutational screening that was completely insured with sequencing data. Considering this new concept, different mutations in the *BRCA1* gene can be standardized for molecular screening and that can be directly applied to the wider population of breast cancer patients with low cost screening. Our findings indicate that additionally this concept could be applied

to a wide range of mutational detection in a number of genes and would be used as a novel low cost precise and robust molecular screening tool for early cancer screening and therapeutic decision making.

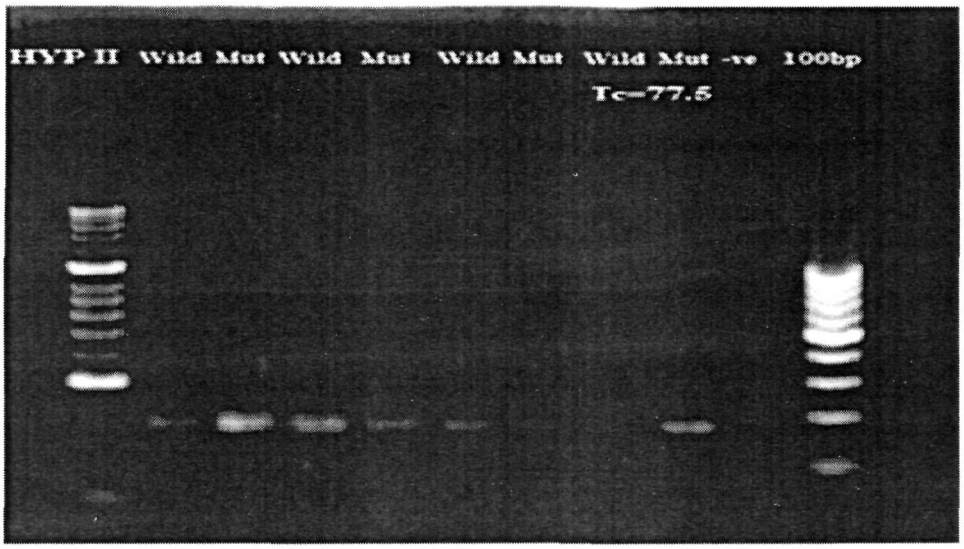


Figure 4.19 Variation in wild and mutant type amplicon produced through Tc concept in 2.5% agarose gel. In 77.5^{0}C the mutant amplicon was amplified and the wild type was not amplified.

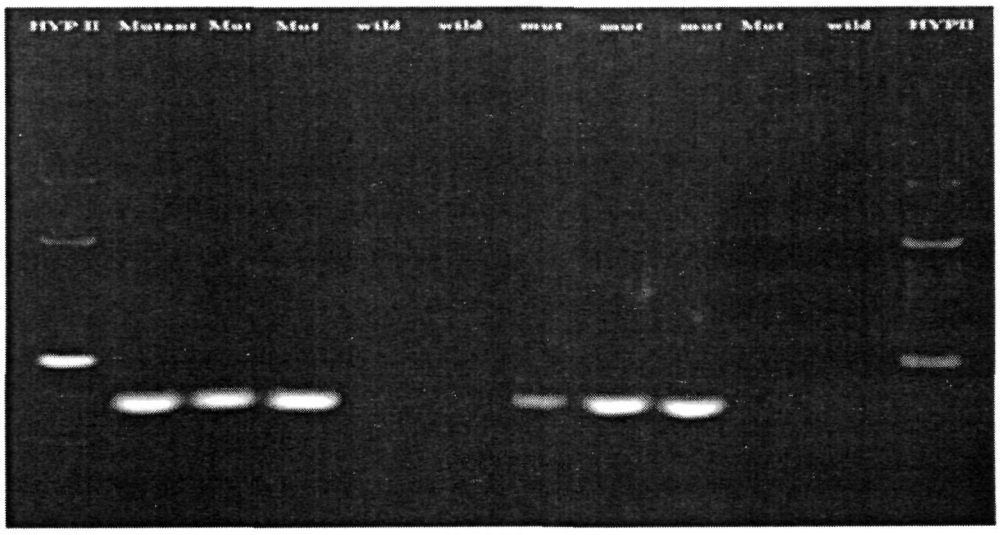


Figure 4.20 PCR amplification through Tc concept, Mutant type of 3889DelAG was amplified, whereas wild type had no amplification, amplicon length approx. 192bp.

Age group of patients	Total number of patients	Mutation Positive cases	Family History and Personal History of patients			
			F.H	No F.H	P.H	No P.H
≤40	3 (9.37%)	1 (33.33%)	1 (3.12%)	2 (6.25%)	0	3 (9.37%)
≥40	29 (90.62%)	12 (37.5%)	3 (9.37%)	26 (81.25%)	2 (6.25%)	27 (84.37%)
Total	32	13 (40.62%)	4 (12.5%)	28 (87.5%)	2 (6.25%)	30 (93.75%)

Table 4.4 Characteristics of sequenced breast carcinoma patients. F.H – family history, No F.H – without reported family history, P.H- Personal history, No P.H- without reported personal history

Gene	Exon	NT	Base change	Codon	AA change	BIC Designation	Variation Type	Reported	Ethnicity	Cases found	Mutation Rate
BRCA1	2	185	DelAG	23	Stop 39	185DelAG	Truncated protein	Ashkenazi Jews, others	Indian	1	3.12%
BRCA1	11a	1014	DelGT	299	Stop 303	1014DelGT	Truncated protein	Various, Pakistani	Indian	3	9.37%
BRCA1	11d	3889	DelAG	1257	Stop 1265	3889DelAG	Truncated protein	Various, Chinese	Indian	9	28.12%

Table 4.5 *BRCA1* Deleterious mutation in North-Eastern breast cancer patients, India. NT- Nucleotide Position, AA Change- Amino Acid Change, BIC Designation - Breast Cancer information core Designation.

4.4 Low penetrance candidate genes involvement in breast cancer

GSTs are a family of enzymes involved in detoxication of benzo(a)pyrene and other carcinogens found in tobacco smoke, cytotoxic drugs, and chemical solvents. Deletions in two GST genes, *GSTM1* and *GSTT1*, occur at frequencies of 15% or greater in human populations. Individuals, who are deletion homozygotes, classified as *GSTM1* null or *GSTT1* null, exhibit absence of enzymatic activity and are hypothesized to be at increased risk for the carcinogenic effects of a wide range of environmental exposures. Associations between *GSTM1* and *GSTT1* null genotypes and cancer of the lung, breast, bladder, and colon have been reported, but results are inconsistent across studies. An amino acid substitution variant in a third GST gene, *CYP1* codon 105 Ile3Val, has been identified recently that encodes an enzyme with reduced catalytic activity. The *CYP1* Val allele is common in human populations but has not been extensively examined in association with cancer. Several previous studies investigated *GSTM1* and *GSTT1* genotypes and breast cancer risk, and one study examined the role of *CYP1* genotype. We examined the relation of *GSTM1*, *GSTT1*, and *CYP1* genotypes and breast cancer risk in Southern Assam with the help of Multiplex PCR (Figure 4.21). To address issues raised by previous studies, we experimented main effects for each GST locus through multiplex PCR techniques. We investigated joint effects for combinations of GST genotypes, as well as joint effects for GST genotypes and *BRCA1* genotype, a gene involved in detoxication of catechol estrogens. DNA was extracted from peripheral blood lymphocytes using standard methods. Genotyping for *GSTM1*, *GSTT1*, and *CYP1* was conducted using multiplex PCR methods, Polymorphism analysis of *GSTT1* and *GSTM1* genes was detected by Multiplex PCR based assays with *CYP1A1* as an internal control gene (Mondal et al., 2012). The CYP1A1 gene primer pair were; Forward- 5'-GAA CTG CCA CTT CAG CTG TCT-3', and Reverse- 5'-GCT GCA TTT GGA AGT GCT C-3'). In addition to the CYP gene, two

sequence specific oligonucleotide primers (GSTT1 or GSTM1) were used for multiplex PCR. For GSTT1, Forward Primer 5'-TTC CTT ACT GGT CCT CAC ATC TC-3' and Reverse Primer 5'-TCA CGG GAT CAT GGC CAG CA-3' and for GSTM1, Forward Primer 5'-GAA CTC CCT GAA AAG CTA AAG C-3' and Reverse Primer 5'-GTT GGG CTC AAA TAT ACG GTG G-3' were used. The genotyping assays for GSTM1 and GSTT1 classify individuals with one or two copies of the relevant gene as "present" and individuals with homozygous deletions as "null." The assay for CYP1 classifies individuals according to the alleles Ile and Val. Assays that were unreadable for each locus are reported as "missing." Approximately half of the samples listed as missing did not amplify for the relevant locus, and the remaining amplified too poorly to assign genotypes. Positive and negative control samples were included with each batch of samples.

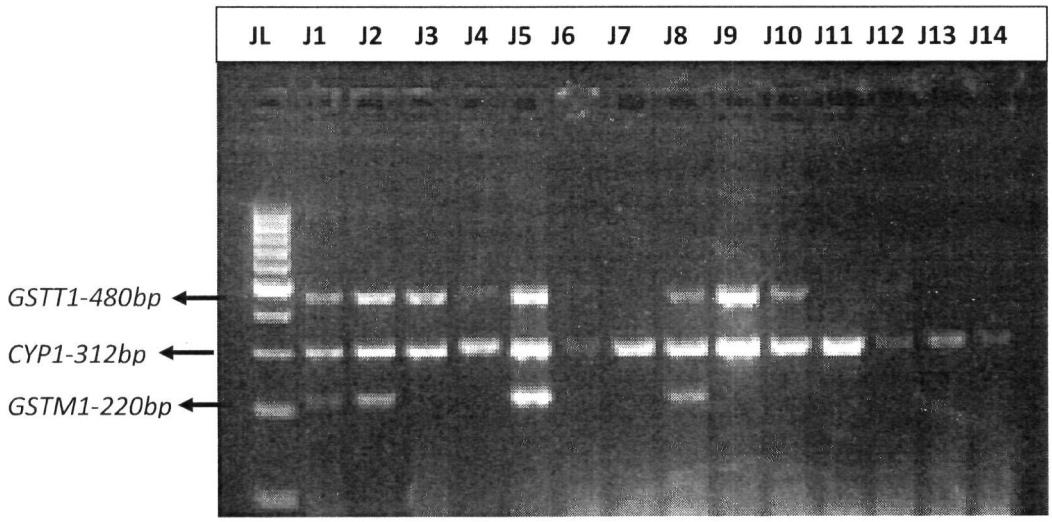


Figure 4.21 Multiplex PCR amplification through *GSTT1, CYP1* and *GSTM1* primer pairs from breast cancer tissue DNA in 2% agarose gel. The product length of *GSTT1, CYP1* and *GSTM1* are approximately 480bp, 312bp and 220bp. Lane J1, J2, J5, J8 contains all the genes. Lane J7, J11, J12, J13, J14 samples were Null for both gene. Lane J3, J4, J6, J7, J9, J10 samples were Null for *GSTM1*.

Gels were scored by two different readers, and discordant samples were repeated. Reliability was assessed by selecting a random sample of 10% of samples from each batch. Batches with, 95% agreement was rerun. In the studied population the GST null genotype examined where it is observed that the frequency of *GSTM1* null genotype was found to be 20% in cases and 13% in controls. The *GSTT1* null genotype is 32% in cases and 20% in controls and both *GSTT1* and *GSTM1* null genotype 28% in cases and 15% in controls (Table 4.6). We observed a risk of 2.05-fold to breast cancer patient (95% CI, 1.23-7.56, P = 0.01) due to null genotypes of GSTM1 and further the risk increases 1.6-fold (95% CI, 1.01-7.02, P = 0.04) due to null *GSTT1*. The risk increases further to 2.63-fold (95% CI, 1.43-10.80, P = 0.006) with both *GSTM1* and *GSTT1* null genotypes.

Polymorphisms in *GSTT1* and *GSTM1* genes and breast carcinoma

The distribution of the *GSTT1* and *GSTM1* genotypes in the patient is shown in Table 4.6. A significant contribution was found in between mutations of BRCA1 and polymorphism *GSTT1* and *GSTM1* in the risk of breast carcinoma (OR=1.60, P=0.02). The lack of the *GSTM1* gene was significantly associated with the poorer mutation rate (OR=2.29, P=0.03). This contribution was significantly higher in patients carrying both null-*GSTT1* and *GSTM1* genotypes. When we stratified the patients according to their mutation status of the *BRCA1* gene, we found that the null-GSTT1 genotype frequency was significantly higher in mutated patients than in controls (0.360 vs 0.260, OR=1.60, P=0.02). Thus, there was an association between the presence of the null-*GSTT1*, null-*GSTM1* genotype and the mutation of the *BRCA1* gene in breast carcinoma patients. Most of the genotypes *GSTT1*, either alone or in combination with *GSTM1* genotypes, were associated with breast carcinoma in the different subgroups. Genotype frequencies and 95% CI for *GSTM1*, *GSTT1*, and *CYP1* were calculated as the proportion of individuals with a given genotype divided by the total number of participants. For *CYP1*,

Patients ID	*GSTT1*	*GSTM1*	Patients ID	*GSTT1*	*GSTM1*
AUBC20	Present	Present	**AUBC40**	Null	Present
AUBC21	Present	Null	**AUBC41**	Present	Present
AUBC22	Present	Present	**AUBC42**	Null	Null
AUBC23	Present	Present	**AUBC43**	Null	Null
AUBC24	Present	Present	**AUBC44**	Present	Null
AUBC25	Null	Present	**AUBC45**	Present	Null
AUBC26	Null	Present	**AUBC46**	Present	Present
AUBC27	Present	Present	**AUBC47**	Null	Present
AUBC28	Null	Null	**AUBC48**	Present	Present
AUBC29	Present	Present	**AUBC49**	Present	Present
AUBC30	Null	Present	**AUBC50**	Present	Present
AUBC31	Null	Null	**AUBC51**	Present	Null
AUBC32	Null	Present	**AUBC52**	Present	Present
AUBC33	Null	Present	**AUBC53**	Present	Present
AUBC34	Present	Present	**AUBC54**	Null	Null
AUBC35	Present	Present	**AUBC55**	Present	Present
AUBC36	Null	Null	**AUBC56**	Null	Present
AUBC37	Null	Present	**AUBC57**	Null	Present
AUBC38	Present	Present	**AUBC58**	Null	Null
AUBC39	Null	Present	**AUBC59**	Null	Null

Table 4.6 GSTs polymorphism in breast cancer patients from Southern Assam.

allele frequencies and 95% CI were calculated as the number of alleles divided by the number of chromosomes, and tests for Hardy-Weinberg equilibrium were conducted by comparing observed and expected genotype frequencies using a x2 test (16). Adjusted OR for breast cancer and 95% CI were calculated from unconditional logistic regression models.

Our results suggest that *GSTM1*, *GSTT1*, and *CYP1* genotypes do play a role in susceptibility to breast cancer, in agreement with most previous studies. However, inability to detect effects for GSTs could result from failure to include relevant environmental exposures or genes that interact with GSTs. The presence of positive associations for GSTs in women with a family history suggests that unknown genetic or environmental exposures may modify the effects of GST genes, a hypothesis that could be investigated further in family-based association studies. Unmeasured genetic or environmental factors that interact with GSTs could also contribute to differences in results across epidemiological studies. A potential role for GST genotypes in breast cancer prognosis and response to treatment, as well as the possibility that *GSTM1* status might modify age at onset for breast cancer, also merit further investigation.