Association between Estrogen Biosynthesis Enzyme CYP17 Gene Polymorphisms and Breast Cancer in Sudanese Women

Mihreteab Ghebretinsae Tekle

B.Sc. in Chemistry, University of Asmara (2003)

A Dissertation

Submitted to University of Gezira in Partial Fulfillment of the Requirements for the Award of the Degree of Master of Science

in

Biochemistry

Department of Biochemistry and Nutrition

Faculty of Medicine

June - 2014

Association between Estrogen Biosynthesis Enzyme CYP17 Gene Polymorphisms and Breast Cancer in Sudanese Women

Mihreteab Ghebretinsae Tekle

Supervision Committee:-

Name	**Position**	**Signature**
Dr. Hani Yousif Zaki Tadros	Main Supervisor	
Dr. Areeg Saad Abdalla Faggad	Co-supervisor	

Date: 07-June-2014

Association between Estrogen Biosynthesis Enzyme CYP17 Gene Polymorphisms and Breast Cancer in Sudanese Women

Mihreteab Ghebretinsae Tekle

Examination Committee:-

Name	**Position**	**Signature**
Dr. Hani Yousif Zaki Tadros	Chairperson	
Dr. Babkir Ali Mohamed Elamin	External Examiner	
Dr. Ahmed Mohammed Elhaj Musa	Internal Examiner	

Date of Examination: 07-June -2014

Dedication

TO MY FAMILY

Acknowledgments

I would like to express my gratitude to my supervisor, Dr. Hani Yousif Zaki, Faculty of Medicine, University of Gezira and Co-supervisor Dr. Areeg Saad Faggad, Faculty of Medicine, University of Gezira, their expertise, understanding, kindness, persistence, enthusiasm, and patience, added considerably to my graduate experience. I would also like to thanks them for encouraging and helping me to shape my interest and ideas. I am forever grateful. Thanks Dr. Hani and Dr. Areeg!

A very special thanks goes out to Dr. Bedreldin Elsonni Abdella for the assistance he provided to finalize this thesis. I appreciate him for kindness and support he has shown during the past two years.

I would also like to thank my family for the love, support, encouragement they provided me through my entire life. They've taught me about hard work and self-respect, about persistence and how to be independent. I am proud of them and love them very much.

I would like to thanks all staff of Department of Biochemistry, Faculty of Medicine, University of Gezira and my beloved classmates Amanuel, Asmerom, Khalda, Nagla, Nisreen, Yasmin, and Zeneb. Thank you for your encouragement and support.

I would like to thank Miss Amel Abdeen, a molecular biology technologist at the Department of Molecular Biology at National Cancer institute –University of Gezira for the technical advice and great help in running the experiments for DNA extraction.

I would like to acknowledge all my friends who supported me during my time here with special thanks to Abduselam, Asmerom, Gossai, Hiba, Mengstu, Muna, Suzan, Um-Salama and Yosuf.

I would like to extend my gratitude and thanks to all those breast cancer patients and healthy individuals participants of this study. I wish them all the best and great success throughout their entire life!

Finally, I am grateful for the funding sources that allowed me to pursue my graduate school studies: The Eritrean National Board for Higher Education and African Development Bank.

Association between Estrogen Biosynthesis Enzyme CYP17 Gene Polymorphisms and Breast Cancer in Sudanese Women

Mihreteab Ghebretinsae Tekle

Abstract

Estrogen hormones are key regulators of a wide variety of biological processes. In addition to their influence on reproduction, cell differentiation and apoptosis, they affect inflammatory response, cell metabolism and most importantly, they regulate physiological breast tissue proliferation and differentiation as well as the development and progression of breast cancer. The purpose of this study was to investigate the association between the estrogen biosynthesis enzyme encoding gene *CYP17A1* polymorphisms (rs743572 and rs6162) and breast cancer in Sudanese females. This case-control study included seventy one breast cancer cases, diagnosed at National Cancer Institute –University of Gezira, Wad-Medani, Sudan in the period from January 2012 to February 2014, and seventy three healthy individuals as a control having no evidence of any personal or family history of cancer were recruited in this study. Polymerase chain reaction-restriction fragment length polymorphism (PCR-RFLP) was used to determine the *CYP17A1* (rs743572 and rs6162) gene polymorphisms. In this study, the frequency distribution of *CYP17A1* (rs743572) genotypes showed a significant association between *CYP17A1* polymorphism and breast cancer (P=0.004). The *CYP17A1* (rs743572) CC genotype was associated with increased risk for breast cancer (OR =1.32, 95% CI: 0.21-8.22). Women carrying C allele in *CYP17A1* (rs743572) gene had 1.48-fold increased breast cancer risk. Genotype was shown to be significantly associated with breast cancer for *CYP17A1* rs6162 SNP (P = 0.034). Both homoyzygous TT and heterozygotes CT genotypes were associated with decreased risk of having breast cancer compared to those with the homoyzygous CC wild type genotype. In conclusion, this study provides support that polymorphic variation in *CYP17A1* rs743572 might play a role in breast cancer susceptibility, and homozygous TT carriers of the *CYP17A1* rs6162 gene is inversely related to breast cancer risk in Sudanese women. The study revealed new associations of *CYP17A1* (rs6162) with breast cancer risk factors which might be used as base-line new additional findings that is recommended to be investigated in a further population-based study.

العلاقة بين تعدد الأشكال لجين انزيم تصنيع هرمون الاستروجين **CYP17** و سرطان الثدي لدى النساء السودانيات

مهراتاب قبرتنسيا تكلي

ملخص الدراسة

هرمون الاستروجين هو المنظم الرئيسي لعدد من العمليات البيولوجية. بالإضافة إلى تأثيره على الإنجاب وتمايز وموت الخلايا ، فإنه يؤثر على الأستجابة الالتهابية ، واستقلاب الخلية و الأهم من ذلك انه يعمل علي تنظم فسيولوجيا تكاثر وتمايز أنسجة الثدي ، وكذلك يعمل علي تطوير سرطان الثدي. الغرض من هذه الدراسة استقصاء العلاقة بين تعدد الأشكال لجين انزيم تصنيع هرمون الاستروجين سايتوكروم ب أ 1 (rs743572 و rs6162) وسرطان الثدي عند النساء السودانيات. شملت هذه الدراسة واحداً وسبعين حالة لسرطان الثدي شخصت في المعهد القومي للسرطان ، جامعة الجزيرة ، ود مدني ، السودان في الفترة من يناير 2012 إلى فبراير عام 2014، و ثلاثة وسبعين من النساء الصحيحات مع عدم وجود أي تاريخ شخصي أو عائلي للإصابة بسرطان الثدي كمجموعة ضابطة. تم استخدام تفاعل البلمرة التسلسلى وتحديد طول الجزء المتعدد الأشكال. أظهرت الدراسة ان التوزيع التكرارى للنمط الوراثى سايتوكروم ب أ 1 rs743572 اوضح وجود علاقة ايجابية ذات دلالة احصائية بين تعدد الأشكال لجين سايتوكروم ب أ 1 وسرطان الثدي (P = 0.004) . النمط الوراثي للجين CC (rs743572) في جين سايتوكروم ب أ 1 له علاقة مع زيادة خطر الإصابة بسرطان الثدي (OR = 1.32، 95٪ CI : 0،-21 22،8) . النساء اللاتي يحملن أليل C في سايتوكروم ب أ 1rs743572 كان لهن 1.48 أضعاف زيادة خطر الاصابة بسرطان الثدي. اوضح النمط الوراثي وجود علاقة ايجابية ذات دلالة احصائية بين الاصابة بسرطان الثدي و جين سايتوكروم ب أ rs6162 (P = 0.034) . النمط الوراثي المتجانس TT والنمط الوراثي غير المتجانسCT لهما علاقة مع انخفاض خطر وجود سرطان الثدي مقارنة مع اللائى لديهن النمط الوراثي المتجانس CC. خلصت هذه الدراسة الى أن تعدد الأشكال لجين انزيم تصنيع هرمون الاستروجين سايتوكروم ب أ 1 rs743572يلعب دورا في الوقاية من الاصابة بسرطان الثدي، والنمط الوراثي المتجانس TT في الجين سايتوكروم ب أ 1- rs6162 يرتبط عكسياً مع خطر الاصابة بسرطان الثدي في النساء السودانيات . كشفت هذه دراسة عن وجود ارتباط جديد بين تعدد الأشكال لجين انزيم تصنيع هرمون الاستروجين سايتوكروم ب أ 1 و عوامل الخطورة للاصابة بسرطان الثدي، هذه النتيجة يمكن استخدامها كنتيجة اولية لدراسات إضافية مستقبلية يوصى في مجموعات أكبر من السكان .

Table of Contents

	Page
Title page	i
Supervisor committee	ii
Examination Committee	iii
Dedication	iv
Acknowledgments	v
Abstract (English)	vi
Abstract (Arabic)	vii
Table of Contents	viii
List of Abbreviations	ix
Chapter one Introduction	1
Literature Review	8
Justification of the Study	11
Objectives of the Study	12
Chapter Two Subjects, Materials and Methods	13
Chapter Three Results	19
Chapter Four Discussion	30
Chapter Five Conclusions	32
Recommendations	32
References	33
Appendices	36

List of Abbreviations

BMI	Body Mass Index
BRCA1	Breast cancer type 1
BRCA2	Breast cancer type 2
CI	Confidence Interval
COMT	Catechol-O-methyltransferase
CYP1A1	Cytochrome P450, family 1, subfamily A, polypeptide 1
CYP1B1	Cytochrome P450, family 1, subfamily B, polypeptide 1
CYP17A1	Cytochrome P450, family 17, subfamily A, polypeptide 1
CYP19A1	Cytochrome P450, family 19, subfamily A, polypeptide 1
DHEA-S	Dehydroepiandrosterone sulfate
ER	Estrogen receptor
ESR1	Estrogen receptor 1
ESR2	Estrogen receptor 2
EREs	Estrogen Response Elements
FNAC	Fine needle aspiration cytology
HER	Human epidermal growth factor receptor
HRT	Hormone Replacement Therapy
HSD17B1	17ß-hydroxysteroid dehydrogenase type 1
GATA3	GATA Binding Protein 3
GSTM1	Glutathione S-Transferase Mu 1
LN	Lymph Node
MCF-7	Michigan Cancer Foundation-7
NCBI	National Center for Biotechnology Information
OR	Odds Ratio
PR	Progesterone Receptor
RFLP	Restriction Fragment Length Polymorphism
PON1	paraoxonase/arylesterase 1
SNP	Single Nucleotide Polymorphism
SULT1A1	Sulfotransferase Family, Cytosolic, 1A, Phenol-Preferring, Member 1
TF	Transcription Factor
TSS	Transcription Start Site
UTR	Untranslated Region
UGT1A1	UDP glucuronosyltransferase 1 family, polypeptide A1

Chapter One

1. Introduction and Literature Review

1.1. Background

1.1.1. The Breast Structure and Function

The female breast and mammary gland are structures derived from skin and are functionally related to the female sex organs. They develop during puberty under hormonal influence and are composed of glandular, adipose, and connective tissues (Figure 1.1). During pregnancy, the duct system and the glands proliferate under the influence of estrogens and progesterone, the breast increasing considerably in size. During the last month of pregnancy, the breast at first secretes a precursor to the milk, the colostrum, which consists of droplets of fat and sloughed cells. After delivery, prolactin— a pituitary hormone leads to milk secretion. Later, a posterior pituitary hormone, oxytocin, promotes lactation by stimulating specialized cells in the gland (Schuenk, 2004).

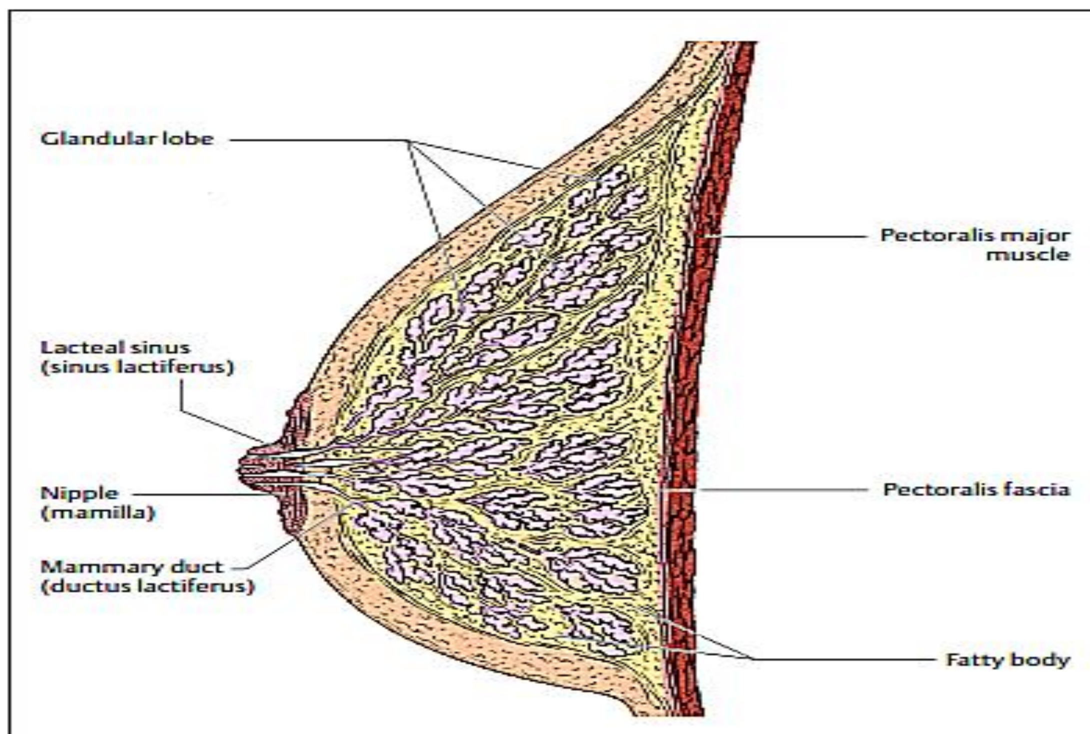

Figure 1.1 Longitudinal section through a female breast (Schuenk, 2004)

1.1.2. Cancer

Cancer remains a major public health problem around the world. It is a leading cause of death in economically developed countries and developing countries (Siegel *et al.*, 2013).

Understanding of cancer has changed over the years, owing to rapid advances in oncology research. The disease itself is not only characterized as a mass of excessive, uncontrolled growth of abnormal cells but is also defined by the dynamic alterations in the genome that cause cancer (Macconaill and Garraway, 2010).

DNA damage response genes play vital roles in the maintenance of a healthy genome. Defects in cell cycle checkpoints and DNA repair genes, especially mutation or aberrant downregulation, are associated with a wide spectrum of human disease, including a predisposition to the development of neurodegenerative conditions and cancer. On the other hand, upregulation of DNA damage response and repair genes can also cause cancer, as well as increase resistance of cancer cells to DNA damaging therapy (Broustas and Lieberman, 2014).

Genomes are transmitted faithfully from dividing cells to their offspring. Changes that occur during DNA repair, chromosome duplication, and transmission or via recombination provide a natural source of genetic variation. They occur at low frequency because of the intrinsic variable nature of genomes, which refer to as genome instability. However, genome instability can be enhanced by exposure to external genotoxic agents or as the result of cellular pathologies (Aguilera and Garcia-Muse, 2013).

1.2. Breast Cancer

1.2.1. Epidemiology of breast cancer

A predicted global burden of 20.3 million new cancer cases by 2030 compared with an estimated 12.7 million cases in 2008, and a predicted 13.2 million cancer-related deaths worldwide by 2030, up from 7.6 million in 2008. Breast cancer continues to be the most lethal malignancy in women across the world. In 2008, approximately 1.4 million women were diagnosed with breast cancer worldwide with corresponding 460000 death (Ferlay *et al.*, 2010).

Breast cancer is the most commonly diagnosed cancer in women and the second leading cause of death in women in Africa, with much geographical variation in incidence and mortality within the continent (Jemal *et al.*, 2012).

In Sudan the commonest cancers in women are those of the breast, cervix and ovary, with breast and cervical cancer accounting for 50% of all cancers among Sudanese women (Hamad, 2006). The frequency of breast cancer among those presenting with a palpable breast lesion is high in Sudan. Most patients with breast cancer present late, because they lack awareness and do not have access to screening programs. Breast cancer awareness and cancer screening helps detect breast cancer at an early stage, and this would improve the outcome (Ahmed *et al.*, 2010).

A descriptive study which was conducted to assess the type, stage and age distribution of breast cancer in female patients living central Sudan revealed that the majority of the patients were less than 50 years old or premenopausal. Invasive ductal carcinoma was the most common pathology, and women presented with stage III or higher tumors that had already metastasized (Elgaili *et al.*, 2010).

1.2.2. Risk factors of breast cancer

The incidence of breast cancer increases with age, doubling about every 10 years until the menopause. Women who start menstruating early in life or who have a late menopause have an increased risk of developing breast cancer. Women who have a natural menopause after the age of 55 are twice as likely to develop breast cancer as women who experience the menopause before the age of 45. At one extreme, women who undergo bilateral oophorectomy before the age of 35 have only 40% of the risk of breast cancer of women who have a natural menopause (McPherson *et al.*, 2000).
High animal fat and red meat or processed meat intake with low physical activity were linked with increased risk of breast cancer. The plant fat dietary pattern was negatively associated with breast cancer in sedentary women (Kruk and Marchlewicz, 2013).

Obesity and its metabolic complications have recently becomes major global issues and are associated with increased risk for cancer, especially with breast cancer in postmenopausal women (Dalamaga, 2013). A significant positive relationship was observed between breast cancer risk and the degree of smoking. Alcohol consumption was also found to be a risk factor (Gao *et al.*, 2013).

A significantly greater risk of breast cancer was associated with long duration of oral contraceptive use (Zhu *et al.*, 2012). The risk of breast cancer is greater for formulations that contain both estrogen and progesterone, compared with estrogen alone. The breast cancer risk associated with hormone replacement therapy (HRT) is higher for estrogen receptor-positive cancers than for estrogen receptor-negative cancers, and for low-grade cancers compared with high-grade cancers. After cessation of HRT the increased risk of breast cancer dissipates within 2 years. The rapidity of the decline suggests that a proportion of breast cancers that are hormone dependent will regress if the hormonal stimulation is removed (Narod, 2011).

Breast cancer can be inherited, only 10% of the cases are heritable and, caused by high-penetrance cancer susceptibility genes, such as *BRCA1* (breast cancer 1) and *BRCA2* (breast cancer 2). However, most sporadic breast carcinomas that have relatively late age onset are likely to be caused by low-penetrance genes acting together with endogenous/lifestyle risk factors. The genes encoding enzymes involved in estrogen metabolism are among these low-penetrance genes (Singh *et al.*, 2008).

1.2.3. Clinical and pathological classification of breast cancer

Current routine clinical management of breast cancer relies on availability of robust clinical and pathologic prognostic and predictive factors to support clinical and patient decision making where potentially suitable treatment options are increasingly available. The three strongest prognostic determinants in operable breast cancer used in routine clinical practice internationally are lymph node (LN) stage, primary tumor size, and tumor histologic grade (Bloom and Richardson, 1957). Histological grading is a combined score based on microscopic evaluation of the morphological and cytological features of tumor cells that reflects the aggressiveness of a tumor. This combined score is then used to stratify breast cancer tumors into three grades: grade 1, slow growing and well differentiated; grade 2, moderately differentiated; and grade 3, highly proliferative and poorly differentiated (Elston and Ellis, 2002). Current classification for solid tumors is based upon characteristics of their extent. Sizes of the primary tumor, presence of metastatic regional lymph nodes and/or of distant metastases (TNM) are the key elements for their categorization. Although many changes have been introduced in the TNM classification during the last decades, the main categories have been maintained unchanged. The size is still divided in T_1 T_2 T_3 categories according to the maximum

diameter of the tumor (less than 2 cm, from 2 to 5 cm and more than 5 cm) plus the T 4 category with special features (Veronesi *et al.*, 2006).

1.2.4. Molecular subtypes of breast cancer

Gene expression studies using DNA microarrays have identified several distinct breast cancer subtypes based on an intrinsic gene list that includes 496 genes that differentiate breast cancers into separate groups based only on gene expression patterns (Perou *et al.*, 2000). These subtypes differ markedly in prognosis and in the repertoire of therapeutic targets they express. The intrinsic subtypes include 2 main subtypes of estrogen receptor (ER)–negative tumors (basal-like and human epidermal growth factor receptor-2 positive/ER- (HER2+/ER-) subtype) and at least 2 types of ER+ tumors (luminal A and luminal B) (Figure 1.2). Basal-like tumors typically show low expression of HER2 and ER and exhibit high expression of genes characteristic of the basal epithelial cell layer, including expression of cytokeratins 5, 6, and 17. The HER2+ (i.e. gene amplified and/or highly over expressed protein) tumors fall into at least 2 distinct expression groups: those that are ER and typically cluster near the basal like tumors (HER2+/ER- subtype), and those that are ER+ (and may also be progesterone receptor positive [PR+]) and cluster with tumors of luminal cell origins as part of the luminal B subtype (Sorlie *et al.*, 2001). The luminal subtype A and B tumors express ER, GATA3, and genes regulated by both ER and GATA3. Compared with luminal B tumors, luminal A tumors express higher levels of ER and GATA3 and show more favorable patient outcomes, whereas luminal B tumors more often express human epidermal growth factor receptor-1 (HER1), HER2, and/or cyclin E1 (Sorlie *et al.*, 2003).

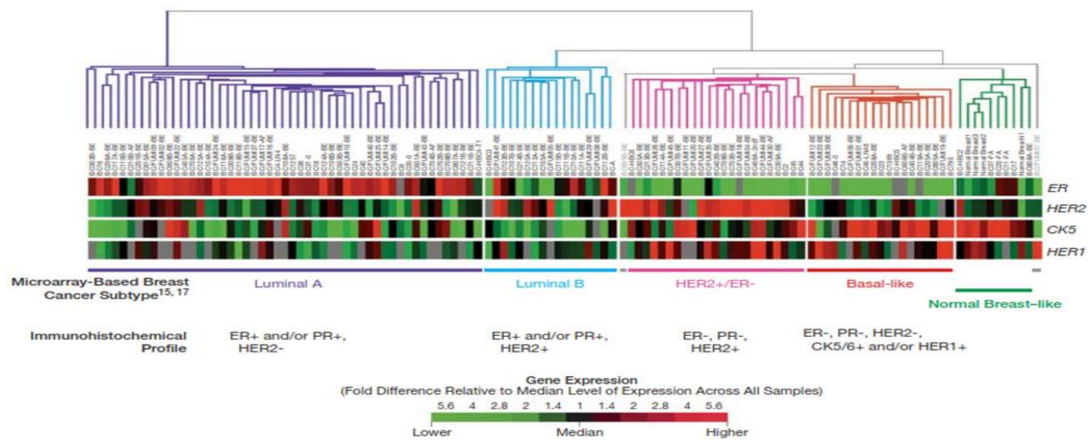

Figure 1.2 Immunohistochemical and molecular identification of breast tumor intrinsic subtypes (Carey *et al.*, 2006).

1.2.5. Diagnosis of breast cancer

The standard diagnosis of breast cancer recommended triple diagnostic investigations, comprising (1) clinical breast examination; (2) breast imaging: ultrasonography, or mammography, (3) cytopathologic with fine needle aspiration cytology (FNAC) or histopathologic findings with a core needle biopsy with cutting needle biopsy. Although mammography is the most commonly used imaging modality for breast screening, its effectiveness may be reduced in women with dense breasts and its sensitivity is poor for young women (Berg *et al.*, 2008). It has been widely reported that ultrasonography is more sensitive than mammography in breast cancer diagnosis (Huang *et al.*, 2012). Fine needle aspiration cytology is sensitive, simple, cost effective less traumatic and rapid method. It can be easily repeated if an adequate sample is not obtained (Nasuti *et al.*, 2002).

1.2.6. Management of breast cancer

Recent information about tumor biology and its behavior suggested that less radical surgery might be just as effective as the more extensive one. Eventually, with the use of adjuvant therapy like radiation and systemic therapy, the extent of surgical resection in the breast and axilla got reduced further and led to an era of breast conservation. The radiation treatment of breast cancer has evolved from 2D to 3D conformal and to accelerated partial breast irradiation, aiming to reduce normal tissue toxicity and overall treatment time. Systemic therapy in the form of hormone therapy, chemotherapy and biological agents is now a well-established modality in treatment of breast cancer. The current perspective of breast cancer management is based on the rapidly evolving and increasingly integrated study on the genetic, molecular, biochemical and cellular basis of disease (Akram and Siddiqui, 2012).

It is generally accepted that cancer arises because of an accumulation of multiple molecular genetic defects that culminate in a cellular phenotype characterized by unregulated growth. Based on the knowledge, a variety of gene therapy strategies have been developed as potential new therapies for cancer (Sharma *et al.*, 2010).

1.3. Estrogens

1.3.1. Biosynthesis and metabolism of estrogens

The naturally occurring estrogens are 17β-estradiol, estrone, and estriol. They are C18 steroids which do not have an angular methyl group attached to the 10 position or a Δ^{4}-3-keto configuration in the A ring. Estrogen biosynthesis, cellular binding and metabolism involve many steps (Figure 1.3), and the genes controlling these steps may contribute to inherent variability in breast cancer susceptibility. Endogenous estrogens are produced predominantly in the ovarian theca cells in premenopausal women and in the breast stromal adipose cells in postmenopausal (Ahsan *et al.*, 2005). Estrogens have been clearly identified as carcinogens, by inducing aneuploidy and structural chromosomal changes, and stimulation of breast cell proliferation has been proposed as the main effect of estrogens in breast carcinogenesis — as the more rapidly cells proliferate, the greater the chance of acquiring a potentially cancer-causing mutation (Zhu and Conney, 1998).

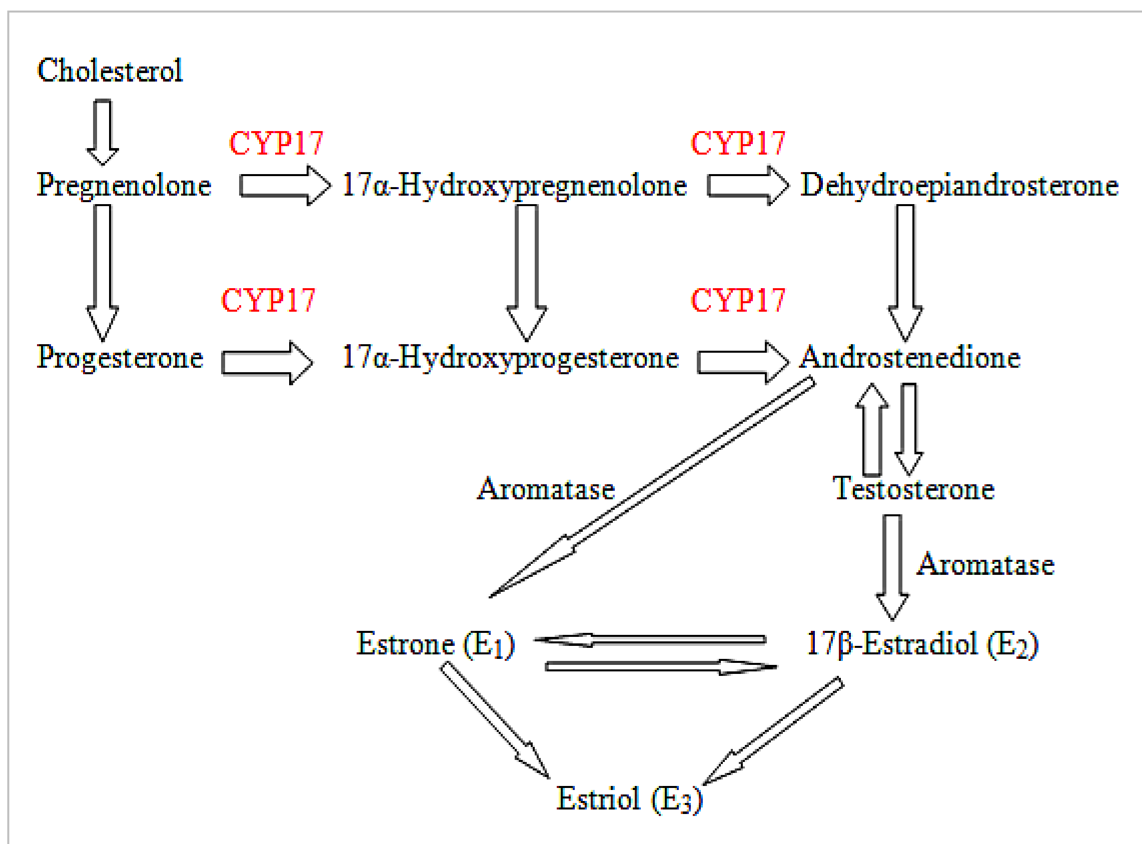

Figure 1.3 Biosynthesis of estrogens

1.3.2. Mechanism of action of estrogens

Cellular effects of estrogens are predominantly brought about by their interaction with estrogen receptors (ER). Upon ligand binding, the ERs in the cytoplasm dissociate from the heat shock proteins, dimerise and translocate to the nucleus. In the nucleus, the ligand bound ER interact directly or indirectly with specific regions of the DNA and act as a transcription factor (TF), recruit other TFs and coactivators/repressors, and thereby modulating the expression of the target genes (Heldring *et al.*, 2007). The short stretches of DNA, where the ligand-bound ER directly binds are called as Estrogen Response Elements (EREs). In addition to the direct binding, ER can form complexes with other TFs, and direct their binding to their respective response elements on the target genes (Klinge, 2001). The regulation of a single gene can be under the control of several TFs, including ER, with the binding regions occurring close (5–10 kb) to the transcription start site (TSS), or as distal elements (Maston *et al.*, 2006).

The cellular response to estrogens is closely linked to their concentration. In breast cancer cells abundant in ERs, such as MCF-7, estrogens cause marked proliferative effect at picomolar concentration (Katzenellenbogen *et al.*, 1987).

1.4. Contribution of genetic changes in cancer

Genetic polymorphism describes the co-occurrence in the same population of more than one allele, or genetic marker, at the same locus with the least frequent allele or marker occurring more frequently than can be accounted for by the mutation alone (Lichtenstein *et al.*, 2000). SNP can be defined as a genomic locus where two or more alternative bases occur with substantial frequency of greater than 1% (Cargill *et al.*, 1999). A total of around 10 million SNPs have been distributed throughout the human genome at a frequency of at least one in 1000 base pairs and exhibit low mutation rate (Botstein and Risch, 2003). The specific causes of cancer are not yet known. However, epidemiological studies suggest the relationship of genetic polymorphism to cancer susceptibility. Genetic association studies with single nucleotide polymorphisms (SNPs) linked to cancer are emerging area of research (Cargill *et al*., 1999).

1.5. Review of literature on estrogen metabolism and signaling in cancer

Many studies have been conducted to investigate the role of estrogen metabolism and estrogen signaling related genes polymorphisms in susceptibility to breast cancer risk. A

study conducted to investigate the role of estrogen signaling and synthesis related genes polymorphism (estrogen receptors, *ESR1* (rs2234693) and *ESR2* (rs2987983); estrogen biosynthesis enzymes, *CYP17A1* (rs743572); and aromatase, *CYP19A1* (rs700519) in susceptibility to breast cancer risk in a case-control approach. All the genetic variants were significantly associated with breast cancer risk among North Indian women (Chattopadhyay *et al.*, 2014).

In African-American women a study was conducted on selected single nucleotide polymorphisms (SNPs) in genes involved in estrogen metabolism or action, specifically, genes which include *CYP 17*, *HSD17B1*, *CYP1A1*, *CYP3A4*, *CYP1B1*, and *ESR2.* It was shown that *CYP1A1* 5639C allele was associated with an increased risk of breast cancer. In addition, African American-predominant *CYP1B1* 432 Valine (Val) allele was significantly more often found in the cases than in the controls and the *HSD17B1* 312 Glycine (Gly) allele was specifically associated with premenopausal breast cancer risk (Kato *et al.*, 2009).

A case-control research done in USA to evaluate whether variation in genes related to estrogen metabolism (*COMT, CYP1A1, CYP1B1, CYP17A1, CYP19A1, ESR1, GSTM1, GSTP1, GSTT1, HSD17B1, SULT1A1,* and *UGT1A1*) contributes to breast cancer risk, no appreciable associations with breast cancer risk were detected for single SNPs or haplotypes in women overall (Reding *et al.*, 2012).

A study was performed to compare genetic polymorphisms of ERα, estrogen metabolism genes (*CYP17A1, CYP19A1, HSD17B1, CYP1B1 and COMT*), and *p53* between ER-positive and ER-negative female Japanese breast cancer patients. The findings revealed that carriers of the G allele of ERα (rs6905370) were more frequent in ER-positive breast cancer than in ER-negative breast cancer especially in those under 50 years old. Pairwise analysis showed that combinations of the ERα G allele with the homozygous Trp genotype of *CYP19A1* codon 39 (rs2236722), the methionine (Met) allele of COMT codon 158 (rs4680) or proline (Pro) allele of *p53* codon 72 (rs1042522) were more frequent in ER-positive than ER-negative breast cancer, especially in patients younger than 50 years old. The frequencies of these combinations were even higher in patients with strongly ER-positive tumors (Hamaguchi *et al.*, 2008).

The association of polymorphisms of genes encoding enzymes involved in estrogen biosynthesis pathway and in the metabolic activation of pro-carcinogens to genotoxic intermediates, such as *CYP17*, *GSTP1* and *PON1* with the risk of breast cancer was undertaken in Italian population. *CYP17* polymorphism had no major effect in breast cancer proneness in the overall population. However, it modified the risk of breast cancer for certain subgroups of patients. As to *GSTP1* and *PON1* 192 polymorphisms, the mutant Valine (Val) and Arginine (Arg) alleles respectively, were associated with a decreased risk of developing breast cancer (Antognelli *et al.*, 2009).

Several studies tested the association between a T/C nucleotide substitution at position 34 in the 5' untranslated region (UTR) promoter of *CYP17A1* gene (rs743572) polymorphism with breast cancer and prostate cancer risk in different populations. However, their findings remain inconclusive. The association between *CYP17A1* (rs743572) polymorphisms with breast cancer risk were mainly evident in premenopausal in Thailand women based on a stratified analysis by menopausal status (Sangrajrang *et al.*, 2009). A polymorphism in the *CYP17A1* gene (rs743572) associated with an increased risk of advanced breast cancer in women carrying C allele in USA (Feigelson *et al.*, 1997). No significant association was found between the *CYP17A1* rs743572 C allele and overall risk of breast cancer in Finnish population. A protective effect of later age at menarche was suggested to be mainly attributable to women with the TT genotype. The effect was seen only among premenopausal women but not the whole population (Mitrunen *et al.*, 2000).

A few studies investigated the association between the C/T nucleotide substitutions at codon 46 in the Exon 1 region of *CYP17A1* rs6162 gene polymorphism with prostate cancer risk in different populations, but the results were inconsistent. Two independent cohorts studies composed of Caucasian men with organ-confined prostate cancer and Taiwanese men on androgen–deprivation therapy were conducted. It was reported that the presence of *CYP17A1* polymorphisms (rs6162) associated with disease progression and mortality in Caucasians and Asians (Levesque *et al.*, 2013). Recently, a research conducted in Japanese men, found that *CYP17A1* rs6162 genetic polymorphisms are associated with the risk of progression to castration-resistant prostate cancer after the initial hormonal therapy for prostate cancer (Yamada *et al.*, 2013).

1.6. Justification of this study

Since breast cancer is the most common cancer among women worldwide. It accounts for the highest morbidity and mortality and the leading cancer in Sudanese women.

Several patient-based studies suggested that estrogens, their metabolic compounds, and the entire biochemical metabolic machinery may play a role in breast carcinogenesis. CYP17 activity determines the local estrogen level. Polymorphisms in this gene can result in either increased or decreased enzyme activity, which directly affects estrogen levels, and may ultimately determines the occurrence, development, treatment, and prognosis of breast cancer.

Recently case–control association studies involving genetic polymorphisms have gained recognition as a potential useful tool for the prediction of breast cancer risk, as these studies can approach the genetic component of the disease in a more subtle and specific manner.

Therefore, testing the association of these variants in Sudanese patients with breast cancer may contribute to better understanding about the role of estrogen biosynthesis enzymes genes in breast cancer.

1.7. Objectives of the Study

1.7.1. General Objective

The aim of this study was to investigate the association between the estrogen biosynthesis enzyme encoding gene *CYP17A1* polymorphisms (rs743572 and rs6162) and breast cancer in Sudanese women.

1.7.2. Specific Objectives

1. To conduct genotyping analysis of SNPs rs743572 and rs6162 in *CYP17A1* gene.
2. To compare the genotype/allel frequencies of both SNPs in breast cancer patients and control groups.
3. To correlate the allel/genotype frequencies of both SNPS with the established disease risk factors: menopausal status, age at menarche, obesity, breast cancer family history, and number of parities.
4. To correlate the allel/ genotype frequencies of both SNPs with the clinical and pathological parameters: stage of breast cancer, lymph node status, estrogen and progesterone receptors status.
5. To conduct haplotype analysis of both SNPs

Chapter Two

2. Subjects, Materials and Methods

2.1. Study design, area and subjects:

This was a case- control study that included seventy one breast cancer patients with age ranged between 30 and 85years, all of them were diagnosed at National Cancer Institute –University of Gezira, Wad-Medani, Sudan in the period from January 2012 to February 2014, and seventy three healthy individuals as a control group (age ranged between 26 and 70 years having no evidence of cancer or family history of cancer were enrolled in this study.

2.2. SNP selection

Two SNPs in the *CYP17A1* gene were selected based upon the allelic or genotypes frequencies of African population information from the National Center for Biotechnology Information (NCBI, http://www.ncbi.nlm.nih.gov/).

2.3. Materials:

2.3.1. Questionnaire:

A questionnaire which included personal information, risk factors, clinical data, and family history was completed for every participant (Appendix-1).

2.3.2. Blood samples collection and storage: 5ml venous blood from each participant was taken into EDTA- tube using 10ml syringe; all the samples were stored at -50°C till DNA extraction.

2.3.3. Chemicals:

The chemicals used in this study included

- DNA extraction kits Lots No.: 15260546, Master Mix, TBE buffer with Cat. No.: IBS-BT04, and Red Safe nucleic acid staining solution with cat. No.: 21141 manufactured by iNtRON BIOTECHNOLOGY, Sujeong-gu, Korea.
- Tris base, EDTA, agarose powder with Cat. No.: MR7734, MR7732C, and PR911685 respectively and Nuclease free water manufactured by CinnaGen Co., Tehran, Iran.

- Forward and Reverse primers, Blue loading dye with lots No.: 111.663, and DNA ladder 50bp-1kbp with lots No.: 111.764 manufactured by Jena Bioscience GmbH, Dortmund, Germany.
- Restriction enzymes TspRI and MscI with lots No.: 0091305 and 0301311 respectively purchased from New England BioLabs, USA.

2.3.4. Instruments:

The following Instruments were used in this study:

- EBA 20 Centrifuge manufactured by Andreas Hettich GmbH & Co. KG, Föhrenstr. 12, D-78532 Tuttlingen, Germany.
- Vortex mixer model No.: G-650E manufactured by scientific industries, INC. Bohemia, N.Y. 11716, USA.
- Electric heating thermostatic water bath range with 37-100°c, model No.: HH.511-25
- PCR TC-3000 Thermal Cycler manufactured by Bar Loworld Scientific LTD, UK.
- Electrophoresis manufactured by Cleaver Scientific, Ltd, UK.
- Automated gel documentation system with Model No: GVMC20, Synoptic, LTD, UK.

2.4. Methods for Laboratory Investigations:

2.4.1. Washing buffer 10X TE preparation

12.1g of Tris base power and 9.3g of EDTA powder were dissolved in 500ml of distilled water.

2.4.2. Washing buffer 1x TE preparation

50ml of 10XTE buffer was taken in 500ml volumetric flask and diluted with distilled water until volume mark.

2.4.3. Preparation of the blood sample for DNA extraction

- 3-4ml of blood was taken into 15ml falcon tube and diluted with 1x TE buffer until the volume mark. The solution was vigorously mixed to homogenize, centrifuged at 5000rpm for 10minutes and then carefully the supernatant was discarded using dropper pipette (washing of WBCs).
- This washing step was repeated 2-3 times, till a clear pellet was obtained.

- 1ml of 1x TE buffer was added into the pellet after discarding the supernatant and mixed well using vortex mixer.
- Finally, the solution was carefully transferred into 1.5ml eppendorf tube, and stored at -50°c to be used later for DNA extraction.

2.4.4. DNA Extraction

- The sample in 1.5ml eppendorf tube was centrifuged at 14,000 rpm for 2 minutes. The supernatant was completely discarded.
- 20µl of Proteinase K and 5µl of RNase A were added into the sample and mixed gently.
- 200µl of Buffer BL was added into the sample and mixed thoroughly in order to assure efficient lysis.
- The lysate was incubated at 56°C for 10 min in a water bath.
- The 1.5ml eppendorf tube was centrifuged at 14,000 rpm for 1minutes to remove suspended drops from the inner side of the lid.
- 500µl of absolute ethanol was added into the lysate, and mixed well by gently inverting the tube 5- 6 times then the tube was centrifuged at 14,000 rpm for 1 minutes to remove drops from inner side of the lid.
- Carefully, the mixture from previous step was transferred to the Spin Column (in a 2 ml Collection Tube) without wetting the rim, the tube was capped and centrifuged at 14,000 rpm for 2min. The filtrate was discarded and the Spin Column was placed in a 2 ml collection tube.
- 700 µl of buffer WA was added to the Spin Column without wetting the rim, and centrifuged for 2 minutes at 14,000 rpm. Then the flow-through was discarded.
- 700 µl of buffer WB was added to the Spin Column without wetting the rim, and centrifuged for 2 minutes at 14,000 rpm. Then the flow-through was discarded and the spin column was placed into a 2.0 ml collection tube. Then again centrifuged for additionally 1 minute to dry the membrane. The flow-through together with the collection tube was discarded.
- The Spin Column was placed into a new 1.5 ml tube, and 100 µl of Buffer CE was added directly onto the membrane. Followed by incubation for 10 minutes at room temperature.
- Finally, a centrifugation for 2 minutes at 14,000 rpm was done to elute the DNA.

2.4.5. Determination of DNA Purity

- 5µL of extracted DNA samples was diluted in 10mM Tris buffer to a final volume of 1ml and mixed well.
- The spectrophotometer was set to zero at 260 nm with buffer, then the absorbance of the diluted DNA samples was measured.
- The spectrophotometer was set to zero at 280 nm with buffer, then the absorbance of the diluted DNA samples was measured.
- Finally the ratio of absorbance at 260nm and 280nm (A260/A280) was calculated to determine the purity of DNA. Pure DNA has an A260/A280 ratio of 1.7-1.9.

2.4.6. Genotyping Analysis of *CYP17A1* SNP (rs743572)

Genotyping analyses were performed by polymerase chain reaction- restriction fragment length polymorphism (PCR–RFLP) based method. *CYP17A1* (rs743572) genotyping was performed using PCR with confronting pair of primers:

Forward: -5-TCCACAAGGCAAGAGATAACACAAA-3

Reverse: -5-TACAGGTTGTTGGAGCTGGAAG-3.

The primers were designed using the Oligo software (Oligo 7 primer analysis software, Molecular Biology Insights, Inc., USA). The PCR reactions were carried out in 25µl reaction volume containing 2µl genomic DNA, 1µl of each primer, 5µl of master mix and 16µl nuclease-free water. The reaction was performed in the thermal cycler, and the DNA was denatured at 94°C for 4 min and amplified during 35 cycles (94°C for 30 seconds, annealing at 59°C for 30 seconds, extension at 72°C for 30 seconds), with final extension at 72°C for 5 min. The PCR product was digested with the restriction enzyme TspRI at 65°C for 2 hrs. The reaction mixture was as follows: 10µL of PCR product, 1µl of digestion buffer, 3µl of nuclease-free water and 0.5µl of TspRI. The digested fragments were visualized by 2% agarose gel electrophoresis with safe staining. Genotypes were distinguished by the presence of a 358-bp band for C allele, a 104-bp and 254 bands for T allele, and a 176-bp band as control.

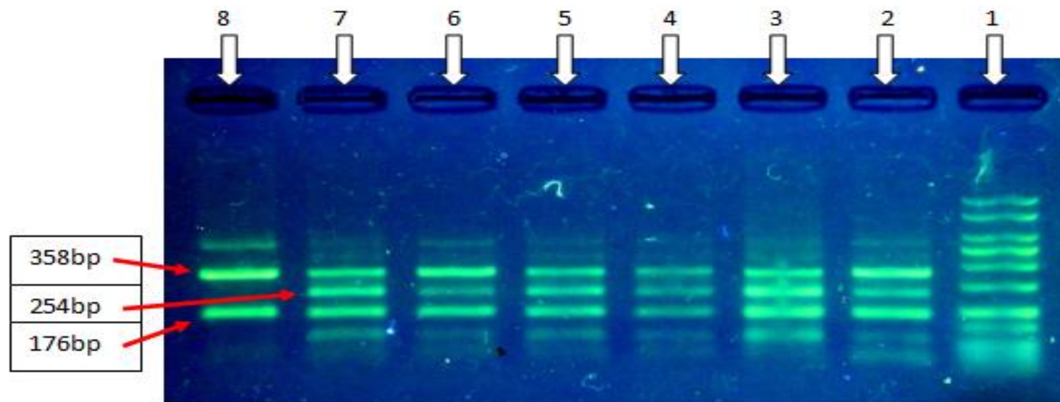

Figure 2-1 TspRI restriction fragment separated by gel electrophoresis
Lane 1: DNA Ladder; Lane: 2, 3, 4, 5, 6, 7, CT genotype;
Lane 8: CC genotype

2.4.7. Genotyping analysis of *CYP17A1* SNP (rs6162)

Genotype analyses were performed by PCR–RFLP based methods. *CYP17A1* genotyping was performed using PCR with confronting two primers (F-5-AGAGGAGTGGAGTGAGGATTC-3 and R-5-TGATCAACTGACCTCCCTTACC-3), designed using the oligo software. The PCR reactions were carried-out in 20µl reaction volume containing 6µl genomic DNA, 0.5µl of each primer, 5µl of master mix and 8µl nuclease-free water. The reaction was performed in the thermal cycler, and the DNA was denatured at 94°C for 4 min and amplified during 40 cycles (94°C for 30 seconds, annealing at 58°C for 30 seconds, extension at 72°C for 40 seconds), with final extension at 72°C for 5 min. The PCR products were digested with the restriction enzyme MscI at 37°C for 2 hrs. The reaction mixture was as follows: 10µL of PCR product, 1µl of digestion buffer, 3µl of nuclease free water and 0.2µl of MscI. The digested fragments were visualized on 2% agarose gel with safe staining. Genotypes were distinguished by the presence of a 520-bp band for C allele, a 104-bp and 416 bands for T allele, and a 334-bp band as control.

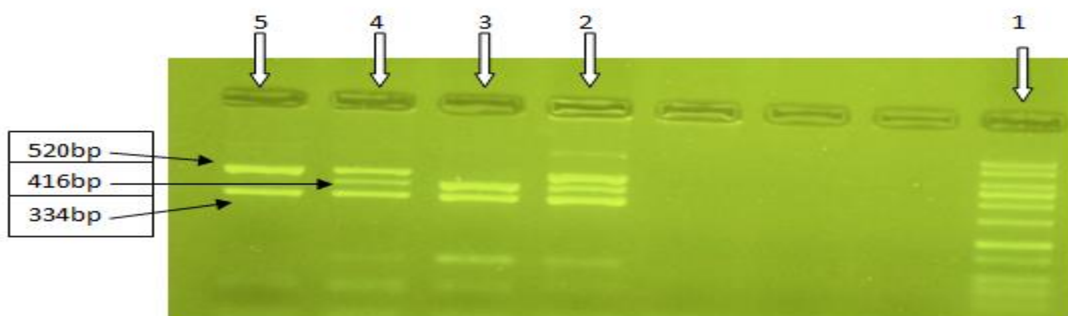

Figure 2-2 MscI restriction fragment separate by gel electrophoresis
Lane 1: DNA marker; Lane 2 and 4: CT genotype
Lane 3: TT genotype; Lane 5: CC genotype

2.5. Data Analysis:

Odds ratios (ORs) and 95% confidence interval were determined using binary logistic regression of statistical package for social science software (SPSS 16.0 for windows, International Business Machines Corporation, New York, USA) to evaluate the association of *CYP17A1* rs743572 and rs6162 genotype frequencies and breast cancer risk. Odds ratios and 95% confidence were also determined using prism software (GraphPad Software, Inc., California, and USA) to test the association of *CYP17A1* rs743572 and rs6162 allel frequencies and breast cancer incidence. Chi-square and the p-values were determined using binary logistic regression to compare the level of significance the genotype frequencies between health individuals and breast cancer patients and also chi-square and the p-values were determined to compare the level of significance of the allel frequencies between health individuals and breast cancer patients using prism software. The haplotype analysis of both SNPs were done with plink an online software (http://pngu.mgh.harvard.edu/purcell/plink/). SPSS was used to analyze data such as characterization of study subjects by age, age at menarche, age at first pregnancy, BMI, number of parities, family history, stages of breast cancer, lymph node status, estrogen and progesterone receptor status.

Chapter Three

3. Results

3.1. Characteristics of study subjects:

The present case-control study included 71 female breast cancer patients (49.3%) and 73 healthy female controls (50.7%).

Age distribution, educational status, occupation, and BMI of study subjects are illustrated in figure 3.1, 3.2, 3.3, and 3.4 respectively.

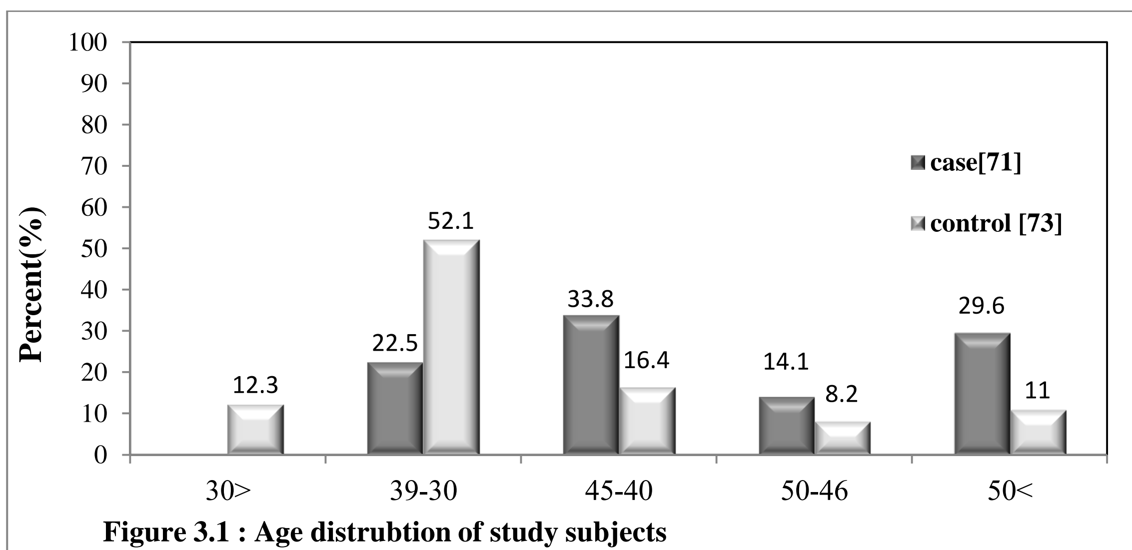

Figure 3.1 : Age distrubtion of study subjects

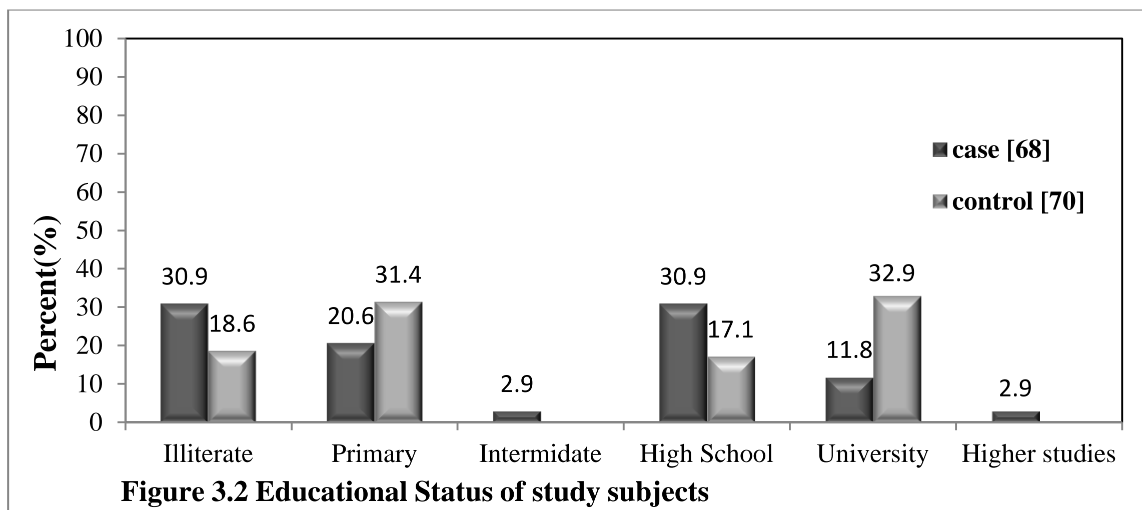

Figure 3.2 Educational Status of study subjects

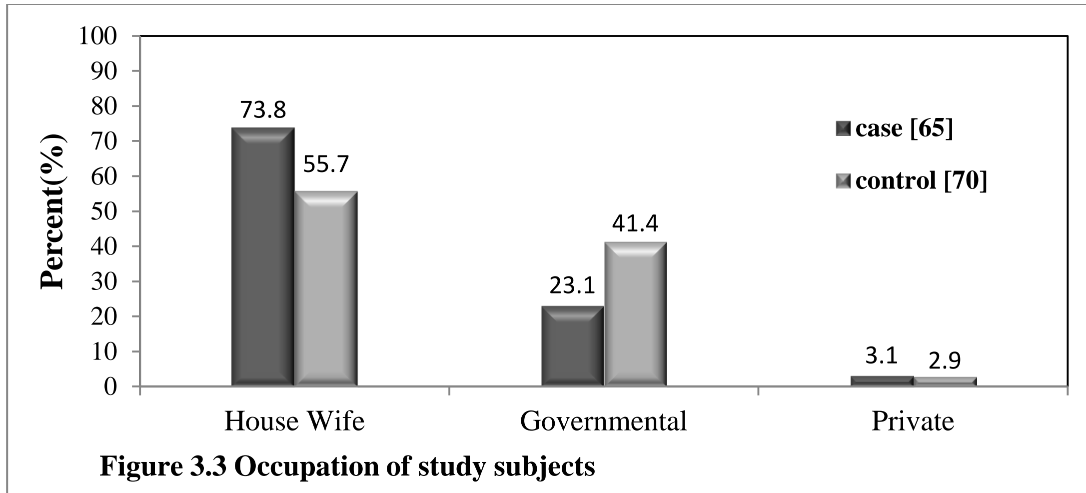

Figure 3.3 Occupation of study subjects

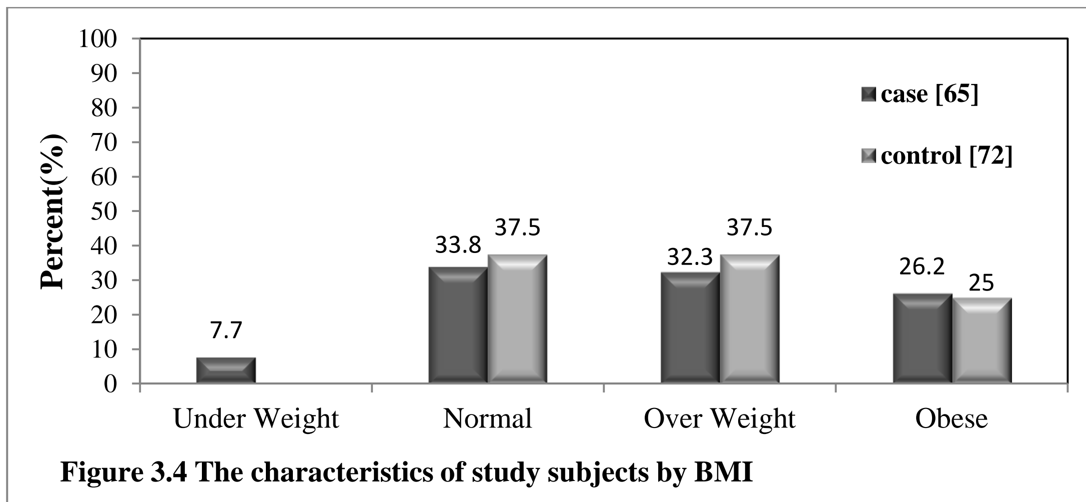

Figure 3.4 The characteristics of study subjects by BMI

3.2. Distribution of reproductive factors

Age at menarche, age at first pregnancy, menopausal status, and use of oral contraceptives of the cases and control groups are presented in figures 3.5, 3.6, 3.7, and 3.8 respectively.

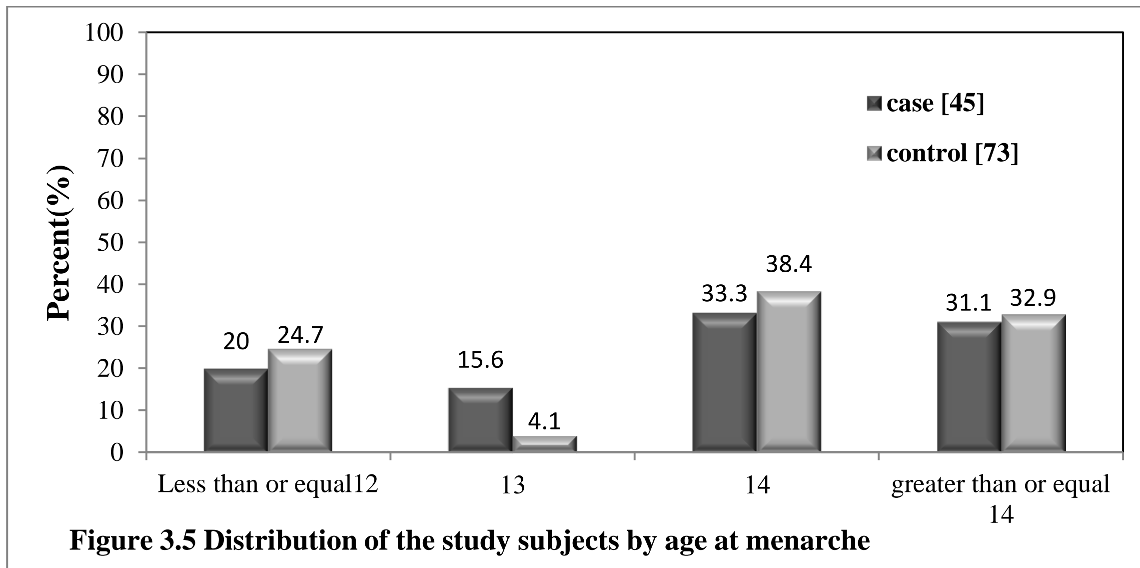

Figure 3.5 Distribution of the study subjects by age at menarche

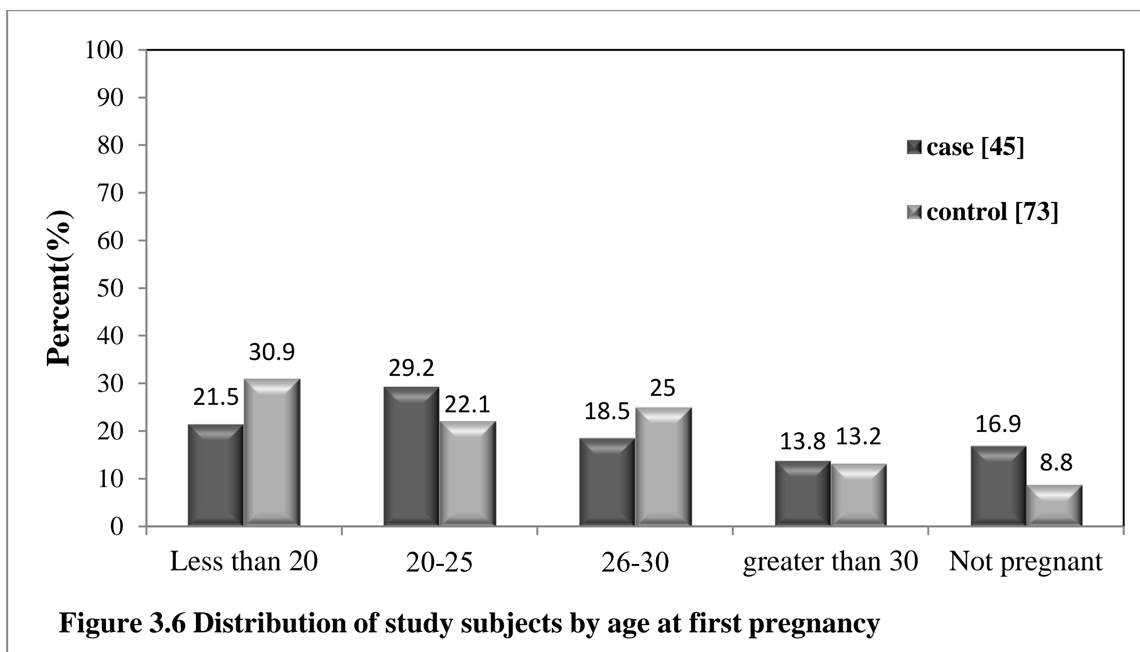

Figure 3.6 Distribution of study subjects by age at first pregnancy

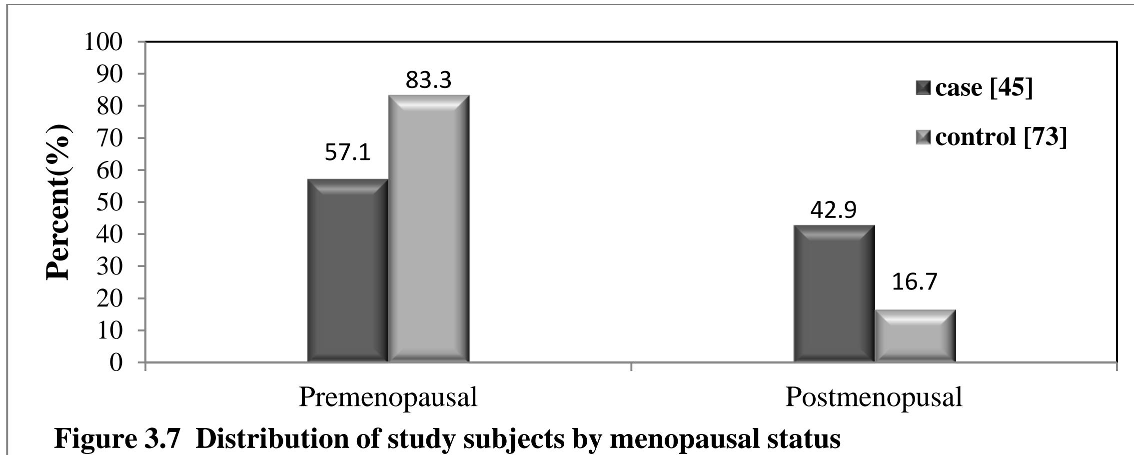

Figure 3.7 Distribution of study subjects by menopausal status

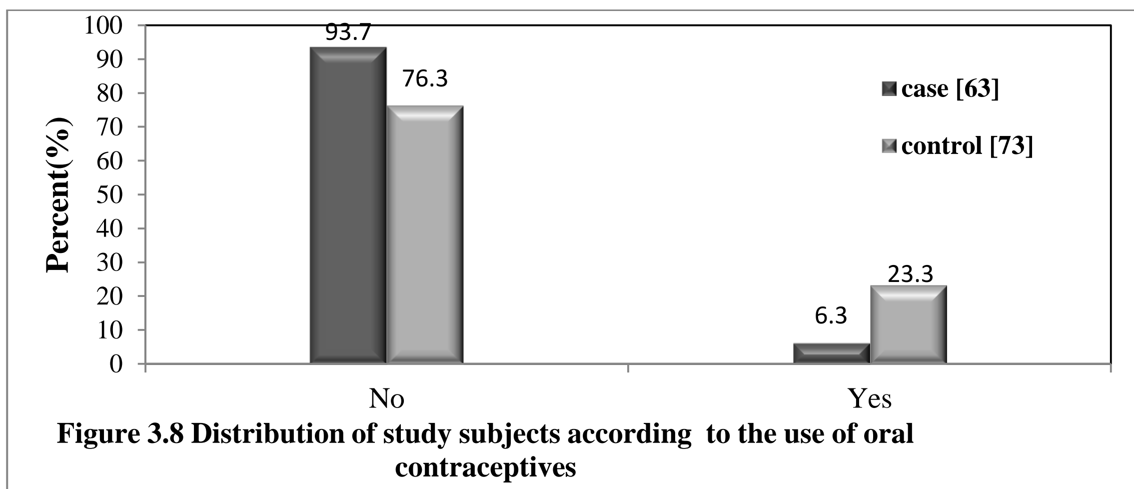

Figure 3.8 Distribution of study subjects according to the use of oral contraceptives

3.3. Clinical and Pathological Features of breast cancer patients

Clinical and pathological findings of breast cancer patients (stage of breast cancer, lymph node status, and estrogen and progesterone receptors status) are shown in the following figures (figure 3.9 to figure 3.11). 46.7% and 66.7% of breast cancer patients have advanced breast cancer and positive lymph node respectively. Patients almost have similar percentage of estrogen positive and negative receptors.

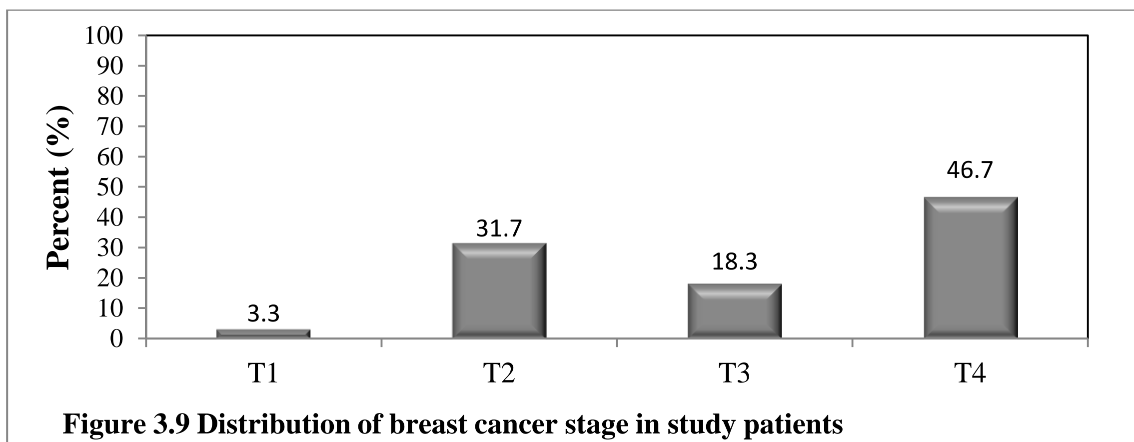

Figure 3.9 Distribution of breast cancer stage in study patients

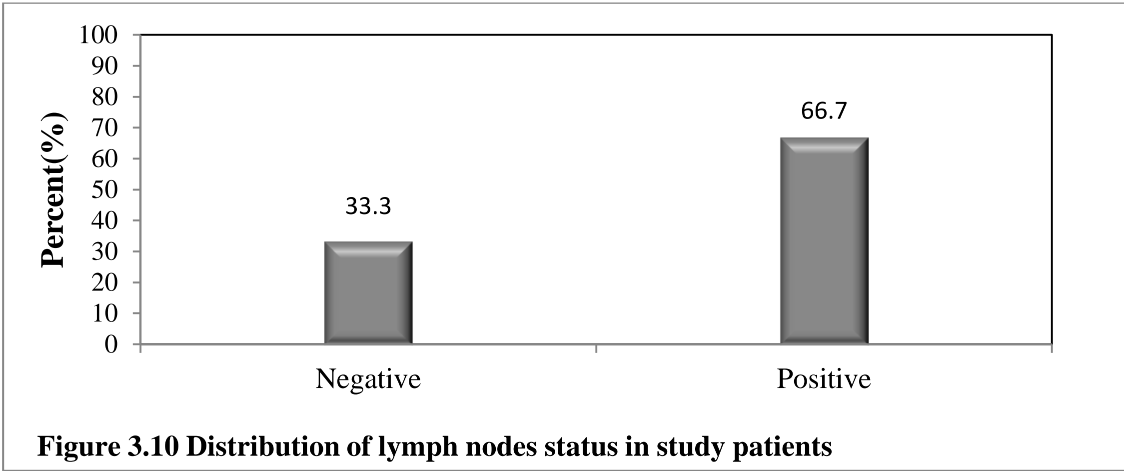

Figure 3.10 Distribution of lymph nodes status in study patients

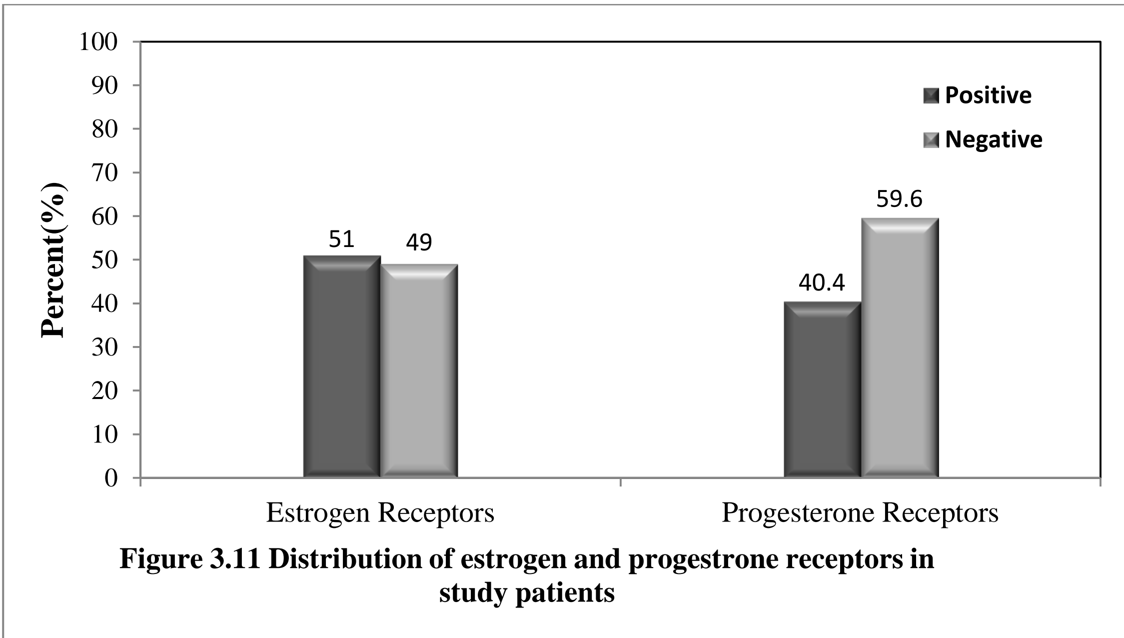

Figure 3.11 Distribution of estrogen and progestrone receptors in study patients

3.4. *CYP17A1* polymorphism (rs743572) with respect to breast cancer

The distribution of genotype and allele frequencies of *CYP17A1* (rs743572) of both cases and controls is shown in table 3.1.

Table 3.1 Distribution of genotype and allel frequencies of *CYP17A* rs743572

Genotype frequency				**Allele frequency**			
Genotypes	Case (%) N=67	Control(%) N=63	OR (95%CI)	Allele	Case N=67	Control N=63	OR(95%CI)
TT	2(3)	3(4.8)	1.00(Ref)	T	0.418	0.516	1.00(Ref)
CT	52(77.6)	59(93.7)	19.5(1.3-292)	C	0.582	0.484	1.48(0.9-2.4)
CC	13(19.4)	1(1.6)	1.32(0.21-8.22)				
Df	X^2	P		df	X^2	p	
2	10.81	0.004		1	2.52	0.11	

The frequency distribution of *CYP17A1* (rs743572) genotype showed a significant association between *CYP17A1* polymorphism and breast cancer in the present study. The CC genotype frequency was elevated in disease group as compared to controls while TT genotype frequency was elevated in controls as compared to disease group. The *CYP17A1* (rs743572) CC genotype is associated with increased risk for breast cancer (OR =1.32, 95% CI: 0.21-8.22).

Women carrying C allele in *CYP17A1* (rs743572) gene had 1.48-fold increased breast cancer risk.

The *CYP17A1* (rs743572) genotype distribution in disease and control did not deviate from Hardy–Weinberg equilibrium.

Since, we obtained low frequency of TT and CC genotypes distribution in both cases and control groups as presented in table 3.1 due to the small number of study subjects. It was difficult to make further association of this SNP with breast cancer risk factors or clinical and pathological features.

3.5. *CYP17A1* polymorphism (rs6162) with respect to breast cancer

The distribution of genotype and allele frequencies of *CYP17A1* (rs6162) of both cases and controls is shown in table 3.2.

Table 3.2 Distributions of genotype and allel frequencies of *CYP17A1* rs6162

	Genotype frequency			Allele frequency			
Genotypes	Case (%) N=57	Control (%) N=55	OR (95%CI)	Allele	Case N=57	Control N=55	OR(95%CI)
CC	28(49.1)	14(25.5)	1.00(Ref)	C	0.684	0.527	1.00(Ref)
CT	22(38.6)	30(54.5)	0.37(0.16-0.85)	T	0.316	0.473	0.51(0.30-0.89)
TT	7(12.3)	11(20)	0.32(0.10-1.0)				
df	X^2	P		df	X^2	P	
2	6.74	0.034		1	5.78	0.016	

The statistical analyses showed that genotype was significantly associated with breast cancer for *CYP17A1* rs6162 SNP (P = 0.034). Both homoyzygous TT and heterozygotes CT genotypes were associated with decreased risk of having breast cancer compared to those with the homoyzygous CC wild genotypes.

There were statistically significant differences in allele frequencies between case and controls for rs6162 SNP (p= 0.016). Women carrying T allele in *CYP17A1* (rs6162) had decreased breast cancer risk.

The *CYP17A1* (rs6162) genotype distribution in disease and control did not deviate from Hardy–Weinberg equilibrium.

3.6. Association of *CYP17A1* rs6162 genotype with breast cancer risk factors:

The association of *CYP17A1* (rs6162) with breast cancer risk factors: menopausal status, BMI, age at menarche among breast cancer cases and control groups are illustrated in table 3.3, 3.4, and 3.5 respectively

Table 3.3 Association of *CYP17A1* rs6162 genotype with menopausal status

	Premenopausal			Postmenopausal		
Genotypes	Cases (%) N =35	Control (%) N=46	OR (95%CI)	Cases (%) N =21	Control (%) N=8	OR (95%CI)
CC	16(45.7)	12(26.1)	1.00(Ref)	11(52.4)	1(12.5)	1.00(Ref)
CT	14(40)	25(54.3)	0.42(0.16-1.14)	8(38.1)	5(62.5)	0.15(0.014-1.5)
TT	5(14.3)	9(19.6)	0.42(0.11-1.57)	2(9.5)	2(25)	0.09(0.005-1.5)

Table 3.4 Association of *CYP17A1* rs6162 genotype with body mass index (BMI)

	Less than 25			Greater than 25		
Genotypes	Cases (%) N =20	Control (%) N=19	OR (95%CI)	Cases (%) N =33	Control (%) N=36	OR (95%CI)
CC	8(40)	4(21.1)	1.00(Ref)	19(57.6)	10(27.8)	1.00(Ref)
CT	9(45)	11(57.9)	0.41(0.09-1.81)	10(30.3)	19(52.8)	0.28(0.09-0.8)
TT	3(15)	4(21.1)	0.38(0.06-2.56)	4(12.1)	7(19.4)	0.30(0.07-1.3)

Table 3.5 Association of *CYP17A1* rs6162 genotype with age at menarche

	Less than or equal to 13			Greater than 13		
Genotypes	Cases (%) N =10	Control (%) N=15	OR (95%CI)	Cases (%) N =24	Control (%) N=40	OR (95%CI)
CC	5(50)	6(40)	1.00(Ref)	8(33.3)	8(20)	1.00(Ref)
CT	4(40)	8(53.3)	0.6(0.11-3.25)	12(50)	22(55)	0.55(0.16-1.8)
TT	1(10)	1(6.7)	1.2(0.06-24.5)	4(16.7)	10(25)	0.4(0.09-1.8)

There was only non-significant decreased risk in carriers of the TT genotype or CT genotype compared to the CC genotype among women with premenopausal, postmenopausal, less than 25 BMI, greater than 25 BMI, ≤13 menarche age and > 13 menarche age. To our knowledge, there was no study conducted before to relate the association of *CYP17A1* rs6162 with breast cancer risk factors.

Table 3.6 Association of *CYP17A1* rs6162 genotype with breast cancer family history

Genotypes	Positive (%) N=13	Negative (%) N=34	OR (95%CI)
CC	8(61.5)	17(50)	1.00(Ref)
CT	4(30.8)	12(35.3)	0.71(0.17-2.9)
TT	1(7.7)	5(14.7)	0.43(0.04-4.26)

Table 3.7 Association of *CYP17A1* rs6162 genotype with number of parities

Genotypes	No (%) N=10	Yes (%) N=45	OR (95%CI)
CC	5(50)	22(48.9)	1.00(Ref)
CT	4(40)	18(40)	0.98(0.23-4.19)
TT	1(10)	5(11.1)	0.88(0.08-9.29)

There was only non-significant decreased risk in carriers of the TT genotype or CT genotype compared with the CC genotype among case women with positive family history of breast cancer.

There was no risk in carriers of the TT genotype or CT genotype compared to the CC genotype among case women with parities.

3.7. Association of *CYP17A1* rs6162 genotype with breast cancer clinical and pathological features:

Tables 3-8, 3-9, 3-10, and 3-11 illustrate the association of *CYP17A1* (rs6162) with breast cancer clinical and pathological features: stages of disease, lymph nodes status, estrogen, and progesterone status among breast cancer cases respectively.

Table 3.8 Association of *CYP17A1* rs6162 genotype with stages of breast cancer

Genotypes	T1+T2 (%) N=16	T3+T4 (%) N=32	OR (95%CI)
CC	10(62.5)	11(34.4)	1.00(Ref)
CT	5(31.2)	15(46.9)	2.73(072-10.3)
TT	1(6.2)	6(18.8)	5.5(0.56-53.53)

Table 3.9 Association of *CYP17A1* rs6162 genotype with lymph nodes status

Genotypes	Positive (%) N=31	Negative (%) N=16	OR (95%CI)
CC	13(41.9)	10(62.5)	1.00(Ref)
CT	12(38.7)	5(31.2)	1.85(0.49-7.0)
TT	6(19.4)	1(6.2)	4.62(0.48-44.7)

Table 3.10 Association of *CYP17A1* rs6162 genotype with estrogen receptors status

Genotypes	Positive (%) N=23	Negative (%) N=18	OR (95%CI)
CC	15(65.2)	6(33.3)	1.00(Ref)
CT	7(30.4)	9(50)	0.32(0.82-12.6)
TT	1(4.3)	3(167)	7.5(0.65-87.2)

Table 3.11 Association of *CYP17A1* rs6162 genotype with progesterone receptors status

Genotypes	Positive (%) N=8	Negative (%) N=23	OR (95%CI)
CC	10(55.6)	11(47.8)	1.00(Ref)
CT	7(38.9)	9(39.1)	1.17 (0.32-4.3)
TT	1(5.6)	3(13)	2.73(0.24-30.7)

There was only non-significant increased risk in carriers of the TT genotype or CT genotype compared with the CC genotype among case women with positive lymph nodes, negative estrogen receptors, negative progesterone receptors and advanced breast cancer.

3.8. Haplotype Analysis

Table 3.12 Linkage Disequilibrium [LD] for SNP pair [rs6162 and rs743572]

Haplotype	Frequency	Expectation under LE
TT	0.094	0.183
CT	0.361	0.283
TC	0.293	0.210
CC	0.253	0.325

R-sq=0.114 D'= 0.466

In phase alleles are TC/CT

The two SNPs were linkage disequilibrium

Table 3.13 Association of haplotype of rs6162/rs743572 among cases and controls

Haplotype	F_U	F_A	Chisq	DF	P	Snps
TT	0.1445	0.04648	6.212	1	0.01269	rs6162/rs743572
CT	0.365	0.357	0.01496	1	0.9027	rs6162/rs743572
TC	0.3178	0.2693	0.6235	1	0.4298	rs6162/rs743572
CC	0.1728	0.3272	6.932	1	0.008465	rs6162/rs743572

We found that TT or CC haplotype frequency of both SNPs significantly associated among breast cancer cases and control groups, but TC or CT haplotype frequency of both SNPs was not significantly associated among breast cancer cases and control groups. Since the CC haplotype frequency is higher in breast cancer cases compared with control groups, it acts as a risk factor for breast cancer considering TT haplotype (higher in control groups compared with breast cancer patients) as a reference.

Chapter Four

4. Discussion

With a key role in estrogen hormone biosynthesis, variation in *CYP17A1* (rs743572 and rs6162) has been hypothesized to be related with the risk of developing breast cancer. In order to examine whether genotype and allelic variants were associated with breast cancer status, we performed genotypes and allelic-frequencies based association analyses in this study.

4.1 Association of *CYP17A1* rs743572 and breast cancer risk

In the present study, women carrying C allele in *CYP17A1* (rs743572) had 1.48-fold increased breast cancer risk. Similar findings were presented previously (Tuzuner *et al.*, 2010). Because the C variant allele creates an additional putative Sp-1 binding site (CCACC) in the promoter region, it is speculated that the C allele enhances the transcription of *CYP17A1*, leading to the enhanced synthesis of estrogens in breast tumors.

The frequency distribution of *CYP17A1* (rs743572) genotype showed a significant association between *CYP17A1* polymorphism and breast cancer in the current study. The CC genotype frequency was elevated in disease group as compared to controls while TT genotype frequency was higher in the controls as compared to disease group. The *CYP17A1* (rs743572) CC genotype is associated with subtle increased risk (OR=1.32) for breast cancer. This result was consistent with the findings observed in prostate cancer patients (Souiden *et al.*, 2011).

On the other hand, our results contradict with a study conducted in South India women (Samson *et al.*, 2009) which reported that the *CYP17A1* (rs743572) gene CC genotype was significantly associated with decreased risk for breast cancer. There was a significant protective association of the CC genotype of the *CYP17A1* (rs743572) gene against the risk of developing breast cancer. Also our results disagree with a study by Einarsdottir which showed that no overall association was found between *CYP17A1* (rs743572) and breast cancer risk for the CC carriers compared to the TT carriers (Einarsdottir *et al.*, 2005).

4.2. Association of *CYP17A1* (rs6162) and breast cancer risk

Genotype frequencies in healthy individuals and breast cancer patients were significantly associated with *CYP17A1* rs6162 SNP. Heterozygous CT genotypes were significantly at decreased risk of having breast cancer compared to those with the homoyzygous CC wild type genotypes. This result is consistent with the findings obtained by Sarma *et al*. in a study conducted in Afican-American prostate cancer men. In this study homozygous TT genotypes were also significantly at decreased risk of having breast cancer compared with those having the homoyzygous CC wild type genotypes. This result disagrees with the results presented by the same group (Sarma *et al.*, 2008).

In the present study, there were statistically significant differences in allele frequencies between case and controls for rs6162 SNP. This result was dissimilar to the findings obtained by Sarma *et al*. (2008).

4.3. Association of haplotype with breast cancer:

In order to examine whether specific phased combinations of allelic variants were associated with breast cancer status, we performed haplotype-based association analyses.

The two SNPs (rs743572 and rs6162) were in weak linkage disequilibrium. This finding agrees with (Zhao *et al.*, 2008) eventhough the extent of linkage disequilibrium was different.

Chapter Five

5. Conclusions and Recommendations

5.1 Conclusions

- This study provides some support that polymorphic variation in *CYP17A1* rs743572 might play a role in breast cancer susceptibility.

- Our data suggest that carrying homozygous TT in *CYP17A1* rs6162 gene is inversely related to breast cancer risk in Sudanese women.

- To our knowledge, this is the first report to explore the possible association between *CYP17A1* rs6162 gene and breast cancer risk factors and clinical parameters. Therefore, additional studies are needed to test our findings.

5.2 Recommendations

- Genetic variations can act as potential biomarkers, helping in evaluation of breast cancer susceptibility.

- Genetic variations in *CYP17A1* can results in either increased or decreased enzyme activity, which directly affects estrogen levels, and may ultimately used by scientists and physicians to determines the development, treatment, and prognosis of breast cancer.

- By investigating the functional aspects of these SNPs in breast cancer patients, and building up more information about the underlying genetic basis for breast cancer, it may be possible to manage the impact on breast cancer.

- These observations need to be tested in larger studies due to the limited statistical power of the study based on a small number of cases and control groups.

References

- Aguilera A. and Garcia-Muse T. (2013). Causes of genome instability. *Annu Rev Genet* **47** 1-32.
- Ahmed H. G., Ali A. S. and Almobarak A. O. (2010). Frequency of breast cancer among sudanese patients with breast palpable lumps. *Indian J Cancer* **47** (1): 23-26.
- Ahsan H., Whittemore A. S., Chen Y., Senie R. T., Hamilton S. P., Wang Q., Gurvich I. and Santella R. M. (2005). Variants in estrogen-biosynthesis genes cyp17 and cyp19 and breast cancer risk: A family-based genetic association study. *Breast Cancer Res* **7** (1): R71-81.
- Akram M. and Siddiqui S. A. (2012). Breast cancer management: Past, present and evolving. *Indian J Cancer* **49** (3): 277-282.
- Antognelli C., Del Buono C., Ludovini V., Gori S., Talesa V. N., Crino L., Barberini F. and Rulli A. (2009). Cyp17, gstp1, pon1 and glo1 gene polymorphisms as risk factors for breast cancer: An italian case-control study. *BMC Cancer* **9** 115.
- Berg W. A., Blume J. D., Cormack J. B., Mendelson E. B., Lehrer D., Bohm-Velez M., Pisano E. D., Jong R. A., Evans W. P., Morton M. J., Mahoney M. C., Larsen L. H., Barr R. G., Farria D. M., Marques H. S. and Boparai K. (2008). Combined screening with ultrasound and mammography vs mammography alone in women at elevated risk of breast cancer. *JAMA* **299** (18): 2151-2163.
- Bloom H. J. and Richardson W. W. (1957). Histological grading and prognosis in breast cancer; a study of 1409 cases of which 359 have been followed for 15 years. *Br J Cancer* **11** (3): 359-377.
- Botstein D. and Risch N. (2003). Discovering genotypes underlying human phenotypes: Past successes for mendelian disease, future approaches for complex disease. *Nat Genet* **33 Suppl** 228-237.
- Broustas C. G. and Lieberman H. B. (2014). DNA damage response genes and the development of cancer metastasis. *Radiat Res*
- Carey L. A., Perou C. M., Livasy C. A., Dressler L. G., Cowan D., Conway K., Karaca G., Troester M. A., Tse C. K., Edmiston S., Deming S. L., Geradts J., Cheang M. C., Nielsen T. O., Moorman P. G., Earp H. S. and Millikan R. C. (2006). Race, breast cancer subtypes, and survival in the carolina breast cancer study. *JAMA* **295** (21): 2492-2502.
- Cargill M., Altshuler D., Ireland J., Sklar P., Ardlie K., Patil N., Shaw N., Lane C. R., Lim E. P., Kalyanaraman N., Nemesh J., Ziaugra L., Friedland L., Rolfe A., Warrington J., Lipshutz R., Daley G. Q. and Lander E. S. (1999). Characterization of single-nucleotide polymorphisms in coding regions of human genes. *Nat Genet* **22** (3): 231-238.
- Chattopadhyay S., Siddiqui S., Akhtar M. S., Najm M. Z., Deo S. V., Shukla N. K. and Husain S. A. (2014). Genetic polymorphisms of esr1, esr2, cyp17a1, and cyp19a1 and the risk of breast cancer: A case control study from north india. *Tumour Biol* **35** (5): 4517-4527.
- Dalamaga M. (2013). Obesity, insulin resistance, adipocytokines and breast cancer: New biomarkers and attractive therapeutic targets. *World J Exp Med* **3** (3): 34-42.
- Einarsdottir K., Rylander-Rudqvist T., Humphreys K., Ahlberg S., Jonasdottir G., Weiderpass E., Chia K. S., Ingelman-Sundberg M., Persson I., Liu J., Hall P. and Wedren S. (2005). Cyp17 gene polymorphism in relation to breast cancer risk: A case-control study. *Breast Cancer Res* **7** (6): R890-896.
- Elgaili E. M., Abuidris D. O., Rahman M., Michalek A. M. and Mohammed S. I. (2010). Breast cancer burden in central sudan. *Int J Womens Health* **2** 77-82.
- Elston C. W. and Ellis I. O. (2002). Pathological prognostic factors in breast cancer. I. The value of histological grade in breast cancer: Experience from a large study with long-term follow-up. *Histopathology* **41** (3A): 154-161.
- Feigelson H. S., Coetzee G. A., Kolonel L. N., Ross R. K. and Henderson B. E. (1997). A polymorphism in the cyp17 gene increases the risk of breast cancer. *Cancer Res* **57** (6): 1063-1065.

- Ferlay J., Shin H. R., Bray F., Forman D., Mathers C. and Parkin D. M. (2010). Estimates of worldwide burden of cancer in 2008: Globocan 2008. *Int J Cancer* **127** (12): 2893-2917.
- Gao C. M., Ding J. H., Li S. P., Liu Y. T., Qian Y., Chang J., Tang J. H. and Tajima K. (2013). Active and passive smoking, and alcohol drinking and breast cancer risk in chinese women. *Asian Pac J Cancer Prev* **14** (2): 993-996.
- Hamad H. M. (2006). Cancer initiatives in sudan. *Ann Oncol* **17 Suppl 8** viii32-viii36.
- Hamaguchi M., Nishio M., Toyama T., Sugiura H., Kondo N., Fujii Y. and Yamashita H. (2008). Possible difference in frequencies of genetic polymorphisms of estrogen receptor alpha, estrogen metabolism and p53 genes between estrogen receptor-positive and -negative breast cancers. *Jpn J Clin Oncol* **38** (11): 734-742.
- Heldring N., Pike A., Andersson S., Matthews J., Cheng G., Hartman J., Tujague M., Strom A., Treuter E., Warner M. and Gustafsson J. A. (2007). Estrogen receptors: How do they signal and what are their targets. *Physiol Rev* **87** (3): 905-931.
- Huang Y., Kang M., Li H., Li J. Y., Zhang J. Y., Liu L. H., Liu X. T., Zhao Y., Wang Q., Li C. C. and Lee H. (2012). Combined performance of physical examination, mammography, and ultrasonography for breast cancer screening among chinese women: A follow-up study. *Curr Oncol* **19** (Suppl 2): eS22-30.
- Jemal A., Bray F., Forman D., O'Brien M., Ferlay J., Center M. and Parkin D. M. (2012). Cancer burden in africa and opportunities for prevention. *Cancer* **118** (18): 4372-4384.
- Kato I., Cichon M., Yee C. L., Land S. and Korczak J. F. (2009). African american-preponderant single nucleotide polymorphisms (snps) and risk of breast cancer. *Cancer Epidemiol* **33** (1): 24-30.
- Katzenellenbogen B. S., Kendra K. L., Norman M. J. and Berthois Y. (1987). Proliferation, hormonal responsiveness, and estrogen receptor content of mcf-7 human breast cancer cells grown in the short-term and long-term absence of estrogens. *Cancer Res* **47** (16): 4355-4360.
- Klinge C. M. (2001). Estrogen receptor interaction with estrogen response elements. *Nucleic Acids Res* **29** (14): 2905-2919.
- Kruk J. and Marchlewicz M. (2013). Dietary fat and physical activity in relation to breast cancer among polish women. *Asian Pac J Cancer Prev* **14** (4): 2495-2502.
- Levesque E., Huang S. P., Audet-Walsh E., Lacombe L., Bao B. Y., Fradet Y., Laverdiere I., Rouleau M., Huang C. Y., Yu C. C., Caron P. and Guillemette C. (2013). Molecular markers in key steroidogenic pathways, circulating steroid levels, and prostate cancer progression. *Clin Cancer Res* **19** (3): 699-709.
- Lichtenstein P., Holm N. V., Verkasalo P. K., Iliadou A., Kaprio J., Koskenvuo M., Pukkala E., Skytthe A. and Hemminki K. (2000). Environmental and heritable factors in the causation of cancer--analyses of cohorts of twins from sweden, denmark, and finland. *N Engl J Med* **343** (2): 78-85.
- Macconaill L. E. and Garraway L. A. (2010). Clinical implications of the cancer genome. *J Clin Oncol* **28** (35): 5219-5228.
- Maston G. A., Evans S. K. and Green M. R. (2006). Transcriptional regulatory elements in the human genome. *Annu Rev Genomics Hum Genet* **7** 29-59.
- McPherson K., Steel C. M. and Dixon J. M. (2000). Abc of breast diseases. Breast cancer-epidemiology, risk factors, and genetics. *BMJ* **321** (7261): 624-628.
- Mitrunen K., Jourenkova N., Kataja V., Eskelinen M., Kosma V. M., Benhamou S., Vainio H., Uusitupa M. and Hirvonen A. (2000). Steroid metabolism gene cyp17 polymorphism and the development of breast cancer. *Cancer Epidemiol Biomarkers Prev* **9** (12): 1343-1348.
- Narod S. A. (2011). Hormone replacement therapy and the risk of breast cancer. *Nat Rev Clin Oncol* **8** (11): 669-676.
- Nasuti J. F., Gupta P. K. and Baloch Z. W. (2002). Diagnostic value and cost-effectiveness of on-site evaluation of fine-needle aspiration specimens: Review of 5,688 cases. *Diagn Cytopathol* **27** (1): 1-4.
- Perou C. M., Sorlie T., Eisen M. B., van de Rijn M., Jeffrey S. S., Rees C. A., Pollack J. R., Ross D. T., Johnsen H., Akslen L. A., Fluge O., Pergamenschikov A., Williams C., Zhu S. X.,

Lonning P. E., Borresen-Dale A. L., Brown P. O. and Botstein D. (2000). Molecular portraits of human breast tumours. *Nature* **406** (6797): 747-752.
- Reding K. W., Chen C., Lowe K., Doody D. R., Carlson C. S., Chen C. T., Houck J., Weiss L. K., Marchbanks P. A., Bernstein L., Spirtas R., McDonald J. A., Strom B. L., Burkman R. T., Simon M. S., Liff J. M., Daling J. R. and Malone K. E. (2012). Estrogen-related genes and their contribution to racial differences in breast cancer risk. *Cancer Causes Control* **23** (5): 671-681.
- Samson M., Rama R., Swaminathan R., Sridevi V., Nancy K. N. and Rajkumar T. (2009). Cyp17 (t34c), cyp19 (trp39arg), and fgfr2 (c906t) polymorphisms and the risk of breast cancer in South Indian women. *Asian Pac J Cancer Prev* **10** (1): 111-114.
- Sangrajrang S., Sato Y., Sakamoto H., Ohnami S., Laird N. M., Khuhaprema T., Brennan P., Boffetta P. and Yoshida T. (2009). Genetic polymorphisms of estrogen metabolizing enzyme and breast cancer risk in thai women. *Int J Cancer* **125** (4): 837-843.
- Sarma A. V., Dunn R. L., Lange L. A., Ray A., Wang Y., Lange E. M. and Cooney K. A. (2008). Genetic polymorphisms in cyp17, cyp3a4, cyp19a1, srd5a2, igf-1, and igfbp-3 and prostate cancer risk in african-american men: The flint men's health study. *Prostate* **68** (3): 296-305.
- Sharma G. N., Dave R., Sanadya J., Sharma P. and Sharma K. K. (2010). Various types and management of breast cancer: An overview. *J Adv Pharm Technol Res* **1** (2): 109-126.
- Siegel R., Naishadham D. and Jemal A. (2013). Cancer statistics, 2013. *CA Cancer J Clin* **63** (1): 11-30.
- Singh V., Parmar D. and Singh M. P. (2008). Do single nucleotide polymorphisms in xenobiotic metabolizing genes determine breast cancer susceptibility and treatment outcomes? *Cancer Invest* **26** (8): 769-783.
- Sorlie T., Tibshirani R., Parker J., Hastie T., Marron J. S., Nobel A., Deng S., Johnsen H., Pesich R., Geisler S., Demeter J., Perou C. M., Lonning P. E., Brown P. O., Borresen-Dale A. L. and Botstein D. (2003). Repeated observation of breast tumor subtypes in independent gene expression data sets. *Proc Natl Acad Sci U S A* **100** (14): 8418-8423.
- Sorlie T., Perou C. M., Tibshirani R., Aas T., Geisler S., Johnsen H., Hastie T., Eisen M. B., van de Rijn M., Jeffrey S. S., Thorsen T., Quist H., Matese J. C., Brown P. O., Botstein D., Lonning P. E. and Borresen-Dale A. L. (2001). Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proc Natl Acad Sci U S A* **98** (19): 10869-10874.
- Souiden Y., Mahdouani M., Chaieb K., Elkamel R. and Mahdouani K. (2011). Cyp17 gene polymorphism and prostate cancer susceptibility in a tunisian population. *Cancer Epidemiol* **35** (5): 480-484.
- Tuzuner B. M., Ozturk T., Kisakesen H. I., Ilvan S., Zerrin C., Ozturk O. and Isbir T. (2010). Cyp17 (t-34c) and cyp19 (trp39arg) polymorphisms and their cooperative effects on breast cancer susceptibility. *In Vivo* **24** (1): 71-74.
- Veronesi U., Viale G., Rotmensz N. and Goldhirsch A. (2006). Rethinking tnm: Breast cancer tnm classification for treatment decision-making and research. *Breast* **15** (1): 3-8.
- Yamada T., Nakayama M., Shimizu T., Nonen S., Nakai Y., Nishimura K., Fujio Y., Okuyama A., Azuma J. and Nonomura N. (2013). Genetic polymorphisms of cyp17a1 in steroidogenesis pathway are associated with risk of progression to castration-resistant prostate cancer in japanese men receiving androgen deprivation therapy. *Int J Clin Oncol* **18** (4): 711-717.
- Zhao Z. Z., Nyholt D. R., Le L., Treloar S. A. and Montgomery G. W. (2008). Common variation in the cyp17a1 and ifit1 genes on chromosome 10 does not contribute to the risk of endometriosis. *Open Reprod Sci J* **1** 35-40.
- Zhu B. T. and Conney A. H. (1998). Functional role of estrogen metabolism in target cells: Review and perspectives. *Carcinogenesis* **19** (1): 1-27.
- Zhu H., Lei X., Feng J. and Wang Y. (2012). Oral contraceptive use and risk of breast cancer: A meta-analysis of prospective cohort studies. *Eur J Contracept Reprod Health Care* **17** (6): 402-414.

Appendices

Appendix-1

Consent Form for Participation in a Research Study

University of Gezira

Title of Study: Association between Estrogen Biosynthesis Enzyme CYP17 Gene Polymorphisms and Breast Cancer in Sudanese Female Patients.

I, the undersigned, confirm that (please tick box as appropriate):

1.	I have read and understood the information about the project, as provided in the Information Sheet dated _________________.	☐
2.	I have been given the opportunity to ask questions about the project and my participation.	☐
3.	I voluntarily agree to participate in the project.	☐
4.	I understand I can withdraw at any time without giving reasons and that I will not be penalized for withdrawing nor will I be questioned on why I have withdrawn.	☐
5.	The procedures regarding confidentiality have been clearly explained to me.	☐
6.	If applicable, separate terms of consent for interviews or other forms of data collection have been explained and provided to me.	☐
7.	I, along with the Researcher, agree to sign and date this informed consent form.	☐

Participant:

______________________	___________________________	________________
Name of Participant	Signature	Date

Researcher:

______________________	___________________________	________________
Name of Researcher	Signature	Date

Questionnaire

Mihreteab Ghebretinsae Tekle

University of Gezira

Tel: 0998634306

Dear Madam,

I, Mihreteab Ghebretinsae Tekle undertaking a research to association between estrogen biosynthesis enzyme CYP17 gene polymorphisms and breast cancer in Sudanese female patients at University of Gezira for the partial fulfillment of MSc in Biochemistry. To this end I kindly request that you complete the following short questionnaire regarding your age, height, weight, family history of breast cancer and Hormonal replacement therapy. It should take no longer than 15 minutes of your time. Although your response is of the utmost importance to me, your participation in this survey is entirely voluntary. Information provided by you remains confidential.

Yours Sincerely,
Mihreteab Ghebretinsae Tekle
Department of Biochemistry and Nutrition
Faculty of Medicine
University of Gezira

Please answer the following questions by crossing(x) the relevant block or writing down your answer in the space provided.

1. Name of Participant: ____

2. Age: ____________________

3. Age at menarche: _______________

4. Age at menopause: _____________

5. Age at first full-term pregnancy: ___________

6. Educational Status: Illiterate ☐ Primary School ☐ Secondary School ☐ Higher Education ☐

7. Occupation: Unemployed ☐ Governmental ☐ Private ☐

8. Residence: Urban ☐ Rural ☐

9: Economic status: Poor ☐ Average ☐ Affluent ☐

10. Height: ______________

11. Weight: _____________

12. BMI: _______________

13. Family history of breast cancer: yes ☐ No ☐

14. Hormonal replacement therapy: Never use ☐ Ever use ☐

15. Parities: __________________

16. Have you ever taken oral contraceptives? Yes ☐ No ☐

If yes, for how long? ________________

Clinical and pathological characteristics

17. Cancer stage: _________________________

18. Tumor Size: __________________________

19. Histological Type: _____________________

20. Tumor Grade: ________________________

21. Receptor Status:

Estrogen Receptors: Positive ☐ Negative ☐

Progesterone Receptor Positive ☐ Negative ☐