POLYMORPHISMS IN *P21* (CODON 31) AND *P53* (CODON 72): ASSOCIATION WITH BREAST CANCER SUSCEPTIBILITY IN THE TURKISH AND GREEK POPULATIONS

A THESIS SUBMITTED TO
THE DEPARTMENT OF MOLECULAR BIOLOGY AND GENETICS
AND
THE INSTITUTE OF ENGINEERING AND SCIENCE OF
BILKENT UNIVERSITY
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR
THE DEGREE OF MASTER OF SCIENCE

BY
GÜLSEN ÇOLAKOĞLU
AUGUST, 2003

I certify that I have read the thesis, and that in my opinion it is fully adequate, in scope and in quality, as a thesis for the Master of Science.

_____

Assoc. Prof. Dr. Uğur ÖZBEK

I certify that I have read the thesis, and that in my opinion it is fully adequate, in scope and in quality, as a thesis for the Master of Science.

_____

Assoc. Prof. Dr. Tayfun ÖZÇELİK

I certify that I have read the thesis, and that in my opinion it is fully adequate, in scope and in quality, as a thesis for the Master of Science.

_____

Asst. Prof. Dr. Işık G. YULUĞ

Approved for the Institute of Engineering and Science

_____

Prof. Dr. Mehmet BARAY

Director of the Institute of Engineering and Science

ABSTRACT

POLYMORPHISMS IN *P21* (CODON 31) AND *P53* (CODON 72): ASSOCIATION WITH BREAST CANCER SUSCEPTIBILITY IN THE TURKISH AND GREEK POPULATIONS

Gülsen ÇOLAKOĞLU
Ms. in Molecular Biology and Genetics
Supervisor: Asst. Prof. Dr. Isık G. YULUĞ
August 2003, 121 pages

The aim of this study was to investigate the potential association of *p53* codon 72 and/or *p21* codon 31 polymorphisms with increased susceptibility for breast cancer either independently or combined in the Turkish and Greek populations. A case-control study was conducted for both populations and the genotypes of the subjects were determined by PCR-RFLP (Turkish; *p53* genotypes for 274 cases and 221 controls, *p21* genotypes for 322 cases and 246 controls, Greek; *p53* genotypes for 138 cases and 138 controls, *p21* genotypes for 156 cases and 136 controls were obtained). Binary logistic regression was used to analyze the data. Although the Greek study population alone did not give statistically significant results, the *p53* codon 72 Arg/Arg inheritance was found to be significantly associated with breast cancer susceptibility in the Turkish study population (OR=2.16; 95% CI=1.08-4.31) as well as in the combined population of Turkish and Greek subjects (OR=2.35; 95% CI=1.25-4.41). This association was further increased with increased BMI (OR=3.86; 95% CI=1.12-13.26) in the Turkish population but the result should be treated with caution because of the wide confidence interval. The inheritance of the combined *p21* codon 31 Arg/Arg or Ser/Arg genotypes increased breast cancer susceptibility in the Turkish study population (OR=1.15; 95% CI=0.75-1.76) although the result is not statistically significant. The most prominent result of this study is that there is an interaction between the *p53* Arg72Arg and *p21* Arg31Arg or Ser31Arg genotypes for breast cancer susceptibility (OR=2.66; 95% CI=1.06-6.66). These results let us to conclude that there is a strong association between the *p53* Arg72Arg genotype and breast cancer risk in the Turkish population and that the combination of high-risk allelic variants of both *p53* and its downstream effector protein *p21* may have a role as a risk factor for breast cancer development.

ÖZET

P21 (CODON 31) VE *P53* (CODON 72) POLİMORFİZMLERİ: TÜRK VE YUNAN POPULASYONLARINDA MEME KANSERİ İLE İLİŞKİSİ

Gülsen ÇOLAKOĞLU
Moleküler Biyoloji ve Genetik Yüksek Lisansı
Tez Yöneticisi: Yrd. Doç. Dr. Işık G. YULUĞ
Ağustos 2003, 121 sayfa

Bu çalışmanın amacı, Türk ve Yunan populasyonlarında, *p53* kodon 72 ve/veya *p21* kodon 31 polimorfizmleri ile meme kanserine yatkınlık arasındaki olası ilişkiyi incelemekti. Her iki populasyon için hasta-kontrol çalışması yapıldı ve tüm örneklerin genotipleri PCR-RFLP yöntemi ile belirlendi (Türklerde; *p53* genotipi için 274 hasta ve 221 kontrol, *p21* genotipi için 322 hasta ve 246 kontrol, Yunanlılarda; *p53* genotipi için 138 hasta ve 138 kontrol, *p21* genotipi için 156 hasta ve136 kontrol elde edildi). Verilerin değerlendirilmesi için ikili lojistik regresyon analizi yöntemi kullanıldı. Yunanistan populasyonu tek başına incelendiğinde, istatistiksel açıdan anlamlı bir sonuç bulunamamasına rağmen, Türk ve Türk-Yunan populasyonları birleştirildiğinde anlamlı sonuçlar elde edildi. *p53* geninin 72. kodonunun Arg/Arg olması durumu meme kanseri riski ile önemli derecede ilişkiliydi (Türk: olasılıklar oranı OR=2.16; %95 güven aralığı (GA)= 1.08-4.31, Türk-Yunan: OR=2.35; 95% GA=1.25-4.41). Bu ilişki, vücut kütle indeksi yüksek Türk kadınları arasında incelenince, olasılıklar oranı önemli derecede artış gösterdi (OR=3.86; 95% GA=1.12-13.26). Ancak, bu sonuç değerledirilirken güven aralığının geniş olduğu dikkate alınmalıdır. *p21* genini incelediğimizde ise, 31. kodonun Arg/Arg ya da Ser/Arg olmasının meme kanserine yakalanma olasılığını arttırdığı görülmüştür, ancak sonuç istatistiksel açıdan anlamlı değildir (OR=1.15; 95% GA=0.75-1.76). Bu çalışmanın belki de en çarpıcı bulgusu, *p53* kodon 72 Arg/Arg genotipini ve *p21* kodon 31 Arg/Arg ya da Ser/Arg genotiplerinden birini aynı anda taşıyan bireylerin meme kanseri riskinin artmasıdır (OR=2.66; 95% GA=1.06-6.66). Sonuç olarak, Türk populasyonunda meme kanserine yakalanma riski ile *p53* Arg72Arg genotipi arasında önemli bir ilişki olduğu ve de ayrıca her iki genin birlikte, belirtilen kodon polimorfizmlerinde yüksek risk genotiplerini taşımasının meme kanserine yakalanma riskini daha da arttırdığı söylenebilir.

ACKNOWLEDGEMENT

I am very grateful to my advisor Asst. Prof. Dr. Isik G. Yulug for her supervision, motivating comments, guidance, and continuous support.

I also would like to thank Prof. Dr. Mehmet Özturk and Assoc. Prof. Dr. Tayfun Özçelik for their valuable advices.

Many thanks to Ebru Demir, Dr. Arzu Atalay, and Bala Gür for supporting me with their knowledge and experience.

I also thank Dr. Betül Bozkurt and Ömer Cengiz from Ankara Numune Research and Teaching Hospital, Prof. Dr. Ali Esat Karakaya, Prof. Dr. Semra Sardaş and Dr. Neslihan Aygün Kocabaş from Gazi University, Faculty of Pharmacy, Prof. Dr. Güven Lüleci from Akdeniz University, Medical School, and Dr. Drakoulis Yannoukakos from National Center for Scientific Research in Greece for providing the samples and clinical data for my work.

I thank all members of MBG department especially İlter Sever, Akın Sevinç and Hilal Çelikkaya for their support and friendship.

Special thanks to my family and Barış Efe for their unconditional support, encouragement and love.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS

AIP1	Apoptosis Induced Protein 1
Arg	Arginine
BMI	Body Mass Index
Bp	Base pairs
BRCA1	Breast Cancer Susceptibility Gene 1
BRCA2	Breast Cancer Susceptibility Gene 2
Cdk	Cyclin-dependent kinase
CI	Confidence Interval
Cip-1	Cdk-interacting protein 1
cm	centimeter
COMT	catechol-O-methyltransferase
CYP1A1	Cytochrome P450 1A1
CYP2D6	Cytochrome P450 2D6
CYP2E1	Cytochrome P450 2E1
CYP17	Cytochrome P450, subfamily XVII
CYP19	Cytochrome P450, subfamily XIX
˚C	Degree celsius
DNA	Deoxyribonucleic Acid
dNTP	Deoxynucleotide triphosphate
EGF	Epithelial Growth Factor
F	Forward
GST	Gluthathione S-Transferase
GSTM1	Gluthathione S-Transferase mu 1
GSTP1	Gluthathione S-Transferase phi 1

GSTT1	Gluthathione S-Transferase theta 1
GTP	Guanosine Triphosphate
HPV	Human Papilloma Virus
hrs	hour(s)
KIP1	Kinase Inhibitory Protein 1
KIP2	Kinase Inhibitory Protein 2
kb	kilobase(s)
kDa	kilo Dalton
LOH	Loss of Heterozygosity
M	Molar
mg	milligram
min	minute(s)
ml	milliliter
mM	millimolar
MSH2	Mut S Homolog 2
µl	microliter
µg	microgram
n	number of subjects
NAT1	N-acetyl transferase Type 1
NAT2	N-acetyl transferase Type 2
ng	nanogram
OR	Odds Ratio
PCR	Polymerase Chain Reaction
pmol	picomol
Pro	Proline
R	Reverse

rpm	revolutions per minute
s	second(s)
Ser	Serine
SNP	Single Nucleotide Polymorphism
TP53	Tumor Protein p53
u	unit(s)
UV	Ultra Violet
V	Volt
WAF-1	Wild type p53 activated protein-1
w/v	weight per volume
X^2	Chi-square

1. INTRODUCTION

1.1. Cancer: A Loss of Normal Growth Regulation

Various disease states can arise when the normal stability of the organization of tissues and organs is disturbed. A **tumor** or **neoplasm** (literally, "new growth") is an example of such tissue in which the control of growth becomes defective. Neoplasms can be defined as benign or malignant based on their likelihood of spreading. Encapsulated nodules of neoplastic tissue that do not spread are called benign tumors. On the other hand, malignant tumors often invade neighboring tissues and even other parts of the body, and thus may become lethal. **Cancer** is the common term for a malignant tumor. The word is from the Latin term for "crab" because early physicians noticed certain cancers had a crablike appearance (Becker *et al*. 1996).

1.1.1. Neoplastic Transformation and Tumor Progression

In most cases, malignant tumors develop from a single progenitor cell. The progenitor cell has undergone a series of irreversible (permanent and heritable) and cumulative changes in a process called **neoplastic transformation**. There are two general characteristics of the transformed cells: they undergo uncontrolled growth and tend to spread. The spread of cancer cells to neighboring tissues is called **invasion;** the spread to distant organs is called **metastasis.** The term *metastases* is used to refer to the tumor nodules that implant at sites distant from the parent tumor (Becker *et al*. 1996).

Tumor progression is the incremental development of increasingly malignant characteristics by a tumor. Typically, tumors are relatively benign, slowly growing, weakly invasive or noninvasive in the early stages of development. With time, however, they can enter a phase of increasingly rapid growth and become highly invasive and metastatic.

1.2. Molecular Genetics of Cancer

Cancer is a genetic disease resulting from mutations in somatic cells (Alberts *et al.* 2002). Several lines of evidence indicate that tumorigenesis in humans is a multistep process and that these steps reflect genetic alterations that drive the progressive transformation of normal human cells into highly malignant derivatives (Hanahan *et al*. 2000).

1.2.1. Genomic Integrity

Genomic instability, which results in an elevated mutation rate, is a fundamental prerequisite of tumorigenesis (Schmute *et al*. 1999). Some cancer cells are defective in the ability to repair local DNA damage or to correct replication errors that affect individual nucleotides. These cells tend to accumulate more point mutations than do normal cells. Other cancer cells cannot maintain the integrity of their chromosomes properly and thus display gross abnormalities in their karyotype (Alberts *et al.* 2002). Cells must protect the integrity of their genome to avoid both the inheritance of deleterious mutations by daughter cells and the accumulation of mutations in genes that control cell proliferation. Although there are many safeguards in cells to protect the genomic integrity, cellular DNA is constantly bombarded by mutagens from endogenous and exogenous sources. DNA replication, gene transcription, DNA repair and cell cycle checkpoints must all interlink to promote cell survival following DNA damage and protect the integrity of chromosomes. A highly coordinated response to DNA damage is the activation of appropriate repair pathways and reversible arrest at cell cycle checkpoints. The cell cycle arrest gives time for repair to be completed (Levitt *et al*. 2002).

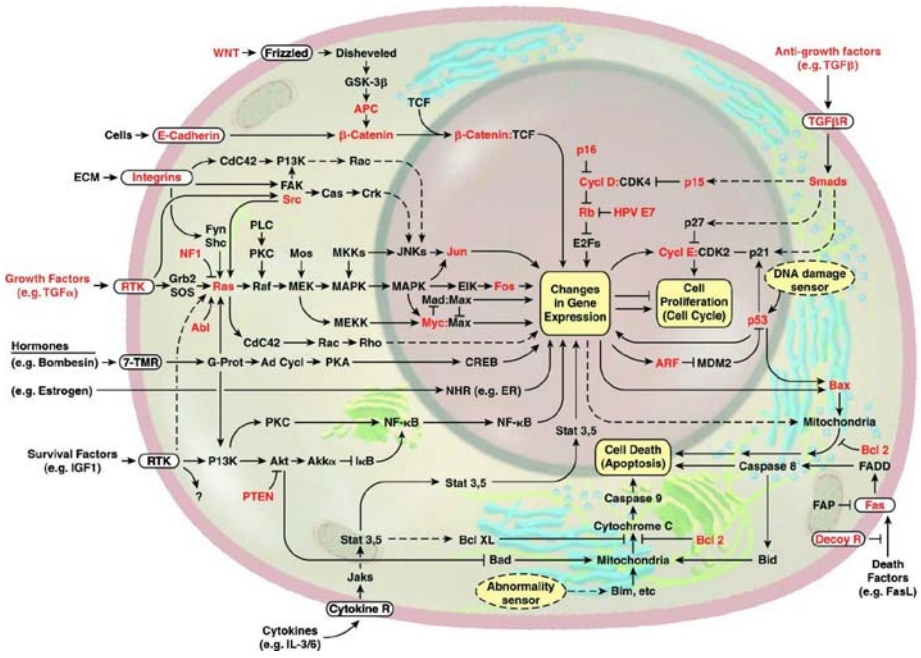
The p53 protein, also known as the "guardian of the genome", responds in several ways to DNA damage in the cell. p53 acts as a transcription factor, stimulating synthesis of a 21-kDa protein that inhibits cyclin-dependent kinase (Cdk)-cyclin complexes. This block stops the cell cycle when DNA damage has occurred, giving the cell time to repair the damage so that genetic errors are not passed on to daughter cells. If the repair fails, p53 can trigger the damaged cells to undergo apoptosis, or programmed cell death, before their genetic abnormalities are inherited. Recent evidence indicates that p53 also stimulates the DNA repair machinery, acting both

directly and indirectly through other proteins. Mutations in the p53 gene not only cause a lack of these protective effects but also stimulate abnormal cell growth (Becker *et al*. 1996).

1.2.2. Cancer-Critical Genes

All genes whose mutation may lead to cancer are cancer-critical genes (Alberts *et al.* 2002). Figure 1.1 summarizes the pathways and genes involved in cancer (Hanahan *et al*. 2000). These genes are classified into two groups according to whether the dangerous mutations they contain are those that cause loss of function or those that lead to gain of function: oncogenes and tumor suppressor genes.

Figure 1.1. Pathways in Cancer



1.2.2.1. Oncogenes

Gain-of-function mutations of proto-oncogenes stimulate cells to increase their ability to proliferate, disseminate, and divide when they should not. These mutations have a dominant effect, and the mutant genes are known as oncogenes (Alberts *et al.* 2002). There are different types of genetic alterations that can activate proto-oncogenes to become oncogenes. The gene may be altered by a small change in sequence such as a point mutation, by a large scale change such as partial deletion, or by a chromosomal translocation that involves the breakage and rejoining of the DNA helix. These changes can occur in the protein-coding region yielding a hyperactive product, or they can occur in adjacent control regions so that the gene is simply expressed at concentrations that are much higher than normal. Alternatively, the cancer-critical-gene may be over-expressed because extra copies are present due to gene amplification events caused by errors in DNA replication.

Oncogenes influence (directly or indirectly) functions connected with cell growth (Lewin B. *et al.* 2000). They may function as growth factors (i.e. *wnt1*- related to wingless), growth factor receptors (i.e. *c-erbB*- EGF receptor kinase), G protein/signal transduction (i.e. *c-ras*- GTP-binding protein), intracellular tyrosine kinases (i.e. c-abl- cytosolic), serine/threonine kinases (i.e. *c-raf*- cytosolic), signaling proteins (i.e. *vav*- SH2 regulator), and as transcription factors (i.e. *c-myc, c-fos, c-jun*). The common feature is that each type of protein is able to direct general changes in cell phenotypes, either by initiating or responding to changes associated with cell growth, or by changing gene expression directly.

1.2.2.2. Tumor Suppressor Genes

Tumor suppressor genes protect cells from dysregulated growth and division. Both of their alleles must be inactivated to observe a phenotypic effect. There are two types of tumor suppressor gene: ‘gatekeepers’ and ‘caretakers’ (Levitt *et al*. 2002). Gatekeeper genes act directly to regulate cell proliferation and are rate limiting for tumorigenesis. The retinoblastoma (*Rb*) and *p53* genes are examples of gatekeeper tumor suppressors. Caretaker genes do not directly regulate proliferation but when mutated lead to accelerated transformation of a normal cell to a neoplastic cell. MSH2 is an example of a caretaker which functions as a repair gene and is well

characterized in hereditary non-polyposis colon cancer. There are many so-called chromosomal instability disorders (i.e. ataxia telangiectasia) in which germ-line mutations in a caretaker gene lead to both genome instability and a predisposition to cancer. This shows the importance of these genes in suppressing neoplastic transformation.

Tumor suppressor genes act mostly in a recessive manner (Oesterreich *et al.* 1999). The classical inactivation of tumor suppressor genes is caused by chromosomal loss of one allele and mutation of the other remaining allele. Functional inactivation of tumor suppressor genes can also be caused by hypermethylation, increased degradation, or mislocalization.

The most important tumor suppressor is p53, the cellular gatekeeper for growth and division. More than half of all human cancers have either lost p53 protein or have cells with *p53* mutations (Lewin *et al.* 2000). The *p53* mutations fall into the category of dominant negative mutations, and the mutants function by overwhelming the wild-type protein and preventing it from functioning. The most common form of a dominant negative mutant is one that forms heteromeric protein containing both mutant and wild-type subunits, in which the wild-type subunits are unable to function. p53 exists as a tetramer. When mutant and wild-type subunits of p53 associate, the tetramer takes up the mutant conformation. The stability of p53 is another important parameter as it usually has a short half-life. The response to DNA damage stabilizes the protein and transactivates it. The cellular oncoprotein Mdm2 inhibits p53 activity. p53 induces transcription of *Mdm2*, so the interaction between p53 and Mdm2 forms a negative feedback loop in which the two components limit each other's activities (Lewin *et al.* 2000).

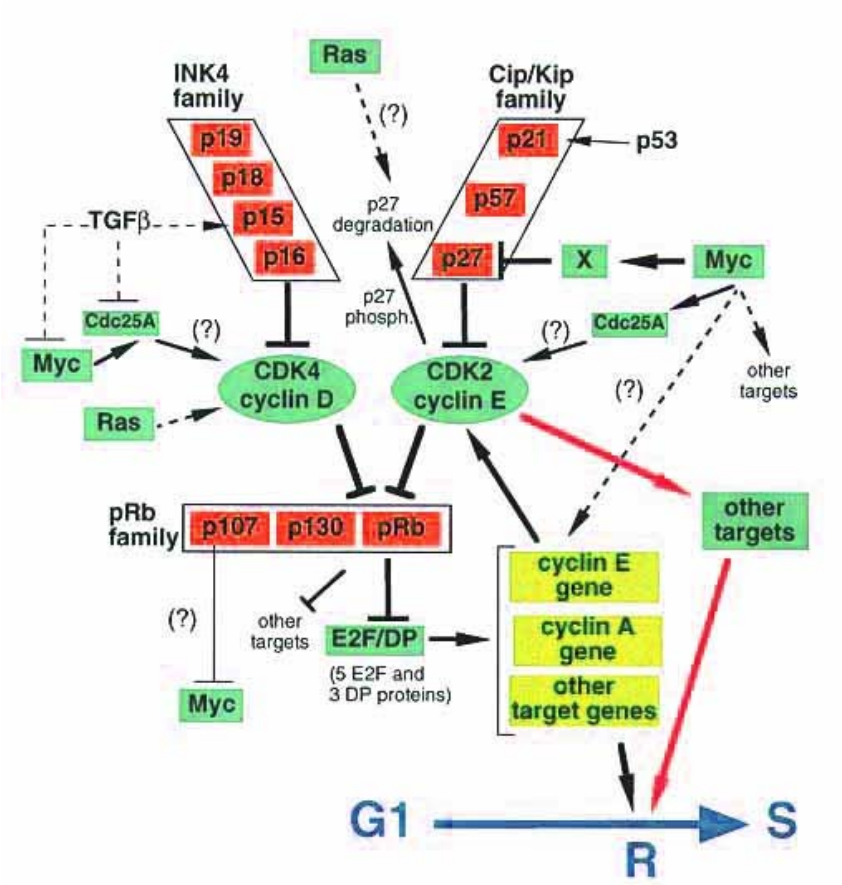
1.3. Cell Cycle Regulation and Cancer

An important component in the maintenance of the genome is the coordination and control of DNA replication, repair, and the distribution of DNA to daughter cells during each division cycle. Regulation at two stages of the cell cycle is critical in response to DNA damage: G_1-S and G_2-M boundaries. Cells delay cell cycle progression in order to facilitate the repair of DNA damage and to ensure that previous steps in the cell cycle are complete before proceeding (Kastan *et al.* 1997).

The p16-cyclin D_1-cdk4-Rb pathway is central to the regulation of the G1-to-S phase transition and to the understanding of human cancers (Figure 2) (Levine *et al.* 1997). One of these four genes is altered or mutated in nearly every cancer examined. p16 is a negative regulator of cyclin D_1-Cdk4, and the gene is heavily methylated in some cancer cells and mutated in other cancers. *Cyclin D_1* is amplified and over-expressed in a number of cancers (about 16% of breast cancers), and cdk4 mutations (no longer sensitive to p16) and gene amplifications have been reported in selected tumors. The retinoblastoma protein (Rb) is the major target of cyclin D_1-Cdk4 for cell cycle regulation and is also present in a mutant form in a number of cancers (such as small-cell lung cancer and osteosarcomas). The Rb protein regulates E2F-DP transcription factor complexes (E2F-1, -2, and –3, and DP-1, -2, and -3), which in turn regulate a number of genes (including those encoding cyclin E, cyclin A, and proliferating cell nuclear antigen) required to initiate or propagate the S phase of the cell cycle. Phosphorylation of Rb by cyclin D_1-Cdk4 releases E2F-DP proteins from the Rb complex, relieving repression of these genes or activating their transcription. The Rb protein regulates the restriction point or start, as a "go- no go" signal for cell cycle progression that is sensitive to the impact of various growth factors (via the regulation of cyclin D_1-Cdk4 and possibly p16).

In response to some forms of DNA damage, p53 is activated and turns on the transcription of one of its downstream genes, *p21 (WAF1, Cip-1)*, for G1 arrest of the cell cycle. p21 binds to a number of cyclin and Cdk complexes: cyclin D_1-Cdk4, cyclin E-Cdk2, cyclin A-Cdk2, and cyclin A-cdc2. One molecule of p21 per complex permits Cdk activity (and may even act as an assembly factor), while two molecules of p21 per complex inhibit kinase activity and block cell cycle progression. p21 also binds to PCNA (Proliferating Cell Nuclear Antigen) (at its C-terminal domain). The available evidence suggests that p21-PCNA complexes block the activity of PCNA in DNA replication, but not its activity in DNA repair. Thus, p21 can act on cyclin-Cdk complexes and PCNA to stop DNA replication (Levine *et al.* 1997).

Figure 1.2. Schematic Representation of Cell Cycle Regulators



1.4. Breast Cancer

1.4.1. Clinical Information

1.4.1.1. Incidence and Mortality

Breast cancer is the most common malignancy in women and constitutes 18% of all cancers in women (Haimov-Kochman R. *et al*. 2002). Approximately 183,000 women are diagnosed with invasive breast cancer each year and nearly 41,000 women die of the disease (DeVita *et al*. 2001). In women aged 40 to 55, breast cancer is the leading cause of all mortality.

1.4.1.2. Histopathology

Breast carcinoma arises from the epithelium of the mammary gland, which includes the milk-producing lobules and the ducts that carry milk to the nipple. Malignant transformation of the stromal, vascular, or fatty components of the breast is not included in this definition and is extremely rare. There is increasing evidence that the breast epithelium undergoes a transformation from normal to hyperplasic, followed by the appearance of atypia in association with hyperplasia, ultimately becoming malignant. Malignant cells continue to evolve from noninvasive carcinoma, typified by ductal carcinoma in situ, to invasive carcinoma, and ultimately, to cells with metastatic potential (Vogelstein *et al.* 2002).

The treatment and prognosis of a woman with breast cancer are strongly influenced by the stage at the time of diagnosis. Multiple staging systems have been proposed, but the most commonly used system is the one adopted by both the American Joint Committee and the International Union against Cancer. This staging system is a detailed TNM (tumor, nodes, metastasis) system but can be summarized as; Stage 0 (carcinoma in situ), Stage I (tumor $\leq$ 2 cm, negative axillary nodes), stage II (tumor size 2-5 cm and/or mobile positive axillary nodes), Stage III (tumor size > 5 cm and/or fixed axillary nodes; inflammatory breast cancer), Stage IV (distant metastases beyond ipsilateral axillary nodes) (Vogelstein *et al.* 2002).

1.4.1.3. Risk Factors

Multiple factors are associated with an increased risk of developing breast cancer, including increasing age, family history, exposure to female reproductive hormones (both endogenous and exogenous), dietary factors, benign breast disease, and environmental factors (DeVita *et al.* 2001).

Family history of breast cancer: The best studied and most significant risk factor is family history of breast cancer. Shared exposure to another risk factor cannot be excluded, but this most commonly represents heritable factors that increase the likelihood of developing breast cancer. The breast cancer susceptibility genes BRCA1 and BRCA2 represent the most dramatic examples, but since they account for only 15 to 20 percent of the breast cancer cases that cluster in families, other less penetrant but more common heritable factors are also considered (Vogelstein *et al.* 2002). The risk of developing breast cancer is increased 1.5- to 3.0- fold if a women has a mother or sister with breast cancer. Family history, however, is a heterogeneous risk factor and depends on the number of relatives with breast cancer, the exact relationship, the age of diagnosis, and the number of unaffected relatives.

Exposure to female reproductive hormones: The development of breast cancer in many women appears to be related to female reproductive hormones. Epidemiological studies have consistently identified a number of weaker breast cancer risk factors, each of which is associated with increased exposure to endogenous estrogens. *Early age at menarche*, nulliparity or *late age at first full term pregnancy*, and *late age at menopause* increase the risk of developing breast cancer (DeVita *et al.* 2001).

Age at menopause: In postmenopausal women, obesity and hormone therapy, both of which are positively correlated with plasma estrogen and estradiol levels, are associated with increased breast cancer risk. The age specific incidence of breast cancer increases steeply with age until menopause. After menopause, although the incidence continues to increase, the rate of increase decreases to approximately one-sixth of that seen in the premenopausal period. This dramatic slowing of the rate of increase in the age specific incidence curve suggests that ovarian activity plays a major role in the etiology of breast cancer (DeVita *et al.* 2001). The relative risk of

developing breast cancer for a women with natural menopause before age 45 is one-half that of a woman whose menopause occurs after age 55.

Age at menarche and the establishment of regular ovulatory cycles are strongly linked to breast cancer risk. Earlier age at menarche is associated with an increased risk of breast cancer; there appears to be a 20% decrease in breast cancer risk for each year that menarche is delayed.

Regarding menarche and menopause, it seems that the total duration of exposure to endogenous estrogen is an important parameter in breast cancer risk.

Pregnancy: The relationship between pregnancy and breast cancer risk appears more complicated. *Age at first full term pregnancy* clearly influences breast cancer risk. Based on epidemiological studies, women whose first full term pregnancy occurs after age 30 have a two- to fivefold increase in breast cancer risk in comparison with women who have their first full term pregnancy before approximately age 18. Additionally, terminal differentiation of breast epithelial cells does not occur until the onset of lactation after the completion of a full term pregnancy. This final stage of differentiation may confer increased resistance to carcinogens (Vogelstein *et al.* 2002).

Body mass index (BMI; kg/m^2): Many studies have examined breast cancer in relation to body weight, height, and overall body size (BMI) (Wrensch *et al.* 2003). Most case-control and cohort studies of increased height, a variable highly correlated with age at menarche, and risk of breast cancer suggest a positive relationship (DeVita *et al.* 2001, Brinton *et al.* 1992). Although being overweight (high BMI) during early adult life has been associated with a lower incidence of premenopausal breast cancer (Wrensch *et al.* 2003, Brinton *et al.* 1992, Franceschi *et al*. 1996), weight gain after age 18 is associated with a significantly increased risk of postmenopausal breast cancer. The protection conferred by increased weight early in life is thought to be secondary to increased irregularity of menstrual cycles in these women, suggesting their exposure to endogenous estrogens is decreased. The increased risk with weight gain in later adult life has been explained by increased estrogen levels in these women secondary to increased production in adipose tissues (DeVita *et al.* 2001).

1.4.2. Genetics of Breast Cancer

Breast cancer is a complex and heterogeneous disease caused by the interaction of both genetic and non-genetic factors. *BRCA1* and *BRCA2* are the two high penetrance breast cancer genes. Breast cancer in families with germ-line mutations in these genes appears as an autosomal dominant trait. In addition, mutations in several other genes such as *TP53*, *MSH2*, and *PTEN* have been identified as rare causes of hereditary breast cancer. It is very likely that other lower-penetrance genes, whose susceptibility inheritance pattern does not fit the classic model of Mendelian inheritance, are also responsible for inherited susceptibility to breast cancer. The presence of breast cancer susceptibility genes is directly responsible for 5 to 10 percent of all breast cancers (Vogelstein *et al.* 2002).

Breast cancer due to inherited susceptibility has several distinctive clinical features: age at diagnosis is considerably lower than in sporadic cases, the prevalence of bilateral breast cancer is higher, and the presence of associated tumors in affected individuals is noted in some families. Associated tumors may include ovarian, colon, prostate, and endometrial cancers and sarcomas. However, breast cancer due to inherited susceptibility does not appear to be distinguished by histologic type, metastatic pattern, or survival characteristics (Vogelstein *et al.* 2002).

Table 1.1 summarizes the inherited defects in somatic genes responsible for hereditary and familial breast cancers (DeVita *et al.* 2001).

Table 1.1. Major Genetic Defects in Breast Cancer

ESTABLISHED FAMILIAL BREAST GENES (ALL TUMOR SUPPRESSORS)		
Gene	**Chromosomal location**	**Disease**
TP53 (p53)	17p13 (mutated, LOH)	Li-Fraumeni syndrome of multiple hereditary cancers
PTEN	10q23 (mutated, LOH)	Cowden's syndrome of multiple hereditary cancers
BRCA-1	17q21 (mutated, LOH)	Familial female breast and ovarian cancers
BRCA-2	13q14 (mutated, LOH)	Familial female and male breast cancers

ESTABLISHED BREAST CANCER PROGRESSION GENES			
Gene	**Chromosomal location**	**Class**	**Function**
C-ERBB2	17q12	Oncogene (amplified)	Growth factor receptor subunit
C-MYC	8q24	Oncogene (amplified)	Cell-cycle/cell death regulator
CCND1 (Cyclin D1)	11q13	Oncogene (amplified)	Cell-cycle G1 regulator
CDKN2 (p16)	9p21	Suppressor gene (methylated, LOH)	Cell-cycle G1 regulator
RB-1	13q14	Suppressor gene (mutated, LOH)	Cell-cycle G1 and G1/S regulator
TP53 (p53)	17p13	Suppressor gene (mutated, LOH)	Cell-cycle/cell death/DNA repair regulator
CDH1 (E-cadherin)	16q22-23	Suppressor gene (methylated, LOH)	Cell-cycle adhesion protein

Another approach to understanding the pathogenesis of breast cancer is the study of non-inherited (sporadic) breast cancers. This is an important complementary approach to the study of germ-line alterations for several reasons. First, the large majority of breast cancers do not arise as a result of inherited mutations in a single breast cancer susceptibility gene, and sporadic tumors may have fundamental molecular genetic differences. Second, genes that are frequently dysregulated or mutated in sporadic breast cancer are candidate genes for susceptibility loci. Third, the study of genetic alterations, such as mutations, deletions, and amplifications, provides clues to the mechanisms that result in the genomic instability in cancer cells. A summary of the genes altered in sporadic breast cancers is given in Table 1.2.

Table 1.2. Somatic Alterations in Breast Cancer

Gene/Region	Modification	Frequency
Growth factors and receptors		
EGFR	Overexpression	20-40%
HER-2/neu	Overexpression	20-40%
FGF1/FGF4	Overexpression	20-30%
TGFα	Overexpression	Not reported
Intracellular signaling molecules		
Ha-ras	Mutation	5-10%
c-src	Overexpression	50-70%
Regulators of cell cycle		
TP53	Mutation/inactivation	30-40%
RB1	Inactivation	20%
Cyclin D	Overexpression	35-45%
TGFβ	Dysregulation	Not reported
Adhesion molecules and proteases		
E-cadherin	Reduced/absent	60-70%
P-cadherin	Reduced/absent	30%
Cathepsin D	Overexpression	20-24%
MMP_s	Increased expression	20-80%
Other genes		
bcl-2	Overexpression	30-45%
c-myc	Amplification	5-20%
nm23 (NME1)	Decreased expression	Not reported

1.4.3. Single Nucleotide Polymorphism Analysis and Importance

Complex diseases do not follow a simple Mendelian mode of inheritance and frequently have an environmental component of causation. Many genes seem to be involved with comparatively low individual impact but, nevertheless, have considerable overall contribution. To understand the genetic contribution to the etiology of complex diseases, model calculations for detection of genes or alleles with modest effect use the approach called "association study" by geneticists or "case-control study" by epidemiologists (Becker *et al.* 2003).

The original Mendelian view of the genome classified alleles as either wild-type or mutant. Subsequently multiple alleles, each with a different effect on the phenotype were recognized. In some cases it may not even be appropriate to define any one allele as "wild-type". The coexistance of multiple alleles at a locus is called genetic polymorphism (Lewin *et al.* 2000). An allele is usually defined as polymorphic if multiple alleles exist as stable components and if it is present at a frequency of >1% in the population. An SNP (single nucleotide polymorphism) marker is just a single base change in a DNA sequence, with an alternative of two possible nucleotides at a given position (Vignal *et al.* 2002). Although in principle any of the four possible nucleotide bases can be present at each position of a sequence stretch, SNPs are usually biallelic in practice. In every 1000 bases along the human chromosomes, on average approximately one nucleotide position is estimated to differ between any two copies of that chromosome (Landegren *et al.* 1998).

There are different reasons why SNPs are currently utilized in epidemiological studies. One is their use in genome-wide scans as markers for disease genes. Another reason is the interest in allele-specific variation on the population level introduced by functionally relevant SNPs or by susceptibility loci in close linkage with them (Becker *et al.* 2003). In this instance, it is usual to start with candidate genes whose functional relevance for a disease is known or strongly assumed; and to consider several genes along well-established functional pathways, since most likely more than a single gene is associated with the disease.

There recently has been a shift away from monogenic disorders toward the analysis of complex multifactorial diseases such as osteoporosis, diabetes, cardiovascular and inflammatory diseases, psychiatric disorders and most cancers, which occur at a

much higher frequency than single gene disorders. There is also increasing interest in the genetics of drug response (pharmacogenetics), an understanding of which may allow the 'tailoring' of therapies on an individual basis (Gray *et al.* 2000). The broadly familial nature of complex diseases clearly indicates a significant genetic component. However, in contrast to monogenic conditions, this genetic element is comprised of multiple gene variants each contributing a small effect. The extent of this problem is likely to be so great that the frequency of any polymorphism contributing to a disease phenotype will be only slightly elevated in a disease group compared with unaffected controls. Association studies with a large sample size, where cases of disease are compared with matched controls from the same population, are likely to provide a greater opportunity to detect small effects. Single nucleotide polymorphisms (SNPs) are the most abundant and stable types of DNA sequence variation in the human genome due to low mutation rates. Many SNPs also have functional consequences if they occur in the coding or regulatory regions of a gene (Gray *et al.* 2000). The SNP markers have gained more and more popularity for their quick, accurate, and inexpensive properties for the genetic analyses of various diseases. The SNP markers provide a new method for identification of complex gene-associated diseases such as breast cancer (Hsieh *et al*. 2001).

A handful of molecular strategies are in use for SNP analysis. All current methods involve target sequence amplification, and this is followed by distinction of DNA sequence variants by short hybridization probes or by restriction endonucleases, discrimination of mismatched DNA substrates by polymerases or ligases, or by observing the template-dependent choice of nucleotide incorporated by a polymerase (Landegren *et al.* 1998).

1.4.4. Genetic Polymorphism and Breast Cancer

Breast cancer is a clinically heterogenous disease, as evidenced by the widely variable morphological appearance and distinctive gene expression profiles. Because of possible effects on protein function or expression, it is reasonable to suspect that polymorphisms in genes involved in carcinogen metabolism, estrogen production, DNA repair and cell-cycle control could predispose individuals to the development of breast cancer, as well as influencing the clinical phenotype of the tumor. Genetic variants associated with an amino acid change can obviously have consequences for protein function, while those that occur in promoter or intronic regions could alter the level of gene expression. Alternatively, the genetic variant may have no direct functional implications but could be linked to other polymorphisms that have altered functions relative to the wild-type sequence (Powell *et al*. 2002).

Although 10-15% of breast cancer cases have some family history of the disease, only 5% can be explained by rare, highly penetrant mutations in genes such as *BRCA1* and *BRCA2*. First degree relatives of breast cancer patients have a two-fold increase in risk over the general population, most of which cannot be accounted for by BRCA1 or BRCA2 (Dunning *et al*. 1999). Apart from shared environmental factors, the remaining familial risk may be due to common, low-penetrance genetic variants that are also referred to as modifier genes. Modifier genes have subtle sequence variants or polymorphisms that are associated with a small to moderate increased relative risk for breast cancer. Such variants are relatively common in the population and may be associated with a much greater proportion of breast cancer risk as a whole than the rare high penetrance genes (Weber *et al.* 2000).

There are different ways of presenting gene polymorphism data in relation to breast cancer risk, depending on the nature of the polymorphism. In the case of simple biallelic polymorphisms, allele frequencies in cases and controls can be compared using the X^2 test to ascertain statistical significance. However, this method does not produce an easily interpretable measure of the magnitude of breast cancer risk and also lacks statistical power compared with some alternatives (Dunning *et al*. 1999). A more appropriate method is to compare genotype frequencies of three possible genotypes among cases and controls. The relative risk of breast cancer for each genotype is then estimated by the odds ratio (OR). The baseline group is usually the common allele homozygotes, which by definition has an OR (and relative risk) of 1.

Depending on the allele frequencies, the number of rare allele homozygotes may be very small, particularly in small studies, and the associated OR will have a wide confidence interval. Under these circumstances, it is common to combine the heterozygotes and rare-allele homozygotes and calculate the rare-allele carrier OR. However, this risk estimate is valid only if the genetic model is dominant, an assumption that should not be made without appropriate evidence (Dunning *et al*. 1999).

Candidate low penetrance gene products have been chosen on the basis of biological plausibility, in that alterations in the protein would affect a pathway involved in carcinogenesis. Low penetrance candidates are found in a wide variety of pathways, ranging from the detoxification of environmental carcinogens to steroid hormone metabolism and DNA damage repair (Weber *et al.* 2000). Candidate modifier genes can be divided into three main groups: genes for proteins with roles in steroid hormone metabolism; genes coding for carcinogen metabolism enzymes; and common alleles of genes that have been identified through family studies such as *TP53* and *BRCA1*. The candidate gene polymorphisms and their possible functional effects are listed in Table 3 (Dunning *et al.* 1999). The existence of low-penetrance genetic polymorphisms may explain why some women are more sensitive than others to environmental carcinogens such as replacement estrogens (Coughlin *et al.* 1999).

Steroid hormone metabolism genes: Several factors, such as age at menarche, age at first pregnancy, number of pregnancies, and age at menopause alter exposure to endogenous hormones and many of these alter breast cancer risk. Hence, genes involved in the metabolism of sex hormones are strong candidates for breast cancer susceptibility genes. Those which take part in the sex hormone biosynthesis pathway may affect production of, and thus exposure to estradiol, the most active estrogen. Genes in this pathway include *CYP17, CYP19*, and the gene for 17β-hydroxysteroid dehydrogenase type 2.

The bioavailability of hormones is partially controlled by catabolism, and catechol estrogens (2 hydroxy-estrogens) are the major breakdown products of estrogens. COMT is a phase II enzyme that methylates catechol-estrogens during their conjugation and inactivation. It has two forms: one membrane-bound and the other cytosolic; both are expressed in breast tissue and share a polymorphism associated with differences in methylation activity.

The sex hormones control the activation of responsive genes by first binding to specific receptors and forming complexes that can in turn bind to sequences in the promoters of downstream, hormone-responsive genes, such as Estrogen Receptor, Progesterone Receptor, and Androgen Receptor, which are candidates for breast cancer susceptibility genes (Dunning *et al*. 1999). The biological role of estrogens, including the growth and differentiation of normal mammary tissue, is mediated through the nuclear receptor protein (ER) that has an estrogen and DNA binding domains (Coughlin *et al.* 1999).

Carcinogen metabolism genes: Several enzymes function in the detoxification of xenobiotic compounds, and their gene expression is induced in response to the presence of the compound (e.g. polycyclic aromatic hydrocarbons found in tobacco smoke). The actions of phase I and phase II enzymes make susceptible compounds more soluble and more readily excreted and consequently reduce cancer risk. However, the more soluble products of some compounds are even more potent carcinogens than the less soluble form. Hence, a genetic change that increases the expression of the gene or the activity of the protein produced may increase the amount of reactive carcinogen formed and, thus, increase the risk of cancer (Dunning *et al.* 1999).

Two phase I enzymes, CYP1A1 and CYP2D6, are induced by, and act on, carcinogens found in tobacco smoke. Both have polymorphic differences in either inducibility or activity. CYP2E1, an enzyme that metabolizes ethanol, is also a candidate because epidemiological studies suggest that breast cancer risk is increased with alcohol consumption.

The GST family are phase II enzymes that are potentially important in regulating susceptibility to cancer because of their ability to metabolize reactive electrophilic intermediates to usually less reactive and more water soluble glutathione conjugates (Mitrunen *et al.* 2001). For both *GSTM1* and *GSTT1*, a high percentage of the Caucasian populations are homozygous for null alleles (up to 60 and 20%, respectively) and have no detoxifying GST activity. Levels of DNA adducts, sister-chromatid-exchange, and somatic genetic mutations may be increased in carriers of *GSTM1* and *GSTT1* null genotypes (Dunning *et al.* 1999).

The N-acetyl transferases, NAT1 and NAT2, are also phase II enzymes, and metabolise aromatic amines, which are present in cigarette smoke and heterocyclic amines in cooked meats (Weber *et al*. 2000). However, the action of NATs on these carcinogens can produce electrophilic ions that may induce point mutations in DNA (Dunning *et al*. 1999). Aryl aromatic amines are mammary carcinogens whose rate of metabolic activation is determined by polymorphisms in *NAT* genes (Weber *et al*. 2000). Polymorphism in both genes results in two phenotypes: slow acetylators who are homozygous for low-activity alleles, and fast acetylators who carry one or more high-activity alleles (Dunning *et al*. 1999).

Common alleles of high-penetrance genes: Mutations in the *TP53* and *BRCA1* genes are associated with a high risk of breast and other cancers. Mutation in the *TP53* gene results in decreased p53 activity, which may lead to failure of cells with DNA damage to arrest and thus to continue to replicate with damaged DNA (Dunning *et al.* 1999). Polymorphisms in the p21 downstream component of p53 pathway are also described. In the case of BRCA1, where the protein function is still uncertain, the majority of confirmed mutations generate truncated proteins that are likely to have severely reduced activity. It has been hypothesized that amino acid substitutions outside the major functional domains may confer more moderate breast cancer risks. The majority of these substitutions are rare, and putative functional effects remain unconfirmed (Dunning *et al.* 1999).

According to the recent review of published case-control studies, polymorphisms in *CYP19, GSTM1, GSTP1* and *TP53* appear to be stronger candidates for low-penetrance breast cancer susceptibility genes, although they too need to be confirmed in larger studies (Dunning *et al.* 1999).

Table 1.3. Genetic Polymorphisms in Relation to Breast Cancer Risk

Gene	Base /Amino acid change	Functional effect
Steroid hormone metabolism genes		
COMT	Exon 4 G → A/ Val158Met	Reduced activity
CYP17	Promoter T → C (T1931C)/ None	Creates a fifth SpI site and might increase transcription
CYP19	Intron 4 $(TTTA)_u$ microsatellite/ None Intron 4 TCT insertion/deletion/ None	Unlikely Unlikely
CYP2D6	2367delA (A allele)/ Frameshift Intron 3G → A (G1934A) (B allele)/ Premature stop at residue 544 Del Lys281 (C allele) 17.5-kb deletion (D allele)	Nonfunctioning enzyme Nonfunctioning enzyme Catalytically normal enzyme, but wrong cellular compartment No enzyme
EDH17B2	Exon 6 A → G/ Ser312Gly	Unlikely
ER	CCC325CCG/ Pro325Pro Intron 1/exon 2 XbaI site	None Unlikely
PR	Alu repeat insertion in introns G/ None	Unlikely

Table 1.3. Continued

Gene	Base /Amino acid change	Functional effect
Carcinogen metabolism genes		
CYP1A1	Exon 7 A → G (A4889G)/ Ile462Val 3' UTR T → C (T6235C)/ None Exon 7 C → A (C4887A)/ Thr461Asp 3' UTR T → C (T5639C)/ None	Uncertain, possible increase in enzyme activity None Unknown None
CYP2E1	Intron 6 unspecified/ None	Unlikely
GSTM1	Gene deletion	Null individuals have no enzyme
GSTP1	A313G/ Ile105Val	Reduced enzyme activity
GSTT1	Gene deletion	Null individuals have no enzyme activity
NAT1	A1088T/ None	Possible increase in enzyme activity
NAT2	G191A C481T G590A G857A	Low activity allele Low activity allele Low activity allele Low activity allele
Other genes		
BRCA1	C2731T/ Pro871Leu	Unknown
BRCA1	G1186A/ Gln356Arg	Unknown
HSP70-2	1267/ Silent	Unknown
HSP70-hom	2437/ Met493Thr	Unknown
TNF-α	-308 G → A/ None	Increased constitutive and inducible levels of TNF-α
TP53	Exon 3 G → C/ Arg72Pro 16-bp insertion in introns 3/ None Intron 6 G → A/ None	Unknown Unlikely Unlikely

1.4.4.1. *p53* and *p21* polymorphisms

Polymorphisms in *TP53* are considered candidate risk factors because of the very important role played by this gene in the maintenance of genomic integrity following genotoxic insult (Powell *et al.* 2002). Highly penetrant germline mutations in *TP53* are very rare, but polymorphisms are quite common and at least 14 polymorphisms have been described (Keshava *et al.* 2002). Five of these are in exons (codons 21, 36, 47, 72, and 213), and 9 are in introns (intron numbers 1-3, 6, 7, and 9) (Keshava *et al.* 2002). Those investigated for association with breast cancer include a 16 bp insertion in intron 3, an Arg72Pro polymorphism in exon 4 and a single nucleotide polymorphism in intron 6. Only the codon 72 polymorphism appears to be significantly associated with the risk of breast cancer (Dunning *et al.* 1999).

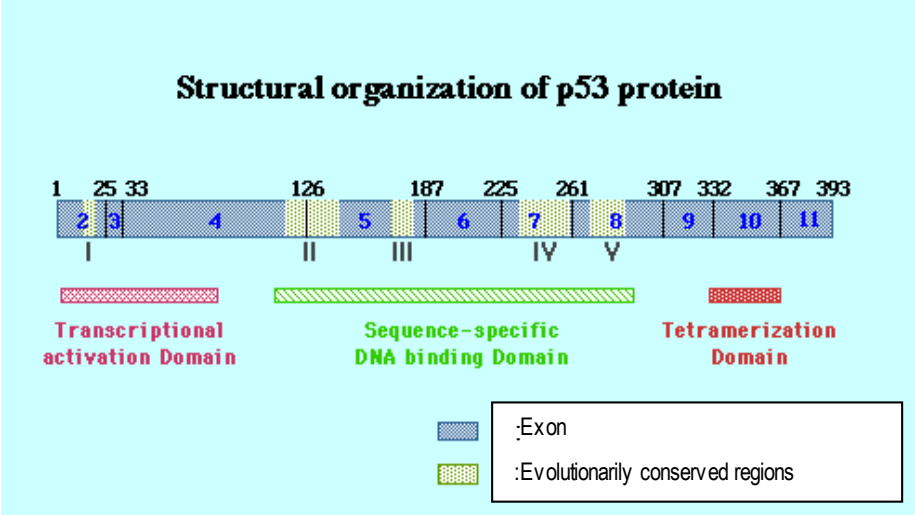
A major downstream component of the TP53 tumor suppressor pathway is the p21 cyclin dependent kinase inhibitor, also known as WAF1 or CIP1 (Powell *et al.* 2002). It was initially thought that somatic mutations in this gene might be involved in tumor formation, particularly for cases having wild type TP53; however, *p21* mutations proved to be extremely rare in a variety of cancer types investigated (Powell *et al.* 2002). Polymorphisms in *p21* have been described, with the two most common being Ser31Arg in exon 2 and a single nucleotide polymorphism in the 3' untranslated region of exon 3, 20 bp downstream from the stop codon (Powell *et al.* 2002). Another polymorphism, Asp149Gly, has also been reported in an Indian population (Powell *et al*. 2002). Interestingly, both codon 31 and 149 polymorphisms appear to occur more frequently in patients whose tumors contain wild type TP53 (Powell *et al*. 2002). Another suspected p21 polymorphism occurs in the 5' region of intron 2 but this remains to be confirmed (Powell *et al.* 2002).

Su *et al.* (2003) reported that different *p53* and *p21* genotypes or their combinations are associated with an altered human gene expression of *p21*. The genotype combination involving both the *p53* codon 72 Pro allele and the *p21* codon 31 Arg allele is associated with a particularly low expression of *p21*.

1.4.4.1.1. p53 Structure-Function Relationship and Polymorphism

The p53 tumor suppressor gene is located at 17p13.1 and encodes a 53-kDa nuclear phosphoprotein whose primary role is to maintain genomic integrity through cell cycle arrest, DNA repair, and apoptosis. The protein consists of 393 amino acids that can be functionally divided into three domains (Figure 3, Ko *et al.* 1996). The NH_2 terminus (amino acids 1-95) controls the transactivational activity of the protein, the central region (amino acids 102-292) controls the DNA binding activity, and the COOH terminus (amino acids 300-393) is responsible for oligomerization, nonspecific DNA binding, and DNA damage recognition (Powell *et al.* 2000).

Figure 1.3. Structural Organization of p53 Protein



Polymorphisms in *TP53* are considered candidate risk factors because of the very important role played by this gene in the maintenance of genomic integrity following genotoxic insult (Powell *et al*. 2002). In human populations, the *p53* gene has a common polymorphism at codon 72. The alleles of the polymorphism at codon 72 (exon 4) are 'G' or 'C' at the nucleotide residue 347. When 'G' is present it encodes an arginine amino acid (CGC; Arg72) with a positive-charged basic side chain, when 'C' is present it encodes a proline residue (CCC; Pro72) with a nonpolar-aliphatic side chain (Langerod *et al.* 2002). Matlashewski *et al.* (1987) concluded that this is a

nonconservative amino acid change, and results in a structural change in the protein, since the $p53_{pro}$ variant migrates more slowly than the $p53_{Arg}$ variant in sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE). It was also noted that the tumors produced by the Pro-72 *p53*-containing cells appeared more slowly and were smaller than the Arg-72 p53 tumors.

The polymorphism occurs in the proline-rich domain of p53, which is required for the growth suppression activity of p53 and also plays an important role in p53-mediated apoptosis but not in cell cycle arrest (Thomas *et al.* 1999). This polyproline region is considered to be an Src homology 3 (SH3) binding domain, and the proline at amino acid 72 constitutes one of the five PXXP SH3 binding motifs defined within this region. As evidenced by monoclonal antibody reactivity, both proteins are structurally wild type, and they exhibit similar levels of affinity for a variety of p53 DNA recognition sequences. However, there are subtle differences in their abilities to interact with basic elements of the transcriptional machinery, and this is reflected in differences in the abilities of each form to induce apoptosis and suppress transformed cell growth. p53-Pro is a stronger inducer of transcription than p53-Arg, whereas a *p53* Arg/Arg genotype induces apoptosis with faster kinetics and suppresses transformation more efficiently than the *p53* Pro/Pro genotype (Thomas *et al*. 1999).

The proline-rich PXXP domain between residues 60-90 of p53 is required for cooperation with anti-neoplastic agents to promote apoptosis of tumor cells, while deleting the C-terminal 30 amino acids of p53 does not have any effect (Baptiste *et al*. 2002).

E6 proteins from both high-risk and low-risk HPV types are able to target $p53_{Arg}$ more efficiently than $p53_{Pro}$ for ubiquitin-mediated degradation. Consistent with this observation, the majority of HPV-associated tumors are homozygous for the $p53_{Arg}$ allele, whereas the majority of the comparable normal population was heterozygous (Thomas *et al.* 1999).

p53 codon 72 polymorphism influences the ability of certain conformational p53 mutants to form stable complexes with p73. When codon 72 encodes Arg, the ability of mutant p53 to bind p73, to neutralize p73-induced apoptosis and to transform cells in cooperation with EJ-Ras is enhanced. Arg-containing allele was preferentially

mutated and retained in squamous cell tumors arising in Arg/Pro germline heterozygotes. Formation of such complexes correlates with a loss of p73 DNA-binding capability, and consequently its ability to serve as a sequence-specific transcriptional activator and an inducer of apoptosis (Marin *et al.* 2000).

Langerod *et al*. (2002) reported that in breast cancer cases, the occurance of a *p53* mutation was significantly more often found on the Arg72 allele than the Pro72 allele. The observed skewed occurance of somatic *p53* mutations on the Arg72 allele in breast carcinomas suggests that this combination gives breast epithelial cells a growth advantage, which may increase the risk of malignant transformation and development of cancer. The coexistance of the Arg72 with a mutation may modify the p53 protein structure in a way that interferes either with the protein's ability to achieve sequence-specific binding to DNA or with the interaction and recruitment of the transcription machinery, causing an altered transcription pattern. Another possibility is that the Arg72 may modify the mutant p53 protein's ability to bind to and interact with other proteins such as p73 (Langerod *et al.* 2002).

p53 recessive mutants carrying the Arg allele can lead to decreased activation of p53 target genes through inactivation of p73. The transdominant p53 mutants achieve this by inactivation of the remaining wild type *p53* allele (Tada *et al*. 2001).

For p53 mutants associated with human tumors, the arginine variant confers greater resistance to p73-dependent apoptosis and cytotoxicity than the equivalent proline form. This correlates with cellular resistance to the apoptotic and cytotoxic activity of specific cancer chemotherapeutic agents (Bergamaschi *et al.* 2003).

In cell lines containing inducible versions of alleles encoding the Pro72 and Arg72 variants, and in cells with endogenous p53, the Arg72 variant induces apoptosis markedly better than does the Pro72 variant. At least one source of this enhanced apoptotic potential is the greater ability of the Arg72 variant to localize to the mitochondria; this localization is accompanied by release of cytochrome c into the cytosol. The two polymorphic variants of p53 are functionally distinct, and these differences may influence cancer risk or treatment (Dumont *et al.* 2003).

Using microsatellite analysis, the frequency of LOH at the *TP53* locus was investigated in patients heterozygous for the codon 72 polymorphism and it was

determined that the Arg allele is preferentially retained in patients heterozygous for this polymorphism (Papadakis *et al*. 2002).

There have been many studies conducted in order to investigate the association of *p53* codon 72 polymorphism and cancer risk. While some of the results are significant, some of them show inconsistency. It is important to note that the results differ according to ethnicity. Table 1.5 summarizes some of the studies.

1.4.4.1.2. p21 Structure-Function Relationship and Polymorphism

Although *p53* mutation is the most common genetic change reported in human cancers, about 50% of cancers do not have *p53* mutations. Mutations or alterations on genes situated upstream or downstream of p53 on the same control pathway might have a similar oncogenic effect. Thus *WAF1/CIP1* alterations could be good candidates to substitute for *p53* mutations (Li *et al.* 1995). Table 1.4 summarizes some of the studies conducted so far in relation to *p21* codon 31 polymorphism and cancer risk.

The human $p21^{waf1/cip1}$ localized to chromosome 6p 21.2 is a cyclin dependent kinase inhibitor (CDKI) upregulated by wild type tumor suppressor protein p53. Wild type p53 binds to a site 2.4 kb upstream of the *p21* coding sequence and stimulates gene expression (Su *et al.* 2003).

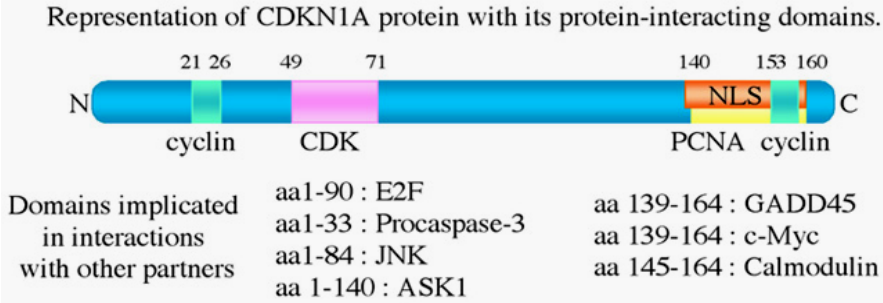
Loss of heterozygosity of the short arm of chromosome 6 where *p21* is situated has been described in cases of colon, lung, ovary and renal cancers, suggesting that *WAF1/CIP1* may be inactivated by a two-hits process in the corresponding tumors (Li *et al.* 1995).

Mutations and deletions of the *p21* gene have been rare in human cancers suggesting that *p21*, if involved in tumorigenesis, may be exerting itself mainly on the expression level rather than on the gene level (Bahl *et al.* 2000). However, *p21* polymorphisms have been observed in various cancers. The polymorphic variants have been reported to occur more frequently in cancer patients than in healthy individuals suggesting a role in increased susceptibility to the development of some types of cancers (Mousses *et al.* 1995). Moreover, the frequency of the *CIP1/WAF1* variants in tumors which contain *p53* gene mutations was found to be significantly less than the frequency of the CIP1/WAF1 variants in tumors without *p53* gene mutations. These data suggest that the variants at codon 31 and/or in the 3`UTR may not be benign polymorphisms, but may possibly be associated with a higher risk of developing cancer (Mousses *et al.* 1995).

The *WAF1/CIP1* gene consists of three exons of 68, 450 and 1600 bp and encodes a 21 kDa protein of 164 amino acids (Figure 1.4). The first exon is non-coding while exon 2 contains 90% of the coding sequence (Li *et al.* 1995). The first ATG codon appears at nucleotide 76 in exon 2, and a termination codon appears at nucleotide

570 in exon 3 (Ralhan *et al.* 2000). The unique ability to associate the proliferating cell nuclear antigen (PCNA), an auxiliary factor for DNA polymerase δ and ε, distinguishes p21 from other cyclin-dependent kinase inhibitors (CDKIs). The Cdk and PCNA inhibitory activities of p21 have been mapped to different domains of the protein. An N-terminal domain which binds and inhibits cyclin-Cdk complexes, and a short sequence near the C-terminus (between amino acid residues 144-151) which binds to PCNA results in the inhibition of DNA replication (Bahl *et al.* 2000).

Figure 1.4. Representation of p21 Protein



http://www.infobiogen.fr/services/chromcancer/Genes/CDKN1AID139.html

The codon 31 polymorphism is found in an area of greater than 90% homology at the protein level with the murine homologue, which is thought to encode a DNA-binding zing-finger domain (Shih *et al.* 2000) in amino acids 13 and 41, and contains a potential nuclear localization signal between amino acids 140 and 163 (Hachiya *et al.* 1999). The sequences between amino acids 13 and 56 are almost perfectly conserved between mouse and human, and there is strong homology between WAF1 and $p27^{KIP1}$ protein, as well as $p57^{KIP2}$ protein. This conservation of the amino acid sequences suggests that this region is important to the function of WAF1 as a CDK inhibitor (Hachiya *et al.* 1999). In addition, serine is an uncharged polar amino acid with a single hydroxy 1 side chain, whereas arginine is a basic, positively charged amino acid with a seven-membered side chain. These observations raise the possibility that this polymorphism encodes functionally distinct proteins. In fact, it has been reported that the Arg variant has been observed in a significant number of cancer cases (Hachiya *et al.* 1999). However, transfection studies have shown no difference in the tumor suppressor abilities of the Ser and Arg alleles in a lung cancer

cell line (Chedid *et al.* 1994). In addition , *in vitro* CDK-cyclin kinase assays have shown that wild type Ser-p21 and the variant Arg-p21 both have similar growth-inhibitory abilities (Sun *et al.* 1995). Moreover, a tumor suppression assay using H1299 cells, which lack p53 protein expression, revealed no functional difference between the protein encoded by the codon 31 Ser and Arg alleles of the *WAF1* gene (Hachiya et al. 1999). However, the assay used constructs which produced high levels of WAF1 protein constitutively. *WAF1* induction is involved in a stress response to DNA damage via transcriptional modulation by p53. Thus, these assays would not reproduce the *in vivo* expression pattern of the WAF1 protein, as neither the cells nor the experimental conditions can be considered equivalent to those of the normal cellular environment. Experiments using endogeneous regulatory sequences of the *WAF1* gene will be required to evaluate the functional difference between wild-type and variant proteins.

Previous studies indicated a significant association of the Arg allele with human malignancies including breast tumor, sarcoma, rhabdomyosarcoma, neuroblastoma, prostate adenocarcinoma, colorectal cancer and lung cancer (Table 1.4). In those studies, the Arg allele was reported to occur more frequently in malignancy groups than in the healthy control, suggesting that the Ser to Arg substitution results in some functional differences which contribute to tumorigenesis (Konishi *et al*. 2000).

The 3' untranslated region of several genes has been shown to be a region important in cellular proliferation, differentiation, tumor suppression, and metastasis suppression. DNA sequences that determine mRNA stability and its rate of degradation exist in this region. Because the two *p21* variants appear to segregate together, a synergistic mechanism may be involved in maintaining an altered cellular phenotype. The 3' untranslated region polymorphism may be located at a site necessary for rapid *p21* message degradation, the variant preventing its timely demise. The *p21* mRNA containing the codon 31 variant may be rendered more stable because of its association with the 3'-UTR polymorphism. This altered message, with an extended half-life, may be sufficient to interfere with DNA damage-induced G_1 cell cycle arrest and increase the sensitivity of cells to DNA damage. This may lead to the observed increase in cancer susceptibility for individuals carrying both polymorphisms, although these findings must await further confirmation by additional studies (Facher et al. 1997).

Probable differences in the p21 function between the Arg and Ser allelic products were also claimed by Sun *et al*. (1995), who reported that there was no significant difference between the two allelic products in their ability to inhibit CDK activity and tumor cell growth. Thus, the association between the Arg allele and the p21 function is questioned.

Although further investigations are required for clarifying the association of the *p21* polymorphism with tumorigenesis, it is important to take the ethnic differences of the allele frequency distribution into account when studying the role of the *p21* polymorphism in carcinogenesis (Konishi *et al.* 2000). The codon 31 polymorphism of *p21* shows distinct differences among major ethnic groups. The frequency of Arg allele ranges from 4% in caucasians, to 16% in Indians, 29% in African Blacks in the USA, and 50% in Chinese (Shih *et al.* 2000).

Table 1.4. *p21* Codon 31 Polymorphism and Cancer Association

Cancer	**Population**	**# of cases**	**# of controls**	**Odds ratio (95% CI)**	**Results**	**Reference**
Breast	Caucasian	93	187	S/S; 1.10 (0.35-3.50)	No association	Keshava *et al.,* 2002
Breast	American	88	21		No correlation	Lukas *et al*., 1997
Breast	African-American	37	65	S/S; 2.32 (0.66-5.60)	Association	Keshava *et al.,* 2002
Breast	Latinas	30	75	S/S; 2.22 (0.71-6.89)	Association	Keshava *et al.,* 2002
Breast	Caucasian	286	81		No association	Powell *et al.,* 2002
Cervical (CIN III, SCCA, Adenocarcinoma)	Korean	212	98	S/S; 3.59 (1.55-8.31)	S/S significant in adenocarcinoma	Roh *et al*., 2001
Colorectal	French	45	70		No association	Li *et al*., 1995
Endometrial	Taiwanese	102	119	R/R; 0.53 (0.25-1.14)	No association	Hsieh *et al.,* 2001
Endometrial	Japanese	54	55	R/R; 2.52 (1.09-5.80)	Arg allele associated, hypertension positive risk	Hachiya *et al*., 1999

Table 1.4. Continued

Cancer	Population	# of cases	# of controls	Odds ratio (95% CI)	Results	Reference
Endometrial	American	47	21		Association between increased p21 protein expression and 31 Arg allele	Lukas *et al.,* 1997
Head and neck carcinoma	American (Pennsylvania)	42	110		Not significant, increased Arg allele frequency	Facher *et al*., 1997
Lung	Swedish	144	761	R; 1.70 (1.00-2.90)	Significant increase in Arg allele frequency	Sjalander *et al*., 1996
Lung	Caucasion	1069	1220	R/R; 0.85 (0.40-2.00)	No association	Su *et al*., 2003
Lung	Taiwanese	155	189	R; 1.15 (0.70-1.86)	No association	Shih *et al*., 2000
Nasopharyngeal	Taiwanese	76	66		No association	Sun *et al.,* 1995
Ovarian	American	53	21		No correlation	Lukas *et al.,* 1997
Prostate adenocarcinoma	American (Pennsylvania)	54	110		Not significant, increased Arg allele frequency	Facher *et al.,* 1997
Skin	Japanese	113	165		No association	Konishi *et al*., 2000

Table 1.5. *p53* Codon 72 Polymorphism and Cancer Association

Cancer	**Population**	**# of cases**	**# of controls**	**Odds ratio (95% CI)**	**Results**	**Reference**
Acute Myeloid Leukemia	Japanese	200			No association	Nakano *et al*., 2000
Bladder	Slovakian	50	145	1.44 (0.82-2.27)	No significance	Biro *et al*., 2000
Breast	Caucasion/ Slavic	448	249	P; 0.96 (0.74-1.25)	No association	Suspitsin *et al*., 2003
Breast	Greek	56	61		Arg/ Arg association	Papadakis *et al.,* 2000
Breast	Caucasion	93	187	3.15 (1.14-8.89)	Association, stronger in postmenopausal breast cancer, no association with p21 codon 31 polymorphism	Keshava *et al*., 2002
Breast	African-American	37	65	1.29 (0.54-3.10)	No association	Keshava *et al*., 2002
Breast	Turkish	115	76		Arg/Arg association	Buyru *et al*., 2003
Breast	Latinas	30	75	0.52 (0.12-2.16)	No association	Keshava *et al*., 2002

Table 1.5. Continued

Cancer	Population	# of cases	# of controls	Odds ratio (95% CI)	Results	Reference
Cervical	Indian	85	29		Arg/Arg significant	Nagpal *et al*., 2002
Cervical	Korean	234	100		No association	Kim *et al*., 2001
Cervical	Indian	232	189	1.12 (0.69-1.82)	No association	Pillai *et al*., 2002
Cervical	Indian	134	131	2.40 (1.89-3.04)	Arg/Arg association	Saranath *et al.,* 2002
Cervical	Israeli Jewish	23	162		Arg/Arg association	Arbel-Alon *et al*., 2002
Cervical	Chinese	15	20		Arg/Arg significant	Qie *et al*., 2002
HPV-associated malignancies-esophageal (HPV+, HPV-)	Chinese	39	23		Arg/Arg significant	Li *et al*., 2002
HPV-associated malignancies-ovarian (HPV+, HPV-)	Chinese	26	13		Arg/Arg significant	Li *et al*., 2002

Table 1.5. Continued

Cancer	Population	# of cases	# of controls	Odds ratio (95% CI)	Results	Reference
HPV-associated malignancies-breast (HPV+, HPV-)	Chinese	19	9		Arg/Arg significant	Li *et al*., 2002
Gastric cardia	Caucasian	32	(non-cardia) 88	3.1 (1.4-7.3)	Arg/Arg association	Zhang *et al.* 2003
Gastric	Japanese	117	116	2.98 (1.07-8.32)	Pro/Pro association	Hiyama *et al.,* 2002
Laryngeal tumors	Greek	37	40		Arg/Arg association	Sourvinos *et al.,* 2001
Lung	American	482	510	1.45 (1.01-2.06)	Pro/ Pro association, Pro/Pro+Arg/Pro also associated, smoking increases the risk	Fan *et al.,* 2000
Lung	Greek	54	99	3.13 (1.46-6.73)	Arg/Arg association	Papadakis et al., 2002
Lung	Japanese	111	170		No association	Pierce et al., 2000
Lung	Caucasian	138	173		No association	Pierce et al., 2000
Lung	Hawaiian	85	103		No association	Pierce et al., 2000

Table 1.5. Continued

Cancer	Population	# of cases	# of controls	Odds ratio (95% CI)	Results	Reference
Squamous Cell Carcinoma	French	45	22		No association in men	Humbey et al., 2003
Vulval cancer	Caucasian	52	246		Lower risk with Arg/Arg	Rosenthal et al., 2000
Vulval intraepithelial neoplasia	Caucasian	48	246		Lower risk with Arg/Arg	Rosenthal et al., 2000
Urothelial	Japanese	112	175	2.28 (1.12-4.66)	No significance in men, Pro/Pro significant in smokers	Kuroda et al., 2003

1.5. AIM

The aim of this study is to test any possible association of the germline polymorphisms of either *p21* codon 31 or *p53* codon 72 with increased susceptibility for breast cancer either independently or combined in the Turkish and Greek populations.

In this study, the questions to be answered are as follows:

1. Does *p53* codon 72 polymorphism constitute a risk for breast cancer development in the Turkish and Greek populations?
2. Does *p21* codon 31 polymorphism constitute a risk for breast cancer development in the Turkish and Greek populations?
3. Is there any relation between the established risk factors for breast cancer and the polymorphisms of *p53* and *p21* genes?
4. Is there a combined effect of *p21* and *p53* polymorphisms for breast cancer susceptibility?

We studied the two different populations separately, because both polymorphisms show considerable differences among populations with different ethnicity.

The *p53* codon 72 polymorphism was studied, since its role as a cancer risk factor was well established in different populations, but its effect on breast cancer was not analyzed thoroughly in the Turkish population.

The *p21* codon 31 polymorphism was studied, since contradictory results have been reported for various cancer types in various populations and its relation with breast cancer susceptibility is less well established.

Analysis for any possible gene-gene interaction between the *p53* and *p21* polymorphisms was included in this study because there have been no reports on the combined analysis of *p53* codon 72 and *p21* codon 31 polymorphisms in breast cancer although they are two critical components of the same pathway for cell cycle regulation.

2. MATERIALS AND METHODS

2.1. MATERIALS

2.1.1. Subjects

Turkish and Greek populations were included in this study. The cases were female patients diagnosed with breast cancer while females with no history of (breast) cancer served as a control group. Blood samples were collected from each individual and used for genomic DNA preparation.

2.1.1.1. Turkish Population

Patients: The Turkish study population consisted of 386 cases (Table 2.1). Blood samples of 301 cases were obtained from Numune Hospital, Ibni Sina Hospital, and Gazi University Medical School, Ankara. Genomic DNA was isolated from blood using the phenol-chloroform extraction method, which is described in Section 2.2.1.

DNA for 85 of the breast cancer patients was obtained from Akdeniz University Medical School, Antalya.

A standardized questionnaire form was used in order to get information about the subject's age, weight, height, age at menarche, age at first full term pregnancy, smoking history, and family history of breast cancer (see; questionnaire form).

Controls: 301 females, with no history of breast cancer, from Numune Hospital, Ibni Sina Hospital, and Gazi University Medical School, Ankara, were used as a control group (Table 2.1). Information about the subject's age, weight, height, age at menarche, age at first full term pregnancy, smoking history, and family history of breast cancer were obtained from standardized questionnaire forms.

2.1.1.2. Greek Population

180 breast cancer and 189 controls were included in this study (Table 2.2). The DNA samples from the Greek population were sent by Dr. Drakoulis Yannoukakos (Molecular Diagnosis Laboratory, National Center for Scientific Research, Athens, Greece).

HASTA ANKET FORMU

1. Adı Soyadı:
2. Yaşı:
3. Medeni hali:
4. Yaşadığı şehir ve süresi:
5. Ağırlığı (kğ):
6. Boyu (cm):
7. Mesleği:
8. İlk mestürasyon periyodunun başlama yaşı:
9. Menapozal durumu:

   Premenapozal ise; son menstürasyon periyodunun kaç gün önce olduğu:

   Postmenapozal ise; son menstürasyon periyodunun kaç gün önce olduğu:
10. Tanı konulduğu zamanki menapozal durumu:
11. Tanının ne zaman konulduğu:
12. Uygulanan tedavi:
13. Daha öne hormon tedavisi gördümü? Ne tip?
14. Oral kontraseptif kullandımı? Nedir?
15. Kaç çocuğu var?
    a. İlk doğumunu yaptığı yaş?
    b. Son doğumunu yaptığı yaş?
16. Daha önce meme ile ilgili operasyon geçirdi mi?
17. Ooferektomi (yumurtalıkların alınması) yapıldı mı? Yapıldı ise kaç yıl önce?
18. Sigara içme alışkanlığı:

    Hiç içmedim ( ) Eskiden içerdim ( )

    1-10 sigara/gün ( ) 11-20 sigara/gün ( ) 20 ve daha fazla/gün ( )

    1 yıldır içiyorum ( ) 2-5 yıldır içiyorum ( ) 5-10 yıldır içiyorum ( )

    10-15 yıldır içiyorum ( ) 15-20 yıldır içiyorum ( ) 20 ve daha fazla yıldır içiyorum ( )
19. Sigara içilen ortamda sıkça bulunuyormusunuz?

    (a) Evet (b) Hayır
20. Alkol kullanıyormusunuz?

    (a) Evet (b) Hayır

    Nadiren Haftada 1 kez Haftada 2-3 kez Haftada 4-5 kez Haftada 6-7 kez
21. Beslenme alışkanlığınızda size en fazla uyan tanım aşağıdakilerden hangisidir?
    a. Kızartma ağırlıklı yağlı diyet
    b. Sebze ağırlıklı yağsız diyet
    c. Dengeli beslenme
21. Radyasyona maruz kaldınız mı? Hangi sıklıkla?

    (a) Evet (b) Hayır
22. Tiroid ile ilgili bir rahatsızlığınız var mı?

    (a) Evet (b) Hayır

    Hipertiroidizm ( ) Hipotiroidizm ( )
23. Aile bireylerinde ve sizde genetik bir rahatsızlık var mı? Tipi?

    (a) Evet (b) Hayır
24. Ailenizde meme kanserli başka bireyler var mı? (Anne, kardeş, anneanne, vb)

    (a) Evet (b) Hayır
25. Tümörün histopatolojisi
26. Tümör grade
27. Tümör stage
28. Östrojen reseptör durumu (+) veya (-)
29. Progesteron reseptör durumu (+) veya (-)

Table 2.1. Selected Characteristics of Breast Cancer Patients and Age-matched Control Subjects in the Turkish Population

Characteristics	Cases	Controls
Age, n; mean (standard deviation)	n=370; 49.30 (12.45)	n=288; 46.59(14.42)
Age, range	20-80	15-83
Family history of breast cancer, n; %	n=297; 14.48%	n=285; 2.46%
Body mass index (BMI), n; mean (standard deviation)	n=331; 27.267 (4.95)	n=274; 27.399 (5.26)
Age at first live birth, n; mean (standard deviation)	n=318; 22.51 (5.14)	n=244; 20.56 (3.93)
Number of children, n; mean (standard deviation)	n=361; 2.70 (1.97)	n=272; 3.14 (2.18)
Age at menarche, n; mean (standard deviation)	n=334;13.58 (1.42)	n=268; 13.83 (1.42)
Premenopausal, n; %	n=160; 43.10%	n=151; 52.61%
Postmenopausal, n; %	n=212; 56.99%	n=136; 47.39%
Smoking status, n; %	n=298; 27.85%	n=275; 30.18%

Table 2.2. Selected Characteristics of Breast Cancer Patients and Age-matched Control Subjects in the Greek Population

Characteristics	Cases	Controls
Age, n; mean (standard deviation)	n=92; 49.66 (14.87)	n=176; 50.301 (12.83)
Premenopausal, n; %	n=92; 56%	n=87; 46.8%
Postmenopausal, n; %	n=72; 44%	n=99; 53.2%

2.1.2. Oligonucleotides

Gene-specific DNA amplifications from genomic DNA were primed using oligonucleotide primers given in Table 2.3.

Table 2.3. List of Primers Used for PCR Experiments

Primer Name	Sequence (5' to 3')	Target	Size of product (bp)	Reference
p21-F	gtc aga acc ggc tgg gga tg	*p21* codon 31	272	Chen WC. *et al*. 2001 Hsieh YY. *et al*. 2001
p21-R	ctc ctc cca act cat ccc gg			
p53-F	tcc ccc ttg ccg tcc caa	*p53* codon 72	279	Storey *et al*. 1998
P53-R	cgt gca agt cac aga ctt			

2.1.3. Chemicals and Reagents

Agarose	Basica LE, EU
Glacial Acetic Acid	Carlo Erba, Milano, Italy
Bromophenol Blue	Sigma, St. Louis, MO, USA
Chloroform	Carlo Erba, Milano, Italy
Ethanol	Merck, Frankfurt, Germany
Ethidium Bromide	Sigma, St. Louis, MO, USA
Ficoll Type 400	Sigma, St. Louis, MO, USA
Gamma Micropor Agarose	Prona LE, EU
Isoamyl Alcohol	Carlo Erba, Milano, Italy
Phenol	Merck, Darmstadt, Germany
Proteinase K	Appligene-Oncor, USA
pUC Mix Marker, 8 (0.5mg DNA/ml)	MBI Fermentas Inc., NY, USA
Sodium Acetate	Carlo Erba, Milano, Italy
Sodium Dodecyl Sulfate(SDS)	Sigma, St. Louis, MO, USA
TrisHCl	Sigma, St. Louis, MO, USA
Xylene Cyanol	Sigma, St. Louis, MO, USA

2.1.4. PCR Materials

All PCR materials were obtained from MBI Fermentas Inc., NY, USA.

- Taq DNA Polymerase (5u/µl)
- 10x PCR Buffer
- 25 mM MgCl2
- 10 mM dNTP mix

2.1.5. Restriction Endonucleases

Both restriction endonucleases were obtained from MBI Fermentas Inc., NY, USA.

- *Bpu* 1102I (EspI), (10u/µl)

  Buffer Y+/Tango (10x), (1x buffer composition, 33mM Tris-acetate pH 7.9 at 37 ˚C, 10mM magnesium acetate, 66mM potassium acetate, 0.1mg/ml BSA)

- *Bsh* 1236I (*Fnu* DII), (10u/µl)

  Buffer R^+(10mM Tris-HCl pH 8.5, 10mM magnesium chloride, 100 mM potassium chloride, 0.1 mg/ml BSA)

2.1.6. Purification Kit

NucleoSpin Extract 2 in 1 for direct purification of PCR products or DNA fragments from agarose gel was obtained from Macherey-Nagel.

2.1.7. Sequencing Kit

DYEnamic ET Terminator Cycle Sequencing Kit was used for sequencing reactions (Amersham Biosciences). The sequences were run on an ABI prism 377 automated sequencing machine.

2.1.8. Solutions

- Agarose Gel Loading Buffer (6x)

  15 % ficoll

  0.05 % bromophenol blue

  0.05 % xylene cyanol

- DNA Extraction Buffer

  10 mM Tris HCl, pH 8.0

  10 mM EDTA, pH 8.0

  0.5 % SDS

- Ethidium Bromide

  10 mg/ml in water (stock solution), 30 ng/ml (working solution)

- Proteinase K; 20 mg/ml (stock), 20 μg/ml (working solution)

- SSC (20X), pH 7.0

  3 M NaCl

  0.3 M Trisodium citrate

- Tris Acetic Acid- EDTA Buffer (TAE) (50X)

  (1 liter), 242 gr Tris base

  57.1 ml glacial acetic acid

  37.2 gr EDTA

- TE Buffer

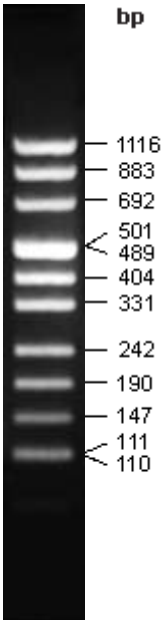
  10 mM Tris-HCl pH 8.0

  1mM EDTA

2.1.9. Standard DNA Size Marker

pUC Mix Marker, 8, (0.5 mg DNA/ml), (MBI, Fermentas)

1.7% Agarose, 0.5µg/lane, 8cm length gel, 1X TBE, 12V/cm



2.2. METHODS

2.2.1. DNA Isolation From Human Blood

Peripheral blood was collected into EDTA containing tubes and stored at 4˚C. Blood samples can be stored at 4˚C for a maximum of five days before being aliquoted and frozen. Before DNA isolation, the blood samples were frozen at –40˚C in 800 µl aliquots. The samples were thawed as required and 800 µl of 1x SSC was added and mixed by vortexing. Centrifugation at 13,000 rpm was carried out for one minute. The supernatant was removed and discarded into disinfectant. It is important not to disturb the cell pellet during this procedure. 1.4 ml of 1x SSC was added over the pellet, vortexed briefly, and centrifuged for one minute at 13,000 rpm. All the supernatant was removed carefully. The washing step was repeated until the pellet became sufficiently clear. 800 µl of DNA extraction buffer and 20 µl of proteinase K (20 mg/ml) were added. After brief vortexing the samples were incubated at 56˚C overnight. Once the cell pellet was dissolved completely, 400 µl of phenol/ chloroform/ isoamyl alcohol (25:24:1) was added and the samples were vortexed for 60 seconds. This step must be carried out under a fume hood. Centrifugation for five minutes was carried out at 13,000 rpm. The upper DNA-containing aqueous layer (~ 700 µl) was collected into a new tube. The extraction step can be repeated by adding 350 µl of phenol/ chloroform/ isoamyl alcohol if the DNA supernatant is sticky and/or if the interface is not clear enough. Then 35 µl of NaOAc (3M, pH 5.2) and 700 µl of cold absolute ethanol (EtOH) were added. The tubes were mixed by inversion and placed at –20˚C for 30 minutes or more (maximum overnight). After centrifugation for five minutes at 13,000 rpm, EtOH was removed with a micropipette and the tubes were left open for air-dry of ethanol. The DNA pellet was dissolved in 200 µl of TE (pH 8.0) or in sterile ddH_20 by incubating at 56˚C for a minimum of 1 hour. Incubation can be performed overnight at 56˚C if it is necessary to dissolve the pellet. The DNA samples were stored at 4˚C or at -20˚C for a long term.

2.2.2. Polymerase Chain Reaction (PCR)

The polymerase chain reaction is a DNA amplification method. The DNA template is amplified enzymatically in the presence of sequence-specific single-stranded DNA oligonucleotide primers (forward and reverse), four deoxyribonucleoside triphosphates (dATP, dGTP, dCTP, dTTP), and DNA Taq polymerase.

All PCR amplifications were carried out on Gene Amp PCR System Perkin Elmer (9600) or Techne (Techgene) PCR machines.

2.2.3. Restriction Endonuclease Digestion

Restriction endonuclease digestions were performed in order to analyze the S31R polymorphism in *p21* and the R72P polymorphism in *p53* genes. The enzyme digestion reactions were carried out in 25 µl reaction volume containing 20 µl of PCR product (diluted if necessary), 10 x buffer Y^+/Tango for *p21* digestion and R^+ buffer for *p53* digestion, 2.5 units of *Bpu* 1102I for *p21* digestion and *Bsh* 1236I for *p53* digestion. The samples were incubated at 37˚C for four hours.

2.2.4. Agarose Gel Electrophoresis

Agarose gel electrophoresis was performed in order to analyze PCR products and restriction fragments. pUC Mix8 was used as the DNA size marker.

PCR products were run on 2% (w/v) agarose gel. 1x TAE buffer was used to prepare the gel and ethidium bromide solution was added (stock: 10mg/ml, final: 30 ng/ml). 5 µl of PCR product was mixed with 1µl 6x loading buffer and the mixture was loaded into the wells of the gel. The products were run at 100 V for 20 minutes. The gel was then visualized under the UV transilluminator and photographs were taken.

The restriction fragments were run on 3% (w/v) Agarose: Gamma Micropor (1:1) gel. 20 µl of digested product and 4 µl of 6x loading buffer containing mix was prepared and then loaded onto the gel. Running procedure was carried out at 80v for 45- 60 minutes. The gel was visualized under UV light and photographs were taken.

2.2.5. Genotyping of Subjects

The *p21* codon 31 and *p53* codon 72 polymorphisms were both analyzed by PCR-based restriction endonuclease enzyme digestion. Genotyping was performed by visualizing the restriction fragments under the UV light.

2.2.5.1. *p21* Codon 31 Genotyping

The *p21* codon 31 polymorphism results from a single base change in codon 31 from AGC to AGA and an amino acid change from serine to arginine. *Bpu*1102I has a restriction recognition site (5`…GC↓TNAGC…3`) at codon 31 when it is AGC (serine) and the change to AGA (arginine) diminishes this site leading to uncut fragments.

Most of the samples were subjected to PCR to a total volume of 25 µl, containing genomic DNA, 10 *p*mol of each primer, 1x PCR buffer (- $MgCl_2$), 1.5 mM $MgCl_2$, 200 µM dNTP mix, and 1U Taq DNA polymerase. For the samples which did not yield enough product with the above concentrations some alternative combinations were used. Increasing the primer concentrations to 12.5 *p*mol and/or $MgCl_2$ concentration to 2 or 4 mM gave better results for these samples. The cycling conditions were as follows: one cycle (denaturing) at 94˚C for 5 min, 35 cycles at 94˚C (denaturing) for 30s, 60˚C (annealing) for 40 s, 72˚C (extension) for 40s, and a final extension cycle at 72˚C for 7 min. The expected amplification product was 272 bp. After digestion with 2.5 U *Bpu*1102I for 3.5 hrs at 37˚C, the cut product gave 89 and 183 bp bands on gel indicating the serine allele, whereas the uncut product gave only a 272 bp band indicating the arginine allele. The products with heterozygous nature gave three bands: 89, 183, and 272 bp. The figure 2.1 shows the schematic representation of the expected digestion patterns.

2.2.5.2. *p53* Codon 72 Genotyping

The *p53* codon 72 polymorphism results from a single base change in codon 72 from CGC to CCC and an amino acid change from arginine to proline. *Bsh*1236I has a restriction recognition site (5`…CG↓CG…3`) at codon 72 when it is CGC (arginine) and the change to CCC (proline) diminishes this site leading to uncut fragments.

Most of the samples were subjected to PCR to a total volume of 25 µl, containing genomic DNA, 10 *p*mol of each primer, 1x PCR buffer (- $MgCl_2$), 1.5 mM $MgCl_2$, 200 µM dNTP mix, and 1U Taq DNA polymerase. The cycling conditions were as follows: one cycle (denaturing) at 94˚C for 5 min, 35 cycles at 94˚C (denaturing) for 30s, 60˚C (annealing) for 30s, 72˚C (extension) for 30s, and a final extension cycle at 72˚ C for 7 min. The expected amplification product was 279 bp. After digestion with 2.5 U *Bsh*1236I for 3.5 hrs at 37˚C, the cut product gave 119 and 160 bp bands on gel indicating the arginine allele, whereas the uncut product gave only 279 bp band indicating the proline allele. The products having both alleles gave three bands: 119, 160, and 279 bp. The figure 2.2 shows the schematic representations of the expected digestion patterns.

2.2.5.3. Sequencing

PCR products were used as templates for sequencing reactions after purification with NucleoSpin Extract 2 in 1. The reactions were carried out with forward and reverse primers for the specific DNA sequence, DYEnamic ET Terminator Cycle Sequencing Kit (Amersham Biosciences), and the reactions were run on the ABI PRISM 377 DNA Sequencer machine.

The sequenced samples for both p21 and p53 were used as controls in every digestion reaction.

Figure 2.1. Schematic Representation of Genotyping of *p21* Codon 31 Polymorphism. Pattern 1,2 and 3 show the *Bpu*1102I restriction enzyme digestion profile.

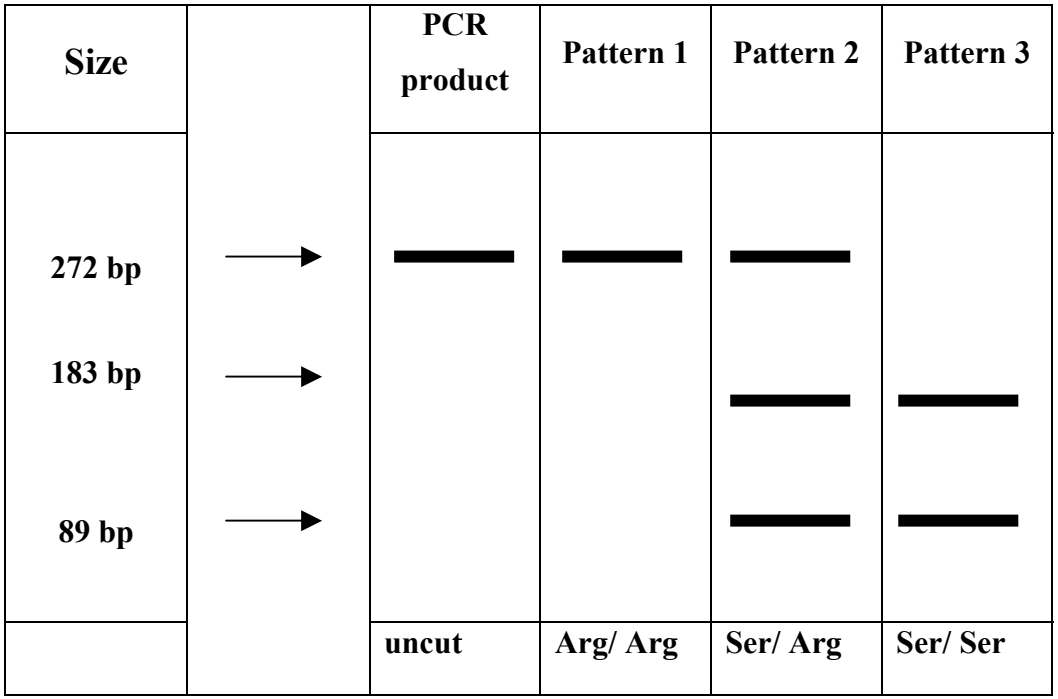
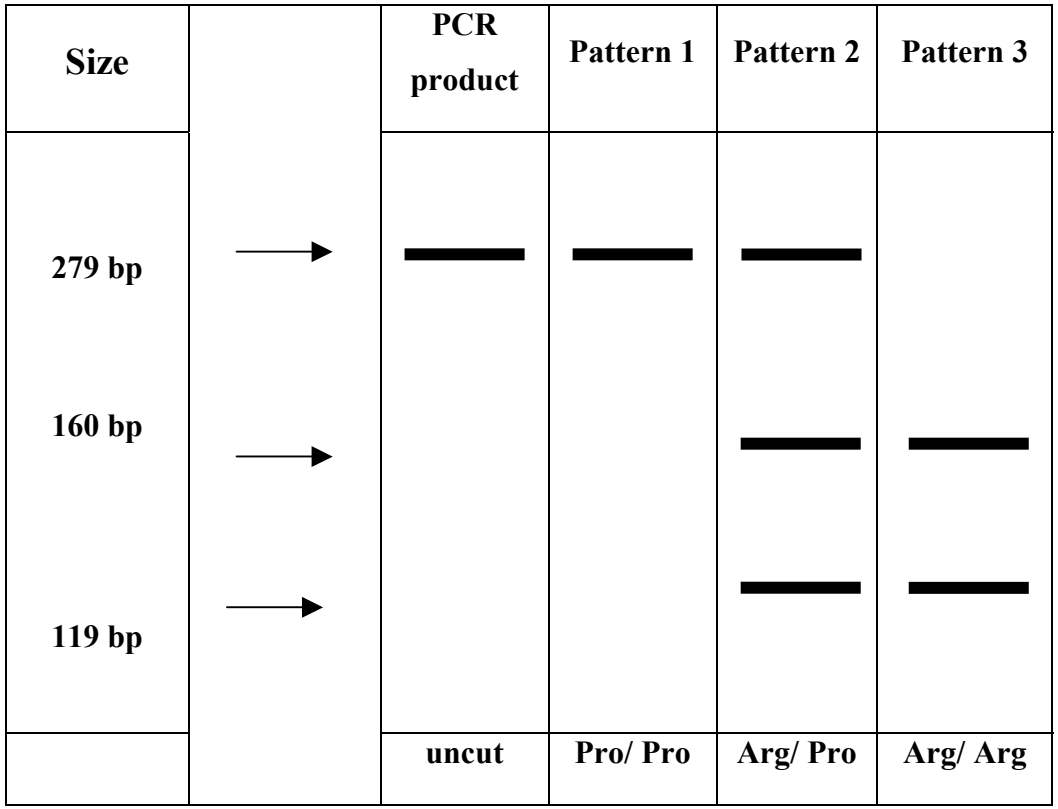


Figure 2.2. Schematic Representation of Genotyping of *p53* Codon 72 Polymorphism. Pattern 1,2 and 3 show the *Bsh*1236 restriction enzyme digestion profile.



2.2.6. Statistical Analyses

The Minitab 13.1 software program was used for statistical analysis of the data. Binary logistic regression analysis was performed for odds ratio and confidence interval calculations. Adjusted odds ratio calculations were carried out with the SPSS software program.

2.2.6.1. Hypothesis Testing

A *hypothesis* is a statement about one or more populations (Daniel 1995). The purpose of hypothesis testing is to aid the clinician, researcher, or administrator in reaching a conclusion concerning a population by examining a sample from that population. Hypothesis testing is carried out on two statistical hypotheses that should be stated separately. The *null hypothesis* (H_0) is the hypothesis to be tested, while the *alternative hypothesis* (H_A) is a statement of what we will believe is true if our sample data causes us to reject the null hypothesis. The null hypothesis is sometimes referred to as a hypothesis of no difference, since it is a statement of agreement with (or no difference from) conditions presumed to be true in the population of interest. Usually the alternative hypothesis and the research hypothesis are the same, and the two terms may be used in an interchangeable manner. For example, in this study, we have tried to investigate whether the cases and the controls can be accepted as statistically different populations based on *p53* codon 72 Arg/Arg genotype frequency. While the alternative hypothesis would be the statement that the cases and the controls are different populations based on *p53* codon 72 Arg/Arg genotype frequency, the null hypothesis would be the opposite statement of the alternative hypothesis. If it is possible to reject H_0 after statistical analyses, the research hypothesis would be accepted.

The decision as to which values go into the rejection region and which ones go into the non-rejection region is made on the basis of the desired level of significance (α). **The level of significance** is the probability of rejecting a true null hypothesis. Because rejecting a true null hypothesis would constitute an

error, a small value of α should be selected. The more frequently selected values of α are 0.01, 0.05, and 0.10. In our study we selected α value of 0.05.

2.2.6.1.1. Types of Errors

The error committed when a true null hypothesis is rejected is called a *type I error*. A *type II error* is the error committed when a false null hypothesis is not rejected. The probability of committing a type II error is designated by β. Table 2.4 summarizes conditions under which type I and type II errors are committed (Daniel 1995).

Table 2.4. Conditions under which Type I and II Errors may be Committed

	CONDITION OF NULL HYPOTHESIS		
POSSIBLE ACTION		*True*	*False*
	Fail to reject H_0	Correct action	Type II error
	Reject H_0	Type I error	Correct action

2.2.6.1.2. P-Values

The *p value* for a hypothesis test is the probability of obtaining, when H_0 is true, a value of the test statistic as extreme as or more extreme (in the appropriate direction) than the one actually computed. The quantity *p* is referred to as the p value for the test (Daniel 1995).

The p value for a test may be defined also as the smallest value of α for which the null hypothesis can be rejected. If the p value is less than or equal to α, the null hypothesis is rejected. If the p value is greater than α, the null hypothesis is not rejected.

2.2.6.1.3. Confidence Interval

100(1-α)% confidence interval (CI) calculation is another method usually used for usual probabilistic and practical interpretations. It is calculated from the standard deviation and sample size values and gives us 100(1-α) percent confidence that the calculated parameter (i.e. odds ratio, mean) is contained in the calculated confidence interval. It is possible to conclude this because in repeated sampling, 100(1-α) percent of the intervals that may be constructed in this manner will include the true parameter. We used 95% CI (if α= 0.05, 100(1-α)% =95%) for odds ratio calculation. For the calculated odds ratio we can say that in repeated data sets, 95% of the cases will have an odds ratio value within the calculated interval but the other 5% will not.

2.2.6.2. The Chi-Square Distribution Test

The chi-square is the most frequently employed statistical technique for the analysis of count or frequency data. The hypothesis testing procedures that use the chi-square distribution are the tests of goodness-of-fit, tests of independence, and tests of homogeneity.

The most frequent use of the chi-square distribution is to test the null hypothesis that the two criteria of classification, when applied to the same set of entities, are independent.

The characteristics of a chi-square test of independence that distinguish it from other chi-square tests are as follows (Daniel 1995):

1. A single sample is selected from a population of interest and the subjects or objects are cross-classified on the basis of the two variables of interest.
2. The rationale for calculating expected cell frequencies is based on the probability law, which states that if two events (the two criteria of classification) are independent, the probability of their joint occurence is equal to the product of their individual probabilities.
3. The hypotheses and conclusions are stated in terms of the independence (or lack of independence) of two variables.

In this study, the chi-square distribution was used to test whether there is a relationship between *p21* codon 31 and/or *p53* codon 72 genotypes and the risk of developing breast cancer among subjects and the control groups.

The classification, according to two criteria, of a set of entities can be shown by a table in which *r* rows represent the various levels of one criterion of classification and the *c* columns represent the various levels of the second criterion. Such a table is called a contingency table. The computed value of X^2 is compared with the tabulated value of X^2 with $k - r$ degrees of freedom, where *k* is equal to the number of groups for which observed and expected frequencies are available, and *r* is the number of restrictions and constraints imposed on the given comparison. X^2 is distributed approximately with (r-1)(c-1) degrees of freedom when the null hypothesis is true. If the computed value of X^2 is equal to or larger than the tabulated value of X^2 for some α, the null hypothesis is rejected at the α level of significance (Daniel 1995).

Sometimes each of two criteria of classification may be broken down into only two categories, or levels. When data are cross-classified in this manner, the result is a contingency table consisting of two rows and two columns. Such a table is commonly referred as a 2 X 2 table. The value of X^2 may be calculated by the following formula:

$$X^2 = \frac{n(ad - bc)^2}{(a + c)(b + d)(a + b)(c + d)}$$

Where *a, b, c*, and *d* are the observed cell frequencies as shown in Table 2.5. When the (r-1)(c-1) rule is applied for finding degrees of freedom to a 2 X 2 table, the result is 1 degree of freedom.

Table 2.5. A 2 X 2 Contingency Table (Daniel 1995)

Second criterion of classification	**First criterion of classification**		
	1	**2**	**Total**
1	*a*	*b*	*a+b*
2	*c*	*d*	*c+d*
Total	*a+c*	*b+d*	*N*

The problems of how to handle small expected frequencies and small total sample sizes may arise in the analysis of 2 X 2 contingency tables. It is suggested that X^2 test should not be used if $n < 20$ or if $20 < n < 40$ and any expected frequency is less than 5. When $n \geq 40$ an expected cell frequency as small as 1 can be tolerated.

2.2.6.3 Relative Risk and Odds Ratio Calculation

An *observational study* is a scientific investigation in which neither the subjects under study nor any of the variables of interest are manipulated in any way.

The term *risk factor* is used to designate a variable that is thought to be related to some outcome variable. The risk factor may be a suspected cause of some specific state of the outcome variable (Daniel W.W. 1995).

In a particular investigation, for example, the outcome variable might be the subjects' status relative to cancer and the risk factor might be their status with respect to cigarette smoking. The model is further simplified if the variables are categorical with only two categories per variable. For the outcome variable the categories might be 'cancer present' and 'cancer absent'. With respect to the risk factor the subjects might be categorized as smokers and nonsmokers.

There are two basic types of observational studies, *prospective* and *retrospective* (Daniel 1995).

A *prospective study* is an observational study in which two random samples of subjects are selected: One sample consists of subjects possessing the risk factor and the other sample consists of subjects who do not possess the risk factor. The subjects are followed into the future (that is, they are followed prospectively) and a record is kept on the number of subjects in each sample who, at some point in time, are classifiable into each of the categories of the outcome variable. For example, if our case were a prospective study, we would collect random samples of subjects, then perform *p53* codon 72 genotyping experiments, categorize the samples into two groups depending on the genotype, and then follow-up the subjects for 30 years to observe which group will develop breast cancer with greater frequency.

A retrospective study is the reverse of a prospective study and is the way that our study was conducted. The samples are selected from those falling into the categories of the outcome variable (breast cancer patients and subjects with no history of breast cancer). The investigator then looks back (that is, takes a retrospective look) at the subjects and determines which ones have (or had) and which ones do not have (or did not have) the risk factor (e.g. *p53* codon 72 Arg/Arg genotype).

Relative risk is the ratio of the risk of developing a disease among subjects with the risk factor to the risk of developing the disease among subjects without the risk factor. A retrospective study is based on a sample of subjects with the disease (cases) and a separate sample of subjects without the disease (controls or noncases). The distribution of the risk factor among the cases and controls is determined retrospectively. Given the results of a retrospective study involving two samples of subjects, cases and controls, the data may be displayed in a 2 X 2 table such as Table 2.5, in which subjects are dichotomized with respect to the presence and absence of the risk factor.

Odds ratio calculation: The appropriate measure for comparing cases and controls in a retrospective study is the *odds ratio* (Daniel 1995). Using probability terminology odds is defined as follows: The odds for success is the ratio of the probability of success to the probability of failure.

Table 2.6. Subjects of a Retrospective Study Classified According to Status Relative to Risk Factor and Whether They Are Cases or Controls.

Risk Factor	**Sample**		
	Cases	**Controls**	**Total**
Present	*a*	*b*	*a+b*
Absent	*c*	*d*	*c+d*
Total	*a+c*	*b+d*	*n*

The definition of odds is used to define two odds that can be calculated from the data displayed as in Table 2.6.

1. The odds of being a case (having the disease) to being a control (not having the disease) among subjects with the risk factor is *[a/(a+b)]/ [b/(a+b)] = a/b.*
2. The odds of being a case (having the disease) to being a control (not having the disease) among subjects without the risk factor is *[c/(c+d)]/ [d/(c+d)] = c/d.*

The odds ratio is computed from the data of a retrospective study. The symbol *OR* is used to indicate that the measure is computed from sample data and used as an estimate of the population odds ratio, *OR* (Daniel 1995).

The estimate of the population odds ratio is:

$$\mathbf{OR} = \frac{a/b}{c/d} = \frac{ad}{bc}$$

$$\mathbf{95\%\ CI} = e^{\ln(OR) \pm 1.96 \text{ times square root of } (1/a+1/b+1/c+1/d)}$$

where *a, b, c,* and *d* are defined in Table 2.6.

The odds ratio can assume values between zero and ∞. A value of zero indicates no association between the risk factor and the disease status. A value less than 1 indicates reduced odds of the disease among subjects with the risk factor. A value greater than 1 indicates increased odds of having the disease among subjects in whom the risk factor is present.

Before calculating the odds ratios and the 95% confidence intervals with Minitab 13.1 software program, the tables for risk and non-risk groups were prepared. While the variables in the risk-groups (i.e. being Arg/Arg at *p53* codon 72) were coded as '1', non-risk variables were coded as '0'. Similarly, breast cancer subjects were coded as '1', meaning that they possess a response under the risk and the control subjects were coded as '0' for the response variable. Appendix shows the coding of every table in the results part.

2.2.6.4. Allele Frequency Calculation

p= [2(number of homozygotes)+ (number of heterozygotes)]/total number of samples

3. RESULTS

Associations between *p53* codon 72 and p21 codon 31 polymorphisms and breast cancer risk in the Turkish and Greek populations were studied.

3.1. Genotyping

Genotyping of both *p53* codon 72 and p21 codon 31 was performed by polymerase chain reaction (PCR) followed by restriction endonuclease digestion.

3.1.1. Genotyping of *p53* codon 72

The number of subjects genotyped for *p53* codon 72 was 274 breast cancer patients and 221 controls. 132 of the cases and 88 of the controls were Arg/Arg, 126 of the cases and 110 of the controls were Arg/Pro, and 16 of the cases and 23 of the controls were Pro/Pro at codon 72 (Table 3.2).

Polymerase chain reaction for *p53* gives a product of 279 bp. An example of the PCR results of *p53* product is shown in **Figure 3.1**. At codon 72, the CGC encodes the Arg allele and has a *Bsh*1236I restriction enzyme site, while CCC encodes the Pro allele and does not have the restriction enzyme site. Therefore, the *Bsh*1236I endonuclease digestion of 279 bp PCR product gives three different band patterns as expected. A single band of 279 bp is observed when a sample is Pro/Pro, two bands of 160 bp and 119 bp are observed for the Arg/Arg genotype, and three bands of 279, 160, and 119 bp are observed in the case of heterozygous samples. **Figure 3.2** shows an example for the *Bsh*1236I enzyme digestion pattern of 279 bp PCR product of *p53* region.

Some of the samples were sequenced in order to control the results of the digestion reactions. After confirming the digestion results, these samples were used as controls for every digestion reaction. **Figure 3.3 A** shows the agarose gel photograph of the *Bsh*1236I enzyme digestion pattern of *p53* PCR products. Lanes 4, 7 and 9 are samples MCK13, MCK32 and MCK18, respectively. **Figure 3.3 B** shows the sequencing reaction results of the *p53* PCR products of the samples MCK13,

MCK32 and MCK18, which were used as control DNAs for the subsequent restriction enzyme digestion experiments. Primers *p53*-F and *p53*-R were used as sequencing primers. These results show that the sample MCK13 is homozygote for Pro allele, MCK32 is homozygote for Arg allele and MCK18 is heterozygote and carries both Arg and Pro alleles.

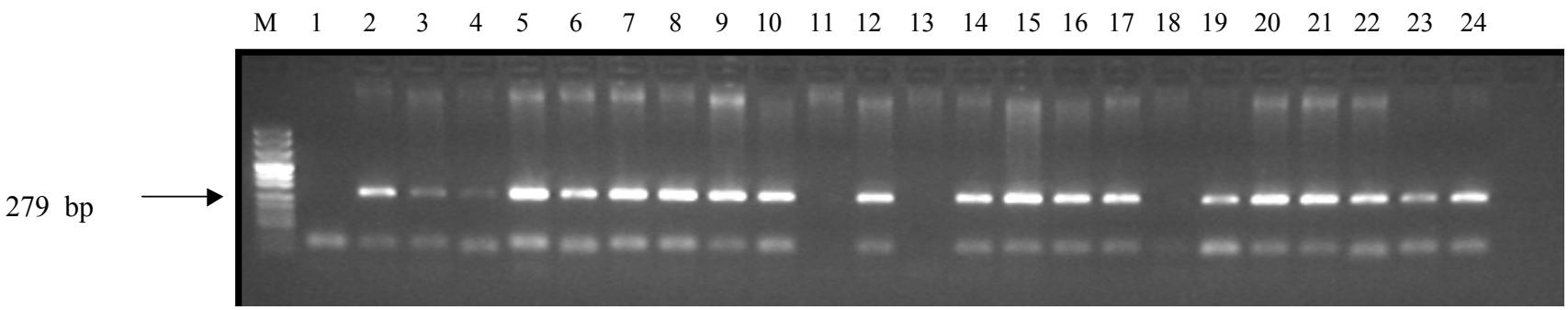


Figure 3.1. Agarose gel photograph of PCR products for *p53* region: primers p53-F and p53-R give 279 bp PCR product. M; DNA size marker (pUC mix 8), lane1; negative control, lanes 2-24 are the samples YB266, YB267, YB270, YB271, YB273-YB289, YB291, and YB292 respectively. The PCR reactions did not result with a product for the samples in lanes 11, 13 and 18.

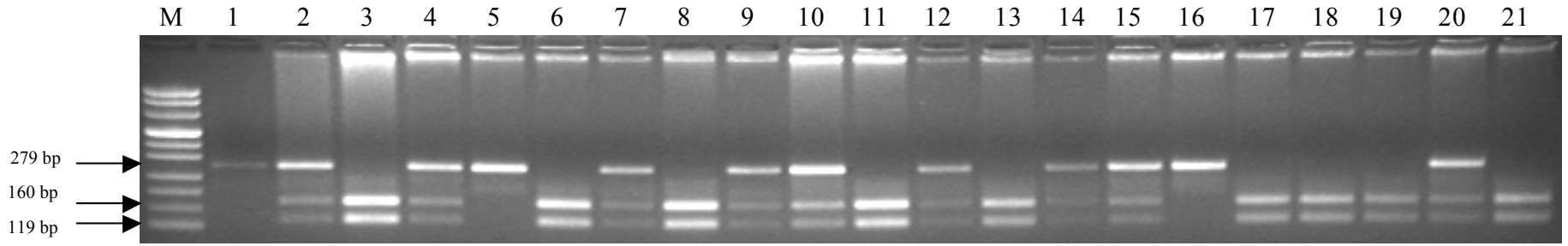


Figure 3.2. Genotyping of the *p53* gene at codon 72. M; DNA size marker (pUC mix 8). 279 bp PCR products were digested with *Bsh*1236I. A single band of 279 bp was observed when the sample is Pro/Pro; lanes 1(MCK 13), 5 (MFN4), 16 (MFN 26). Two bands of 160 bp and 119 bp were observed for the Arg/Arg genotype; lanes 3, 6, 8, 11, 13, 17, 18, 19, 21 (MCK 22, MFN5, -13, -17, -22, -27, -28, -29, and -34). Three bands of 279, 160, and 119 bp were observed in the case of heterozygous samples; lanes 2, 4, 7, 9, 10, 12, 14, 15, 20 (MCK19, GK100, MFN12, -14, -16, -18, -23, -25, and -33).

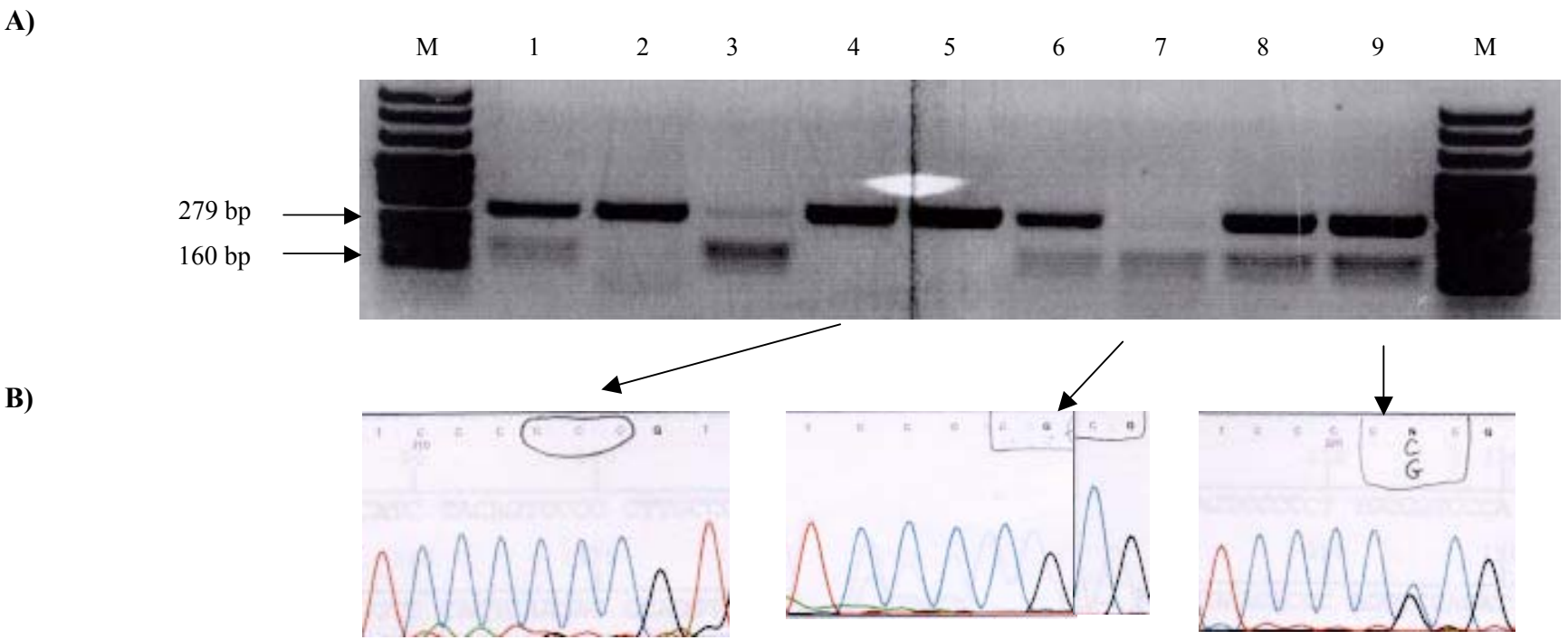


Figure 3.3. Restriction enzyme digestion and sequencing reaction results for p53 samples. **A)** Agarose gel photograph of *Bsh*1236I enzyme digested samples for *p53* codon 72. M; DNA size marker. The sample (MCK13) at lane 4 was homozygous for Pro allele, the sample (MCK32) at lane 7 was homozygous for the Arg allele, and the sample at lane 9 (MCK18) was heterozygous at codon 72. **B)** The PCR products of the samples in lanes 4, 7 and 9 were used for the sequencing reactions. The genotyping results with the digestion patterns were confirmed with the sequencing reactions. Lane 4: CCC (homozygous for Pro allele), lane 7: CGC (homozygous for Arg allele), lane 9: CCC or CGC (heterozygous; contains both Arg and Pro alleles).

3.1.2. Genotyping of *p21* codon 31

The number of subjects genotyped for *p21* codon 31 was 322 breast cancer patients and 246 controls. 259 of the cases and 203 of the controls were Ser/Ser, 59 of the cases and 41 of the controls were Ser/Arg, and 4 of the cases and 2 of the controls were Arg/Arg at codon 31.

Polymerase chain reaction for *p21* gives a product of 272 bp. An example of the PCR results of *p21* product is shown in Figure 3.4. At codon 31, the AGA encodes the Arg allele and does not have a restriction site for *Bpu*1102I enzyme, while AGC encodes the Ser allele and has a restriction site for the enzyme. Therefore, the restriction endonuclease digestion of 272 bp PCR product gives three different band patterns as expected. A single band of 272 bp is observed when a sample is Arg/Arg, two bands of 183 and 89 bp are observed for the Ser/Ser genotype, and three bands of 272, 183, and 89 bp are observed in the case of heterozygous samples. Figure 3.5 shows an example for the *Bpu*1102I enzyme digestion pattern of 272 bp PCR product of *p21* region.

Some of the samples were sequenced in order to control the results of the digestion reactions. After confirming the digestion results, these samples were used as controls for every digestion reaction. Figure 3.6 A shows the agarose gel photograph of the *Bpu*1102I enzyme digestion patterns of *p21* PCR products. Lanes 3, 4 and 6 are samples MCK68, MFN21 and MFN45, respectively.

Figure 3.6 B shows the sequencing reaction results of the *p21* PCR products of the samples in lanes 3, 4 and 6 which were used as control DNAs for the subsequent restriction enzyme digestion experiments. Primers *p21*-F and *p21*-R were used as sequencing primers. These results show that the sample MCK68 is homozygous for Arg allele, MFN21 is homozygous for Ser allele and MFN45 is heterozygous and carries both Arg and Ser alleles.

All the breast cancer patient and control DNA samples were subjected to the restriction enzyme analysis for genotyping. The results were used for statistical analyses to evaluate the p53 codon 72 and *p21* codon 31 genotypes and their association with various risk factors for breast cancer susceptibility.

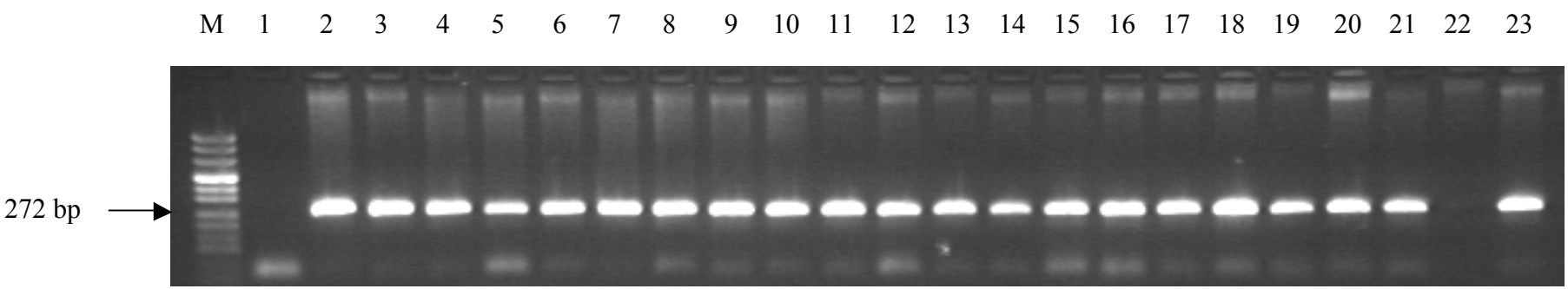


Figure 3.4. Agarose gel photograph of PCR products for *p21* region: primers p21-F and p21-R give 272 bp PCR product. M; DNA size marker (pUC mix 8), lane1; negative control, lanes 2-24 are the samples YB252-262, YB266, YB267, YB270, YB271,YB273-279 respectively. The PCR reactions did not result with a product for the sample in lane 22.

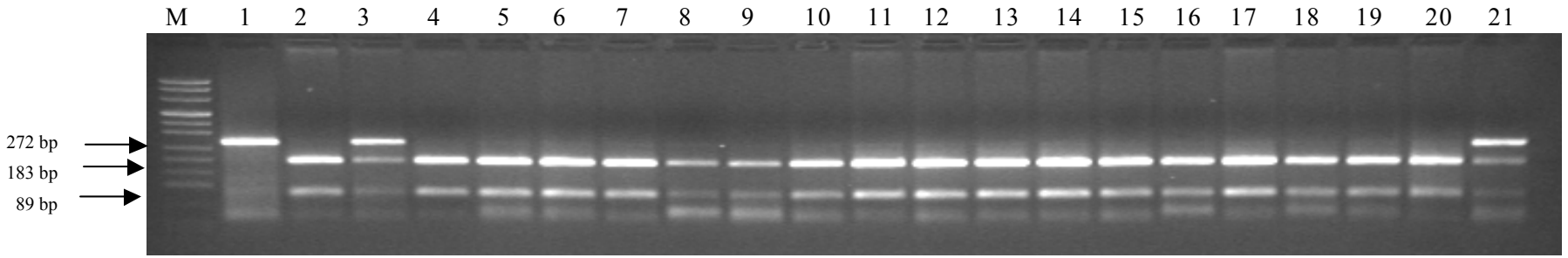


Figure 3.5. Genotyping of the *p21* gene at codon 31. M; DNA size marker (pUC mix 8). 272 bp PCR products were digested with *Bpu*1102I. A single band of 272 bp was observed when the sample is Arg/Arg, lanes 1(MCK 68). Two bands of 183 bp and 89 bp were observed for the Ser/Ser genotype; lanes 2, 4-20 (MFN 21, AB1-5, AB8-19). Three bands of 272, 183, and 89 bp were observed in the case of heterozygous samples; lanes 3 and 21 (MFN45 and AB20).

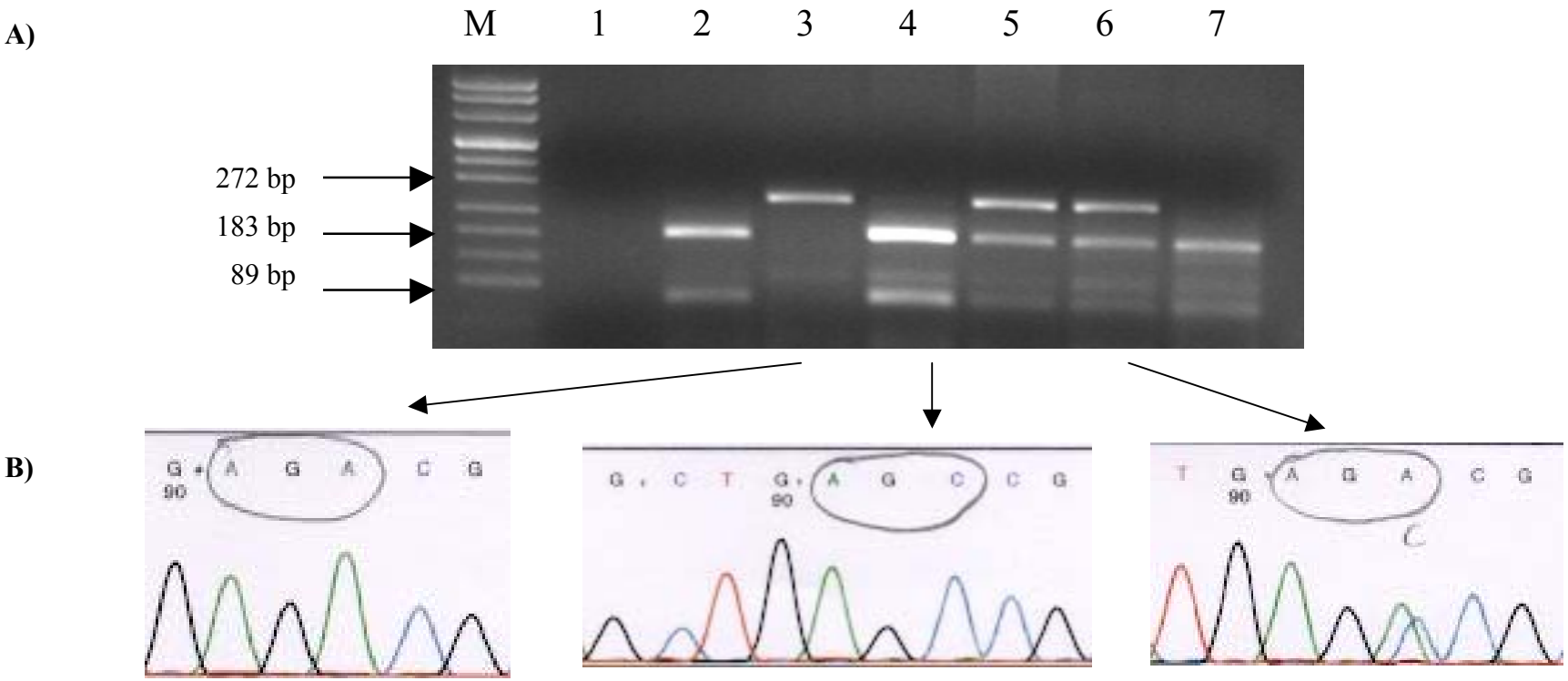


Figure 3.6. Restriction enzyme digestion and sequencing reaction results for p21 samples. **A)** Agarose gel photograph of *Bpu*1102I enzyme digested samples for *p21* codon 31. M; DNA size marker. The sample (MCK68) at lane 3 was homozygous for the Arg allele, the sample (MFN21) at lane 4 was homozygous for the Ser allele, and the sample at lane 6 (MFN45) was heterozygous at codon 31. **B)** The PCR products of the samples in lanes 3, 4 and 6 were used for sequencing reaction. The genotyping results with the digestion patterns were confirmed with the sequencing reaction. Lane 3: AGA (homozygous for Arg allele), lane 4: AGC (homozygous for Ser allele), lane 6: AGA or AGC (heterozygous; contains both Arg and Ser alleles).

3.2. Statistical Analysis

3.2.1. Characteristics of the Subjects in the Turkish Population

The genotyping results of breast cancer patients and controls were subjected to statistical analyses. The statistical analyses was performed by using the Minitab and SPSS software programs. **Table 3.1** summarizes the general characteristics of the subjects in the Turkish population. Results were analyzed whenever information for subjects was available. For the group of cases, the mean age was 49.30 (SD: 12.45, range: 20-80), the mean age at menarche was 13.58 (SD: 1.42, range: 11-20), the mean age at first live birth was 22.51 (SD: 5.14, range: 15-42), the mean number of children was 2.70 (SD: 1.99, range: 0-12), and the mean BMI (kg/m^2) was 27.27 (SD: 4.95, range: 16.02-42.22). For the control group, the mean age was 46.59 (SD: 14.42, range: 15-83), the mean age at menarche was 13.83 (SD: 1.42, range: 10-19), the mean age at first live birth was 20.56 (SD: 3.93, range: 14-37), the mean number of children was 3.14 (SD: 2.18, range: 0-12), and the mean BMI (kg/m^2) was 27.40 (SD: 5.26, range: 14.69-44.06). When the subjects were grouped according to their menopausal status, 43.10% of the cases and 52.61% of the controls were premenopausal, while 56.99% of the cases and 47.39% of the controls were postmenopausal. The higher frequency of postmenopausal cases compared to postmenopausal controls resulted in a higher risk of breast cancer for women who were postmenopausal (OR= 1.47, 95% CI= 1.08-2.00).

We examined whether the age at menarche being greater than 14 is a risk factor for breast cancer. 54.19% of the cases and 43.28% of controls were in this group and this gives us a significant association between age at menarche greater than 14 and breast cancer risk (OR=1.55, 95% CI=1.12-2.14).

Another parameter that we examined for breast cancer risk association was the BMI. Having a BMI greater than the mean BMI of the control group (27.40) was considered to be a risk factor, but no significant results were obtained (OR=0.90, 95% CI=0.65-1.24).

When we analyzed the subjects according to their smoking status, we again could not find any significant association between smoking status and breast cancer risk (OR=0.89, 95% CI=0.62-1.28).

While the frequency of cases with a family history of breast cancer was 14.48%, the frequency for the control group was only 2.46%. OR calculation from these frequencies show that women with a family history of breast cancer have 6.72 times higher risk for the development of breast cancer compared to women without a family history of breast cancer (OR=6.72, 95% CI= 2.97-22) (Table 3.1).

Table 3.1. General Characteristics of the Subjects in the Turkish Study Population

Characteristics	**Cases (n; mean or %)**	**Controls (n; mean or %)**	**OR (95%CI)**
Age	n=370; 49.30	n=288; 46.59	
Family history of breast cancer	n=297; 14.48%	n=285; 2.46%	***6.72 (2.97-15.22)***
Body mass index (BMI)	n=331; 27.267	n=274; 27.399	
BMI≥27.40	n=331; 44.71%	n=274; 47.45%	0.90 (0.65-1.24)
Age at first live birth	n=318; 22.51	n=244; 20.56	
Number of children	n=361; 2.70	n=272; 3.14	
Age at menarche	n=334;13.58	n=268; 13.83	
Age at menarche<14	n=334; 54.19%	n=268; 43.28%	***1.55 (1.12-2.14)***
Premenopausal	n=160; 43.10%	n=151; 52.61%	0.68 (0.50-0.93)
Postmenopausal	n=212; 56.99%	n=136; 47.39%	***1.47 (1.08-2.00)***
Smoking	n=298; 27.85%	n=275; 30.18%	0.89 (0.62-1.28)

3.2.2. Genotype Distributions in the Turkish Population

p53 codon 72 and *p21* codon 31 genotype distributions in the case and control groups were examined and OR calculations were carried out in order to determine the genotypes of the risk groups and the degree of association between the risk groups and breast cancer incidence. ORs were adjusted according to menopausal status, age, smoking status, body-mass-index (BMI), age at menarche, age at first live birth, number of children, and family history of breast cancer. Genotype distributions were also classified according to menopausal status and BMI. The results are summarized in Tables 3.2, 3.3, and 3.4.

3.2.2.1. Distribution of the *p53* Codon 72 Genotype

Table 3.2 shows the distribution of the *p53* codon 72 genotypes in the age-matched controls and breast cancer patients. The *p53* codon 72 Arg/Arg genotype was determined to be a risk factor for breast cancer with a crude odds ratio of 2.16 (95% CI=1.08-4.31) for all subjects. The crude odds ratios for other genotypes were 1.65 (95% CI 0.83-3.27) for the Arg/Pro genotype, 0.46 (95% CI=0.23-0.93) for the Pro/Pro genotype (serves as the reference group), and 0.71 (95% CI=0.50-1.02) for the combined Pro/Pro and Arg/Pro genotypes. The crude odds ratios for premenapousal women were 2.59 (95% CI=0.97-6.93) for the Arg/Arg genotype, 2.30 (95% CI=0.87-6.09) for the Arg/Pro genotype, 0.39 (95% CI=0.14-1.03) for the Pro/Pro genotype, and 0.78 (95% CI=0.46-1.32) for the combination of Pro/Pro and Arg/Pro genotypes. The crude odds ratios for postmenopausal women were 1.37 (95% CI=0.48-3.95) for the Arg/Arg genotype, 1.02 (95% CI=0.36-2.91) for the Arg/Pro genotype, 0.73 (95% CI=0.25-2.09) for the Pro/Pro genotype, and 0.74 (95% CI=0.45-1.22) for the combination of Pro/Pro and Arg/Pro genotypes.

The adjusted odds ratios were 1.62 (95% CI=0.60-4.35) for the Arg/Pro genotype and 1.88 (95% CI=0.69-5.10) for the Arg/Arg genotype when both premenopausal and postmenopausal subjects were considered. We observed an important decrease in OR for the risk group Arg/Arg genotype. When subjects were classified according to their menopausal status, the menopausal status was excluded for the adjusted OR calculation. In the premenopausal group, the adjusted odds ratios were 2.26 (95%

CI=0.53-9.61) for the Arg/Pro genotype and 2.60 (95% CI=0.59-11.38) for the Arg/Arg genotype. In the postmenopausal group, the adjusted odds ratios were 1.27 (95% CI=0.23-6.87) for the Arg/Pro genotype and 1.57 (95% CI=0.29-8.62) for the Arg/Arg genotype.

3.2.2.2. Distribution of *p53* Genotypes According to BMI

Table 3.3 shows the analysis of the *p53* codon 72 genotypes stratified according to BMI based on the median value for controls. Two main groups of analysis are the subjects with BMI value below 27.40 and the subjects with BMI value equal to or above 27.40. Among the subjects with high BMI (≥27.40), the Arg/Arg genotype had a significantly increased risk for breast cancer (OR=3.86, 95% CI=1.12-13.26). The odds ratios for pre- or postmenopausal status had not shown significant increase for Arg/Arg or the combination of Arg/Arg and Arg/Pro genotypes.

Table 3.2. Distribution of the *p53* Codon 72 Genotypes Stratified According to Menopausal Status in the Age-matched Control and Breast Cancer Patients.

Menopausal status	**Genotype**	**Case (%), n=274**	**Control (%), n =221**	**Crude OR 95% CI**	**Adjusted OR[a] 95% CI**
All	R/R	132 (48.18)	88 (39.82)	***2.16 (1.08-4.31)***	1.88 (0.69-5.10)
	R/P	126 (45.99)	110 (49.77)	1.65 (0.83-3.27)	1.62 (0.60-4.35)
	P/P	16 (5.84)	23 (10.41)	0.46 (0.23-0.93)	1.00[b]
	R/P+P/P	142 (51.82)	133 (60.18)	0.71 (0.50-1.02)	
Menopausal status	**Genotype**	**Case (%), n=116**	**Control (%), n=111**	**Crude OR 95% CI**	**Adjusted OR 95[c]% CI**
Pre	R/R	52 (44.83)	43 (38.74)	2.59 (0.97-6.93)	2.60 (0.59-11.38)
	R/P	57 (49.14)	53 (47.75)	2.30 (0.87-6.09)	2.26 (0.53-9.61)
	P/P	7 (6.03)	15 (13.51)	0.39 (0.14-1.03)	1.00[b]
	R/P+P/P	64 (55.17)	68 (61.26)	0.78 (0.46-1.32)	
Menopausal status	**Genotype**	**Case (%), n=153**	**Control (%), n=102**	**Crude OR 95% CI**	**Adjusted OR 95[c]% CI**
Post	R/R	76 (49.67)	43 (42.16)	1.37 (0.48-3.95)	1.57 (0.29-8.62)
	R/P	68 (44.44)	52 (50.98)	1.02 (0.36-2.91)	1.27 (0.23-6.87)
	P/P	9 (5.88)	7 (6.86)	0.73 (0.25-2.09)	1.00[b]
	R/P+P/P	77 (50.33)	59 (57.84)	0.74 (0.45-1.22)	

a) Adjustment is according to menopausal status, age, age at menarche, age at full-term pregnancy, number of children, family history of breast cancer, and smoking.

b) Reference group

c) Adjustment is according to age, age at menarche, age at full-term pregnancy, number of children, family history of breast cancer, and smoking.

Table 3.3. Distribution of the *p53* Codon 72 Genotypes Stratified According to BMI in Cases and Controls.

Low BMI (<27.40)					**High BMI (≥27.40)**		
Menopausal status	**Genotype**	**Cases; n (%) n=131**	**Controls; n (%) n=108**	**OR (95% CI)**	**Cases; n (%) n=107**	**Controls; n (%) n=97**	**OR (95% CI)**
All	P/P	10 (7.63)	12 (11.11)		4 (3.74)	10 (10.31)	
	R/R	57 (43.51)	47 (43.52)	1.46 (0.58-3.67)	54 (50.47)	35 (36.08)	***3.86 (1.12-13.26)***
	R/P+R/R	121 (92.37)	96 (88.89)	1.51 (0.63-3.65)	103 (96.26)	87 (89.69)	2.96 (0.90-9.77)
Menopausal status	**Genotype**	**Cases; n (%) n=65**	**Controls; n (%) n=70**	**OR (95% CI)**	**Cases; n (%) n=37**	**Controls; n (%) n=38**	**OR (95% CI)**
Pre	P/P	5 (7.69)	9 (12.86)		2 (5.41)	6 (15.79)	
	R/R	29 (44.62)	30 (42.86)	1.74 (0.52-5.81)	15 (40.54)	12 (31.58)	3.75 (0.64-22.04)
	R/P+R/R	60 (92.31)	61 (87.14)	1.77 (0.56-5.59)	35 (94.59)	32 (84.21)	3.28 (0.62-17.44)
Menopausal status	**Genotype**	**Cases; n (%) n=66**	**Controls; n (%) n=38**	**OR (95% CI)**	**Cases; n (%) n=70**	**Controls; n (%) n=59**	**OR (95% CI)**
Post	P/P	5 (7.58)	3 (7.89)		2 (2.86)	4 (6.78)	
	R/R	28 (42.42)	17 (44.74)	0.99 (0.21-4.67)	39 (55.71)	23 (38.98)	3.39 (0.58-19.99)
	R/P+R/R	61 (92.42)	35 (92.11)	1.05 (0.24-4.64)	68 (97.14)	55 (93.22)	2.47 (0.44-14.01)

3.2.2.3. Distribution of the *p21* Codon 31 Genotype

Table 3.4 shows the distribution of the *p21* codon 31 genotypes in the age-matched controls and breast cancer patients. Because the frequency of the *p21* codon 31 Arg/Arg genotype was very low (<5) the combination of Arg/Arg and Ser/Arg genotypes was used and determined to have an increased risk for breast cancer with a crude odds ratio of 1.15 (95% CI=0.75-1.76) for all subjects. The crude odds ratios for the whole group were 0.64 (95% CI=0.12-3.52) for the Ser/Ser genotype (serves as the reference group), 1.13 (95% CI=0.73-1.75) for the Ser/Arg genotype, and 1.57 (95% CI=0.28-8.64) for the Arg/Arg genotype. The crude odds ratios for premenapousal women were 0.55 (95% CI=0.05-6.17) for the Ser/Ser genotype, 0.95 (95% CI=0.50-1.81) for the Ser/Arg genotype, and 0.99 (95% CI=0.53-1.85) for the combination of Arg/Arg and Ser/Arg genotypes. The crude odds ratios for postmenopausal women were 0.76 (95% CI=0.07-8.47) for the Ser/Ser genotype, 1.22 (95% CI=0.66-2.25) for the Ser/Arg genotype, and 1.22 (95% CI=0.67-2.23) for the combination of Arg/Arg and Ser/Arg genotypes.

The adjusted odds ratios for the whole group were 3.70 (95% CI=0.33-41.66) for the Arg/Arg genotype, 1.03 (95% CI=0.58-1.82) for the Ser/Arg genotype and 1.10 (95% CI=0.63-1.92) for the combination of Arg/Arg and Ser/Arg genotypes. In the premenopausal or postmenopausal groups, the odds ratios for the Arg/Arg genotype were not meaningful, and for the combination of Arg/Arg and Ser/Arg genotypes the odds ratio was 1.40 (95% CI=0.66-2.96) and 0.62 (95% CI=0.24-1.63) for post and premenopausal groups respectively.

3.2.2.4. Distribution of *p21* Genotypes According to BMI

Table 3.5 summarizes the results of the distribution of the *p21* codon 31 genotypes stratified according to BMI in cases and controls. Although there is increased association between breast cancer risk and combined Arg/Arg and Ser/Arg genotype frequencies among the subjects with BMI value below 27.40, the association is not significantly high (OR=1.32, 95% CI=0.71-2.45). A stronger association between breast cancer risk and the combination of Arg/Arg and Ser/Arg genotypes is present in the postmenopausal group (OR=2.59, 95% CI=0.96-7.01). Among the subjects

with BMI value above 27.40, there was no association between breast cancer risk and the combination of the Ser/Arg and Arg/Arg genotypes.

Table 3.4. Distribution of the *p21* Codon 31 Genotypes Stratified According to Menopausal Status in the Age-matched Control and Breast Cancer Patients.

Menopausal status	**Genotype**	**Cases; n (%), n=322**	**Controls; n (%), n=246**	**Crude OR 95% CI**	**Adjusted OR[a] 95% CI**
All	S/S	259 (80.43)	203 (82.52)	0.64 (0.12-3.52)	1.00[b]
	S/R	59 (18.32)	41 (16.67)	1.13 (0.73-1.75)	1.03 (0.58-1.82)
	R/R	4 (1.24)	2 (0.81)	1.57 (0.28-8.64)	3.70 (0.33-41.66)
	S/R+R/R	63 (19.57)	43 (17.48)	1.15 (0.75-1.76)	1.10 (0.63-1.92)
Menopausal status	**Genotype**	**Cases; n (%), n=133**	**Controls; n (%), n=121**	**Crude OR 95% CI**	**Adjusted OR[c] 95% CI**
Pre	S/S	108 (81. 20)	98 (80.99)	0.55 (0.05-6.17)	1.00[b]
	S/R	23 (17.29)	22 (18.18)	0.95 (0.50-1.81)	0.58 (0.21-1.56)
	R/R	2 (1.50)	1 (0.82)	1.81 (0.16-20.33)	1.28 (0.091-17.96)
	S/R+R/R	25 (18.80)	23 (19.01)	0.99 (0.53-1.85)	0.62 (0.24-1.63)
Menopausal status	**Genotype**	**Cases; n (%), n=184**	**Controls; n (%), n=117**	**Crude OR 95% CI**	**Adjusted OR[c] 95% CI**
Post	S/S	147 (79.89)	97 (82.91)	0.76 (0.07-8.47)	1.00[b]
	S/R	35(19.02)	19 (16.24)	1.22 (0.66-2.25)	1.34 (0.63-2.86)
	R/R	2 (1.09)	1 (0.85)	1.32 (0.12-14.75)	128.298 (0.00-Exp)[d]
	S/R+R/R	37 (20.11)	20 (17.09)	1.22 (0.67-2.23)	1.40 (0.66-2.96)

a) Adjustment is according to menopausal status, age, age at menarche, age at full-term pregnancy, number of children, family history of breast cancer, and smoking.

b) Reference group

c) Adjustment is according to age, age at menarche, age at full-term pregnancy, number of children, family history of breast cancer, and smoking.

d) Exp means exponential. R/R genotype carrying individuals were too low to calculate adjusted OR.

Table 3.5. Distribution of the *p21* Codon 31 Genotype Stratified According to BMI in Cases and Controls

Low BMI (<27.40)					**High BMI (≥27.40)**		
Menopausal status	**Genotype**	**Cases; n (%) n=153**	**Controls; n (%) n=120**	**OR (95% CI)**	**Cases; n (%) n=124**	**Controls (%) n=108**	**OR (95% CI)**
All	S/S	121 (79.08%)	100 (83.33%)		104 (83.87%)	87 (80.56%)	
	S/R+R/R	32 (20.92%)	20 (16.67%)	1.32 (0.71-2.45)	20 (16.13%)	21 (19.44%)	0.80 (0.41-1.57)
Menopausal status	**Genotype**	**Cases; n (%) n=78**	**Controls; n (%) n=74**	**OR (95% CI)**	**Cases; n (%) n=39**	**Controls (%) n=43**	**OR (95% CI)**
Pre	S/S	67 (85.90%)	60 (81.08%)		31 (79.49%)	34 (79.07%)	
	S/R+R/R	11 (14.10%)	14 (18.92)	0.70 (0.30-1.67)	8 (20.51%)	9 (20.93%)	0.97 (0.33-2.84)
Menopausal status	**Genotype**	**Cases; n (%) n=75**	**Controls; n (%) n=46**	**OR (95% CI)**	**Cases; n (%) n=85**	**Controls (%) n=65**	**OR (95% CI)**
Post	S/S	54 (72.00%)	40 (86.96%)		73 (85.88%)	53 (81.54%)	
	S/R+R/R	21 (28.00%)	6 (13.04%)	2.59 (0.96-7.01)	12 (14.12%)	12 (18.46%)	0.73 (0.30-1.74)

3.2.2.5. Combined Analysis of *p21* and *p53* for Breast Cancer Risk

Combined analysis of the *p53* codon 72 Arg/Arg genotype and the *p21* codon 31 Arg/Arg or Ser/Arg genotypes was carried out (Table 3.6). The presence of the *p53* codon 72 Pro/Pro genotype together with the *p21* codon 31 Ser/Ser genotype was designated as the reference group (OR=1.00). The combination of the relative risks of both genotypes showed a significantly increased risk for breast cancer (OR=2.66, 95% CI=1.06-6.66).

Table 3.6. Combination of the *p21* Codon 31 Genotype with the *p53* Codon 72 Genotype for Breast Cancer Risk

Genotype at risk	***P21***	***P53***	**Cases; n (%) n=267**	**Controls; n (%) n=214**	**Crude OR**
None[a]	S/S	P/P	13 (4.87%)	21 (9.81%)	1.00
One	S/S	R/R	103 (38.57%)	68 (31.78%)	***2.45 (1.15-5.21)***
	R/R+S/R	P/P	3 (1.12%)	2 (0.93%)	2.42 (0.36-16.50)
Two	R/R+S/R	R/R	28 (10.49%)	17 (7.94%)	***2.66 (1.06-6.66)***

a) None group is a reference group

3.2.3. Characteristics of the Subjects in the Greek Study Population

The general characteristics of the Greek population are shown in **Table 3.7**. The only information available for Greek population is age and menopausal status. The mean age is 49.66 for the cases and 50.30 for the controls. 56% of the cases are premenopausal, while 44% are postmenopausal. For the control group the frequencies for pre- and postmenopausal subjects are 46.8% and 53.2%, respectively. Compared to the Turkish population, the premenopausal status of the Greek population OR is higher (OR=1.45, 95% CI=0.95-2.22) than the postmenopausal status value (OR=0.69, 95% CI=0.45-1.05), but the results are not statistically significant.

3.2.4. Genotype Distributions in the Greek Population

The distribution of the *p53* codon 72 and *p21* codon 31 genotypes in the Greek population is summarized in Tables 3.8 and 3.9, respectively.

3.2.4.1. Distribution of the *p53* Codon 72 Genotype

The odds ratios for the *p53* codon 72 genotypes are 7.93 (95% CI=0.95-65.98) for the Arg/Arg genotype, 6.50 (95% CI=0.77-54.64) for the Arg/Pro genotype, and 0.13 (95% CI=0.02-1.05) for the Pro/Pro genotype (Table 3.8). Because the Pro/Pro frequency is less than 5%, the Pro/Pro and Arg/Pro genotypes were combined and the odds ratio is 0.74 (95% CI=0.46-1.20). From the OR calculations it is obvious that the genotype under the risk for breast cancer is Arg/Arg but the results are not statistically significant as in the Turkish population. When the subjects were grouped according to their menopausal status, the results were not statistically meaningful. For the premenopausal group, none of the odds ratios were significant. For the postmenopausal group, even the calculation of odds ratios was not possible because there was no postmenopausal case with Pro/Pro genotype. The odds ratio was calculated only for the combination of Arg/Pro and Pro/Pro genotypes was calculated but the result was not significant (OR= 0.57; 95% CI=0.28-1.17).

3.2.4.2. Distribution of the *p21* Codon 31 Genotype

The results for the *p21* codon 31 genotype distributions in the Greek population are different from the Turkish population. Although the results are not significant, the Ser/Ser genotype seems to be a risk factor for breast cancer (OR=1.12, 95% CI=0.60-2.09). Because there was no case with the Arg/Arg genotype in the population, we needed to combine the Arg/Arg genotype with the Ser/Arg genotype while calculating the odds ratio for the Ser/Ser genotype. In the case of Ser/Arg, the combination of Arg/Arg and Ser/Ser was taken as the non-risk group in order not to contradict the logic of the OR calculation of Ser/Ser genotype. The combination of the Arg/Arg and Ser/Arg genotypes had an odds ratio of 0.89 (95% CI=0.48-1.67) (**Table 3.9**).

Table 3.7. Characteristics of Subjects from the Greek Population

Characteristics	Cases (n; mean or %)	Controls (n; mean or %)	OR 95% CI
Age	n=92; 49.66	n=176; 50.301	
Premenopausal	n=92; 56%	n=87; 46.8%	1.45 (0.95-2.22)
Postmenopausal	n=72; 44%	n=99; 53.2%	0.69 (0.45-1.05)

Table 3.8. Distribution of the *p53* Codon 72 Genotype Stratified According to Menopausal Status in the Age-matched Control and Breast Cancer Patients in the Greek Population.

Menopausal status	**Genotype**	**Cases; n (%), n=138**	**Controls; n (%), n=138**	**Crude OR 95% CI**
All	R/R	85 (61.59)	75 (54.35)	7.93 (0.95-65.98)
	R/P	52 (37.68)	56 (40.58)	6.50 (0.77-54.64)
	P/P	1 (0.72)	7 (5.07)	0.13 (0.02-1.05)
	R/P+P/P	53 (38.41)	63 (45.65)	0.74 (0.46-1.20)
Menopausal status	**Genotype**	**Cases; n (%), n=74**	**Controls; n (%), n=63**	**Crude OR 95% CI**
Pre	R/R	44 (59.46)	37 (58.73)	3.57 (0.36-35.76)
	R/P	29 (39.19)	23 (36.51)	3.78 (0.37-38.82)
	P/P	1 (1.35)	3 (4.76)	0.28 (0.03-2.81)
	R/P+P/P	30 (40.54)	26 (41.27)	0.97 (0.49-1.92)
Menopausal status	**Genotype**	**Cases; n (%), n=72**	**Controls; n (%), n=74**	**Crude OR 95% CI**
Post	R/R	35 (48.61)	37 (50.00)	*
	R/P	37 (51.39)	33 (44.59)	*
	P/P	0 (0.00)	4 (5.41)	*
	R/P+P/P	37 (51.39)	37 (50.00)	0.57 (0.28-1.17)

* Because the frequency of the P/P genotype for cases was 0.00%, we could not calculate the odds ratios for R/R, R/P, and P/P.

Table 3.9. Distribution of the *p21* Codon 31 Genotypes Stratified According to Menopausal status in the Age-matched Control and Breast Cancer Patients in the Greek Population.

Menopausal status	**Genotype**	**Cases; n (%), n=156**	**Controls; n (%), n=136**	**Crude OR 95% CI**
All	S/S	132 (84.62)	113 (83.09)	1.12 (0.60-2.09)[a]
	S/R	24 (15.38)	20 (14.71)	1.03 (0.54-1.96)[b]
	R/R	0 (0.00)	3 (2.21)	*
	S/R+R/R	24 (15.38)	23 (16.91)	0.89 (0.48-1.67)
Menopausal status	**Genotype**	**Cases; n (%), n=81**	**Controls; n (%), n=57**	**Crude OR 95% CI**
Pre	S/S	66 (81.48)	47 (82.46)	0.94 (0.39-2.26)[a]
	S/R	15 (18.52)	10 (17.54)	1.07 (0.44-2.58)[b]
	R/R	0 (0.00)	0 (0.00)	*
	S/R+R/R	15 (18.52)	10 (17.54)	1.07 (0.44-2.58)
Menopausal status	**Genotype**	**Cases; n (%), n=66**	**Controls; n (%), n=77**	**Crude OR 95% CI**
Post	S/S	58 (87.88)	65 (84.42)	1.34 (0.51-3.50)[a]
	S/R	8 (12.12)	9 (11.69)	1.00 (0.36-2.75)[b]
	R/R	0 (0.00)	3 (3.90)	*
	S/R+R/R	8 (12.12)	12 (15.58)	0.75 (0.29-1.96)

a) Because the frequency of R/R genotype was 0.00%, the combination of R/R and S/R was taken as non-risk group in calculating the OR for S/S genotype

b) Non-risk group was taken as the combination of S/S and R/R, in calculating the OR for S/R genotype

***** Because the frequency of R/R genotype was 0.00% for the cases, we could not calculate the OR for R/R genotype

3.2.5. Genotype Distributions in the Turkish and Greek Populations

The distribution of the *p53* codon 72 and *p21* codon 31 genotypes in the population containing both Turkish and Greek subjects is summarized in Tables 3.10 and 3.11, respectively.

3.2.5.1. Distribution of the *p53* Codon 72 Genotype

The odds ratios for the *p53* codon 72 genotypes are 2.35 (95% CI=1.25-4.41) for the Arg/Arg genotype, 1.89 (95% CI=1.01-3.56) for the Arg/Pro genotype, 0.43 (95% CI= 0.23-0.80), and 0.75 (95% CI=0.56-0.99) for the combination of the Pro/Pro and Arg/Pro genotypes. From the odds ratio calculations it is obvious that Arg/Arg genotype is strongly associated with breast cancer risk in a population of two Eastern Mediterranean countries. The Arg/Pro genotype was also found to be a significant risk factor for breast cancer in the combined population. When the subjects were grouped according to their menopausal status, premenopausal women with the Arg/Arg or Arg/Pro genotypes had a significantly increased risk for breast cancer susceptibility with odds ratio values of 2.70 (95% CI=1.12-6.54) and 2.55 (95% CI=1.05-6.19), respectively. The postmenopausal group did not give statistically significant results.

3.2.5.2. Distribution of the *p21* Codon 31 Genotype

The results for the *p21* codon 31 genotype distributions are inconsistent in the combined population of the Turkish and Greek subjects when compared to the results of the Turkish and Greek populations alone (Table 3.11).

Table 3.10. Distribution of the *p53* Codon 72 Genotypes Stratified According to Menopausal status in the Age-matched Control and Breast Cancer Patients in the Turkish and Greek Populations.

Menopausal status	**Genotype**	**Cases (%), n=412**	**Controls (%), n=359**	**OR (95% CI)**
All	R/R	217 (52.67)	163 (45.40)	***2.35 (1.25-4.41)***
	R/P	178 (43.20)	166 (46.24)	***1.89 (1.01-3.56)***
	P/P	17 (4.13)	30 (8.36)	0.43 (0.23-0.80)
	R/P+P/P	195 (47.33)	196 (54.60)	0.75 (0.56-0.99)
Menopausal status	**Genotype**	**Cases (%), n=190**	**Controls (%), n=174**	**OR (95% CI)**
Pre	R/R	96 (50.53)	80 (45.98)	***2.70 (1.12-6.54)***
	R/P	86 (45.26)	76 (43.68)	***2.55 (1.05-6.19)***
	P/P	8 (4.21)	18 (10.34)	0.37 (0.15-0.90)
	R/P+P/P	94 (49.47)	94 (54.02)	0.83 (0.55-1.26)
Menopausal status	**Genotype**	**Cases (%), n=225**	**Controls (%), n=176**	**OR (95% CI)**
Post	R/R	111 (49.33)	80 (45.45)	1.70 (0.67-4.28)
	R/P	105 (46.67)	85 (48.30)	1.27 (0.50-3.21)
	P/P	9 (4.00)	11 (6.25)	0.59 (0.23-1.49)
	R/P+P/P	114 (50.67)	96 (54.55)	0.73 (0.49-1.09)

Table 3.11. Distribution of the *p21* Codon 31 Genotypes Stratified According to Menopausal status in the Age-matched Control and Breast Cancer Patients in the Turkish and Greek Populations.

Menopausal status	**Genotype**	**Cases; n (%), n=478**	**Controls; n (%), n=382**	**OR (95% CI)**
All	S/S	391 (81.80)	316 (82.72)	1.55 (0.41-5.81)
	S/R	83 (17.36)	61 (15.97)	1.10 (0.77-1.58)
	R/R	4 (0.84)	5 (1.31)	0.65 (0.17-2.43)
	S/R+R/R	87 (18.20)	66 (17.28)	1.07 (0.75-1.52)
Menopausal status	**Genotype**	**Cases; n (%), n=214**	**Controls; n (%), n=178**	**OR (95% CI)**
Pre	S/S	174 (81.31)	145 (81.46)	0.60 (0.05-6.68)
	S/R	38 (17.76)	32 (17.98)	0.99 (0.59-1.66)
	R/R	2 (0.93)	1 (0.56)	1.67 (0.15-18.57)
	S/R+R/R	40 (18.69)	33 (18.54)	1.01 (0.61-1.68)
Menopausal status	**Genotype**	**Cases; n (%), n=250**	**Controls; n (%), n=194**	**OR (95% CI)**
Post	S/S	205 (82.00)	162 (83.51)	2.53 (0.46-13.99)
	S/R	43 (17.20)	28 (14.43)	1.21 (0.72-2.04)
	R/R	2 (0.80)	4 (2.06)	0.40 (0.07-2.18)
	S/R+R/R	45 (18.00)	32 (16.49)	1.11 (0.68-1.83)

3.2.6. P-value Calculation

In this study, the chi-square test was used to calculate p-values and to test whether there is a relationship between *p21* codon 31 and/or p53 codon 72 genotypes and the risk of developing breast cancer among subjects and the control groups. Table 3.12 shows the p-values for the whole group and separately for the Turkish and Greek populations.

Table 3.12. P-values for the *p53* Codon 72 and *p21* Codon 31 in the Turkish and Greek Populations (df=1).

Population	**P-value**	
	p53 codon 72	***p21 codon 31***
Turkish	0.027	0.527
Greek	0.025*	0.723
Whole	0.007	0.725

***** 2 cells with expected counts less than 5,0.

4. DISCUSSION

Both p53 and p21 play important roles in cell cycle regulation. p53 is a transcription factor for p21 and p53 induces cell cycle arrest at the G1 phase by increasing p21 expression upon DNA damage. Increased *p21* concentration in cells mediates cell cycle arrest through cyclin-cdk inhibition. In order to understand human cancers it is critical to understand the potential roles and the interactions of these two key mediators of the cell cycle.

Several factors make it important to analyze the effects of *p53* and *p21* polymorphisms in breast cancer. First, although 20-40% of breast cancers have *p53* mutations, highly penetrant mutations are rare. Second, *p21* mutations are also extremely rare in breast cancer and lastly, polymorphisms in both genes are quite common. It is important to analyze a potential gene-gene association between polymorphisms in the two genes because p21 is the major downstream component of the p53 tumor suppressor pathway. Such associations are reported for *p53* codon 72 and *p21* codon 31 polymorphisms (Su *et al.* 2003). It was found that the genotype combination involving both the *p53* codon 72 Pro allele and the *p21* codon 31 Arg allele is associated with a particularly low expression of *p21*.

For *p53* only the codon 72 polymorphism appears to be significantly associated with the risk of breast cancer. The most common polymorphism in *p21* is the Ser31Arg polymorphism in exon 2. This polymorphism has been studied in various cancers including breast cancer, but its significant role has not been investigated yet because of the contradictory results.

The two polymorphic variants of *p53* at codon 72 (Arg and Pro) are functionally distinct, and these differences may influence cancer risk or treatment (Dumont *et al.* 2003). The *p53* codon 72 polymorphism occurs in the proline-rich domain of p53, which is required for the growth suppression activity of p53 and also plays an important role in p53-mediated apoptosis (Thomas *et al*. 1999). The two polymorphic variations in codon 72 of *p53* encodes an arginine amino acid with a positive-charged basic side chain or a proline residue with a non-polar aliphatic side chain (Langerod *et al.* 2002). Matlashewski *et al.* (1987) concluded that this is a nonconservative amino acid change and results in a structural change in the protein. Furthermore, Pro-72 p53-containing tumor cells appeared to grow more slowly and

were smaller than Arg-72 p53-containing tumor cells. Thomas *et al*. (1999) showed that p53-Pro is a stronger inducer of transcription than p53-Arg, whereas a p53 Arg/Arg genotype induces apoptosis with faster kinetics and suppresses transformation more efficiently than the *p53* Pro/Pro genotype. Moreover, E6 proteins from both high-risk and low-risk HPV types are able to target $p53_{Arg}$ more efficiently than $p53_{Pro}$ for ubiquitin-mediated degradation and the majority of HPV-associated tumors are homozygous for the $p53_{Arg}$ allele, whereas the majority of the comparable normal population was heterozygous.

Further evidence for the importance of codon 72 containing region in *p53* comes from Baptiste *et al*. (2002). It was shown that the proline-rich PXXP domain between residues 60-90 of p53 is required for cooperation with anti-neoplastic agents to promote apoptosis of tumor cells.

p53 codon 72 polymorphism is important for the interaction of certain p53 mutants to form stable complexes with p73. When codon 72 encodes Arg, the ability of mutant p53 to bind p73 is enhanced. Formation of such complexes correlates with a loss of p73 DNA-binding capability, and consequently its ability to serve as a sequence-specific transcriptional activator and an inducer of apoptosis (Marin *et al.* 2000).

It was reported for squamous cell tumors and breast cancer cases that the occurrence of a *p53* mutation was significantly more often found on the Arg72 allele than the Pro72 allele (Marin *et al*. 2000, Langerod *et al.* 2002). Tada *et al*. (2001) reported that *p53* recessive mutants carrying the Arg allele can lead to decreased activation of p53 target genes through inactivation of p73. In addition, the arginine variant confers greater resistance to p73-dependent apoptosis and cytotoxicity than the equivalent proline form (Bergamaschi *et al*. 2003).

Besides studies which try to understand the functional differences of two variants in *p53* codon 72, there have been many studies conducted in order to investigate the association of *p53* codon 72 polymorphism and cancer risk.

The codon *31* of p21 is found in the area (sequences between amino acids 13 and 56), which is almost perfectly conserved between mouse and human. This conservation of the amino acid sequences suggests that this region is important to the function of WAF1 as a CDK inhibitor (Hachiya *et al.* 1999). In addition, serine is an uncharged polar amino acid with a single hydroxy-1 side chain, whereas arginine is a

basic, positively charged amino acid with a seven-membered side chain. These observations raise the possibility that this polymorphism encodes functionally distinct proteins.

Previous studies indicated a significant association of the Arg allele with human malignancies including breast tumor, sarcoma, rhabdomyosarcoma, neuroblastoma, prostate adenocarcinoma, colorectal cancer and lung cancer (Table 1.4). In those studies, the Arg allele was reported to occur more frequently in malignancy groups than in the healthy control group, suggesting that the Ser to Arg substitution results in some functional differences which contribute to tumorigenesis (Konishi *et al*. 2000). However, further investigations are required for clarifying the association of the p21 polymorphism with tumorigenesis.

In this study, we tried to investigate the effects of *p53* codon 72 and *p21* codon 31 polymorphisms for increased susceptibility for breast cancer either independently or together. We performed the study in two different Mediterranean populations, Turkish and Greek. *p53* and *p21* polymorphisms were analyzed in case and control groups in both populations.

***p53* codon 72 polymorphism:**

Allelic frequencies of arginine (R) and proline (P) are in the Hardy-Weinberg equilibrium in the Turkish (R: 0.6, P: 0.4) and Greek (R: 0.7, P: 0.3) populations as well as in the whole group for the control subjects (R: 0.7, P: 0.3).

In the Turkish population, *p53* codon 72 the Arg/Arg genotype was found to be a significant risk factor for breast cancer development (OR=2.16; 95% CI=1.08-4.31). When we analyzed the subjects according to their menopausal status, it was shown that premenopausal cases with the Arg/Arg genotype have more risk for breast cancer (OR=2.59; 95% CI=0.97-6.93) than the postmenopausal cases (OR=1.37; 95% CI=0.48-3.95). It was also shown that *p53* Arg72Arg genotype is significantly associated with breast cancer risk in women with a high BMI (OR=3.86; 95% CI=1.12-13.26), but not in women with a low BMI (OR=1.46; 95% CI=0.58-3.67). When we grouped the subjects according to low or high BMI, their menopausal status did not show any significant effect on the association of breast cancer risk and the *p53* Arg72Arg genotype.

In the Greek population, it is obvious that the genotype under risk for breast cancer is also Arg/Arg (OR=7.93; 95% CI=0.95-65.98) but the results are not statistically significant as in the Turkish population. The confidence interval was too wide to conclude that the result is statistically significant. Such limitation is related to the low number of subjects in the study population. Because of the low number of subjects, analyses according to menopausal status did not produce any statistically meaningful results.

The Turkish and Greek populations are Eastern Mediterranean populations. We combined these populations and analyzed the *p53* codon 72 polymorphism in relation to breast cancer susceptibility. It was found that the Arg/Arg genotype is a very strong risk factor for breast cancer for the combined study population (OR=2.35; 95% CI=1.25-4.41). Even, the heterozygous genotype carrying only one Arg allele was significantly associated with breast cancer susceptibility (OR=1.89; 95% CI=1.01-3.56). When we analyzed the subjects according to their menopausal status, it was shown that premenopausal cases carrying the Arg allele at codon 72 have more risk for breast cancer (OR=2.70; 95% CI=1.12-6.54 for the Arg/Arg genotype and OR=2.55; 95% CI=1.05-6.19 for the Arg/Pro genotype) than the postmenopausal cases.

The Turkish population can be accepted as of Caucasian origin and the results should be considered accordingly. Our observation of the association between *p53* Arg/Arg genotype and breast cancer risk is in parallel with a previous study conducted in the Turkish population (Buyru *et al*., 2003). Although this study was carried on *p53* codon 72 polymorphism and breast cancer susceptibility, the number of subjects of our study is greater. In this study the authors analyzed only the genotype frequency and association of the *p53* codon 72 polymorphism. But the number of breast cancer cases and control group was very low compared to our study. They also did not looked at all the established risk factors for breast cancer.

Our result of increased association between Arg/Arg genotype and breast cancer risk in Greek population is in parallel with the study of Papadakis *et al.* (2000) on breast cancer in Greek population. Our study contradicts the lack of association in breast cancer that have been observed in a population of Caucasian-Slavic origin (Suspitsin *et al*. 2003).

p53 codon 72 polymorphism and cancer risk association have been studied in cancers other than breast (acute myeloid leukemia, bladder, cervical, gastric, laryngeal, lung, squamous cell carcinoma, vulval, and urothelial) and in different populations (Table 1.5). The most studied cancer is cervical cancer. Positive results for Arg/Arg association were observed in Chinese, Indian, and in Israeli Jewish populations (Qie *et a*l. 2002, Li *et al*. 2002, Nagpal *et al*. 2002, Saranth *et al.* 2002, and Arbel-Alon *et al*. 2002). There are also studies that show no association with cervical cancer in the Korean and Indian populations (Kim *et al*. 2001, Pillai *et al*. 2002).

***p21* codon 31 polymorphism:**

Allelic frequencies of serine (S) and arginine (R) are in the Hardy-Weinberg equilibrium in the Turkish (R: 0.9, P: 0.1) and Greek (R: 0.9, P: 0.1) populations as well as in the whole group for the control subjects (R: 0.9, P: 0.1).

For the *p21* codon 31, the Arg/Arg or Ser/Arg genotypes were accepted as risk groups for breast cancer in our Turkish study population (OR=1.15; 95% CI=0.75-1.76). Although the result is not statistically significant, there is a slight increase in association for breast cancer. When we consider the effect of BMI and menopausal status, we see that postmenopausal women with low BMI and Arg/Arg or Ser/Arg genotypes have an increased association for breast cancer (OR=2.59; 95% CI=0.96-7.01), which is different from the *p53* case.

The results for the *p21* codon 31 genotype distributions in the Greek population are different from the Turkish population. Although the results are not significant, the Ser/Ser genotype seems to be a risk factor for breast cancer (OR=1.12, 95% CI=0.60-2.09).

In the study group of both Turkish and Greek subjects, the analysis of *p21* codon 31 genotype distributions did not reveal consistent results. Neither we could determine the risk group for breast cancer nor we obtained significant associations between breast cancer susceptibility and *p21* codon 31 genotypes.

Keshava *et al*. (2002) conducted a study on the association between *p21* codon 31 polymorphism and breast cancer risk in three different populations; Caucasians, African-American, and Latinas. In this study, the Ser allele was found to be the minor allele and the individuals carrying Ser at *p21* codon 31 were analyzed as the risk group for breast cancer risk. Although, an association for Ser variant in African-

Americans (OR=2.32; 95% CI=0.66-5.60) and Latinas (OR=2.22; 95% CI=0.71-6.89) was reported, no association was found in Caucasians. Ser variant association seems parallel to our results for the Greek population (OR=1.12; 95% CI=0.60-2.09). Another point that should be discussed is the confidence interval that they accepted as significant. In our study we accepted the results as statistically significant if the confidence interval does not include 1.00 and this is the accepted method for most of the studies. Lukas *et al.* (1997) also reported no association of *p21* codon 31 polymorphism and breast cancer risk but in American population.

Risk association studies for *p21* codon 31 polymorphism have been studied in different cancers and populations (Table 1.4). Association for the Arg allele was found to be significant for endometrial cancer in the Japanese and American populations (Hachiya *et al*. 1999, Lukas *et al.* 1997) and non-significant in the Taiwanese population (Hsieh *et al.* 2001). Significant increase in Arg allele frequency was also reported for lung cancer in the Swedish population (Sjalander *et al*. 1996). Non-significant increase in Arg allele was reported for head and neck carcinoma and prostate adenocarcinoma in American population (Facher *et al*. 1997). The *p21* codon 31 polymorphism is not correlated to cancer susceptibility in all previous studies (Li *et al*. 1995, Su *et al.* 2003, Shih *et al*. 2000, Sun *et al.* 1995, Lukas *et al*. 1997, Konishi *et al*. 2000).

Combined analysis of *p21* (codon 31) and *p53* (codon 72) polymorphisms:

In our study, the combination of *p53* Arg72Arg and *p21* Arg31Arg or Ser31Arg genotypes showed a statistically significant potential association between the two genes and a risk for breast cancer development (OR=2.66; 95% CI=1.06-6.66). Interestingly, there was statistically significant association for the *p53* high risk genotype Arg72Arg and the *p21* non-risk genotype Ser31Ser (OR=2.45; 95% CI=1.15-5.21).

Previously, two studies examined a potential association between common polymorphisms in *p53* and *p21* in relation to breast cancer (Powell *et al*. 2002, Keshava *et al*. 2002). Although, Powell *et al*. (2002) studied the effects of different polymorphisms of both *p53* and *p21* in breast cancer, the combined effect of the polymorphisms was not analysed. Keshava *et al*. (2002) looked for association between *p21* Ser variant and *p53* $^{1\text{-}2\text{-}1}$ haplotype, but no gene-gene association was found.

P-value calculation:

P-values smaller than 0.05 (α=0.05) were accepted as significant. Consistent with the odds ratio and confidence interval calculations, it was found that there is a relationship between the *p53* codon 72 genotypes and the risk of developing breast cancer among subjects and the control groups in the Turkish and whole populations. P-value for the *p53* codon 72 genotypes in the Greek population was also less than 0.05 (0.025), but there was count values less than 5. This means that even if the p-value is significant for the given data, the results should be interpreted with caution.

Established risk factors:

Family history of breast cancer is the best studied and most significant risk factor for breast cancer development (Vogelstein *et al.* 2002). When stratification according to family history of breast cancer was carried out in our Turkish study population, it was shown that family history of breast cancer is very strongly associated with breast cancer risk (OR=6.72; 95% CI=2.97-15.22).

Early age at menarche, late age at first full term pregnancy, and late age at menopause, which are all associated with increased exposure to endogenous estrogens, increase the risk of developing breast cancer. Generally, cancer is a disease of late age, reflecting accumulation of mutations that eventually lead to tumorigenesis. Therefore, increasing age is also an important parameter for breast cancer risk. The case group had slightly earlier age at menarche than the control group. In order to determine the effect of early age at menarche on breast cancer risk in our population, we selected the group of subjects having age at menarche less than 14 years (Vogelstain *et al.* 2002) as a risk group and carried out regression analysis. Age at menarche less than 14 significantly increased the susceptibility for breast cancer (OR=1.55; 95% CI=1.12-2.14). When we analyzed the Turkish study population for the other established risk factors, the results were as expected. The breast cancer patients were older than the control subjects. In addition, the cases had later age at first live birth and less number of children compared to the controls. In addition, when compared with the control group, a statistically significant number of breast cancer patients were postmenopausal (OR=1.47; 95% CI=1.08-2.00).

In our Turkish study population, we tried to investigate whether smoking status has any implications on breast cancer risk. Consistent with the earlier studies (Braga *et al*. 1996, Garcia-Closas *et al*. 1999) no risk assessment was found related to smoking for breast cancer susceptibility (OR=0.89; 95% CI=0.62-1.28). However, Terry *et al.* (2002) suggested that smoking does not decrease the risk of breast cancer and indeed suggested that there may be an increased breast cancer risk with smoking of long duration, smoking before a first-full term pregnancy, and passive smoking. These findings require confirmation in future studies.

Another parameter that might affect the susceptibility for breast cancer is BMI (Ursin *et al.* 1995). Although being overweight (high BMI) during early adult life has been associated with a lower incidence of premenopausal breast cancer (Wrensch *et al.* 2003), weight gain after age 18 is associated with a significantly increased risk of postmenopausal breast cancer (DeVita *et al*. 2001, Radimer *et al*. 1993). In parallel with the positive association studies, we found that p53 Arg72Arg genotype is significantly associated with breast cancer risk in women with a high BMI (OR=3.86; 95% CI=1.12-13.26).

We performed direct analysis for menopausal status, BMI, age at menarche, family history of breast cancer, and smoking status on breast cancer risk. It was possible to analyze the overall effect of these factors together with the first pregnancy age and the number of children by examining the crude and adjusted odds ratios. Adjusted odds ratio is the odds ratio of only a defined parameter and the effect of the above factors is excluded. For p53 Arg/Arg genotype the crude and adjusted odds ratios together with 95% CI are 2.16 (1.08-4.31) and 1.88 (0.69-5.10), respectively. When adjusted we see a considerable decrease in the OR, which means the effect of variables other than Arg/Arg inheritance may have an important effect on breast cancer development.

To our knowledge, this is the first genetic study to show associations between the genotype frequencies of *p21* codon 31 and *p53* codon 72 polymorphisms and the established breast cancer risk factors in the Turkish population. This is also the first study to show a combined effect of *p21* codon 31 and *p53* codon 72 polymorphisms on breast cancer risk.

5. CONCLUSION AND FUTURE PERSPECTIVES

The findings of our study are as follows:

1. In the Turkish study population, Arg72Arg genotype of *p53* is in strong association with breast cancer development, especially among the subjects with increased BMI.
2. The combined effect of *p21* codon 31 Arg/Arg and Ser/Arg genotypes has a slightly increased susceptibility for breast cancer in the Turkish population.
3. There is a prominent increase in breast cancer risk for the individuals carrying both high-risk allelic variants of *p53* codon 72 and *p21* codon 31 polymorphisms in the Turkish population.
4. In the combined population of the Turkish and Greek subjects, individuals, especially the premenopausal subjects, carrying Arg allele at codon 72 of *p53* have a significantly increased risk for breast cancer susceptibility.
5. Generally recognized breast cancer risk factors such as family history of breast cancer, earlier age at menarche, and postmenopausal state contributed to a higher risk for breast cancer in the Turkish population. Smoking status does not seem to have any effect on breast cancer susceptibility. Although BMI increased the risk of *p53* Arg72Arg genotype for breast cancer susceptibility, it does not have any contribution alone or together with *p21* codon 31 polymorphism. Other established risk factors for breast cancer such as increasing age, late age at first full term pregnancy, less number of children might be also involved in the increased susceptibility for breast cancer.
6. The studies we performed with Greek population did not reveal a significant association with the studied polymorphisms and breast cancer risk.

Further analysis of these polymorphisms in large and diverse populations is necessary to confirm the previous results.

p53 and *p21* mutational status in relation to *p53* and *p21* expression levels and polymorphisms could be evaluated in order to understand better the interaction between *p53* and *p21* genes.

The effects of allelic differences in *p53* codon 72 and *p21* codon 31 on the gene-gene interactions can be investigated to understand the underlying mechanisms.

Because p53 is a critical regulator of apoptosis, effect of *p53* polymorphism on drug response and treatment can be studied.

Possible interactions with other genetic variations such as polymorphisms in steroid hormone metabolism genes and carcinogen metabolism genes and with DNA damage responsive elements (i.e. BRCA genes) can be analyzed and the results of the combined effects of these variations can be evaluated with breast cancer risk.

Better defined groups for the relation of the polymorphisms and breast cancer etiology (e. i. staging, progression, receptor status) may give us valuable results. Furthermore, such kind of studies may help for early diagnosis of breast cancer by determining the risk factors that individuals carry.

6. REFERENCES

Alberts B., Johnson A., Lewis J., Raff M., Roberts K., and Walter P. (2002) Molecular Biology Of The Cell, Garland Science, Taylor & Francis Group, NY.

Arbel-Alon S., Menczer J., Feldman N., Glezerman M., Yeremin L., Friedman E. (2002) Codon 72 polymorphism of p53 in Israeli Jewish cervical cancer patients and healthy women. Int. J. Gynecol. Cancer 12(6), 741-744.

Bahl R., Arora S., Nath N., Mathur M., Shukla K., Ralhan R. (2000) Novel polymorphism in $p21^{waf1/cip1}$ cyclin dependent kinase inhibitor gene: association with human esophageal cancer. Oncogene 19, 323-328.

Baptiste N., Friedlander P., Chen X., Prives C. (2002) The proline-rich domain of p53 is required for cooperation with anti-neoplastic agents to promote apoptosis of tumor cells. Oncogene 21(1), 9-21.

Becker N., Nieters A., Rittgen W. (2003) Single Nucleotide Polymorphism – disease relationships: statistical issues for the performance of association studies. Mutation Research 525, 11-18.

Becker W. M., Reece J. B., and Poenie M. F. (1996) The World Of The Cell, The Benjamin/Cummings Publishing Company, Inc.

Bergamaschi D., Gasco M., Hiller L., Sulli A., Syed N., Trigiante G., Yulug I., Merlano M., Numico G., Comino A., Attard M., Reelfs O., Gusterson B., Bell A. K., Heath V., Tavasso M., Farrel P. J., Smith P., Lu X., Crook T. (2003) p53 polymorphism influences response in cancer chemotherapy via modulation of p73-dependent apoptosis. Cancer Cell 3, 387-402.

Biro E., Kalina I., Salagovic J., Habalova V., Hriv ak M., Valansky L. (2000) p53 single nucleotide polymorphisms and bladder cancer. Neoplasma 47(5), 303-306.

Braga C., Negri E., La Vecchia C., Filiberti R., Franceschi S. (1996) Cigarette smoking and the risk of breast cancer. European Journal of Cancer Prevention 5(3), 159-164.

Brinton L. A., Swanson C. A. (1992) Height and weight at various ages and risk of breast cancer. Annual Epidemiology 2(5), 597-609.

Buyru N., Tigli H., Dalay H. (2003) p53 codon 72 polymorphism in breast cancer. Oncol. Rep. 10(3), 711-714.

Chedid M., Michieli P., Lengel C., Huppi K., Givol D. (1994) A single nucleotide substitution at codon 31 (Ser/Arg) defines a polymorphism in a highly conserved region of the p53 inducible gene WAF1/CIP1. Oncogene 9, 3021-3024.

Chen WC., Lu HF., Chen HY., Hsu CD., Tsai FJ. (2001) Arginine form of p21 gene codon 31 is less prominent in patients with calcium oxalate stone. Urol. Res. 29, 94-97.

Coughlin S. S., Piper M. (1999) Genetic polymorphisms and risk of breast cancer. Cancer Epidemiology, Biomarkers & Prevention 8, 1023-1032.

Daniel W.W. (1995) Biostatistics: A Foundation for Analysis in Health Science, 6th Edition, NY, John Wiley and Sons Inc.

Dumont P., Leu J. IJ., Pietra III A. C. D., George D. L., Murphy M. (2003) The codon 72 polymorphic variants of p53 have markedly different apoptotic potential. Nature Genetics 33, 357-365.

Dunning A. M., Healey C. S., Pharoah P. D. P., Teare M. D., Ponder B. A. J., Easton D. F. (1999) A systematic review of genetic polymorphisms and breast cancer risk. Epidemiology, Biomarkers & Prevention 8, 843-854.

DeVita V. T., Jr., Hellman S., Rosenberg S. A. (2001) Cancer, Principles & Practice of Oncology, 6th Edition, Lippincott Williams & Wilkins.

Facher E. A., Becich M. J., Deka A., Law J. C., (1997) Association Between Human Cancer and Two Polymorphisms Occurring Together in the $p21^{Waf1/Cip1}$ Cyclin-Dependent Kinase Inhibitor Gene. Cancer 79, 2424-2429.

Fan R., Wu MT., Miller D., Wain JC., Kelsey K. T., Wiencke J. K., Christiani D. C. (2000) The p53 codon 72 polymorphism and lung cancer risk. Cancer Epidemiology, Biomarkers & Prevention 9, 1037-1042.

Franceschi S., Favero A., La Vecchia C., Baron A. E., Negri E., Giacosa A., Montella M., Conti E., Amadori D. (1996) Body size indices and breast cancer risk before and after menopause. Int. J. Cancer 67(2), 181-186.

Garcia-Closas M., Kelsey K. T., Hankinson S. E., Spiegelman C. W., Springer K., Willett W. C., Speizer F. E., Hunter D. J. (1999) Glutathione S-Transferase Mu and Theta Polymorphisms and Breast Cancer Susceptibility. Journal of the National Cancer Institute 91(22), 1960-1964.

Gray I. C., Campbell D. A., Spurr K. (2000) Single nucleotide polymorphisms as tools in human genetics. Human Molecular Genetics 9(16), 2403-2408.

Hachiya T., Kuriaki Y., Ueoka Y., Nishida J., Kato K., Wake N., (1999) WAF1 genotype and Endometrial Cancer Susceptibility. Gynecologic Oncology 72, 187-192.

Haimov-Kochman R., Lavy Y., Hochner-Celinkier D. (2002) Review of risk factors for breast cancer-what`s new? Harefuah 141(8), 702-708, 761.

Hanahan D. and Weinberg R. A. (2000) The Hallmarks of Cancer. Cell 100, 57-70.

Hiyama T., Tanaka S., Kitadai Y., Ito M., Sumii M., Yoshihara M., Shimamoto M., Haruma K., Chayama K. (2002) p53 codon 72 polymorphism in gastric cancer susceptibility in patients with *Helicobacter pylori*-associated chronic gastritis. Int. J. Cancer 100, 304-308.

Hsieh YY., Tsai FJ., Chang CC., Chen WC., Tsai CH., Tsai HD., Lin CC. (2001) p21 gene codon 31 Arginine/Serine polymorphism: Non-association with endometriosis. Journal of Clinical Laboratory Analysis 15, 184-187.

Humbey O., Cairey-Remonnay S., Guerrini J. S., Algros M. P., Mougin C., Bittard H., Aubin F. (2003) Detection of the human papillomavirus and analysis of TP53 polymorphism of exon 4 at codon 72 in penile squamous cell carcinoma. European Journal of Cancer 39, 684-690.

Kastan M. B. (editor)(1997) Genetic Instability and Tumorigenesis, Springer-Verlag. Berlin Heidelberg.

Keshava C., Frye B. L., Wolff M. S., McCanlies E. C., Weston A. (2002) Waf-1 (p21) and p53 Polymorphisms in Breast Cancer. Cancer Epidemiology, Biomarkers & Prevention 11, 127-130.

Kim JW., Roh JW., Park NH., Song YS., Kang SB., Lee HP. (2001) Polymorphism of TP53 codon 72 and the risk of cervical cancer among Korean women. Am. J. Obstet. Gynecol. 184(2), 55-58.

Ko L. J., Prives C. (1996) p53: puzzle and paradime. Genes and Development 10, 1054-1072.

Konishi R., Sakatani S., Kiyokane K., Suzuki K., (2000) Polymorphisms of p21 cyclin-dependent kinase inhibitor and malignant skin tumors. Journal of Dermatological Science 24, 177-183.

Kuroda Y., Tsukino H., Nakao H., Imai H., Katoh T. (2003) p53 codon 72 polymorphism and urothelial cancer risk. Cancer Letters 189, 77-83.

Landegren U., Nilsson M., Kwok P. (1998) Reading bits of genetic information: Methods for single-nucleotide polymorphism analysis. Genome Research 8, 769-776.

Langerod A., Bukholm I. R. K., Bregard A., Lonning PE., Andersen T. I., Rognum T. O., Meling G. I., Lothe R. A., Borresen-Dale AL. (2002) the TP53 codon 72 polymorphism may affect the function of TP53 mutations in breast carcinomas but not in colorectal carcinomas. Cancer Epidemiology, Biomarkers & Prevention 11, 1684-1688.

Lewin B. (2000) Genes VII, Oxford University Press Inc., NY.

Levine A. J. (1997) p53, The Cellular Gatekeeper for Growth and Division. Cell 88, 323-331.

Levitt N. C. and Hickson I. D. (2002) Caretaker tumor suppressor genes that defend genome integrity. TRENDS in Molecular Medicine 8-4, 179-186.

Li Y. J., Laurent-Puig P., Salmon R. J., Thomas G., Hamelin R. (1995) Polymorphisms and probable lack of mutation in the WAF1-CIP1 gene in colorectal cancer. Oncogene 10, 599-601.

Li T., Lu ZM., Guo M., Wu GJ., Chen KN., Xing HP., Mei Q., Ke Y. (2002) p53 codon 72 polymorphism (C/G) and the risk of human papillomavirus-associated carcinomas in China. Cancer, 95, 2571-2576.

Lukas J., Groshen S., Saffari B., Niu N., Reles A., Wen WH., Felix J., Jones LA., Hall FL., Press MF. (1997) WAF1/Cip1 gene polymorphism and expression in carcinomas of the breast, ovary, and endometrium. Am J Pathol. 150(1), 167-175.

Marin M. C., Jost C. A., Brooks L. A., Irwin M. S., O'Nions J., Tidy J. A., James N., McGregor J. M., Harwood C. A., Yulug I. G., Vousden K. H., Allday M. J., Gusterson B., Ikawa S., Hinds P. W., Crook T., Kaelin Jr W. G. (2000) A common polymorphism acts as an intragenic modifier of mutant p53 behaviour. Nature Genetics 25, 47-54.

Matlashewski GJ., Tuck S., Pim D., Lamb P., Schneider J., Crawford LV. (1987) Primary structure polymorphism at amino acid residue 72 of human p53. Molecular and Cellular Biology 7(2), 961-963.

Mitrunen K., Jourenkova N., Kataja V., Eskelinen M., Kosma VM., Benhamou S., Vainio H., Uusitupa M., Hirvonen A. (2001) Glutathione S-Transferase M1, M3, P1, and T1 genetic polymorphisms and susceptibility to breast cancer. Cancer Epidemiology, Biomarkers & Prevention 10, 229-236.

Mousses S., Ozcelik H., Lee P. D., Malkin D., Bull S. B., Andrulis I. L. (1995) Two variants of the CIP1/WAF1 gene occur together and are associated with human cancer. Human Molecular Genetics 4(6), 1089-1092.

Nagpal JK., Sahni S., Das BR. (2002) p53 codon 72 polymorphism and susceptibility to development of human papilloma virus-associated cervical cancer in Indian women. Eur. J. Clin. Invest. 32(12), 943-948.

Nakano Y., Naoe T., Kunishima S., Minami S., Miyawaki S., Asou N., Kuriyama K., Saito H., Ohno R. (2000) Poor clinical significance of p53 gene polymorphism in acute myeloid leukemia. Leukemia Research 24(4), 349-352.

Oesterreich S., Fuqua SAW. (1999) Tumor suppressor genes in breast cancer. Endocrine-related Cancer 6, 405-419.

Papadakis E. D., Soulitzis N., Spandidos D. A. (2002) Association of p53 codon 72 polymorphism with advanced lung cancer: the Arg allele is preferentially retained in tumors arising in Arg/Pro germline heterozygotes. British Journal of Cancer 87, 1013-1018.

Papadakis E. N., Dokianakis D. N., Spandidos D. A. (2000) p53 codon 72 polymorphism as a risk factor in the development of breast cancer. Molecular Cell Biology Research Communications 3, 389-392.

Pierce L. M., Sivaraman L., Chang W., Lum A., Donlon T., Seifried A., Wilkens L. R., Lau A., F., Marchand L. L. (2000) Relationships of TP53 codon 72 and HRAS1 polymorphisms with lung cancer risk in an ethnically diverse population. Cancer Epidemiology, Biomarkers & Prevention 9, 1199-1204.

Pillai M. R., Sreevidya S., Pollock B. H., Jayaprakash P. G., Herman B. (2002) Polymorphism at codon 72 of p53, human papillomavirus, and cervical cancer in South India. J. Cancer Res. Clin. Oncol. 128, 627-631.

Powell B. L., Soong R., Iacopetta B., Seshandri R, Smith D. R. (2000) Prognostic significance of mutations to different structural and functional regions of the p53 gene in breast cancer. Clinical Cancer Research 6, 443-451.

Powell B. L., Staveren I.L., Roosken P., Grieu F., Berns E. M. J. J., Iacopetta B. (2002) Association between common polymorphisms in TP53 and p21WAF1/Cip1 and phenotypic features of breast cancer. Carcinogenesis 23(2), 311-315.

Qie M., Zhang Y., Wu J. (2002) Study on the relationship between cervical cancer and p53 codon 72 polymorphism. Hua Xi Yi Ke Da Xue Xue Bao 33(2), 274-275.

Ralhan R., Agarwal S., Mathur M., Wasylyk B., Srivastava A. (2000) Association between polymorphism in $p21^{Waf1/Cip1}$ cyclin-dependent kinase inhibitor gene and human oral cancer. Clinical Cancer Research 6, 2440-2447.

Radimer K., Siskind V., Bain C., Schofield F. (1993) Relation between anthropometric indicators and risk of breast cancer among Australian women. American Journal of Epidemiology 138(2), 77-89.

Roh JW., Kim MH., Kim JW., Park NH., Song YS., Kang SB., Lee HP., (2001) Polymorphisms in codon 31 of p21 and cervical cancer susceptibility in Korean women. Cancer Letters 165, 59-62.

Rosenthal AN., Ryan A., Hopster D., Jacobs IJ. (2000) p53 codon 72 polymorphism in vulval cancer and vulval intraepithelial neoplasia. British Journal of Cancer 83(10), 1287-1290.

Saranath D., Khan Z., Tandle A. T., Dedhia P., Sharmaa B., Contractor R., Shrivastava S., Dinshaw K. (2002) HPV16/18 prevalence in cervical lesions/cancers and p53 genotypes in cervical cancer patients from India. Gynecologic Oncology 86, 157-162.

Schmute C. and Fishel R. (1999) Genomic Instability: First Step to Carcinogenesis. Anticancer Research 19, 4665-469.

Shih CM.,Lin PT., Wang HC., Huang WC., Wang YC. (2000) Lack of evidence of association of $p21^{WAF1/CIP1}$ polymorphism with lung cancer susceptibility and prognosis in Taiwan. Jpn. J. Cancer Res. 91, 9-15.

Sjalander A., Birgander R., Rannug A., Alexandrie AK., Tornling G., Beckman G., (1996) Association between p21 codon 31 A1(arg) allele and lung cancer. Human Heredity 46 (4), 221-225.

Sourvinos G., Rizos E., Spandidos D. A. (2001) p53 codon 72 polymorphism is linked to the development and not the progression of benign and malignant laryngeal tumours. Oral Oncology 37, 572-578.

Storey A., Thomas M., Kalita A., Harwood C., Gardiol D., Mantovani F., Breuer J., Leigh IM., Matlashewski G., Banks L. (1998) Role of a p53 polymorphism in the development of human papilloma-virus-associated cancer. Nature 393, 229-234.

Su L., Liu G., Zhou W., Xu L. L., Miller D. P., Park S., Lynch T. J., Wain J. C., Christiani D. C. (2003) No Association between the p21 codon 31 Serine-Arginine Polymorphism and Lung Cancer Risk. Cancer Epidemiology, Biomarkers & Prevention 12, 174-175.

Su L., Sai Y., Fan R., Thurston S. W., Miller D. P., Zhou W., Wain J. C., Lynch T., Liu G., Christiani D. C. (2003) p53 (codon 72) and p21 (codon 31) polymorphisms alter in vivo mRNA expression of p21. Lung Cancer 40, 259-266.

Sun Y., Hildesheim A., Li H., Li Y., Chen JY., Cheng YJ., Hayes RB., Rothman N., Bi WF., Cao Y. et al., (1995) No point mutation but a codon 31 ser→arg polymorphism of the WAF-1/CIP-!/p21 tumor suppressor gene in nasopharyngeal carcinoma(NPC): the polymorphism distinguishes Caucasions from Chinese. Cancer Epidemiology, Biomarkers & Prevention 4(3), 261-267.

Suspitsin E. N., Buslov K. G., Grigoriev M. Y., Ishutkina J. G., Ulibina J. M., Gorodinskaya V. M., Pozharisski K. M., Berstein L. M., Hanson K., P. (2003) Evidence against involvement of P53 polymorphism in breast cancer predisposition. Int. J. Cancer 103, 431-433.

Tada M., Furuuchi K., Kaneda M., Matsumoto J., Takahashi M., Hirai A., Mitsumoto Y., Iggo R. D., Moriuchi T. (2001) Inactivate the remaining p53 allele or the alternate p73? Preferential selection of the Arg72 polymorphism in cancers with

recessive p53 mutants but not transdominant mutants. Carcinogenesis 22(2), 515-517.

Terry P. D. and Rohan T. E. (2002) Cigarette smoking and the risk of breast cancer in women: a review of the literature. Cancer Epidemiology, Biomarkers & Prevention 11, 953-971

Thomas M., Kalita A., Labrecque S., Pim D., Banks L., Matlashewski G. (1999) Two polymorphic variants of wild type p53 differ biochemically and biologically. Molecular and Cellular Biology 19(2), 1092-1100.

Ursin G., Longnecker M. P., Haile R. W., Greenland S. (1995) A meta-analysis of body mass index and risk of premenopausal breast cancer. Epidemiology 6(2), 137-141.

Vignal A., Milan D., SanCristobal M., Eggen A. (2002) A review on SNP and other types of molecular markers and their use in animal genetics. Genet. Sel. Evol. 34, 275-305.

Vogelstein B., Kinzler K. W. (2002) The Genetic Basis of Human Cancer, McGraw-Hill, Medical Publishing Division.

Weber B. L., Nathanson K. L. (2000) Low penetrance genes associated with increased risk for breast cancer. European Journal of Cancer 36, 1193-1199.

Wrensch M., Chew T., Farren G., Barlow J., Belli F., Clarke C., Erdmann C. A., Lee M., Moghadassi M., Peskin-Mentzer R., Quesenberry Jr. C. P., Souders-Mason V., Spence L., Suzuki M., Gould M. (2003) Risk factors for breast cancer in population with high incidence rates. Breast Cancer Research 5(4), 88-102.

Zhang ZW., Newcomb P., Hollowood A., Feakins R., Moorghen M., Storey A., Farthing M. J. G., Alderson D., Holly J. (2003) Age-associated increase in codon 72 Arginine frequency in gastric cardia and non-cardia adenocarcinoma. Clinical Cancer Research 9, 2151-2156.

7. APPENDIX

The coding of the risk and non-risk groups of every table in the results part (Chapter 3). Cases: Breast cancer patients. Controls: Non-breast cancer subjects.

Table 3.1. General Characteristics of the Subjects in the Turkish Study Population

Characteristics	**Variable 1*** **Case/Control**	**Variable 2**	**OR (95%CI)**
Family history of breast cancer(FHBC)	Case: 1, Control: 0	FHBC present: 1, FHBC absent: 0	6.72 (2.97-15.22)
BMI≥27.40	Case: 1, Control: 0	[BMI≥27.40]: 1, [BMI<27.40]: 0	0.90 (0.65-1.24)
Age at menarche(AM)<14	Case: 1, Control: 0	[AM<14]: 1, [AM≥14]: 0	1.55 (1.12-2.14)
Premenopausal	Case: 1, Control: 0	Pre: 1, Post: 0	0.68 (0.50-0.93)
Postmenopausal	Case: 1, Control: 0	Post: 1, Pre: 0	1.47 (1.08-2.00)
Smoking	Case: 1, Control: 0	Smoker: 1, Non-smoker: 0	0.89 (0.62-1.28)

* Variable 1; cases: 1, controls: 0

Table 3.2. Distribution of the *p53* Codon 72 Genotypes in the Age-matched Control and Breast Cancer Patients

*

Menopausal status	**Genotype**	**Variable 2**	**Crude OR 95% CI**	**Variable 2**	**Adjusted OR[a] 95% CI**
All	R/R	R/R: 1, P/P: 0	2.16 (1.08-4.31)	R/R: 1, P/P: 0	1.88 (0.69-5.10)
	R/P	R/P: 1, P/P: 0	1.65 (0.83-3.27)	R/P: 1, P/P: 0	1.62 (0.60-4.35)
	P/P	P/P: 1, R/R:0	0.46 (0.23-0.93)	P/P: 0 (reference)	1.00
	R/P+P/P	R/P, P/P: 1, R/R:0	0.71 (0.50-1.02)		
Pre	R/R	R/R: 1, P/P: 0	2.59 (0.97-6.93)	R/R: 1, P/P: 0	2.60 (0.59-11.38)
	R/P	R/P: 1, P/P: 0	2.30 (0.87-6.09)	R/P: 1, P/P: 0	2.26 (0.53-9.61)
	P/P	P/P: 1, R/R:0	0.39 (0.14-1.03)	P/P: 0 (reference)	1.00
	R/P+P/P	R/P, P/P: 1, R/R:0	0.78 (0.46-1.32)		
Post	R/R	R/R: 1, P/P: 0	1.37 (0.48-3.95)	R/R: 1, P/P: 0	1.57 (0.29-8.62)
	R/P	R/P: 1, P/P: 0	1.02 (0.36-2.91)	R/P: 1, P/P: 0	1.27 (0.23-6.87)
	P/P	P/P: 1, R/R:0	0.73 (0.25-2.09)	P/P: 0 (reference)	1.00
	R/P+P/P	R/P, P/P: 1, R/R:0	0.74 (0.45-1.22)		

* Variable 1; cases: 1, controls: 0

Table 3.3. Distribution of the *p53* Codon 72 Genotypes Stratified According to BMI in Cases and Controls

*

		Low BMI (<27.40)		High BMI (≥27.40)	
Menopausal status	**Genotype**	**Variable 2**	**OR (95% CI)**	**Variable 2**	**OR (95% CI)**
All	P/P	Reference		reference	
	R/R	R/R: 1, P/P: 0	1.46 (0.58-3.67)	R/R: 1, P/P: 0	3.86 (1.12-13.26)
	R/P+R/R	R/P, R/R: 1, P/P:0	1.51 (0.63-3.65)	R/P, R/R: 1, P/P:0	2.96 (0.90-9.77)
Pre	P/P	Reference		reference	
	R/R	R/R: 1, P/P: 0	1.74 (0.52-5.81)	R/R: 1, P/P: 0	3.75 (0.64-22.04)
	R/P+R/R	R/P, R/R: 1, P/P:0	1.77 (0.56-5.59)	R/P, R/R: 1, P/P:0	3.28 (0.62-17.44)
Post	P/P	Reference		reference	
	R/R	R/R: 1, P/P: 0	0.99 (0.21-4.67)	R/R: 1, P/P: 0	3.39 (0.58-19.99)
	R/P+R/R	R/P, R/R: 1, P/P:0	1.05 (0.24-4.64)	R/P, R/R: 1, P/P:0	2.47 (0.44-14.01)

* Variable 1; cases: 1, controls: 0

Table 3.4. Distribution of the *p21* Codon 31 Genotypes in the Age-matched Control and Breast Cancer Patients

*

Menopausal status	Genotype	Variable 2	Crude OR 95% CI	Variable 2	Adjusted OR[a] 95% CI
All	S/S	S/S: 1, R/R: 0	0.64 (0.12-3.52)	S/S: 0 (reference)	1.00
	S/R	S/R: 1, R/R: 0	1.13 (0.73-1.75)	S/R: 1, S/S: 0	1.03 (0.58-1.82)
	R/R	R/R: 1, S/S: 0	1.57 (0.28-8.64)	R/R: 1, S/S: 0	3.70 (0.33-41.66)
	S/R+R/R	S/R, R/R: 1, S/S: 0	1.15 (0.75-1.76)	S/R, R/R: 1, S/S: 0	1.10 (0.63-1.92)
Pre	S/S	S/S: 1, R/R: 0	0.55 (0.05-6.17)	S/S: 0 (reference)	1.00
	S/R	S/R: 1, R/R: 0	0.95 (0.50-1.81)	S/R: 1, S/S: 0	0.58 (0.21-1.56)
	R/R	R/R: 1, S/S: 0	1.81 (0.16-20.33)	R/R: 1, S/S: 0	1.28 (0.091-17.96)
	S/R+R/R	S/R, R/R: 1, S/S: 0	0.99 (0.53-1.85)	S/R, R/R: 1, S/S: 0	0.62 (0.24-1.63)
Post	S/S	S/S: 1, R/R: 0	0.76 (0.07-8.47)	S/S: 0 (reference)	1.00
	S/R	S/R: 1, R/R: 0	1.22 (0.66-2.25)	S/R: 1, S/S: 0	1.34 (0.63-2.86)
	R/R	R/R: 1, S/S: 0	1.32 (0.12-14.75)	R/R: 1, S/S: 0	128.298 (0.00-EXP)
	S/R+R/R	S/R, R/R: 1, S/S: 0	1.22 (0.67-2.23)	S/R, R/R: 1, S/S: 0	1.40 (0.66-2.96)

* Variable 1; cases: 1, controls: 0

Table 3.5. Distribution of the *p21* Codon 31 Genotype Stratified According to BMI in Cases and Controls

*

		Low BMI (<27.40)		**High BMI (≥27.40)**	
Menopausal status	**Genotype**	**Variable 2**	**OR (95% CI)**	**Variable 2**	**OR (95% CI)**
All	S/S	reference		reference	
	S/R+R/R	S/R, R/R: 1, S/S: 0	1.32 (0.71-2.45)	S/R, R/R: 1, S/S: 0	0.80 (0.41-1.57)
Menopausal status	**Genotype**		**OR (95% CI)**		**OR (95% CI)**
Pre	S/S	reference		reference	
	S/R+R/R	S/R, R/R: 1, S/S: 0	0.70 (0.30-1.67)	S/R, R/R: 1, S/S: 0	0.97 (0.33-2.84)
Menopausal status	**Genotype**		**OR (95% CI)**		**OR (95% CI)**
Post	S/S	reference		reference	
	S/R+R/R	S/R, R/R: 1, S/S: 0	2.59 (0.96-7.01)	S/R, R/R: 1, S/S: 0	0.73 (0.30-1.74)

* Variable 1; cases: 1, controls: 0

Table 3.6. Combination of the *p21* Codon 31 Genotype with the p53 Codon 72 Genotype for Breast Cancer Risk
*

Genotype at risk	**P21**	**P53**	**Variable 2**	**Crude OR**
None	S/S	P/P	reference	1.00
One	S/S	R/R	S31S/ R72R: 1; S31S/ P72P:01	2.45 (1.15-5.21)
	R/R+S/R	P/P	R31R+S31R/ P72P:1; S31S/ P72P:01	2.42 (0.36-16.50)
Two	R/R+S/R	R/R	R31R+S31R/ R72R: 1; S31S/ P72P:01	2.66 (1.06-6.66)

* Variable 1; cases: 1, controls:0

Table 3.7. Characteristics of Subjects from the Greek Population

Characteristics	Variable 1 (cases/controls)	Variable 2	OR 95% CI
Premenopausal	Cases: 1, controls: 0	Pre: 1, post: 0	1.45 (0.95-2.22)
Postmenopausal	Cases: 1, controls: 0	Post: 1, pre: 0	0.69 (0.45-1.05)

Table 3.8. Distribution of the *p53* Codon 72 Genotype in the Age-matched Control and Breast Cancer Patients in the Greek Population

Menopausal status	**Genotype**	**Variable 1 (cases/controls)**	**Variable 2**	**Crude OR 95% CI**
All	R/R	Cases: 1, controls: 0	R/R: 1, P/P: 0	7.93 (0.95-65.98)
	R/P	Cases: 1, controls: 0	R/P: 1, P/P: 0	6.50 (0.77-54.64)
	P/P	Cases: 1, controls: 0	P/P: 1, R/R:0	0.13 (0.02-1.05)
	R/P+P/P	Cases: 1, controls: 0	R/P, P/P: 1, R/R:0	0.74 (0.46-1.20)
Pre	R/R	Cases: 1, controls: 0	R/R: 1, P/P: 0	3.57 (0.36-35.76)
	R/P	Cases: 1, controls: 0	R/P: 1, P/P: 0	3.78 (0.37-38.82)
	P/P	Cases: 1, controls: 0	P/P: 1, R/R:0	0.28 (0.03-2.81)
	R/P+P/P	Cases: 1, controls: 0	R/P, P/P: 1, R/R:0	0.97 (0.49-1.92)
Post	R/R	*	*	*
	R/P	*	*	*
	P/P	*	*	*
	R/P+P/P	Cases: 1, controls: 0	R/P, P/P:1, R/R:0	0.57 (0.28-1.17)

* Because the frequency of P/P genotype for cases was 0.00%, we could not calculate the odds ratios for R/R, R/P, and P/P.

Table 3.9. Distribution of the *p21* Codon 31 Genotypes in the Age-matched Control and Breast Cancer Patients in the Greek Population

Menopausal status	**Genotype**	**Variable 1 (cases/controls)**	**Variable 2**	**Crude OR 95% CI**
All	S/S	Cases: 1, controls: 0	S/S: 1; R/R, S/R: 0	1.12 (0.60-2.09)[a]
	S/R	Cases: 1, controls: 0	S/R: 1; S/S, R/R: 0	1.03 (0.54-1.96)[b]
	R/R	*	*	*
	S/R+R/R	Cases: 1, controls: 0	S/R, R/R: 1, S/S: 0	0.89 (0.48-1.67)
Pre	S/S	Cases: 1, controls: 0	S/S: 1; R/R, S/R: 0	0.94 (0.39-2.26)[a]
	S/R	Cases: 1, controls: 0	S/R: 1; S/S, R/R: 0	1.07 (0.44-2.58)[b]
	R/R	Cases: 1, controls: 0	*	*
	S/R+R/R	Cases: 1, controls: 0	S/R, R/R: 1, S/S: 0	1.07 (0.44-2.58)
Post	S/S	Cases: 1, controls: 0	S/S: 1; R/R, S/R: 0	1.34 (0.51-3.50)[a]
	S/R	Cases: 1, controls: 0	S/R: 1; S/S, R/R: 0	1.00 (0.36-2.75)[b]
	R/R	Cases: 1, controls: 0	*	*
	S/R+R/R	Cases: 1, controls: 0	S/R, R/R: 1, S/S: 0	0.75 (0.29-1.96)

a) Because the frequency of R/R genotype was 0.00%, the combination of R/R and S/R was taken as non-risk group in calculating the OR for S/S genotype

b) Non-risk group was taken as the combination of S/S and R/R, in calculating the OR for S/R genotype

* Because the frequency of R/R genotype was 0.00% for the cases, we could not calculate the OR for R/R genotype

Table 3.10. Distribution of the *p53* Codon 72 Genotypes in the Age-matched Control and Breast Cancer Patients in the Turkish and Greek Populations.

Menopausal status	**Genotype**	**Variable 1 (cases/controls)**	**Variable 2**	**OR (95% CI)**
All	R/R	Cases: 1, controls: 0	R/R: 1, P/P: 0	***2.35 (1.25-4.41)***
	R/P	Cases: 1, controls: 0	R/P: 1, P/P:0	***1.89 (1.01-3.56)***
	P/P	Cases: 1, controls: 0	P/P: 1, R/R:0	0.43 (0.23-0.80)
	R/P+P/P	Cases: 1, controls: 0	R/P: 1, P/P: 1, R/R:0	0.75 (0.56-0.99)
Pre	R/R	Cases: 1, controls: 0	R/R: 1, P/P: 0	***2.70 (1.12-6.54)***
	R/P	Cases: 1, controls: 0	R/P: 1, P/P:0	***2.55 (1.05-6.19)***
	P/P	Cases: 1, controls: 0	P/P: 1, R/R:0	0.37 (0.15-0.90)
	R/P+P/P	Cases: 1, controls: 0	R/P: 1, P/P: 1, R/R:0	0.83 (0.55-1.26)
Post	R/R	Cases: 1, controls: 0	R/R: 1, P/P: 0	1.70 (0.67-4.28)
	R/P	Cases: 1, controls: 0	R/P: 1, P/P:0	1.27 (0.50-3.21)
	P/P	Cases: 1, controls: 0	P/P: 1, R/R:0	0.59 (0.23-1.49)
	R/P+P/P	Cases: 1, controls: 0	R/P: 1, P/P: 1, R/R:0	0.73 (0.49-1.09)

Table 3.11. Distribution of the *p21* Codon 31 Genotypes in the Age-matched Control and Breast Cancer Patients in the Turkish and Greek Populations.

Menopausal status	**Genotype**	**Variable 1 (cases/controls)**	**Variable 2**	**OR (95% CI)**
All	S/S	Cases: 1, controls: 0	S/S: 1; R/R: 0	1.55 (0.41-5.81)
	S/R	Cases: 1, controls: 0	S/R: 1; S/S: 0	1.10 (0.77-1.58)
	R/R	Cases: 1, controls: 0	R/R: 1, S/S:0	0.65 (0.17-2.43)
	S/R+R/R	Cases: 1, controls: 0	S/R, R/R: 1; S/S: 0	1.07 (0.75-1.52)
Pre	S/S	Cases: 1, controls: 0	S/S: 1; R/R: 0	0.60 (0.05-6.68)
	S/R	Cases: 1, controls: 0	S/R: 1; S/S: 0	0.99 (0.59-1.66)
	R/R	Cases: 1, controls: 0	R/R: 1, S/S:0	1.67 (0.15-18.57)
	S/R+R/R	Cases: 1, controls: 0	S/R, R/R: 1; S/S: 0	1.01 (0.61-1.68)
Post	S/S	Cases: 1, controls: 0	S/S: 1; R/R: 0	2.53 (0.46-13.99)
	S/R	Cases: 1, controls: 0	S/R: 1; S/S: 0	1.21 (0.72-2.04)
	R/R	Cases: 1, controls: 0	R/R: 1, S/S:0	0.40 (0.07-2.18)
	S/R+R/R	Cases: 1, controls: 0	S/R, R/R: 1; S/S: 0	1.11 (0.68-1.83)