

ANKARA ÜNİVERSİTESİ
BİYOTEKNOLOJİ ENSTİTÜSÜ

DOKTORA TEZİ

**SPORADİK KOLOREKTAL KANSER VAKALARINDA GENOM EBADINDA
TEK NÜKLEOTİT POLİMORFİZM PROFİLİNİN BELİRLENMESİ İLE YENİ
GENETİK YATKINLIK GENLERİNİN VE KANSERİN GELİŞMESİNDE ETKEN
OLAN GENLERİN BELİRLENMESİ**

DİLAY ÇIĞLIDAĞ DÜNGÜL

Danışman Öğretim Üyesi

Doç.Dr.Hilal ÖZDAĞ

ANKARA

2011

ONAY SAYFASI

Doç.Dr.Hilal ÖZDAĞ danışmanlığında, Dilay ÇIĞLIDAĞ DÜNGÜL tarafından hazırlanan bu çalışma 06/07/2011 tarihinde aşağıdaki jüri tarafından Temel Biyoteknoloji Anabilim Dalı'nda doktora tezi olarak kabul edilmiştir.

Başkan: Yrd.Doç.Dr.Ali Osmay GÜRE İmza:

Üye: Prof.Dr.Nurten AKARSU İmza:

Üye: Prof.Dr.Arzu Ensari İmza:

Üye: Doç.Dr.Hilal ÖZDAĞ İmza:

Üye: Yard.Doç.Dr.Özlem İLK İmza:

ÖZET

Kolorektal karsinomalar, Batı'da, akciğer ve meme karsinomalarından sonra en sık ölüme neden olan malignite grubunu oluşturmakta ve insidansı birçok ülkede artış göstermektedir. Kolorektal kanserde beş yıllık yaşam şansı yaklaşık %40 kadardır. Olguların büyük kısmının (%85-90) sporadik olduğu, % 10-15'lik küçük bir grubun ise herediter sendromlar kapsamında geliştiği bilinmektedir.

Kanser hastalığının yaklaşık olarak %10-15'i ailesel özellik gösterirken kalan %85-90'ı sporadik olarak ortaya çıkmaktadır. Ailesel kanserlerin ortaya çıkışına neden olan genlerin birçoğu bağlantı analizleri ile saptanmıştır. Ancak kanser vakalarının büyük bir çoğunluğunu oluşturan sporadik kanserin görüldüğü bireylerde, genetik yatkınlığa neden olan gen bölgelerinin saptanmasında diğer kompleks, multifaktöryel hastalıklarda olduğu üzere asosiyasyon çalışmaları yürütülmektedir. Asosiyasyon çalışmalarının önündeki en büyük sıkıntılardan biri olan popülasyon stratifikasyonu ise aile bazlı çalışma tasarımları ile çözümlenmektedir.

Aile temelli genom boyu asosiyasyon çalışması analizi için en az 200 trionun gerekliliğini vurgulayan TDT yaklaşımının aksine 200'den daha az sayıda trio ile aile temelli genom boyu asosiyasyon çalışması analizlerinin yapılabileceği ekibimiz tarafından daha önce geliştirilen SPRT analizi ile gösterilmiştir. Bu tez kapsamında elde edilen 51 trio dan oluşan veri SPRT yaklaşımı ile analiz edilmiştir.

Tez çalışması kapsamında sporadik kolorektal kanser vakaları ve herbir vakanın anne ve babasının periferik kanlarından elde edilen DNA örnekleri ile genom boyu asosiyasyon çalışması yürütülmüştür. Çalışma sonunda toplanmış olan 51 trio ve 33 duo elde edilen bilginin analizi sonucunda kolorektal kanser ile asosiyasyon gösteren aday genetik yatkınlık gen bölgeleri saptanmıştır. Bu bölgeler SPRT ile tespit edilen 1p36.12, 3p22.3, 5q11.2, 11q22.1, 20p12.1 ve C-TDT ile tespit edilen 2p16.3, 2p22.3, 2q14.1, 2q22.1-2q22.2, 3p26.1, 4q28.3, 4q34.3, 5p15.2, 5q34, 6q16.3, 8q24.3, 9p21.3, 11q14.1, 11q22.1, 14q12, 14q31.3, 14q32.2 kromozom bölgeleridir.

Ayrıca yapılan kopya sayısı deęişiklikleri (Copy Number Variation, CNV) analizi ile 4q13.2, 9p12, 15q11.1-15q11.2 kromozom lokuslarında çeşitli baz çifti uzunluklarındaki bölgelerde delesyonlar ve 2q11.1, 17q21.31, 19p13.2, Xq21.2, Xq28, Xp11.23 lokuslarında ise yine farklı uzunluklardaki bölgelerde amplifikasyonlar saptanmıştır. Saptanan delesyon bölgelerinin tümünün, amplifikasyon bölgelerinin ise bazılarının daha önce kopya sayısı varyasyonu olarak tanımlandığı belirlenmiştir.

Anahtar Kelimeler: Aile temelli genom boyu asosiyasyon çalışması, C-TDT, SPRT, kolorektal kanser

ABSTRACT

Colorectal cancers show the third highest mortality rate after lung and breast cancer in Western countries. Its incidence increases in most of the countries. Five year survival rate of colorectal cancer is 40%. It is known that while 85 to 90 % of the colorectal cancer cases are sporadic the remaining 10 to 15% are familial.

Cancer, a genetic disease, arises in any tissue because of many factors through a multistep process. Because of these characteristics cancer has a heterogenous profile which makes its mechanism difficult to reveal and explain. While 10-15% of this disease is hereditary the remaining 85-90% is sporadic. Most of the genes causing hereditary cancers have been identified through linkage analysis. Meanwhile genetic predisposition to sporadic cancer cases which constitutes most of cancer cases has been studied by association studies as in other complex and multifactorial diseases. Population stratification, one of the most important problem in association studies, has been solved using family based designs.

Within the frame of this project primarily a novel approach previously developed by our group adapting SPRT analysis to family based association analysis was used. SPRT contrary to TDT which requires at least 200 trios in order to conduct a family based association study, has the ability to conduct a reliable association analysis using less than 200 trios. 51 trios which were obtained in this thesis were analysed by SPRT.

In this study genome-wide association studies were conducted using DNA samples extracted from peripheral blood samples of sporadic colorectal cancer cases and their healthy parents. As a result of the data obtained from 51 number of trios and 33 number of duos novel candidate genetic susceptibility loci to colorectal cancer were identified. The loci detected by SPRT are 1p36.12, 3p22.3, 5q11.2, 11q22.1, 20p12.1 and loci detected by C-TDT are 2p16.3, 2p22.3, 2q14.1, 2q22.1-2q22.2, 3p26.1, 4q28.3, 4q34.3, 5p15.2, 5q34, 6q16.3, 8q24.3, 9p21.3, 11q14.1, 11q22.1, 14q12, 14q31.3, 14q32.2 respectively.

Also various length of deletions and amplifications are detected by copy number variation (CNV) analysis. The amplified loci are 2q11.1, 17q21.31, 19p13.2, Xq21.2, Xq28, Xp11.23 and deleted loci are 4q13.2, 9p12, 15q11.1-15q11.2. All of the these deleted loci and some of the amplified loci are previously defined as CNV.

Key words: Family based genome-wide association study, C-TDT, SPRT, colorectal cancer

ÖNSÖZ VE TEŞEKKÜR

Bu tezin oluşması ve sonuçlanmasında emeğini ve zamanını harcamaktan kaçınmayan sevgili hocam Doç. Dr. Hilal Özdağ'a,

Tez projemin her basamağında kıymetli bilgilerini paylaşan ve emeğini esirgemeyen sevgili Hocam Prof. Dr. Nurten Akarsu'ya,

Bu çalışmanın biyoinformatik analizlerinde emeğini esirgemeyen sevgili hocalarım Doç. Dr. Hakkı Gökhan İlk, Yrd. Doç. Dr. Özlem İlk ve arkadaşlarım Dr. Farid Rajablı, Blm.Uzm Gül İnan ve Blm.Uzm Mehmet Korkmaz'a,

Doktora eğitimim boyunca hem bilim adamı hem de pediatrist kimliği ile verdiği destek ve güven için sevgili hocam Prof. Dr. Nejat Akar'a,

Tez çalışmamı 107S348 no'lu proje olarak destekleyen Tübitak Sağlık Bilimleri Araştırma Grubu'na,

107S348 no'lu Tübitak projemize ve dolayısı ile bu teze katkılarından dolayı Prof. Dr. Ayhan Kuzu, Prof. Dr. Arzu Ensari ve Doç.Dr. Berna Savaş'a,

Yüksek lisans ve doktora öğrenimim boyunca her zaman dostluğu ile hep yanımda olan sevgili arkadaşım Blm.Uzm Uğur Akpulat'a,

Doktora çalışmamda aynı laboratuvarı paylaştığım çalışma arkadaşlarıma,

Doktora projem boyunca her konuda destek olan çok sevgili arkadaşım, Blm.Uzm Nevin Belder'e,

Her zaman her konuda hiç düşünmeden yanımda olan sevgili halam Av.Sevgi Çığlıdağ'a,

Uzun süren eğitim hayatım boyunca gösterdiği anlayış için canım kardeşim Cenk Çığlıdağ'a,

Dođru ve mutlu bir insan olmam için verdiđi emek, gösterdiđi anlayış ve sonsuz sevgisi için canım babam Muzaffer Cengiz ıđlıdađ' a,

Őu an yanımda olamasa da varlığını her zaman içimde hissettiđim, varolma nedenim canım annem Tünay ıđlıdađ' a,

Sonsuz sabrı ve sevgisi için, sevgilim, hayat ve yol arkadaşım Ömür Döngöl'e,

Bu tez çalışmasının tam ortasında dünyaya gelen, hayatımın en deđerli varlığı, ışığı, yaşama sebebim canım kızım Tünay Ada Döngöl'e,

sonsuz teşekkürlerimi sunarım.

Dilay IĐLIDAĐ DÖNGÖL

Ankara, Haziran 2011

Annem Tünay Çıglıdağ ve kızım Tünay Ada Düngül'e

İÇİNDEKİLER

ONAY SAYFASI.....	i
ÖZET	ii
ABSTRACT	iv
ÖNSÖZ VE TEŞEKKÜR.....	vi
İÇİNDEKİLER.....	ix
SİMGELER DİZİNİ	xiii
ŞEKİLLER DİZİNİ	xv
ÇİZELGELER DİZİNİ.....	xvi
1.GİRİŞ.....	1
2. KURAMSAL TEMELLER.....	2
2.1. Kanser ve Moleküler Biyolojisi.....	2
2.2. Kolorektal Kanser.....	6
2.2.1. Kolorektal Kanserin Moleküler Biyolojisi	9
2.2.2. Kolorektal Kanserin Kalıtımı	12
2.3. Hastalık Yatkınlık Genlerinin Lokalizasyonun Tespiti.....	14
2.3.1. Moleküler Belirteçler ve Genlerin Lokalizasyonlarının Tespiti.....	15
2.3.1.1. Tek Nükleotid Polimorfizmi.....	16
2.3.1.2. Kopya Sayısı Varyasyonları	17
2.4. Tek Gen Hastalıkları ve Bağlantı Analizleri	18

2.5. Kompleks Hastalıklar ve Asosiyasyon Çalışmaları	22
2.6. Asosiyasyon Çalışmaları	24
2.6.1. Aday Gen Assosiyasyon Çalışmaları	25
2.6.2. Genom Boyu Asosiyasyon Çalışmaları.....	26
2.6.2.1. Popülasyon Temelli Genom Boyu Asosiyasyon Çalışmaları.....	29
2.6.2.2. Aile Temelli Genom Boyu Asosiyasyon Çalışmaları ve İstatistiksel Yaklaşımlar.....	30
2.7. Kanser Araştırmalarında Genom Boyu Asosiyasyon Çalışmaları	35
2.8. Kolorektal Kanser Araştırmalarında Genom Boyu Asosiyasyon Çalışmaları	37
2.9. Amaç ve Mantıksal Temeller	42
3. MATERYAL VE YÖNTEM.....	44
3.1. Materyal.....	44
3.2. Yöntem	44
3.2.1. DNA İzolasyonu	44
3.2.2. DNA'nın Spektrofotometrik Analizi.....	46
3.2.3. DNA'nın Agaroz Jel Elektroforezi.....	46
3.3. 250K DNA Mikrodizin.....	47
3.3.1. Restriksiyon Enzim ile DNA'nın kesilmesi	47
3.3.2. Ligasyon	48
3.3.3. PCR ve PCR Sonrası Jel Elektroforezi.....	48
3.3.4. PCR Purifikasyon ve Elüsyonu	50
3.3.5. Purifiye PCR Ürünlerinin Miktar Ölçümü	50
3.3.6. Fragmantasyon ve Fragmantasyon Sonrası Jel Elektroforezi	50

3.3.7. İşaretleme.....	51
3.3.8. Hibridizasyon	52
3.3.9. Yıkama, Boyama ve Tarama	53
3.4. Biyoinformatik Analizler.....	56
3.4.1. Ardışık Hipotez Testleri (Sequential Probability Ratio Test, SPRT).....	56
3.4.2. Birleşik Dengesiz Aktarım Testi (C-TDT) ve Permütasyon Testleri	56
3.4.2.1. Birleşik Dengesiz Aktarım Testi (C-TDT) ve Permütasyon Testleri için Benzetim Çalışması	60
3.5. Biyoinformatik Analiz Sonuçlarının Doğrulanması için Yapılan Çalışmalar.....	65
3.5.1. SNP Genotiplendirme Çalışmaları	65
3.5.2. Kopya Sayısı Değişiklikleri Analizi (Copy Number Variation, CNV).....	65
3.5.3. Haplotip Analizi	66
3.5.3.1. DNA Dizi Analizi.....	66
4. ARAŞTIRMA BULGULARI	68
4.1. DNA Bütünlük, Kalite ve Miktar Tayini.....	68
4.2. 250K DNA Mikrodizin Analizi.....	69
4.2.1. Restriksiyon Endonükleaz Kesim Sonuçları	69
4.2.2. PCR Elektroforez Sonuçları	69
4.2.3. PCR Pürifikasyon Sonrası Miktar Tayini.....	70
4.2.4. Fragmentasyon ve İşaretleme	71
4.2.5. Hibridizasyon, Yıkama, Boyama ve Tarama	71
4.3. Biyoinformatik Bulgular.....	73
4.3.1. SPRT Analizi Sonuçları	73

4.3.3. C-TDT Analizi İçin Yapılan Benzetim Çalışmasının Sonuçları	76
4.3.4. C-TDT Analizi Sonuçları	77
4.4. Biyoinformatik Analiz Sonuçlarının Doğrulanması için Yapılan Çalışmaların Sonuçları.....	82
4.4.1. SNP Genotiplendirme Çalışmalarının Sonuçları.....	82
4.4.2. Kopya Sayısı Değişiklikleri Analizi Bulguları.....	85
4.4.3. Haplotip Analizi	87
4.4.3.1. DNA Dizi Analizi Sonuçları.....	89
5. TARTIŞMA ve SONUÇ	90
KAYNAKLAR.....	100
ÖZGEÇMİŞ.....	108

SİMGELER DİZİNİ

APC	Adenomatous polyposis coli
BSA	Bovine Serum Albumin
CADM1	Cell Adhesion Molecule 1
MYC	v-Myc Myelocytomatosis Viral Oncogene Homolog
CNV	Copy Number Variation
C-TDT	Combined Transmission Disequilibrium Test
DMSO	Dimethyl sulfoxide
dNTP	Deoksiribonükleotit trifosfat
FAP	Familial adenomatous polyposis
HNPCC	Hereditör Nonpolyposis Colorectal Cancer
HSDNA	Herring Sperm DNA
MAPK	MAP Kinaz
MgCl ₂	Magnezyum klorür
PCR	Polimeraz Zincir Reaksiyonu
PRKX	Protein kinase, X-linked
PTPN5	Protein Tyrosine Phosphatase, Non-Receptor Type 5
SAPE	Streptavidin Phycoerythrin
SNP	Single Nucleotide Polymorphism
SPRT	Sequential Probability Ratio Test
TdT	Terminal Deoksinükleotid Transferaz

TDT	Transmission Disequilibrium Test
TMACL	Tetra Metil Amonyum Klorit
μ l	Mikrolitre

ŞEKİLLER DİZİNİ

Şekil 2.1: Kadınlarda yerleşim yerlerine göre en sık görülen kanserler (Türkiye 2004-2006)	7
Şekil 2.2: Kadınlarda yaşa göre en sık görülen kanserler (Türkiye 2004-2006).....	7
Şekil 2.3: Erkeklerde yerleşim yerlerine göre en sık görülen kanserler (Türkiye 2004-2006)	7
Şekil 2.4: Erkeklerde yaşa göre en sık görülen kanserler (Türkiye 2004-2006)	8
Şekil 2.5: Kolorektal kanserde adenom karsinom sekansı	10
Şekil 2.6: Kolorektal karsinogenezin çok basamaklı modeli (vogelogram).....	11
Şekil 2.7: Kolorektal kanserin heterojenitesi.....	12
Şekil 2.8: Tek gen hastalıkları ve kompleks hastalıklara genetik ve çevresel katkılar (A) Tek gen hastalığı (B) Kompleks hastalık	22
Şekil 2.9: Düşük frekanslı alleller ve hastalığa yatkınlık	23
Şekil 2.10: Genom boyu asosiyasyon yaklaşımı ile genetik asosiyasyonun belirlenmesi ..	27
Şekil 4.1: Genomik DNA'ların agaroz jel elektroforezindeki görüntüleri.....	68
Şekil 4.2: DNA örneklerinin restriksiyon endonükleaz ile kesimi sonrası %1'lik agaroz jel elektroforezi görüntüsü.....	69
Şekil 4.3: PCR jel görüntüleri.....	70
Şekil 4.4: Fragmantasyon görüntüleri	71
Şekil 4.5: SNP'lerin yoğun bulunduğu kromozom bölgeleri.....	81

ÇİZELGELER DİZİNİ

Çizelge 3.1: Kesim karışımı hazırlanışı.....	47
Çizelge 3.2: Ligasyon karışımı hazırlanışı	48
Çizelge 3.3: PCR karışımı hazırlanışı.....	49
Çizelge 3.4: Fragmentasyon karışımı hazırlanışı	51
Çizelge 3.5: İşaretleme karışımı hazırlanışı	51
Çizelge 3.6: Hibridizasyon karışımı hazırlanışı	52
Çizelge 3.7: Boyama karışımının hazırlanışı (1.tüp).....	55
Çizelge 3.8: Boyama karışımı hazırlanışı (2.tüp).....	55
Çizelge 3.9: Herhangi bir SNP’te anne-baba-çocuk üçlüsü genotiplerinin, C-TDT test istatistiğinde, T ve NT terimlerine olan katkıları	61
Çizelge 3.10: Herhangi bir SNP’te anne veya babanın genotipinin gözlemlenmesi üzerine, aynı SNP’te çocuğun genotipinin alabileceği değerler ve bu genotiplerin C-TDT test istatistiğinde, $N_M > C$ veya $F > C$, $N_M < C$ veya $F < C$, $N_M = C$ veya $F = C$ terimlerine katkıları .	63
Çizelge 3.11: MİR128-2 geni için dizayn edilen PCR ve sekans primerleri.....	66
Çizelge 3.12: ARPP21 geni için dizayn edilen PCR ve sekans primerleri.....	66
Çizelge 4.1: Genomik DNA’ların spektrofotometrik okumaları.....	68
Çizelge 4.2: PCR ürünlerinden bazılarının spektrofotometrik okumaları.....	70
Çizelge 4.3: SPRT analizi hastalıkla ilişkili bulunan SNP’ler	73
Çizelge 4.4: Örneklem büyüklüklerine göre güç ve hesaplanmış anlamlılık düzeyler	77
Çizelge 4.5: Aynı lokusta ≥ 30 SNP < 35 bulunması esas alındığında belirlenen lokuslar ..	78

Çizelge 4.6: Aynı lokusta ≥ 35 SNP < 40 bulunması esas alındığında belirlenen lokuslar .	79
Çizelge 4.7: Aynı lokusta < 40 SNP esas alındığında belirlenen lokuslar	80
Çizelge 4.8: Doğrulama analizleri için genotiplendirilmek üzere seçilen SNP'ler.....	83
Çizelge 4.9: rs8793 (TMBIM4 geni) lokusunda genotiplerin hasta ve kontrol gruplarında dağılımı.....	83
Çizelge 4.10: rs2229080 (DCC geni) lokusunda genotiplerin hasta ve kontrol gruplarında dağılımı.....	84
Çizelge 4.11: rs3753841(COL11A1) lokusunda genotiplerin hasta ve kontrol gruplarında dağılımı.....	84
Çizelge 4.12: rs3792594 (PDCD6IP) lokusunda genotiplerin hasta ve kontrol gruplarında dağılımı.....	84
Çizelge 4.13: rs56400844(CXCR4) lokusunda genotiplerin hasta ve kontrol gruplarında dağılımı.....	85
Çizelge 4.14: Kopya sayısı değişiklikleri analizinde delete olduğu saptanan bölgeler.....	86
Çizelge 4.15: Kopya sayısı değişiklikleri analizinde amplifiye olduğu saptanan bölgeler.	86
Çizelge 4.16: Haplotip Analizi (Homozigosite Haritalaması).....	88

1.GİRİŞ

Kanser, dünya çapında çok sık görülen, mortalitesi yüksek bir hastalıktır. Kanserın ortaya çıkışı ve ilerlemesi birçok genin mutasyona uğradığı multigenik ve çok basamaklı bir süreçtir. Bu hastalığın gelişiminde hem genetik değişikliklerin hem de çevresel etkiler gibi birçok faktörün etkisi bulunmaktadır.

Kanserın multigenik bir hastalık olması karsinogenezin anlaşılmasında genomik bilginin hayatiyetini ortaya koymaktadır. Bu çerçevede İnsan Genom Projesinin sonuç ve çıktılarının kanser araştırmaları üzerindeki etkileri bu sahada son derece önemli adımların atılmasını sağlamıştır. Bu yoldan hareketle biri ABD “Cancer Genome Anatomy Project” (<http://cgap.nci.nih.gov/>) diğeri İngiltere merkezli “Cancer Genome Project” (<http://www.sanger.ac.uk/genetics/CGP/>) projeleri kanser genomunun ortaya çıkarılmasını hedefleyerek yola çıkmış ve önemli yol almıştır.

Genom projelerinin çıktılarının başlarında genom üzerindeki rezolüsyonu yüksek (birkaç yüz bazda bir) moleküler belirteçlerin (Tek Nükleotid Polimorfizmler/Single Nucleotide Polymorphism-SNP) dökümünün çıkarılması gelmektedir. Moleküler belirteçler ile hastalıklardan sorumlu genlerin keşfedilip tanımlanması 1980’li yıllardan bu yana uygulanan bir yaklaşımdır. Ancak gerek bu öncül moleküler belirteçlerinin rezolüsyonlarının (birkaç bin bazda bir) düşük olması, gerek bu belirteçleri tarama için kullanılan teknolojilerin işlem hacminin düşüklüğü ile uygulamalarındaki teknik sıkıntı ve zorluklar “gen avlama” çalışmalarının hızında ivmeyi düşürmüştür.

Genom projesinin büyük ölçüde tamamlanıp, bu projenin paralelinde yeni genetik teknolojilerin gelişmesi (Mikrodizin, Yeni Nesil Sekanslama, Biyoinformatik Analiz) “gen avlama” çalışmalarına büyük bir ivme kazandırmıştır. Artan bilgi ve gelişen teknoloji sayesinde aynı anda yüzbinler hatta milyonlarca moleküler belirteç taranabilmekte, bu analiz paralel olarak yüzler hatta binlerce kişi için gerçekleştirilebilmekte ve analiz sonuçları biyoinformatik araçlarla yorumlanabilmektedir.

İşte bu kapsamda multigenik, multistep ve multifaktöryel bir hastalık olan kanserde de karsinogenezden sorumlu genlerin tanımlanmasında bir çığır açılmıştır. Günümüzde karsinogenezde rol alan, kanserin ortaya çıkmasında yatkınlık nedeni olan genlerin tanımlanmasında genomda rezolüsyonu en yüksek moleküler belirteç olan SNP'lerin mikrodizin teknolojisi ile taranabilmesi genom boyu asosiyasyon ve kopya sayısı değişim analizlerinin yüksek işlem hacmi ile gerçekleştirilebilmesini mümkün kılmıştır.

Bu genomik analizler ile kanserin erken teşhisi ve tedavisinde etkin olan yeni moleküler biyobelirteçlerin tanımlanması gerçekleştirilebilmektedir. 2000'li yıllardan itibaren özellikle gelişmiş ülkelerde yoğun olarak yürütülmekte olan bu tip analizlerin ülkemiz popülasyonu için öncelikli olan kanserlerde geniş hasta gruplarında uygulanıp, popülasyonumuz için doğrulanmış biyobelirteçlerin tanımlanması ülkemiz insanın sağlığı, ülkemiz ekonomisi ve ülkemizin ilerlemesi için büyük önem arz etmektedir.

Sunulan bu tez çalışmasında sporadik kolorektal kanserin ortaya çıkması ve/veya ilerlemesinde etken olan genlerin aile temelli asosiyasyon analizleri ile belirlenmesi hedeflenmiştir.

2. KURAMSAL TEMELLER

2.1. Kanser ve Moleküler Biyolojisi

Kanser dünyada en önemli insan sağlığı problemlerinden birisidir. Kanseri hücrelerimizin çoğalmasını (proliferasyonunu), farklılaşmasını (differansiyasyonunu) ve ölümünü kontrol eden genlerin hastalığı olarak tanımlamak mümkündür.

Kompleks bir hastalık olan kanser hem çevresel hem de kalıtsal faktörlerden kaynaklanmaktadır (Lichtenstein et al. 2000). Başka bir deyişle kanser multistep ve multifaktöryel bir hastalıktır. Günümüzde elimizde olan verilere göre, kanser gelişiminin çok kompleks bir süreç olduğu açıktır. Diyet, çevresel faktörler, kalıtılan mutasyonlar ve somatik mutasyonlar gibi faktörler kanserin açığa çıkma nedenleri arasında yer almaktadır.

Normal hücrelerin kanser hücrelerine dönüşmesi genomun hasara uğramasına neden olan birçok mutasyonu gerektirmektedir. Genomdaki bu hasar, DNA replikasyonunda meydana gelen hataların sonucu olabileceği gibi, metabolizma sırasında meydana gelen serbest radikallerin bazı DNA bazlarına saldırması gibi nedenlerden de kaynaklanabilmektedir. DNA hasarı, ionize radyasyon, UV radyasyon ve kimyasal karsinojen gibi ajanların etkileşimi ile de gerçekleşebilmektedir (Bertram 2001). DNA'nın birçok hasarı tamir etme mekanizması vardır. Ancak DNA tamir mekanizmasının işlevsiz veya yetersiz kaldığı durumlarda onarılamayan bu hasarlar, mutasyon adı verilen DNA dizisindeki kalıcı değişiklikleri meydana getirmektedir.

Hanahan ve Weinberg kanser hücrelerini karakterize eden altı özelliği, büyüme sinyallerinde kendi kendine yetebilme, büyümeyi durdurabilecek sinyallere duyarsızlık, programlanmış hücre ölümünden (apoptozis) kaçabilme, sınırsız bölünme, çoğalma potansiyeline sahip olma, anjiogenezi sürekli destekleyebilme, doku invazyonu ve metastaz yapabilme olarak tanımlamışlardır (Oláh 2005).

Yukarıdaki paragrafta özetlenen kanser hücrelerinin altı özelliğinin ortaya çıkmasında ön planda olan iki gen grubu onkogenler ve tümör baskılayıcı genlerdir.

Bir hücrenin doğal hayat sürecinde proto-onkogenler normal hücrel faaliyetleri idare eden genlerdir. Bu genler herhangi bir etkenle mutasyona uğradıklarında "işlev kazanımı" geçirerek onkogene dönüşmekte ve hücrelerin kontrolsüz çoğalmasına neden olarak ilerleyen süreçte hücrelerin kanserleşmesine neden olmaktadır (Pitot 1993; Oláh 2005).

Proto-onkogenler, nokta mutasyonları, kromozomal translokasyonlar, yeniden düzenlenmeler, gen amplifikasyonları ve hipometilasyonları gibi birçok farklı yolla onkogenlere dönüşüp kanser oluşturma özelliği kazanırlar. Onkogenler genel olarak büyüme faktörleri ve reseptörleri, transkripsiyon faktörleri, G proteinleri, serin/treonin kinazlar, reseptör olmayan tirozin kinazlar, sitoplazmik proteinler benzeri gen gruplarında yer almaktadır (Martinez et al. 2003).

Karsinogenezde önemli tetik rolünü üstlenen diğer bir gen olan tümör baskılayıcı genler hücrede proliferasyonu baskılamakta ve apoptozisin (programlı hücre ölümünün) indüklenmesinde görev alırlar. Bu genlerin kaybı veya inaktive olması sonucu hücre

proliferasyonunda artma, hücresel ölüm sürecinde aksamalar gözlenir. Bu nedenle tümör baskılayıcı genlerde meydana gelen mutasyonlar genel olarak “işlev kaybı” sonucunu ortaya çıkarır (Kleinsmith 2006; Martinez et al.2003).

Tümör baskılayıcı genler onkogenlerin tersine hücre büyümesini doğrudan veya dolaylı olarak inhibe edebilirler.

Hücre büyümesini doğrudan inhibe eden veya hücre ölümünü doğrudan yönlendiren genler koruyucu (gatekeeper) genler olarak ifade edilmektedirler (Kinzler et al. 1997). Diğer bir deyişle, koruyucu genler direk olarak tümör gelişimi üzerine sınırlayıcı etkileri bulunan genlerdir. Tümör oluşumu için bu genin her iki kopyasının da inaktive olması gerekmektedir (Knudson 1971; Martinez et al. 2003). Proliferasyonu doğrudan baskılamayan fakat genetik dengesizliği (instabilite) engellemekle sorumlu olan tümör baskılayıcı genler ise bakıcı (caretaker) genler olarak tanımlanmaktadır. Bakıcı genlerdeki ortaya çıkacak olan herhangi bir işlev kaybı mutasyonu genomu global olarak mutasyonlara açık hale getirmektedir.Yapılan çalışmalar birçok sporadik tümörde bu genlerin mutasyon sonucu inaktive olduğunu göstermektedir (Vogelstein et al. 1988).

Meme ve yumurtalık kanserinin gelişmesinde önemli rol oynayan BRCA1 ve BRCA2 genleri bakıcı (caretaker) genlere örnek olarak verilebilir. Bu genler DNA'nın çift zincir kırıklarının tamir edilmesinde görev olmaktadır. Bu genlerdeki oluşacak bir mutasyon olduğunda DNA çift zincir kırıklarının onarımı engellenecek ve bu nedenle de kanser gelişimi olası olacaktır (Zhang et al 2005).

İlk olarak karakterize edilen tümör baskılayıcı gen, kalıtsal retinoblastoma hastalığından sorumlu olan RB genidir. RB geni koruyucu (gatekeeper) bir gen olarak tanımlanmaktadır. Bu gen tarafından kodlanan Rb proteini, büyüme faktörlerinin yokluğunda hücre farklılaşmasını baskılayan bir proteindir. Bu gende meydana gelen bir mutasyon sonucu retinada tümör oluşmaktadır (Martinez et al. 2003; Ashktorab et al. 2010).

RB geninin tanımlanmasından sonra *P53*, *APC*, *PTEN*, *TGF-β* gibi birçok tümör baskılayıcı gen belirlenmiştir. Pek çok kanser türünün oluşumunda rolü olduğu bilinen *P53* bu tümör baskılayıcı genler içinde ön plana çıkan gen olmuştur (Kleinsmith 2006). *P53* geni genomda herhangi bir hasar meydana geldiği zaman hücre döngüsünü durdurup bu

hasarın onarılmasını sağlar. Ancak hasar onarılamayacak derecede ise bu durumda hücreyi apoptozise yönlendirir. Bu yüzden *P53* geni “genomun gardiyanı” olarak adlandırılmıştır (Martinez et al. 2003).

Normal hücrel işleyişte hücrenin yaşam süresi dolduğunda veya onarılamayacak düzeyde bir DNA hasarı meydana geldiğinde hücre programlı bir şekilde ölür. Bu mekanizma apoptozis olarak adlandırılır. Bu mekanizmanın herhangi bir nedenle bozulması sonucu başta kanser olmak üzere çeşitli hastalıklara neden olmaktadır (Martinez et al. 2003). p53 apoptozis sürecinin sağlıklı olarak işlemesi için gerekli olan temel bir genidir.

p53 ve pRB gibi tümör baskılayıcı genler tarafından kontrol edilen diğer bir mekanizma ise hücrel yaşlanma (senesens) olarak tanımlanmaktadır. Senesens, DNA hasarı, onkogenik uyarılar ve mitojenik sinyaller ile indüklenebilen bir mekanizma olup hücrenin proliferasyon yeteneğinin geri dönüşümsüz olarak kaybedilmesi olayıdır. Kanser hücrelerinin senesensden kaçtığı düşünülmektedir. Bu nedenle senesens potansiyel bir anti-kanser mekanizma olarak tanımlanabilmektedir (Campisi 2005).

Özetle, genetik bir hastalık olan kanser birçok faktörün etkisi altında birçok basamaktan geçerek ortaya çıkmaktadır. Bu basamakların gerçekleşmesinde etken olan moleküler değişiklikler binlerce araştırmanın konusu olmuştur (Takayama et al. 2006). Bu araştırmalarla kanserin ortaya çıkış ve gelişim mekanizmasına dair önemli bilgiler açığa çıkarılmış olmasına rağmen 2010 yılında, 1,529,560 kişinin (789,620 erkek ve 739,940 kadın) kanser tanısı alacağı ve aynı yıl 569,490 kişinin kanserden öleceği tahmin edilmektedir. Dünya sağlık örgütü, 2030 yılında dünya çapında 12 milyon kişinin kanser nedeni ile hayatını kaybedeceğini bildirmektedir. Kanser nedenli ölümlerin başında akciğer, mide, kolorektal, karaciğer ve meme kanseri gelmektedir ve bu ölümlerin %70’inden fazlası düşük ve orta gelirli ülkelerde gerçekleşmektedir. Kolorektal kanser ise dünya çapında yılda yaklaşık 639,000 kişinin ölümüne neden olmaktadır (Altekruse et al. 2010).

2.2. Kolorektal Kanser

Kolorektal kanser, sanayileşmiş ülkelerde hem erkek hem de kadınlarda kanser nedenli ölümlerde en sık görülen nedenlerden biridir (Garcea et al. 2003). Bu kanser tipinin yıllık insidansı yaklaşık olarak 100.000 de 50 yeni olgudur. Kolorektal kanserin ölüm oranı gelişmiş tedavi, etkili görüntüleme ve izlemeye bağlı olarak azalmasına rağmen prevalansı sabit bir biçimde giderek artmaktadır (Naccarati et al. 2007).

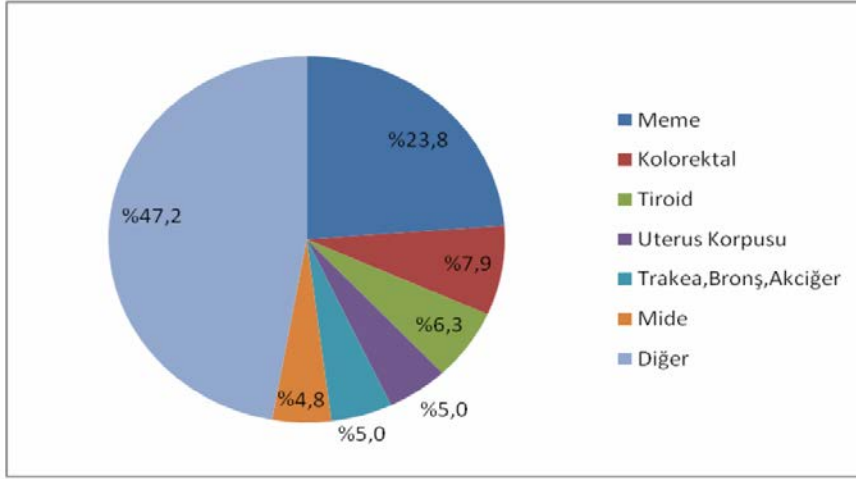
Dünyada 2003-2007 yılları arasında kolon ve rektum kanseri teşhisi alma yaşı ortalama 70 olarak bildirilmiştir. Kolon ve rektum kanseri olgularının ortalama %0.1'i 20 yaşın altında, %1.1'i 23-34 yaşları arasında, %3.8'ü 35-44 yaşları arasında, %12.4'ü 45-54 yaşları arasında, %19.2'si 55-64 yaşları arasında, %24.4'ü 65-74 yaşları arasında, %26.8'i 75-84 yaşları arasında ve %12.2'si 85 yaş üstünde teşhis edilmektedir.

Ocak 2007'de açıklanan verilere göre, Amerika'daki yaklaşık olarak 540,636 erkek ve 571,857 kadının kolon ve rektum kanser öyküsüne sahip olduğu bildirilmiştir (Altekruse et al. 2010).

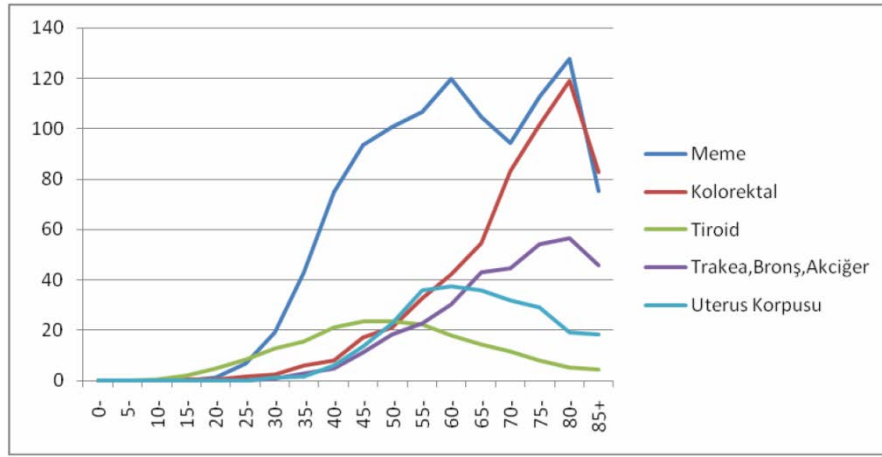
Türkiye'de Sağlık Bakanlığının Kansерle Savaş Dairesi Başkanlığının verilerine göre tüm kanserlerin görülme sıklığı ülkemizde giderek artmaktadır. 2004-2006 yılları için 41,438 erkek, 27,709 kadın olmak üzere toplam 69,147 kanser vakası değerlendirilerek elde edilen verilere göre, kadınlarda 2004 yılında tüm kanserler için kaba hız yüz binde 152.7, 2005 yılında 164.6, 2006 yılında ise 175.9 olarak belirlenmiştir.

Kadınlarda 2004, 2005 ve 2006 yıllarında en sık görülen ilk üç kanser meme, kolorektal ve tiroid olarak sıralanmaktadır (Şekil 2.1). Kadınlarda kanserin görülme sıklığının özellikle 43-45 yaşından sonra arttığı gösterilmiştir (Şekil 2.2).

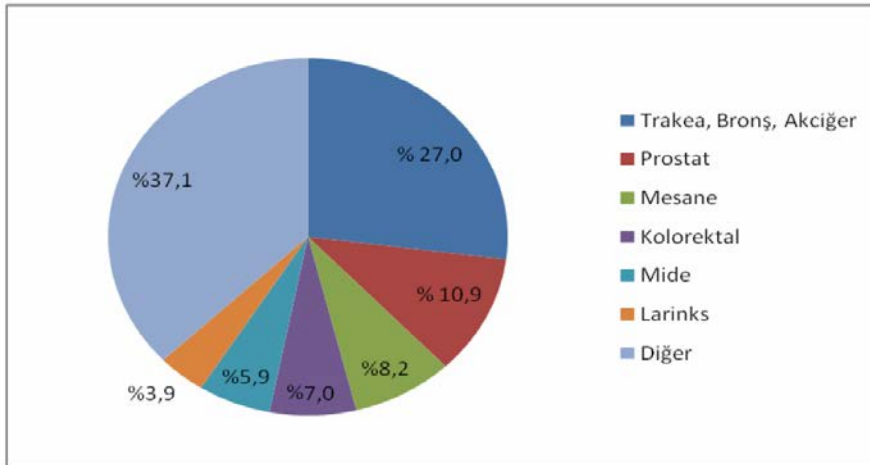
Erkeklerde ise 2004 yılında tüm kanserler için kaba hız yüz binde 229.4, 2005 yılında 246, 2006 yılında ise 261.6 olarak hesaplanmıştır. Erkeklerde en sık görülen kanserler akciğer, prostat, mesane, kolorektal ve mide kanserleri olarak sıralanmaktadır (Şekil 2.3). Ayrıca en sık görülen beş kanserin sıralamasının üç yılda da değişmediği görülmektedir. Erkeklerde kanserin ortaya çıkışı yaşa göre izlendiğinde özellikle 55-60 yaşlarında arttığı belirlenmiştir (Şekil 2.4).



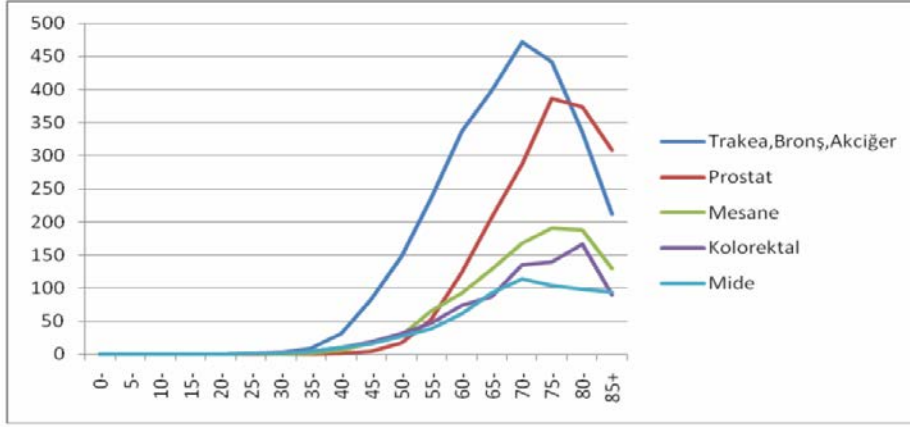
Şekil 2.1: Kadınlarda yerleşim yerlerine göre en sık görülen kanserler (Türkiye 2004-2006)



Şekil 2.2: Kadınlarda yaşa göre en sık görülen kanserler (Türkiye 2004-2006)



Şekil 2.3: Erkeklerde yerleşim yerlerine göre en sık görülen kanserler (Türkiye 2004-2006)



Şekil 2.4: Erkeklerde yaşa göre en sık görülen kanserler (Türkiye 2004-2006)

Elde edilen bu bulgulara göre, ülkemizde erkeklerde akciğer, mesane ve larinks gibi sigara kullanımı ile ilişkili kanserler ilk sıralarda yer almaktadır. Kadınlarda ise meme kanseri en sık görülen kanserdir. Kolorektal kanserler, hem erkeklerde hem de kadınlarda üst sıralarda bulunmaktadır.

Bir bireyin yaşamı boyunca kolorektal kanser geliştirme riski %5 olarak tespit edilmiştir. Birinci derece akrabalarında kolon kanseri veya adenomatöz polip öyküsü olan bireylerin ise riski 2-3 kat daha fazladır (Garcea et al. 2003).

Kolorektal kanserin gelişim nedenleri arasında yaş, beslenme, kalıtım, önceki kanser öyküsü, polip, virus (human papilloma virus), dışarıdan alınan hormonlar, alkol, düşük selenyum ve çevresel faktörler sayılabilmektedir. Birçok epidemiyolojik çalışma kolorektal kanserde diyetin ve yaşam tarzının rolünün altını çizmektedir. Yağ, kırmızı et, alkol ve sigara tüketiminin pozitif korelasyon gösterdiği, sebze ve lifli gıdaların tüketiminin ise ters korelasyon gösterdiği rapor edilmiştir. Ayrıca yüksek fiziksel aktivite, kadınlarda menopaz öncesi hormon tedavisi, aspirin gibi antiinflamatuvarların düzenli kullanımı gibi faktörlerin ise koruyucu etkisi olduğu bildirilmiştir.

2.2.1. Kolorektal Kanserin Moleküler Biyolojisi

Geçtiğimiz 40 yılda kanserin moleküler temelini anlaşılmada büyük gelişmeler kaydedilmiştir (Martinez et al.2003). Moleküler biyolojisinin aydınlatılması için çok sayıda çalışma yapılmış kanserlerden biri olan kolorektal kanserin, genetik olarak heterojen ve karmaşık bir hastalık olduğu gösterilmiştir (Takayama et al. 2006).

Kolorektal kanserin başlaması ve ilerlemesi de diğer kanserler gibi, anahtar düzenleyici genlerin fonksiyonlarındaki değişikliklerin birikmesi ve genetik dengesizlik ile ilgilidir (Ashktorab et al. 2010).

Geçmişte kolorektal kanserde iki temel genetik dengesizlik tanımlanmıştır. Bunlardan sayısal veya yapısal kromozomal anormalliklerin (anöploidi) artması yolu ile oluşan kromozomal dengesizlik (chromosomal instability, CIN) en yaygın görülenidir. Diğer ise DNA yanlış eşleşme (mismatch) tamirindeki bozuklukların neden olduğu mikrosatellit dengesizliğidir (microsatellite instability, MIN).

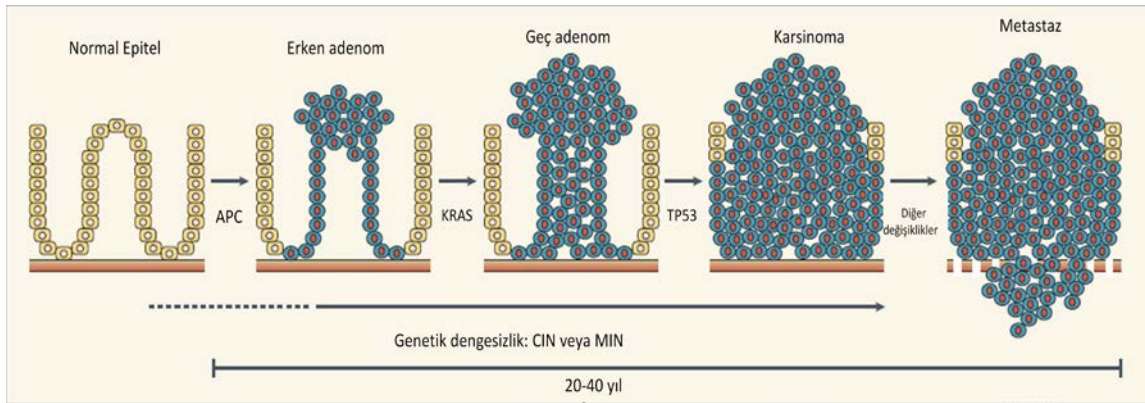
Yakın zamanda moleküler biyolojideki gelişmeler kolorektal kanserin bu iki yolakla incelenmesinin doğru olmadığını, sürecin daha karmaşık olduğunu göstermiştir. Günümüzde kolorektal kanser gelişimi için genetik dengesizliğin üç ana formu tanımlanmaktadır. Bunlardan biri kolorektal kanser olgularının yaklaşık %13'üne neden olan mikrosatellit dengesizliğidir. HNPCC sendromlarının %90'ında rol oynayan mikrosatellit dengesizliği, sporadik kolorektal kanserlerin %20'sinde görülmektedir (Mutch 2007). Mikrosatellit dengesizliği hMSH2, hMLH1, hPMS1, hPMS2 ve hMSH6'yı içeren DNA yanlış eşleşme tamir enzimlerinde tanımlanmıştır. Ayrıca, mikrosatellit dengesizliği görülen tümörlerde, TGF BR11, BAX, IGFIIR, hMSH6, hMSH3, PTEN, E2F-4 gibi mikrosatellit dengesizlik yolağındaki genlerin mikrosatellitlerindeki mutasyonlar da tanımlanmıştır.

Kolorektal kanser tümörlerinin % 40'ında etkili olan diğer genetik dengesizlik formu epigenetik değişikliklerle karakterizedir. Epigenetik mekanizma ile kodlayan DNA dizisinin 5' ucunda yer alan CpG zengin bölgeler hipermetilasyona giderek bazı genlerin sessizleşmesine neden olmaktadır. hMLH1, P16INK4A, MGMT, ER, APC, COX2 genlerinin insan kolorektal kanserlerinde hipermetile olarak sessizleştiği rapor edilmiştir (Takayama et al. 2006; Worthley et al. 2007).

Kolorektal kanserlerin geri kalan % 47'sinde ise kromozomların büyük parçalarının kazanımı veya kaybına neden olan kromozomal dengesizlik yolağı rol oynamaktadır (Ashktorab et al. 2010) Sporadik kolorektal kanserlerinin yaklaşık olarak %80'inde etkin olan kromozomal dengesizlik yolağında tümör baskılayıcı genlerden APC, P53, SMAD2, SMAD4, DCC ve onkogenlerden ise K-ras, β -katenin rol oynamaktadır (Takayama et al. 2006; Cheah 2009).

Kolorektal kanser, kanserin çok basamaklı karakterinin en iyi ve ayrıntılı olarak çalışıldığı kanser modeli olmuştur.

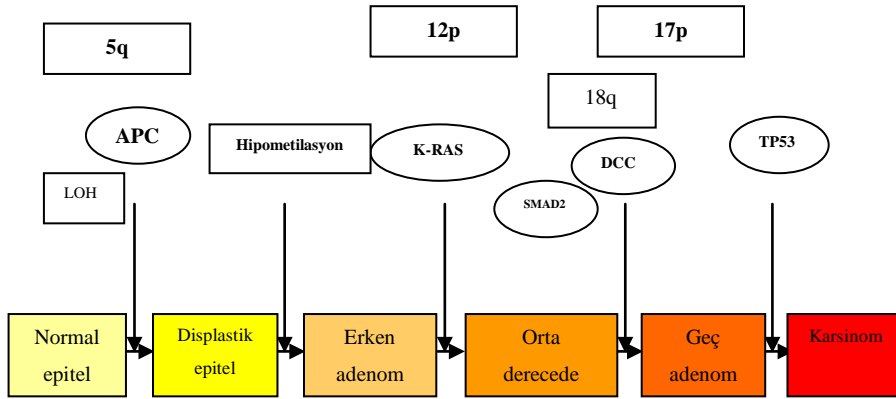
Kolorektal kanserin, *KRAS* gibi onkogen ve *APC*, *TP53* gibi tümör baskılayıcı genlerdeki birçok mutasyonun sonucunda 20 ila 40 yılda geliştiği bilinmektedir (Şekil 2.5). Bu mutasyonlar tümörde karakteristik adenom karsinom sekansı ile birlikte ortaya çıkmaktadır. Bölünen kanser hücrelerinde mutasyon oranını hızlandıran genetik dengesizliğin, kolorektal karsinogenez sürecinde meydana geldiği düşünülmektedir. Tümörlerin küçük çapta olması ve genç olan lezyonlarda kromozomal değişikliklerin henüz meydana gelmemesinden dolayı erken tümörlerde kromozomal dengesizliğin (CIN) bulunduğu dair deneysel kanıtlar sınırlıdır. Bununla birlikte CIN'in tümörjenik süreçte erken meydana geldiğine dair güçlü kanıt bulunmaktadır. CIN'in, tümörigenezde ilk olay olup olmadığı, ve bu yüzden APC mutasyonundan önce gelip gelmediği de hala birçok tartışmanın konusunu oluşturmaktadır.



Şekil 2.5: Kolorektal kanserde adenom karsinom sekansı

Kolorektal karsinogenezin son basamağında yer alan invazyon ve metastaz süreçlerinde proteolizis, adezyon, anjiogenez ve hücre büyümesi ile ilişkilendirilmiş birçok gen tanımlanmıştır. Bu genler arasında MMP7, MMP2, MMP1, MMP8, MMP13, MMP3, TIMP1, uPAR, integrinler, kaderinler, CD44, CEA, VEGF, PD-ECGF, TRAIL-R, CXCR4, Drg-1, c-Met bulunmaktadır (Takayama et al. 2006).

Çok basamaklı karsinogeneze iyi bir örnek teşkil eden kolorektal kanserde geleneksel olarak adlandırabilecek kromozomal dengesizlik yolağında ilk mutasyon normal epitel hücresinde meydana gelir (Takayama et al. 2006; Cheah 2009). Normal epitel hücresinde *APC* geninin kaybı ile başlayan süreci hücre çoğalmasındaki artış ve DNA hipometilasyonu takip etmektedir. *APC* geninin her iki allelinin kaybı ya da inaktivasyonu erken adenom oluşumuna neden olmaktadır. Erken adenom hücrelerindeki 12p, 18q, 17p kromozom bölgelerinde daha sonra oluşan mutasyonlar kolon kanserine neden olacak malignant transformasyona yol açmaktadırlar. Bu genetik olayların ilerlemesi ve fenotipik korelasyonları ilk defa Vogelstein ve ark tarafından tanımlanmıştır. (Vogelstein et al. 1988). Birbirine bağlı olarak gelişen kolorektal kanser adenom karsinom sekansı şekil 2.6 da gösterilmiştir.



Şekil 2.6: Kolorektal karsinogenezin çok basamaklı modeli (vogelogram)

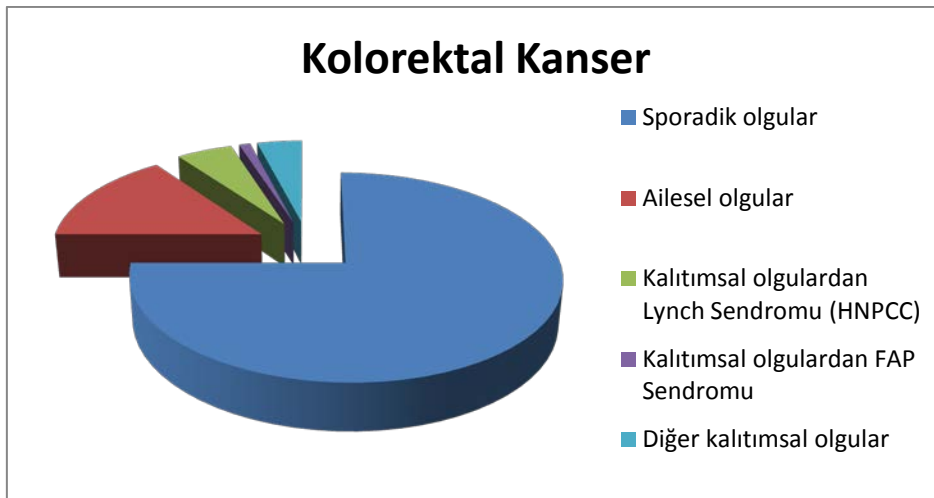
Kanser oluşum mekanizması en iyi aydınlatılmış kanser tiplerinin başında yeralan kolorektal kanserin moleküler biyolojisinin yanısıra kalıtımı ile ilgili de pek çok çalışma bulunmaktadır.

2.2.2. Kolorektal Kanser Kalıtımı

Geçtiğimiz yıllarda, kanserin genetik temelini aydınlatılması için hem kalıtsal hem somatik mutasyonların tanımlanması amacıyla güçlü moleküler biyolojik tekniklerle birçok çalışma yapılmıştır.

Kolorektal kanser, farklı genetik ve çevresel faktörlerin birlikte etki etmesiyle gelişen kompleks bir hastalıktır. Kolorektal kanserin yaklaşık %15'i ailesel, %10'u kalıtımsal ve %75 ise sporadik olarak ortaya çıkmaktadır (Lynch et al. 2006).

Bütün kolorektal kanser olgularının % 2-5'ini oluşturan Lynch sendromu (herediter nonpolyposis colorectal cancer, HNPCC) en yaygın olarak görülen kalıtımsal sendromdur. Yine tüm kolorektal kanser olgularının yaklaşık % 1'inden sorumlu olan ailesel adenomatöz polip (familial adenomatous polyposis, FAP) sendromu ise ikinci en yaygın görülen kalıtımsal kolorektal kanser sendromudur. HNPCC ve FAP olgularından başka kalıtımsal kolorektal kanser olgularının arasında Peutz-Jeghers sendromu (PJS), ailesel juvenil polipozis (familial juvenile polyposis, FJP), Cowden Hastalığı (Cowden Disease, CD) ve Bannayan-Ruvalcaba-Riley Sendromu (BRRS) yer almaktadır (Lynch et al. 2009) (Şekil 2.7).



Şekil 2.7: Kolorektal kanserin heterojenitesi

Sporadik kolorektal kanser gelişimi genetik ve spesifik çevresel/yaşam tarzını risk faktörlerinin kompleks etkileşimini farklı derecelerde iç içe geçmesini gerektirmektedir. Sporadik kolorektal kanser olgularında ilgili aday genlerin sayısı ile yüksek frekanslı düşük penetranslı aleller geniş kapsamlıdır (Naccarati et al. 2007).

Ortaya çıkış mekanizmalarına göre farklılık gösteren kanser tiplerinin kendilerine ait birçok ayırıcı özellikleri bulunmaktadır (Burt et al. 2006).

Kalıtımsal kanser tipinin birçok karakteristiği vardır.

- a. Belli kanser tipleri otozomal dominant geçişli görünümündedir.
- b. Bireyde birden fazla primer kanser oluşumu gözlenmektedir.
- c. Kalıtımsal kanserler bilateral (meme, böbrek, yumurtalık kanserleri gibi) ve çok odaklı kanserlerdir.
- d. Kalıtımsal kanserlerde kanserli bireyin birinci derece akrabasının aynı mutasyonu taşıma riski %50 dir. Eksik penetrans ve değişken ifade düzeyi nedenleri ile ailesel mutasyonu taşıyanlar kansere yakalanmayabilir ve akrabalar arasında kansere yakalanma yaşında değişkenlikler görülebilir. Ailesel mutasyonu taşımayanlar kanser için genel popülasyonunki kadar risk taşırlar.

Ailesel kanser tipi karakteristikleri arasında,

- a. Ailede belirli bir kanserde istatistiksel olarak beklenenden daha fazla olgu görülmesi fakat spesifik bir kalıtım paterni olmaması,
- b. Sporadik olguların kümelenmesi sonucu ortaya çıkabilmesi,
- c. Aynı genetik altyapı, benzer çevre ve/veya yaşam tarzı faktörleri sonucu ortaya çıkabilmesi ve
- d. Çoğunlukla kalıtımsal kanserin klasik özelliklerini göstermemesi sayılabilir.

Sporadik kanser tipinde ise ailedeki kanserler kalıtımsal olmayan nedenler yüzünden ortaya çıkar. Kanser tipik bir yaşta ortaya çıkmaktadır. Ailede birden fazla olgu varsa belirli bir kalıtım paterni yoktur.

Kolorektal kanserin %75'i ise sporadik olgular olduđu ve bu olguların yukarıda da söylendiđi gibi belirli bir kalıtım paterni olmaması nedeni ile genetik yatkınlığa neden olan genlerin tanımlanmasında çeşitli yaklaşımlar kullanılmaktadır.

2.3. Hastalık Yatkınlık Genlerinin Lokalizasyonun Tespiti

Kanserlerin bir kısmının kalıtsal özellik göstermesi, bu grup kanserlere neden olan genlerin tanımlanmasında genetik haritalama çalışmalarına başvurulmasına neden olmuştur.

Genlerin lokalizasyonlarının bulunmasında sitogenetik analiz ve genetik haritalama metodları kullanılmaktadır.

Genetik haritalama özellikle kalıtsal hastalıklara yol açan genlerin kromozomlar üzerindeki lokalizasyonlarının saptanmasında yaygın olarak kullanılmaktadır. Lokalizasyonu aranan gen ile lokalizasyonu bilinen bir genetik, moleküler belirtecin (marker) kuşaklar arasında birlikte kalıtılmasının test edilmesi esasına dayanan genetik haritalamada moleküler biyolojik yöntemler ve istatistiksel analizler kullanılmaktadır.

Genetik haritalama da nirengi noktası olarak kullanılan genetik belirteçler, kromozomda ya da DNA dizisi üzerinde tanımlanabilen (mikrosatellitler gibi) ve hastalık genlerinin tespitinde ve bağlantı analizinde kullanılabilen genomik varyasyonlardır.

Mayoz bölünme sırasında kromozomlar crossing over olarak adlandırılan olay esnasında karşılıklı parça değişimine uğrarlar. Crossing overda birbirine yakın genler çoğunlukla bir arada giderken birbirinden uzak olan genler ise bağımsız düzenlenme kuralına göre ayrı giderler. Yukarıdaki bilgiler ışığında geliştirilen aranan (hastalıkla ilişkili olabileceđi düşünölen) gen, lokalizasyonu bilinen moleküler belirtece çok yakınsa birlikte kalıtılacaklarından, aranan genin yeri de bulunmuş olacaktır gerçeđi ile araştırmacılar, hastalıklarla ilgili hedef genleri bulabilmektedirler.

Hastalıklarda yatkınlık geninin lokalizasyonunun tespiti için farklı fakat ilgili iki esas yaklaşım kullanılmaktadır. Bunlar bağlantı (linkage) analizleri ve ilişkilendirme (asosiyasyon) çalışmalarıdır.

Hastalık geni lokalizasyonu tanımlanmasında kullanılan yaklaşımlardan biri olan bağlantı analizlerinde genomda belirli bir bölgenin kalıtımı ile hastalığa yatkın ailelerde hastalığın kalıtımı arasındaki istatistiksel ilişki ile tanımlanır.

Hastalık genlerinin haritalanmasında bağlantı analizlerine bir alternatif asosiyasyon çalışmalarıdır. Asosiyasyon çalışmaları; kalıtım modeli bilinmeyen, genetik etkenlerin işe karıştığı kompleks hastalıkların incelenmesinde ve bu hastalıklarla ilişkili olabilecek yeni aday genlerin lokalize edilmesinde kullanılmaktadır. Bu çerçevede asosiyasyon çalışmalarında, genetik varyasyonun, hastalıklı bireydeki frekansı ile kontrol bireydeki frekansı karşılaştırılmaktadır (Pharoah et al. 2004).

2.3.1. Moleküler Belirteçler ve Genlerin Lokalizasyonlarının Tespiti

Moleküler belirteçler genomda lokalizasyonu bilinen, mutasyon veya polimorfizm kapsamında meydana gelen varyasyonlardır. Sıklıkla kullanılan moleküler belirteçler arasında, mikrosatellitler, restriksiyon parça uzunluk polimorfizmleri (Restriction Fragment Length Polymorphism, RFLP), değişken sayıda ardarda tekrarlayan DNA dizileri (Variable Number of Tandem Repeat, VNTR), kısa tekrarlayan DNA dizileri (short tandem repeat, STR), tek nükleotid polimorfizmleri (Single-Nucleotide Polymorphism, SNP) ve kopya sayısı varyasyonları (copy number variation, CNV) sayılabilirler.

Organizmaların genetik haritalarının oluşturulmasında ve filogenetik çalışmalarda kullanılmakta olan moleküler belirteçler, bir ailenin nesiller boyunca kalıtılan özelliklerin izleme olanağı vermektedir. Ayrıca moleküler belirteçler bir hastalıkla ilgili bu hastalığın genetik nedeni arasındaki ilişkiyi açıklamada da kullanılmaktadır. Kromozom üzerinde birbirine yakın olan DNA dizileri beraber kalıtıldıklarından dolayı moleküler belirteçler, asosiyasyon çalışmalarında, hastalıkla daha önce ilişkilendirilmemiş bir geni lokalize etmek için kullanılabilir potansiyeline sahiptir. Araştırmacılar kompleks hastalıklarda, hedef genleri aramak için genellikle ilgili genlere yakın olan ve birlikte kalıtılan genetik belirteçleri tanımlamaktadırlar. Keza, tek genin sorumlu olduğu kalıtsal hastalıklarda sorumlu genin lokalizasyonunda kullanılan bağlantı analizlerinde yine moleküler belirteçlerden faydalanılmaktadır.

Genomun haritasının çıkarılması, genlerin hastalıklarla ilişkilendirilmesi ve lokalizasyonlarının belirlenmesi gibi amaçlarla yukarıda adı geçen (mikrosatellitler, RFLP, VNTR, STR) DNA temelli moleküler belirteçler uzun yıllardır kullanılmaktadır. Ancak bu belirteçlerin gerek rezolüsyonlarının sınırlı olması, gerek yüksek işlem hacimli analizlere uygun olmamaları nedeniyle yeni nesil moleküler belirteçlerin keşfi için çalışmalar devam etmiştir. Genomda çok yaygın olarak bulunan ve rezolüsyonları oldukça yüksek olan SNP'ler keşiflerini takiben çok değerli belirteçler haline gelmişlerdir.

Genomda 5 Mb ve daha büyük boyuttaki kayıp ve kazanımlar 1959'dan bu yana klasik sitogenetik analizlerle tanımlanmıştır. Ancak pratik olarak 13.20 Mb'ın altındaki dengesizliklerin tespit edilmesi çok da kolay olmamaktadır. Moleküler sitogenetikteki ilerlemelerle birlikte submikroskopik genomik değişkenler de tespit edilir hale gelmiş ve genom boyu tarama tekniklerinin de gelişimi ile birlikte kopya sayısı varyasyonları (CNV) adı verilen yeni varyasyon tipi 6. nesil genetik moleküler belirteç olarak belirlenmiştir (Rodriguez-Revenge et al. 2007; Ewens et al. 2008).

İnsan Genom Projesi, HapMap Projesi, 1000 Genom Projeleri ışığında tespit edilen yeni nesil moleküler belirteçler ile gen tespiti genom ebadında yürütülmekte ve bu kapsamda genoma dağılmış CNV'ler ve SNP'ler, paralel olarak genotiplenebilmekte ve hastalık ile ilişkilendirilebilmektedir.

2.3.1.1. Tek Nükleotid Polimorfizmi

DNA dizisi üzerinde bireysel varyasyonlar görülmektedir. DNA dizisini oluşturan adenin, guanin, sitozin, timin baz çiftinden birinin diğer baz çifti ile değişmesi tek nükleotid polimorfizmi (SNP) olarak adlandırılır. SNP'ler insan genomunda en sık bulunan genetik varyasyonlardır. Bir genetik varyasyonun polimorfizm olarak tanımlanabilmesi için bir popülasyondaki sıklığının en az %1 olması gerekmektedir.

İnsan genomunda genler tüm DNA dizisinin çok az bir kısmını kapsamaktadır ve genler arasındaki DNA dizisinin %90'dan fazlası belirli bir ürün kodlamamaktadır. İnsan genomunda SNP'ler her 1000 baz çiftinde bir bulunurlar ve birçok SNP protein yapısında ve fonksiyonunda değişikliğe neden olmaz (Miller et al. 2001). Eğer SNP'ler amino

asitlerde deęişikliğe neden oluyorsa non-synonymous veya missense (yanlış anlamlı) SNP olarak adlandırılırlar.

Promotor bölgedeki SNP'ler protein yapısına etkileri olmamakla birlikte transkripsiyon faktörlerinin bağlanmasını etkileyerek protein ekspresyonunu deęiştirebilmektedirler. Fakat çoęu SNP hem kodlamayan dizide bulunmaları hem de aynı amino asidi kodlayan varyantlar oluşturdıkları için fenotip üzerinde herhangi bir etkisi yoktur.

Kodlamayan dizideki SNP'lerden 5' veya 3' transle olmayan bölgede olanların gen ekspresyonunun transkripsiyon sonrası regülasyonunda kritik rol oynayabildięi bilinmektedir.

Aday gen analizinde SNP'ler seçilirken bu farklılıkların anlaşılması önemlidir. Genel olarak promoter bölgedeki SNP'ler ve yanlış anlamlı SNP'ler kodlamayan bölgedekilere göre muhtemelen daha önemlidir (Yende et al. 2006).

Son yıllarda yapılan çalışmalarla SNP'lerin tanımlanması ve genomda sık bulunmaları ile genomda kararlı bir şekilde dağılım göstermeleri nedenleri ile SNP'ler tercih edilen moleküler belirteç olmuşlardır. Yüksek işlem hacimli genotipleme yöntemlerinin gelişmesi ile bir defa da bir bireye ait milyonun üzerinde SNP genotiplendirilebilir ve bir manada ilgili bireyin genom haritası ortaya koyulabilir hale gelmiştir. Bu nedenle SNP (DNA) mikrodizinleri, hem bağlantı hem de asosiyasyon analizi çalışmalarında yaygın olarak kullanılmaktadırlar (Gupta et al. 2008).

2.3.1.2. Kopya Sayısı Varyasyonları

Kopya sayısı varyasyonları, popülasyonlarda farklı frekansta gözlenen, çeşitli büyüklüklerdeki DNA segmentlerinin delesyonu veya amplifikasyonlarıdır. Kopya sayısı varyasyonları, daha ziyade tekrarlayan DNA sekanslarının bulunduğu bölgelerde ortaya çıkma eğilimindedirler. Bazı kopya sayısı varyasyonları yaygındır ve nesilden nesile aktarılmaktadırlar. Bazıları ise de novo olarak yani yeni bir varyasyon olarak ortaya çıkmaktadırlar. Seyrek görülen SNP ler gibi seyrek görülen CNV'ler de zararlı olmaya daha yatkındır. Genom boyu kopya sayısı varyasyonları taramaları CNV'lerin daha önce kabul edilenin aksine, oldukça yaygın olduklarını göstermiştir. Yakın bir zamana kadar

yapısal varyasyonlar, kopya sayısı varyasyonlarının daha az doğrulukla belirlenebilmesi, biyolojik konfirmasyonunun oldukça maliyetli olması ve küçük kopya sayısı varyasyonlarının (100.000 bp den küçük) daha az güvenilir olarak belirlenebilmesi gibi nedenlerle SNP'ler kadar kapsamlı çalışılmamıştır. Fakat günümüzde, hızla gelişen teknolojilerle önemli sayıda kopya sayısı varyasyonu rapor edilmektedir (Cichon et al. 2009).

2.4. Tek Gen Hastalıkları ve Bağlantı Analizleri

Tek bir gendeki mutasyon sonucu ortaya çıkan hastalıklar tek gen hastalığı olarak ifade edilmektedir. Tek gendeki bu mutasyon hastalığın kalıtsal olarak gelecek nesillere aktarılma riskini oluşturmaktadır. Mendelyan kurallara göre kalıtılan tek gen hastalıkları otozomal dominant, otozomal resesif, X'e bağlı dominant, X'e bağlı resesif ve Y'ye bağlı kalıtım olarak gruplandırılabilir.

Tek gen hastalıklarının oluşmasında, modifiye edici genlerin ve/veya çevrenin katkıları da bulunmaktadır.

Tek gen hastalıklarına neden olan mutasyonların tanımlanmasında aile düzeyinde uygulanan bağlantı analizleri kullanılmaktadır. Bağlantı analizleri bir genetik belirtecin, ailelerde, kuşaklar boyunca hastalık ile birlikte kalıtılıp kalıtılmadığını gözlemlemeye dayanmaktadır.

Bağlantı kavramı, mayoz bölünme esnasında kromozomdaki iki lokasyon arasındaki olası rekombinasyon ile ilgilidir. İnsan genomu 22 otozomal, 1 tane de seks kromozomu olmak üzere 23 kromozomu kapsamaktadır. Sperm veya yumurta hücreleri genomun bir kopyasını (haploit), döllenmiş yumurta ise iki ebeveynden gelen birer kopya yani iki kopyasını (diploit) içermektedir. Bundan dolayı diploit bir hücre 22 homolog çift otozomal kromozom ve bir çift de seks (XX veya XY) kromozomundan oluşmaktadır. Haploit sperm ve yumurta hücrelerinin üretimi için genomun diploitten haploite indirgenmesi için mayoz bölünme süreci gerekmektedir. Mayoz esnasında maternal ve paternal kökenli homolog kromozomlar dört zincirli bir yapı olan bivalent oluşturmak üzere hizalanırlar. Homolog kromatit çiftleri kiazma oluşturmak üzere bağlanırlar. Mayotik rekombinasyon, iki kromozom arasında genetik materyalin değiş tokuşu bu kiazma bölgelerinde meydana

gelmektedir. Rekombinasyon frekansının ve dağılımının anlaşılması genetik bağlantı haritalarında oldukça büyük bir önem arz etmektedir. Rekombinasyonun frekansı morgon veya sentimorgan olmak üzere genetik uzaklığa çevirebilmektedir. 1 sentimorgan %1 rekombinasyona karşılık gelmektedir. Harita uzaklıklarının hesaplanması, tüm kromozom ve genom boyu genetik bağlantı haritalarının yapılmasına olanak tanımaktadır (Collins 2007).

Kısaca özetlemek gerekirse, mayoz bölünme sırasında homolog kromozomlar karşı karşıya gelerek birbirleri ile parça değişimine uğramaktadırlar. Bu parça değişimi olayına crossing over adı verilmektedir. Bu değişimler sonucunda oluşan yavru kromozomlar ebeveynlerinin kromozomları ile aynı genetik yapıda olabilecekleri gibi, farklı bir genetik yapıda, diğer bir deyişle rekombinant da olabilmektedirler.

Eğer aynı kromozom üzerindeki gen lokasyonları birbirine çok yakınsa crossing over gerçekleşmesi fiziksel olarak zor olmaktadır ve rekombinasyon olasılığı da düşmektedir. Bu nedenle aynı kromozomda yakın yerleşimli genler mayozda rekombinasyona uğramazlar ve bir sonraki kuşağa daima birlikte aktarılırlar. Diğer bir deyişle eğer iki lokus bağlıysa (bağlantılı kalıtım, linkage) nadiren rekombine olurlar ve belli aleller yavru kuşağa tek bir birim olarak aktarılırlar. Aynı kromozom üzerinde birbirine uzak olan ya da farklı kromozomlarda bulunan genlerin yavru kuşaklara beraber aktarımı ise yalnızca şansa bağlı olarak olmaktadır.

Bağlantı analizlerinde kullanılan tipik olarak genetik belirteçler, duplikasyon polimorfizmleri (mikrosatellitler) veya SNP'lerdir. Bağlantı analizlerinde, geleneksel olarak, bir lokusta 13-20 farklı alele sahip olabilen yaklaşık 400 mikrosatellit kullanılmaktadır. Son yıllarda yapılan bağlantı ve asosiyasyon analizlerinde yaygın olarak SNP'ler kullanılmaya başlanmıştır. SNP'ler de bir lokus için sadece 2 farklı alel olmasına karşın, mikrosatellit panellerine göre, SNP panellerinin daha yoğun olması daha büyük bir rezolüsyon sağlamaktadır. Bu rezolüsyon genomun çok daha ayrıntılı bir şekilde taranmasına ve aranılan genlerin daha net bir şekilde lokalize edilmelerine imkan vermektedir.

Bağlantı analizleri, çalışılmak istenen hastalığın kalıtım modeli mendelyan kalıtım gösteriyorsa, tek bir gendeki mutasyon bu hastalıktan sorumlu ise ve bu hastalık erken yaş hastalığı ise (aile bireylerine ulaşmak daha kolay olacağından) başarılı olmaktadır.

Ayrıca bağlantı analizleri bazı yaygın kanserlerde örneğin, meme ve yumurtalık kanserinde BRCA1 ve BRCA2, kolorektal kanserde MSH2 ve MLH1, melonomada CDKN2A gibi yüksek penetranslı genlerin haritalanmasında da kullanılmıştır (Pharoah et al. 2004). Fakat genel olarak bağlantı analizleri, kalıtım modeli net olmayan kanser, şizofreni, otizm, major depresyon, bipolar hastalıklar, otoimmün hastalıklar gibi hastalıklarda büyük oranda başarısız olmaktadır.

Bağlantı Analizinin İstatistiksel Temelleri

Bağlantı analizleri parametrik ve parametrik olmayan bağlantı analizleri olarak ikiye ayrılmaktadır.

Eğer hastalığın alel frekansı ve genetik kalıtım modeli doğru olarak belirlenebiliyorsa hastalık geninin tanımlanabilmesinde parametrik bağlantı analizleri güçlü bir yöntem olarak karşımıza çıkmaktadır. Buna karşılık, kalıtım modelinin yanlış belirlenmesi bağlantı analiz sonuçlarının gücünü düşürebilmekte ve yanlış sonuçların alınmasına neden olabilmektedir. Parametrik bağlantı analizlerinde istatistiksel değerlendirme, tahmini göreceli risk (Odds) oranının logaritması (LOD skor) hesaplanarak yapılmaktadır. LOD skor bağlantı analizi yapmak için kullanılan bir istatistik testtir. Bu istatistik test özetle şöyle yapılır:

1. Bir soyağacı oluşturulur.
2. Rekombinasyon sıklığı çeşitli noktalar için hesaplanır .
3. Her nokta için bir LOD skoru hesaplanır.
4. En yüksek LOD skoruna sahip nokta bağlantı olasılığının en yüksek olduğu yer sayılır.

LOD skoru ařağıdaki formül ile hesaplanır:

$$\begin{aligned} LOD = Z &= \log_{10} \frac{\text{Belli bir baęlantı deęeri ile doęum olasılıęı}}{\text{Baęlantı olmadan doęum olasılıęı}} \\ &= \log_{10} \frac{(1 - \theta)^{NR} \times \theta^R}{0.5^{(NR+R)}} \end{aligned}$$

Bu formüldeki NR rekombinasyon yapmamıř (rekombinant olmayan) R ise rekombinasyon yapmıř çocukların sayısıdır. θ , rekombinasyon sıklıęıdır. Tamamen baęlantısız alellerin (yani farklı kromozmlarda bulunan alellerin) rekombinasyon oranı %50 olduęundan paydada 0.5 vardır.

Özetle, baęlantı analizlerinde LOD skor, hastalık alel ile belirteç alel arasında baęlantı olması olasılıęının, baęlantı olmaması olasılıęına oranıdır. LOD deęerinin 3'e eřit ya da büyük olması baęlantı olduęunu ifade etmektedir.

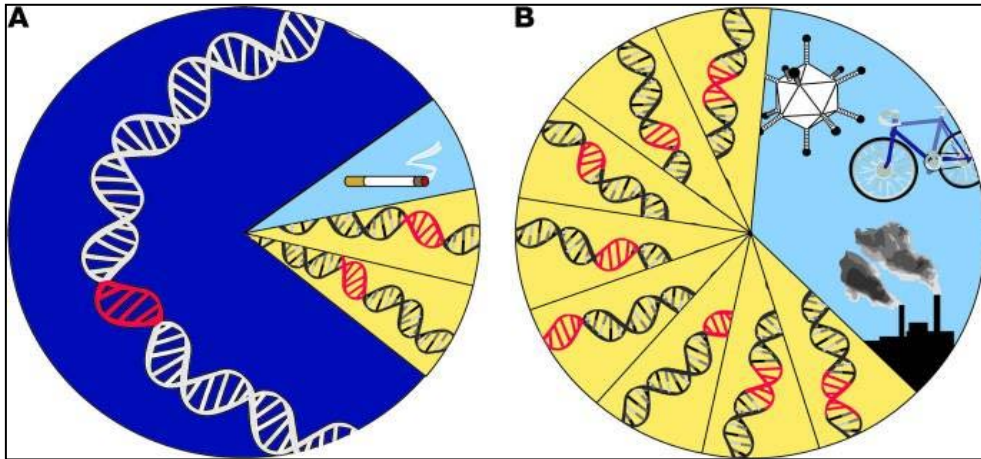
Parametrik olmayan baęlantı analiz yaklařımı ise belirli bir kalıtım modeli özellięine dayanmamaktadır. Bu nedenle model baęımsız yaklařım olarak da tanımlanabilmektedir. Bu yaklařımda da aynı istatistiksel yorumla parametrik baęlantı analizlerindeki gibi parametrik olmayan LOD skor hesaplanmaktadır. Parametrik olmayan baęlantı analiz yaklařımının model baęımsız olması bir avantaj olarak deęerlendirilirken, parametrik baęlantı analiz metodu ile kıyaslandığında metodun gücündeki etkili azalma dezavantaj olarak ortaya çıkmaktadır (Benke et al. 2010).

Tek bir genin baskın olduęu ailesel kanser sendromlarından farklı olarak sporadik kanserlere olan yatkınlık multifaktöryel özellik göstermektedir. Bu çerçevede düşük penetranslı tümör yatkınlık genlerinin baęlantı analizi ve pozisyonel klonlama ile tanımlanmasına insan popülasyonunun genetik ve çevresel heterojenitesi engel oluşturmaktadır (Pharoah et al. 2004). Bu noktada genom boyu asosiyasyon çalışmaları devreye girmektedir.

2.5. Kompleks Hastalıklar ve Asosiyasyon Çalışmaları

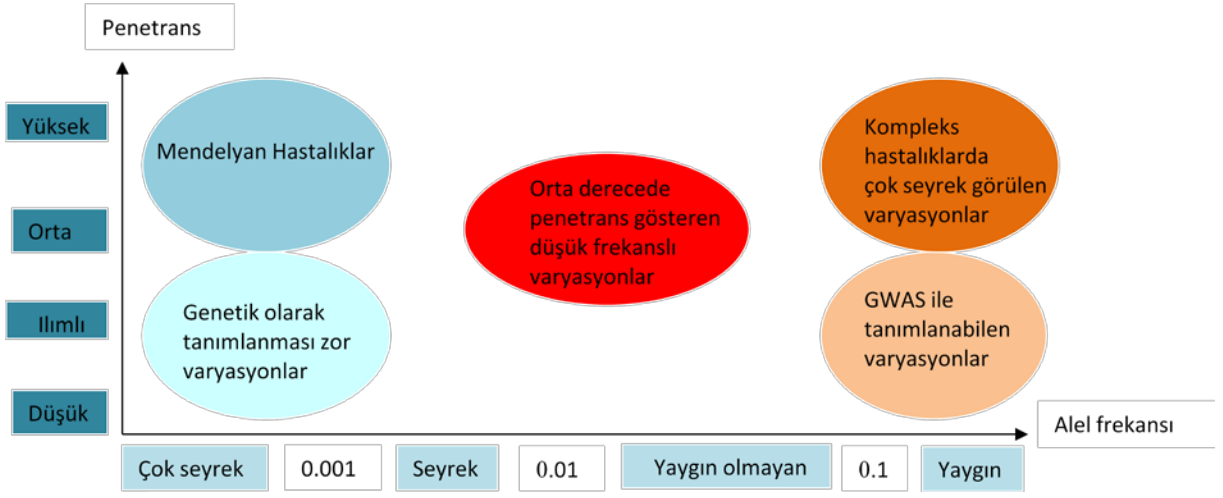
Birçok kalıtsal ve çevresel faktörün sorumlu olduğu hastalıklar kompleks veya mutifaktöryel hastalıklar olarak tanımlanmaktadır (Schork 1997).

Geçmiş 30 yılda Mendelyen kurallara göre kalıtılan hastalıkların moleküler genetiğinin anlaşılması büyük gelişme göstermiştir. Mendelyen kurallara göre tek bir gendeki mutasyonlar hastalık üzerinde geniş etkiye sahiptir. Günümüzde, birçok genin birlikte etkisinden kaynaklanan, poligenik kalıtım gösteren kompleks insan hastalıklarının moleküler ve genetik temelini çözülmesi ile mücadele edilmektedir. Bu mücadeleye genomik teknolojiye gelişmeler katkıda bulunmaktadır (Pharoah et al. 2004) (Şekil 2.8).



Şekil 2.8: Tek gen hastalıkları ve kompleks hastalıklara genetik ve çevresel katkılar (A) Tek gen hastalığı (B) Kompleks hastalık

Düşük penetranslı alleller birçok kompleks hastalık için yatkınlığa katkıda bulunmaktadır. Bu gibi alleller ailelerde seyrek olarak çoklu olgu yaratacağından bu allellerin genetik bağlantı analizleri ile tanımlanması zor hatta imkansızdır. Bu nedenle düşük penetranslı allellerin aranması, belirlenmesi, hasta ve kontrol gruplarında spesifik allel frekanslarının kıyaslanması ile yapılan asosiyasyon çalışmaları ile mümkün olmaktadır (Webb et al. 2007; McCarthy et al. 2008) (Şekil 2.9).



Şekil 2.9: Düşük frekanslı alleller ve hastalığa yatkınlık

Asosiyasyon çalışmaları geniş populasyonlardaki bireylerde genetik varyantlar ve hastalık arasındaki asosiyasyonu belirler. Birçok asosiyasyon çalışması populasyon temellidir fakat ebeveynler ve etkilenmiş çocuk üçlülerini kullanan aile bazlı çalışmalar da yürütülmektedir. Aile bazlı çalışma dizaynı çocuktaki spesifik allel ve hastalık arasındaki asosiyasyonu, heterozigot ebeveynlerin bu alleli etkilenmiş çocuğa beklenenden daha sık aktarıp aktarmadığı test ederek gösterir. Asosiyasyon çalışmaları kohort (aynı yıl doğma, aynı işte çalışma, aynı bölgede oturma vb istatistiksel olarak ortak özellikler gösteren bireysel topluluk) veya olgu-kontrol olarak yapılabilir. Kohort çalışmalarının çok fazla zaman alıcı olması ve seyrek görülen hastalıklarda kullanılmaya elverişsiz olması gibi sorunları varken olgu-kontrol dizaynı ise seçimde veya bilgide önyargıya (bias) neden olabilmektedir.

Çalışma tasarımı ne olursa olsun genetik varyasyonlar incelenecekse iki genel yaklaşım vardır. Bunlar genom boyu asosiyasyon çalışmaları ve aday gen asosiyasyon çalışmalarıdır. Aday gen asosiyasyon çalışmaları biyolojik yolak ile ilgili en olası bir veya daha fazla gendeki genetik varyasyonun rolünü araştırmaktadır. Genom boyu asosiyasyon çalışmalarında araştırmacı yatkınlık lokusu ile ilgili önceden bir bilgiye sahip değildir, ilgilenilen hastalıkla asosiyasyon gösteren kromozomal bölgeyi tanımlamak için

çalışmaktadır. Genotipleme platformlarının fiyatlarının düşmesine bağlı olarak genom boyu asosiyasyon çalışmalarının kullanımı gittikçe yaygınlaşmaktadır (Yende et al. 2006).

2.6. Asosiyasyon Çalışmaları

İnsan genom sekansının taslağının 2001 yılında çıkarılması ve 2003 yılında tam olarak netleştirilmesinden sonra hastalıkların genetik öğelerini inceleyen araştırmalarda patlama olmuştur. Sosyodemografik ve klinik risk faktörleri belirli bir bireyin neden hastalık geliştirdiği ya da öldüğünü tam olarak açıklayamamaktadır. Bu nedenle bilim adamları genetik varyasyonların hastalıklara yatkınlığa neden olup olmadığını ve hastalıkların sonucunu etkileyip etkilemediğini anlamaya çalışmaktadırlar (Yende et al. 2006; Manolio et al. 2008).

İnsan genomunun netleşmesini takiben insan genomik varyasyonlarının tanımlanması amacıyla yapılan projelerden biri HapMap projesidir. HapMap, 4 farklı populasyonda yaklaşık 3 milyon yaygın SNP'in frekanslarını ve bu SNP'lerin arasındaki asosiyasyon paternlerini belirlemek üzere dizayn edilmiş bir projedir (Manolio et al. 2008). HapMap projesi ile genomik seviyede tag SNP setleri tanımlanmıştır. Tag SNP, genomun bir bölgesindeki yüksek bağlantı dengesizliğinde (LD) olan tüm SNP'leri temsil eden SNP'dir. Bir kromozom bölgesindeki tüm SNP'ler genotiplendirilmesine gerek duyulmaksızın, sadece tag SNP'lerin genotiplendirilmesi ile genetik varyasyonun tanımlanması mümkün olabilmektedir. Bu bilgiler aracılığı ile HapMap projesi, genom boyunca asosiyasyon çalışmalarına yön vermek için gerekli temeli sağlamıştır.

Kalıtım modeli bilinmeyen, genetik etkenlerin işe karıştığı astım, kanser, diabet, kalp hastalıkları gibi kompleks ve yaygın hastalıkların araştırılmasına ve bu gibi hastalıklarda rol oynayan genlerin tanımlanması için kullanılan asosiyasyon çalışmaları, olgular (belirli bir hastalığı olan kişiler veya belli bir özelliğe sahip bireyler) ve kontroller arasındaki allel frekansı farkını değerlendirmektedir (Kruglyak 2008). Asosiyasyon çalışmaları alelleri incelemektedir ve belirteç lokustaki belirli bir allelin (moleküler marker), hastalığa yüksek risk oluşturan yatkınlık alleli ile nasıl birlikte kalıtıldığı hakkında bilgi vermektedir.

Kompleks hastalıkları etkileyen genleri ortaya çıkarmak ve tanımlamak için iki temel asosiyasyon çalışması yaklaşımı vardır. Bunlar aday gen asosiyasyon çalışmaları ve genom boyu asosiyasyon çalışmalarıdır (Schork 1997).

2.6.1. Aday Gen Assosiyasyon Çalışmaları

Gen tanımlanması ve karakterizasyonu için kullanılan aday gen asosiyasyon yaklaşımı, çok sayıdaki genom projeleri ve biyoinformatikteki gelişmeler ile daha net sonuç verebilir hale gelmiştir.

Aday gen analizi, belirli bir moleküler belirteç ile bir hastalık arasındaki asosiyasyonu test etmeyi amaçlayan bir yaklaşımdır (Schork 1997).

Aday gen asosiyasyon çalışmalarında, hastalığın biyolojisi ile ilgili varsayımlardan yola çıkarak belirlenen ilgili aday genlerle çalışılmaktadır. Aday gen asosiyasyon yaklaşımında araştırılacak hastalığa ait eldeki bütün veriler göz önüne alınarak aday genler belirlenmektedir.

Aday gen asosiyasyon yaklaşımı yeni genlerin izolasyonunu gerektirmemektedir ancak aday genlerin fonksiyonları ile ilişkili bilgilerin elde edilebilmesi için bu genlerin genomdaki lokalizasyonuna gerek duymaktadır. Genomda belirlenen aday genlerin bulunduğu bölgelere sıkı bağlantı gösteren genetik belirteçler öncelikli olarak test edilmektedir. Test edilen bu genetik varyasyonlar, hastalıklı kişilerde hastalıklı olmayanlara göre daha sık görülüyorsa, o varyasyon ve hastalık arasında bir ilişki olabileceği çıkarımı yapılabilmektedir (Schork 1997).

Aday gen asosiyasyon çalışmaları birçok hastalığın araştırılmasında kullanıldığı gibi kanser ve kolorektal kanser araştırmalarında da sıklıkla kullanılmaktadır (Engel et al. 2002; Hildesheim et al. 2002; Sachse et al. 2002; Landi et al. 2003; Slattery et al. 2010).

2.6.2. Genom Boyu Asosiyasyon Çalışmaları

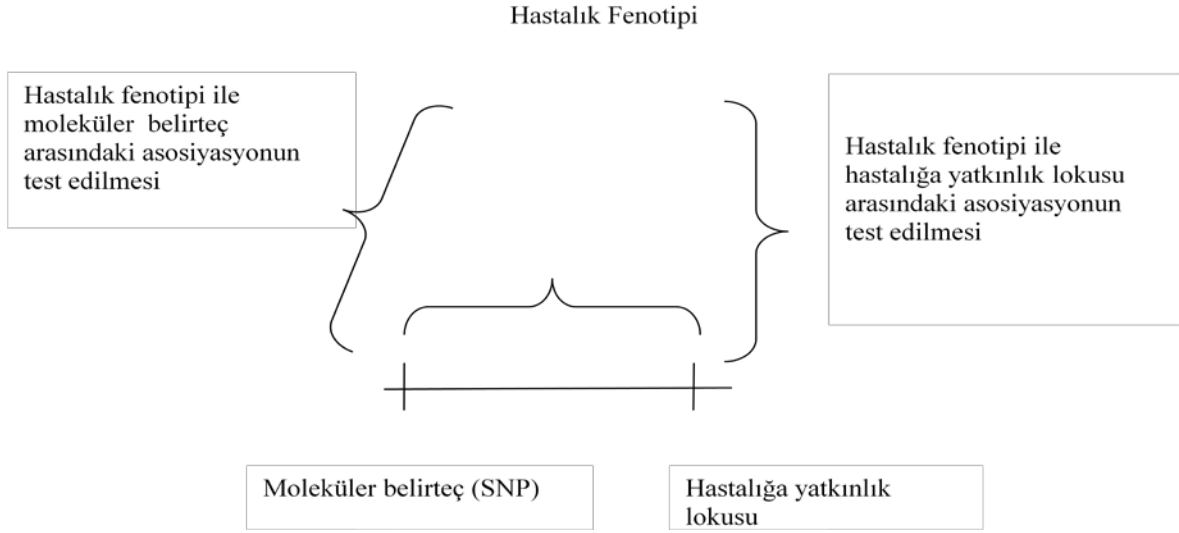
Hastalığa yatkınlığın genetik temelini açığa çıkarılması için kullanılan güncel strateji genom boyu asosiyasyon çalışmalarıdır (genom wide association study, GWAS) (Moore et al. 2010).

Daha önce de bahsedildiği üzere asosiyasyon çalışmalarında genomda en yaygın olarak bulunan genomik varyantlar olan SNP'ler kullanılmaktadır. SNP'ler binlerce jenerasyon önce oluşan ve doğal seleksiyon veya şansla yayılan mutasyonlardır. Daha önce varolan SNP'e çok yakın ikinci bir SNP oluştuğunda (aralarında onbinlerce baz çiftinden de fazla mesafe olabilmektedir) bu iki varyant alel her zaman bir sonraki jenerasyondaki aynı oğula aktarılmaktadır. İki allelin tesadüfi olmayan bu birlikteliği bağlantı dengesizliği (linkage disequilibrium, LD) olarak tanımlanmaktadır (Cichon et al. 2009).

LD genom haritalama veya hastalık genlerinin tanımlanmasında kullanılmaktadır. Bu bilgiyi esas alan HapMap projesi ile milyonlarca yaygın SNP ve ilgili lokuslardaki alellerin bağlantı dengesizliği (linkage disequilibrium, LD) ortaya çıkarılmıştır (Moore et al. 2010).

Eğer bir SNP yaygın bir hastalığın riskini artırıyorsa, hastalık ile bu SNP (direk asosiyasyon) ve birkaç yakın SNP (LD den kaynaklanan indirect asosiyasyon) arasında istatistiksel olarak anlamlı bir ilişki, asosiyasyon olmaktadır (Cichon et al. 2009). Farklı lokuslardaki polimorfizmler arasındaki tesadüfi olmayan asosiyasyon, LD derecesine göre ölçülmektedir. LD derecesi, sayısal olarak, gözlenen ve beklenen allel frekansı arasındaki farktır. LD'nin derecesi genetik bağlantı, seleksiyon, rekombinasyon ve mutasyon oranı, genetik kayma, tesadüfi olmayan çiftleşme ve popülasyon yapısı gibi birçok faktörden etkilenmektedir.

Hastalık fenotipi ile hastalığa yatkınlık lokusu arasındaki ilişkinin test edilmesini amaçlayan genom boyu asosiyasyon çalışmalarında, hastalık fenotipi ile asosiyasyon gösteren moleküler belirteç amaca ulaşmada ipucu olarak kullanılmaktadır. Bu sayede hastalık fenotipi ile asosiyasyon gösteren moleküler belirteç ile yüksek LD gösteren lokus hastalığa yatkınlık lokusu olarak tanımlanabilmektedir (Şekil 2.10).



Şekil 2.10: Genom boyu asosiyasyon yaklaşımı ile genetik asosiyasyonun belirlenmesi

Genom boyu asosiyasyon çalışmalarında genler değil lokuslar tanımlanmaktadır (Hardy et al. 2009). Genom boyu asosiyasyon yaklaşımı, insan hastalıklarının yaygın varyasyon-yaygın hastalık hipotezine uyan lokusları bulmak üzere tasarlanmıştır. Yaygın varyasyon-yaygın hastalık hipotezi birçok yaygın hastalığın yaygın allellerden kaynaklandığı belirtmektedir (Hardy et al. 2009; Moore et al. 2010). Daha açık söylemek gerekirse, yaygın varyasyon-yaygın hastalık hipotezi zayıftan orta dereceye kadar değişen aralıkta etkileri olan birkaç yaygın allelik varyasyonun kombinasyonunun hastalığa yatkınlıktan sorumlu olduğunu söylemektedir (Benke et al. 2010). Bu hipoteze karşılık öne sürülen diğer bir hipotez olan seyrek varyasyon-yaygın hastalık hipotezi ise farklı bireylerde, güçlü etkiye sahip birçok seyrek, bağımsız risk allelinin incelenilen hastalığa neden olduğunu söylemektedir (Benke et al. 2010). Bugünkü literatür, her iki hipotezin de ilgililenen gen veya hastalığa bağlı olarak doğru olduğunu öne sürmektedir (Iyengar et al. 2007).

Ayrıca bazı durumlarda genom boyu asosiyasyon çalışmaları ile tanımlanan düşük penetranslı, yaygın (düşük-risk) varyantların bağlantı dengesizliğinde buldukları yüksek penetranslı, seyrek (yüksek-risk) varyantların belirteçleri olabilirler. Bu bağlı varyantlar aynı genetik lokusta olabilirler fakat bazen bağlantı daha uzak lokuslara kadar uzayabilir.

Genom varyasyonları ile ilgili bu bilgiler ve yeni biyomühendislik yöntemlerinin birlikte kullanımı ile bir milyondan fazla SNP'i analiz edebilen çipler dizayn edilebilmiştir. Bu çiplerin kullanımı arttıkça maliyeti düşmüş ve böylece çalışmaların sayısı da gittikçe artmıştır. 2008 yılında yayınlanan iki yayına göre genom boyu asosiyasyon çalışmaları ile 70 hastalık için yüzlerce yatkınlık lokusu rapor edilmiştir (Moore et al. 2010).

Tipik bir genom boyunca asosiyasyon çalışması 4 aşamadan oluşmaktadır. Bunlar sırası ile olgu/kontrol grubunun seçilmesi, DNA izolasyonu ve genotiplendirme, hastalık ile eşik değerini geçen SNP'ler arasındaki ilişkinin istatistiksel olarak değerlendirilmesi ve son olarak verileri güçlendirmek ve doğrulamak için tanımlanan ilişkinin bağımsız populasyon örneklerinde tekrarlanması veya işlevsel anlamın deneysel olarak incelenmesidir.

Günümüzde çok yaygın olarak kullanılan genom boyu asosiyasyon çalışmalarının yararları, hataları ve kısıtlamaları da bulunmaktadır. Yararları arasında bir başlangıç hipotezine ihtiyaç duymaması, işlenebilen dijital (elektronik yolla sayısal olarak gösterilebilen) bilgi kullanması, hem sekans hem de kopya sayısı varyasyonları hakkında bilgi sağlaması sayılabilmektedir. Ayrıca genom boyu asosiyasyon çalışmaları çoklu interaksiyon gösteren hastalık genlerini ve bu genlerin ayrı ayrı yollarını tanımlayabilerek, hastalığın etiyolojisinin kapsamlı olarak anlaşılmasını sağlamaktadır.

Genom boyu asosiyasyon çalışmalarında çok sayıda örneğe gereksinim duyulması çalışmalara önemli bir kısıtlama getirmektedir. Genom boyu asosiyasyon çalışmalarının bir diğer kısıtlayıcı özelliği ise popülasyonda sadece yaygın olan, yani popülasyonda %5'ten daha fazla görülen alelleri belirleyebilmeleridir. Bu nedenle genom boyu asosiyasyon çalışmalarının, hastalık ile ilişkili bütün genetik değişkenlikler hakkında bilgi sağladığı düşünülmesi önemli bir hata olacaktır (Hardy et al. 2009). Genom boyu asosiyasyon çalışmalarından elde edilecek verilerin doğrulanması için yine çok sayıda örnek ile tekrarlanması gerekliliği de bu çalışmalara önemli bir kısıtlama getirmektedir (Hardy et al. 2009).

Yukarıda yararları ve kısıtlamalarından bahsedilen genom boyu asosiyasyon çalışmalarının kompleks hastalıkların araştırılmasında birçok çalışmaya ışık tuttuğu ve tutmaya devam edeceği açıktır.

Birçok arařtırmada vazgeçilmez bir araç olarak kullanılan genom boyu asosiyasyon çalıřmaları popülasyon veya aile temelli olarak iki farklı şekilde dizayn edilebilmektedir.

2.6.2.1. Popülasyon Temelli Genom Boyu Asosiyasyon Çalıřmaları

Günümüzde popülasyon temelli genom boyu asosiyasyon çalıřmaları yaygın olarak kullanılmaktadır. Literatürde, Haziran 2010 itibari ile 165 özellik için ($p < 5 \times 10^{-8}$) 904 tane yayınlanmış genom boyu asosiyasyon çalıřması bulunmaktadır (Freedman et al. 2011).

Popülasyon temelli genom boyu asosiyasyon çalıřmalarının tasarımında olgu ve kontrol grupları ile çalıřılmaktadır. Olgu ve kontrol gruplarını oluřturan bireyler bağımsız seçilen bireyler olduklarından popülasyon katmanlařması (stratification, stratifikasyon) ortaya çıkmaktadır (Cardon et al. 2003). Popülasyondaki katmanlařma gerçekte olmayan iliřkilerin varmış gibi görünmesine neden olabilmektedir. Popülasyon katmanlařması

popülasyondaki alt popülasyonlar arasında allel frekansında sistematik bir farkın olması durumudur. Bu fark alt popülasyonların farklı soydan gelmelerinden kaynaklanabilir. Popülasyon katmanlařması,

- a. karıřma,
- b. test istatistiğinde aşırı dağılmaya yolaçan ilintisizlik (relatedness),
- c. seçmede önyargı (bias) olarak üç önemli probleme yol açmaktadır (Duncan et al. 2005).

Populasyon temelli asosiyasyon çalıřmalarının kullanımını kısıtlayan etkenlerden biri de bu çalıřmalarda kullanılan testlerin toplumlarda çiftleşmenin genetik açıdan rastgele olması ve test yapılan tarihten en az iki kuřak öncesi dönemden beri toplumda genetik karıřım olmaması kořulu ile deęerlendirmede kullanılabilir olmasıdır.

Yukarıda bahsedildięi gibi popülasyon stratifikasyonu popülasyon temelli asosiyasyon çalıřmalarında önemli bir sorun teşkil etmektedir. Bu nedenle son yıllarda yapılan arařtırma ve incelemeler sonucunda popülasyon temelli asosiyasyon çalıřmalarında

güvenilir sonuç alabilmenin tek yolunun binlerce olgu-kontrol üzerinde çalışmak olduğu hususunda karar birliğine ulaşılmıştır (Duncan et al. 2005; Bowcock 2007).

Binlerce hatta onbirlerce örnek analiz edilerek yapılabilen popülasyon temelli asosiyasyon çalışmalarında, bu kadar çok sayıda materyal (olgu-kontrol) bulabilmenin ve çalışmanın oldukça zor ve maliyetli olduğu açıkça görülmektedir. Örnek toplamadaki sıkıntının aşılabilmesi için çok merkezli ulusal veya uluslararası konsorsiyumlar oluşturulmaktadır. Araştırmacılar lojistik şart ve imkanlarına göre çalıştıkları hastalıkların ve/veya kendi çalışma gruplarının özelliğine göre popülasyon temelli ve/veya aile temelli asosiyasyon çalışması yaklaşımını tercih etmektedirler.

2.6.2.2. Aile Temelli Genom Boyu Asosiyasyon Çalışmaları ve İstatistiksel Yaklaşımlar

Aile temelli genom boyu asosiyasyon çalışma tasarımında etkilenmiş birey, hasta (çocuk) ile onun sağlıklı anne ve babası ile çalışılmaktadır.

Popülasyon temelli asosiyasyon çalışmalarının önündeki en büyük engel olan popülasyon katmanlaşması aile bazlı tasarımla büyük oranda aşılabilmektedir. Bu nedenle son yıllarda asosiyasyon çalışmalarında aile temelli yaklaşıma ilgi artmıştır.

Aile temelli yaklaşım ile yapılan asosiyasyon çalışmalarının istatistiksel güç hesaplamaları bu şekilde tasarlanan çalışmalarda ulaşılması gereken olgu sayısının yüzlü rakamlarla sınırlandırılabilceğini ortaya koymuştur (Yang et al. 2003).

Genom boyu asosiyasyon çalışmalarında kullanılan testler her aileden sadece bir etkilenmiş (hasta) çocuk çalışmasını gerektirir ve birden çok hastanın olduğu ailelerde geçerliliği düşüktür. Bu nedenle, asosiyasyonu, ilişkiyi değerlendirmek ve eş zamanlı olarak hastalık ve belirteç alelin kuşaklar arası birlikteliğini değerlendiren bağlantı hakkında da bilgi edinebilmek için kalıtımda dengesiz aktarım (transmission disequilibrium, TDT) testi kullanıma girmiştir (Spielman et al. 1993); (Ewens et al. 1995). Bu test toplumdaki genetik yapılaşmadan etkilenmez. İlgilenilen alel açısından heterozigot ebeveynler ve hasta çocukları inceleyen TDT analizinde çocuğa (hasta) kalıtılan alellerin, kalıtılmayan aleller ile karşılaştırılması yapılmaktadır. Aile temelli asosiyasyon

çalışmalarında kullanımı çok yaygın olan bu TDT testinin kabul edilebilir bir güçte olabilmesi için en az 200 trio'ya ihtiyaç vardır.

TDT analizinin yapılabilmesi için hasta çocuğa aktarılan ve aktarılmayan alellerin karşılaştırıldığı bir çizelge oluşturulmasına ihtiyaç duyulmaktadır (Dracopoli et al. 1994). Bu çizelge ile her bir SNP için her bir anne-baba-çocuk üçlüsü incelenerek çocuğun sözkonusu SNP için sahip olduğu genotipi oluşturan durum Çizelge 2.2'de gösterilen şekilde ortaya konur. Söz konusu SNP için incelenen bütün ailelerde Çizelge 2.2'de tanımlanan genotip olasılıklarının (a, b, c, d) bu ailelerde kaç kez gözlemlendiği bu tabloya işlenir. Her iki ebeveynen de kalıtılan ve kalıtılmayan alelleri aynı çizelge içinde belirtilmesi gerektiği için dört gözlü bir TDT tablosu oluşturulmaktadır. Ayrıca bu dört gözlü bir TDT tablosu anne-babada belirli bir genotip görüldüğü durumda çocukların genotiplerini gösterecek şekilde ebeveyn-çocuk karşılaştırılarak oluşturulmaktadır (Çizelge 2.1 ve 2.2).

Çizelge 2.1: TDT için dört gözlü tablo

		Çocuğa Aktarılmayan Alleller	
		Allel A	Allel B
Çocuğa Aktarılan Alleller	Allel A	a	b
	Allel B	c	d

Çizelge 2.2: Anne ve babanın genotiplerine göre olası çocuk genotipleri

	Anne	Baba	Çocuk
1	AA	AA	AA
2	AA	AB	AA, AB
3	AA	BB	AB
4	AB	AA	AA, AB

5	AB	AB	AA, AB, AB, BB
6	AB	BB	AB, BB
7	BB	AA	AB
8	BB	AB	AB, BB
9	BB	BB	BB

Aile temelli asosiyasyon çalışmalarında, TDT'den başka TDT benzeri testler de kullanılmaktadır. Bunlardan biri kardeş çiftlerini kullanan sib-TDT (discordant-sibship test for disequilibrium, S-TDT)'dir (Spielman et al. 1998). S-TDT analizi ebeveynlerin genotip bilgisine gerek duymadan yapılabilir. Ebeveynlerin genotip bilgisinin olduğu ve olmadığı durumlarda TDT ve S-TDT testleri birlikte de kullanılabilir (Horvath et al. 1998; Martin et al. 2000).

Ayrıca aile temelli asosiyasyon çalışmalarında, bir ebeveynin genotipine ulaşılabilirken, diğerine ulaşamadığı durumlarda yeni bir test olan 1-TDT kullanılabilir. 1-TDT, hastalık lokusunu, etkilenmiş bireyler (hasta) ve onların sadece bir ebeveynlerini kullanarak belirleyebilen bir testtir. 1-TDT 'nin ilişkisi, asosiyasyonu TDT ve S-TDT'den daha az güçle de olsa saptayabildiği gösterilmiştir (Sun et al. 1999).

Özellikle geç yaşlarda ortaya çıkan hastalıklarda anne ve babaya ait genotip bilgisi elde edilemeyebilir. Bunun sebebi, genellikle, anne veya babadan bir tanesinin ya da her ikisinin çoktan ölmüş olmalarıdır (Cervino et al. 2000; Weinberg 1999). Aile temelli asosiyasyon çalışmalarında, anne-baba-çocuk üçlüsünden oluşan ailelerin yanında, anne veya babadan sadece bir tanesinin olduğu yani anne-çocuk ikilisi veya baba-çocuk ikilisinden oluşan aileler de bulunabilir. Çalışmanın böyle farklı aile gruplarını içermesi durumunda, hasta çocuk ve sadece bir ebeveyninden oluşan aileleri çalışmanın dışında bırakıp, çalışmayı sadece anne-baba-çocuk üçlüsünden oluşan aileler üzerine yöneltmek bilgi ve veri kaybına yol açacaktır. Bu durum ise, verinin analizi sırasında istatistiksel güç kaybına sebep olacaktır (Weinberg 1999; Cervino et al. 2000; Hu et al. 2010). Bu sebeple, maksimum fayda elde edebilmek için, yani daha güvenilir sonuçlar elde edebilmek için,

hem anne-baba-çocuk üçlüsünden hem de anne-çocuk ve baba-çocuk ikililerinden gelen bilgiyi ortak olarak değerlendirmek çok önemlidir (Hu et al. 2010). Bu bağlamda Allen ve arkadaşları hem anne-baba-çocuk üçlüsüne hem de anne-çocuk ve baba-çocuk ikililerine ait genotip bilgisi içeren asosiyasyon çalışmalarında C-TDT test istatistiğini önermektedir (Allen et al. 2003). Aslında bu test istatistiği, yukarıda bahsi geçen 1-TDT istatistiğinin, anne-baba-çocuk üçlüsünden ve anne-çocuk ve baba-çocuk ikililerinden gelen bilginin katkısı eşit olacak şekilde ayarlanmış bir uzantısıdır. Daha sonra C-TDT'nin permütasyon tabanlı randomize edilmiş versiyonu da aile bazlı asosiyasyon çalışmalarında kullanılmak üzere tanımlanmıştır (Hu et al. 2010).

Çalışmanın anne-baba-çocuk üçlüsü ve çocuk-ebeveyn ikilisinden oluşması durumunda, C-TDT istatistiğinin, sadece anne-baba-çocuk üçlüsü kullanan TDT'ye göre daha güçlü sonuçlar çıkardığını göstermektedir. Diğer bir taraftan, bu test istatistiği asimptotik bir dağılım yerine, permütasyon yöntemine dayandığı için, küçük örneklem büyüklüklerinde de etkili bir şekilde kullanılabilir (Good 2001). Yalnız bu yöntem birden fazla taşıyıcı çocuktan oluşan ailelere veya tamamıyla sadece anne-çocuk ikililerinden ya da tamamıyla sadece baba-çocuk ikililerinden oluşan çalışmalara uygulanamamaktadır (Hu et al. 2010).

Ayrıca, aile temelli asosiyasyon analizlerinde güvenilir sonuç alınabilmesi için TDT analizinin öngördüğü 200 trio'luk örneklem boyutuna ulaşamama problemini aşabilmek için, ekibimiz tarafından önerilen ardışık hipotez testleri ('sequential hypothesis tests' (Wald 1947) veya sequential probability ratio test, SPRT) (Ilk et al. 2011) alternatif bir yaklaşım olarak aile temelli asosiyasyon analizlerinde kullanılabilir. Ekibimiz tarafından geliştirilen bu testin, SPRT'nin TDT ye göre en önemli üstünlüklerinden biri SNP'leri "ilişkili", "ilişkili değil" veya "ilişkili veya ilişkili değil diye belirleyebilmek için daha fazla trio toplamaya devam edilmesi gereken" SNP'ler diye üç grupta toplayabilmesidir. TDT analizinde ise SNP'ler sadece "ilişkili" veya "ilişkili olduğunu kanıtlayacak delil yok" olarak ikiye ayrılmaktadır.

2011 yılında yayınlanan makalede SPRT analizi bir simülasyon (benzetim) çalışması ile incelenmiştir (Ilk et al. 2011). Simülasyon verisi üretildikten sonra TDT ve SPRT uygulanmış ve karşılaştırılmıştır. 100'den ufak trio sayısı için, SPRT %87.9 ila %93.8 arasında doğruluk değeri alırken, TDT'nin ise sadece %39.9 ila %74.8 arasında değerler

aldığı belirlenmiştir. 50 trio için doğruluk değerlerine bakıldığında SPRT %92.7’lerde iken TDT’nin ise %57.9’larda kaldığı görülmüştür (Ilk et al. 2011) (Çizelge 2.3).

Çizelge 2.3: SPRT ve TDT’nin değişen trio sayısına göre doğruluğu (%) ($\alpha=0.1\%$ ve $\beta=20\%$)

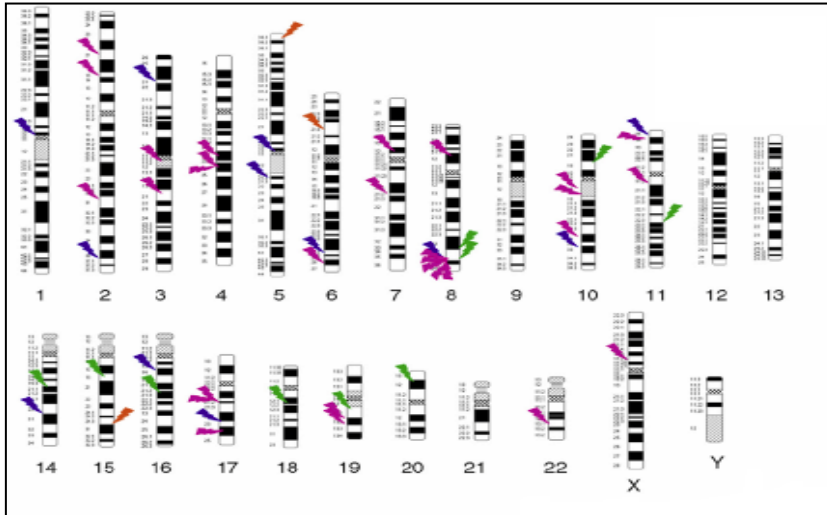
Trio sayısı	Doğruluk %	
	SPRT	TDT
30	87.9	39.9
40	91.5	50.0
50	92.7	57.9
60	93.3	63.3
70	93.6	67.9
80	93.7	71.9
90	93.8	74.8
100	93.9	77.6
110	94.0	79.8
120	94.1	81.7
130	94.1	83.3
140	94.1	84.7
150	94.1	85.9
160	94.1	86.8
170	94.1	87.8
180	94.1	88.5
190	94.1	89.2
200	94.1	89.8

SPRT’de, SNP’leri iki hipotezden birine atamak için çok az sayıda trio’ya ihtiyaç olduğu görünmüştür. Sadece 50 trio ile bile, bu SNP’lerin %60’ı hakkında karar verilebilmiştir. Geriye kalan %40’ı içinse ‘daha fazla trio toplamaya devam et’ kararı çıkmıştır. Trio sayısı 80’e çıktığında, bu SNP’lerin %90’ından fazlası için SPRT testleri sonuçlanmıştır.

2.7. Kanser Araştırmalarında Genom Boyu Asosiyasyon Çalışmaları

Geçtiğimiz 20 yılda kanser genetiği araştırmalarındaki ilerlemeler ile mendelyan kalıtım paterni gösteren bir grup kanserden sorumlu olan ve yüksek penetrans göstererek kansere yatkınlığa neden olan mutasyonlu genler tanımlanmış ve karakterize edilmiştir. Kansere yatkınlığa neden olan bu genler arasında, kalıtsal meme ve rahim kanserinde *BRCA1* ve *BRCA2* genleri, Lynch sendromunda yanlış eşleşme tamir genleri, ailesel adenomatoz polipte *APC* geni, Li-Fraumeni Sendromunda *TP53* geni sayılabilir (Stadler et al. 2010)

Yakın geçmişte genom boyu asosiyasyon çalışmaları ile başta meme, prostat, akciğer kanseri ve kolorektal kanser olmak üzere birçok kanserde 100 den fazla yaygın, düşük penetranslı lokus tanımlanmıştır (Varghese et al. 2010) (Şekil 2.11).



Şekil 2.11: Bazı yaygın kanserler için genom boyu asosiyasyon çalışmaları ile tanımlanan lokuslar (Mor ile işaret edilen bölgeler prostat kanseri, yeşil ile işaret edilen bölgeler meme kanseri, kırmızı ile işaret edilen bölgeler kolorektal kanser ve siyah ile işaret edilen bölgeler ise akciğer kanseri için tanımlanan yaygın kanserler için genom boyu asosiyasyon çalışmaları ile tanımlanan lokusları göstermektedir.)

Yüksek penetrans ve seyrek popülasyon frekansı gösteren meme, prostat, akciğer kanseri ve kolorektal kanser yatkınlık genleri arasında *P53*, *APC*, *CDH1*, *BRCA1*, *BRCA2*, *MLH1*, *MSH2*, *PTEN*, *STK11*, *CDKN2A*, *MSH6* ve *PMS2*, orta derecede risk gösteren alleller arasında *ATM*, *CHEK2*, *PALB2*, *BRIP1*, *CYP1A1*, *GSTM1*, düşük penetrans ve yüksek popülasyon frekansı gösteren risk allelleri arasında ise 8q24 lokusu, *FGFR2*, *MSMB*, *CHRNA5*, *CHRNA3*, *CHRNA4* genleri sayılabilmektedir (Stadler et al. 2010).

Meme kanserinde yapılan genom boyu asosiyasyon çalışmaları ile bugüne kadar 13 bağımsız yatkınlık lokusu tanımlanmıştır. Tanımlanan yatkınlık lokuslarındaki en güçlü asosiyasyon gösteren SNP, *FGFR2* geninde bulunmaktadır. Diğer yatkınlık lokuslarında bulunan ve daha önce kanserle ilişkilendirilmemiş, sırası ile 16q12 ve 11p15 bölgelerinde bulunan *TOX3*, ve *LSP1* gibi genler ile ilgili doğrulama çalışmaları halihazırda yapılmakta olup karsinogenezdeki olası potansiyel rolleri araştırılmaktadır (Varghese et al. 2010; Stadler et al. 2010).

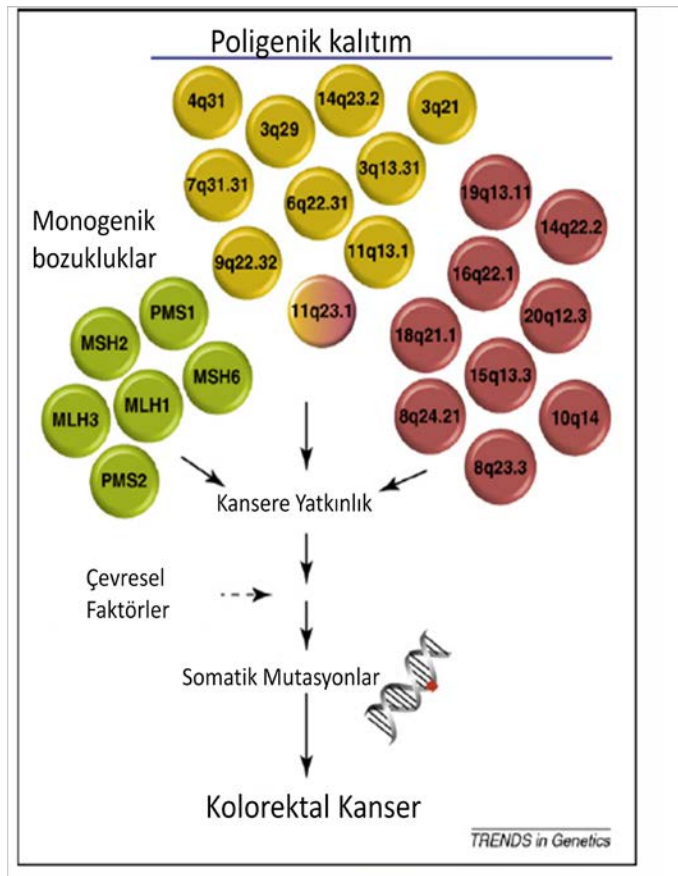
Prostat kanseri için 30 yaygın genetik risk lokusu tanımlanmıştır. Bunların başında, kolorektal kanser, mesane ve meme kanserinde de tanımlandığı gibi 8q24 lokusu gelmektedir. Haiman ve arkadaşlarının yaptığı bir çalışmada 8q24 bölgesinde birbirinden bağımsız 7 tane risk varyantı gösterilmiştir. Bu risk SNP'lerin içinde bulunduğu ya da yakınlarındaki genlerin (*MSMB*, *KLK3*, *KLK2* vs) prostat kanseri oluşumunda rolleri olduğu da ispatlanmıştır (Stadler et al. 2010). 8q24 den başka 17q ve 10q lokusları prostat kanser riski için tanımlanmış en önemli lokuslardır.

Ayrıca akciğer (15q25, 5q15.33), tiroid (9p22.33, 14q13.3), rahim (9p22), pankreas (5q15.33, 1q32.1), mide kanseri (8q24.3) ile ilişkilendirilmiş birçok lokus ve SNP bulunmaktadır (Stadler et al. 2010).

2.8. Kolorektal Kanser Araştırmalarında Genom Boyu Asosiyasyon Çalışmaları

Ailesel yatkınlığın vakaların yaklaşık %25' ine yol açtığı kolorektal kanserde, bilinen kalımsal yatkınlık sendromları (Lynch sendromu, FAP, Peutz-Jeghers sendromu, Juvenil Polipozis vb) hastalığın %10 'undan sorumludur. Geriye kalan %75 ise sporadik olgulardan oluşmaktadır.

Genetik heterojenite açısından güzel bir örnek teşkil eden kolorektal kanserde tek gendeki mutasyonlar (monogenik bozukluklar) ve bağlantı analizleri (poligenik kalıtım) veya genom boyu asosiyasyon çalışmaları (poligenik kalıtım) ile tanımlanan birçok lokustaki varyasyonlar kanser riskini artırmaktadır. Somatik mutasyona yol açabilen çevresel faktörlerin etkisiyle de bireysel kanser riski oluşabilir ve sonunda kanser gelişimi gerçekleşir (Galvan et al. 2010) (Şekil 2.12).



Şekil 2.12: Kolorektal kansere yatkınlıkta genetik heterojenitenin rolü

Son yıllarda kolorektal kanser arařtırmalarında genom boyunca asosiyasyon alıřmaları yaygın olarak kullanılmaktadır. Bu alıřmaların ortak zellikleri populasyon temelli asosiyasyon alıřmaları olmalarıdır.

Hem sporadik hem de aile yküsü olan kolorektal kanser olgularında yapılan genom boyu asosiyasyon alıřmaları ile kolorektal kanser geliřimi riskini etkileyen birok lokus tanımlanmıřtır (izelge 2.4). Bugüne kadar İngiltere, İskoya, Japonya, İsrail ve Kanada'da yapılan asosiyasyon alıřmaları ile 10 farklı yatkinlık lokusu (eřik deęeri $p < 13.5$) tanımlanmıřtır. Bu lokuslar 8q24.21, 18q21.1 (*SMAD7*), 15q13.3 (*GREM1*), 11q23.1, 8q23.3 (*EIF3H*), 10p14, 19q13.11 (*RHPN2*), 20p12.3, 14q22.2 (*BMP4*) ve 16q22.1 (*CDH1*) dir. Bu yatkinlık lokuslarından 11q23.1 lokusu hem baęlantı analizleri hem de genom boyu asosiyasyon analizleri ile tanımlanabilmiřtir (Tomlinson et al 2010.; Le Marchand 2009; Galvan et al. 2010).

izelge 2.4: Kolorektal kanserde yapılan genom boyu asosiyasyon alıřmaları ve tanımlanan tüm lokuslar

İlk Yazar	Yayınlandığı Tarih	Yayınlandığı Dergi	Makale Bařlığı	İlk ařamada alıřılan rnek sayısı	Doęrulama basamaęında alıřılan rnek sayısı	Rapor edilen blge	Rapor edilen gen/genler	Rapor edilen SNP
Cui	17.01.2011	BMJ	Common variant in 6q26-q27 is associated with distal colon cancer in an Asian population	1583 Japon olgu/1898 Japon kontrol	4584 Japon olgu/ 2596 Japon kontrol ve 225 Koreli olgu/377 Koreli kontrol	6q25.3, 8q24	SLC22A3	rs7758229, rs6983267, rs7837328

Houlston	24.10.2010	Nature Genetics	Meta-analysis of three genome-wide association studies identifies susceptibility loci for colorectal cancer at 1q41, 3q26.2, 12q13.13 and 20q13.33	3334 Avrupa kökenli olgu/ 4628 Avrupa kökenli kontrol	14851 Avrupa kökenli olgu/ 15569 Avrupa kökenli kontrol	1q41, 3q26.2, 12q13.12, 20q13.33,	LARP4,DIP2, LAMA5, DUSP10	rs6691170, rs6687758, rs10936599, rs11169552, rs7136702, rs4925386
Lascorz	07.07.2010	Carcinogenesis	Genome-wide association study for colorectal cancer identifies risk polymorphisms in German familial cases and implicates MAPK signalling pathways in disease susceptibility	371 Alman olgu/ 1263 Alman kontrol	4121 Alman, 794 Çek olgu/ 7344 Alman, 815 Çek kontrol	7p14.1	GLI3, INHBA	rs12701937
COGENT Study	16.11.2008	Nature Genetics	Meta-analysis of genome-wide association data identifies four new susceptibility loci for colorectal cancer	1902 olgu/1929 kontrol	4878 olgu/4914 kontrol	20p12.3, 14q22.2, 19q13.11, 16q22.1	BMP4, RHPN2, CDH1	rs4444235, rs9929218, rs10411210rs 961253
Tenesa	30.03.2008	Nature Genetics	Genome-wide association scan identifies a colorectal cancer susceptibility locus on 11q23 and replicates risk loci at 8q24 and 18q21	981 olgu, 1,002 kontrol	16,476 olgu, 15,351 kontrol	8q24.21, 18q21.1, 11q23.1	POU5FIP1, HsG57825, DQ515897, SMAD7,genler arası bir bölge	rs7014346 rs4939827 rs3802842

Tomlinson	30.03.2008	Nature Genetics	A genome-wide association study identifies colorectal cancer susceptibility loci on chromosome 10p14 and 8q23.3	922 olgu, 927 kontrol	17,872 olgu, 17,526 kontrol	8q23.3, 10p14	EIF3H, Genler arası bir bölge	rs16892766 rs10795668
Jaeger	30.01.2008	Nature Genetics	Common genetic variants at the CRAC1 (HMPS) locus on chromosome 15q13.3 influence colorectal cancer risk	730 olgu/960 kontrol	4500 olgu/3860 kontrol	15q13.3	CRAC1	rs4779584
Broderick	14.10.2007	Nature Genetics	A genome-wide association study shows that common alleles of SMAD7 influence colorectal cancer risk	940 olgu, 965 kontrol	7,473 olgu, 5,984 kontrol	18q21.1	SMAD7	rs4939827
Tomlinson	08.07.2007	Nature Genetics	A genome-wide association scan of tag SNPs identifies a susceptibility variant for colorectal cancer at 8q24.21	930 olgu, 960 kontrol	7,334 olgu, 5,246 kontrol	8q24.21	Genler arası bir bölge	rs6983267

Zanke	08.07.2007	Nature Genetics	Genome-wide association scan identifies a colorectal cancer susceptibility locus on chromosome 8q24	1,257 olgu, 1,336 kontrol	6,223 olgu, 6,443 kontrol	8q24.21	Açık okuma çerçevesi DQ515897	rs10505477
-------	------------	-----------------	---	---------------------------	---------------------------	---------	-------------------------------	------------

Kolorektal kanserde tanımlanan yakınlık lokuslarının neredeyse yarısı bağlantı dengesizliğinde (LD) veya daha önce karsinogenez ile ilişkilendirilmiş TGF- β sinyal yolağı genlerine yakın bir mesafede bulunmaktadır. Genom boyu asosiyasyon çalışmaları ile tanımlanan ve TGF- β sinyal yolağında bulunan genler arasında SMAD7, RHPN2, BMP4, BMP2 ve GREM1 sayılabilmektedir.

Genom boyu asosiyasyon çalışmaları ile tanımlanan bu varyantların toplu etkisi kolorektal kanser riskinin sadece küçük bir parçasını açıklayabilmektedir (Varghese et al. 2010; Galvan et al. 2010; Stadler et al. 2010). Bu veya diğer risk varyantlarının kolorektal kanser riski açısından, çevresel etkenlerle nasıl etkileşime girebileceği hala tam olarak bilinmemektedir.

2.9. Amaç ve Mantıksal Temeller

Gelişmiş ülkelerde kanser nedenli ölümlerin yaygın nedeni olan kolorektal kanserde kalıtsal yatkınlık, bütün olguların yaklaşık %30'undan sorumlu iken, APC, MMR, MYH, SMAD4, BMPR1A ve STK11/LKB1 genlerindeki yüksek risk gösteren eşey hattı mutasyonları ise, yine yaklaşık olarak tüm kolorektal kanser olgularının %5'inden sorumludur.

Genetik heterojenitenin çok net bir şekilde izlendiği kolorektal kanserde, tek bir gendeki bozukluk kansere yatkınlığın oluşmasına neden olabileceği gibi poligenik durumlar da yatkınlığa neden olabilmektedir. Ayrıca çevresel faktörlerin etkisi ile oluşabilecek somatik mutasyon veya mutasyonlar bireysel kanser riski oluşturabilmektedir.

Kolorektal kanser araştırmalarında yaygın olarak kullanılan genom boyu asosiyasyon çalışmalarında, hem sporadik hem de aile öyküsü olan kolorektal kanser olguları kullanılmıştır. Bugüne kadar yapılan genom boyu asosiyasyon çalışmalarının ortak özellikleri ise popülasyon temelli asosiyasyon çalışmaları olmalarıdır.

Bugüne kadar İngiltere, İskoçya, Japonya, İsrail ve Kanada'da yapılan asosiyasyon çalışmaları ile eşik değerleri $p < 10^{-5}$ olan ve aşağıda adı geçen 10 farklı yatkınlık lokusu tanımlanmıştır (Şekil 2.13).

Yakın zamanda yapılan genom boyu asosiyasyon çalışmaları, kolorektal kansere yatkınlık oluşturan 8q24.21, 8q23.3, 10p14, 11q23.1, 14q22.2, 15q13.3, 16q22.1, 18q21.1, 19q13.11 ve 20p12.3 lokuslarını tanımlayarak, kolorektal kanserde kalıtsal riskin yaygın ve düşük riskli varyantlardan kaynaklandığı hipotezini doğrulamışlardır.

Bugüne dek popülasyon yapısı, bulunduğu coğrafya temel alındığında özgünlük taşıyan Türk popülasyonunda herhangi bir kanser tipinde yapılmış olan genom boyu asosiyasyon çalışması bulunmamaktadır.

Popülasyon temelli bir asosiyasyon çalışmasının lojistik ve ekonomik olarak ülkemizde gerçekleştirilmesi mümkün görünmemektedir.

Bu bilgilerden de anlaşılacağı üzere genetik heterojenitesi çok yüksek olan kolorektal kanserin Türk popülasyonunda da araştırılması çok büyük önem taşımaktadır.

T.C Saęlık Bakanlıęı Kanserle Savař Daire Bařkanlıęının 2004-2006 yılları kanser istatistięine gre, kadınlarda grlen kanserler iinde 2. sırada yer alan kolorektal kanser erkeklerde ise 4. sırada yer almaktadır. lkemiz iin nemli saęlık problemlerinin bařında gelen kolorektal kanserin aydınlatılmasının nemi aıktır.

Bu tez alıřması sporadik kolorektal kanserde aile temelli yaklařımla yapılmıř olması nedeni ile bir ilk teřkil edecek olmanın yanısıra Trk poplasyonuna zg verilerin ortaya ıkarılmasına da olanak saęlayacaktır.

3. MATERYAL VE YÖNTEM

3.1. Materyal

Bu tez çalışması için, ailesinde (birinci derece akrabalarında) kanser hikayesi bulunmayan kolorektal kanser vakalarından ve bu vakalarının sağlıklı (kansere geçirmemiş) anne ve babalarından (hasta-anne-baba olmak üzere üçlü) periferik kan örnekleri toplanmıştır. Materyal bulmanın kısıtlılığı nedeni ile tek ebeveyni hayatta olan hastaların (hasta-anne veya hasta-baba olmak üzere ikili) kan örnekleri de çalışmaya dahil edilmiştir. Bu çalışma için etik kurul onayı ve kan örnekleri alınan bireylerden bilgilendirilmiş onam formu alınmıştır.

Projede, 41 kadın, 43 erkek toplam 84 sporadik kolorektal kanser hastası ile çalışılmıştır. Kadınlarda yaş ortalaması 37 iken erkeklerde yaş ortalaması 46 olarak gözlenmektedir.

Proje süresince toplamda 51 trio ve 33 duo (Hasta-Anne veya Hasta-Baba) materyaline ulaşılmıştır.

Ayrıca çalışmanın doğrulama basamağında kullanılmak üzere toplam 61 sporadik kolorektal kanser hastasından ve 80 kontrol bireyden periferik kan toplanmıştır.

3.2. Yöntem

3.2.1. DNA İzolasyonu

DNA izolasyonu hastalar ile sağlıklı anne ve/veya babalarından alınan periferik kandan fenol-kloroform-izoamilalkol izolasyon yöntemi kullanılarak gerçekleştirildi.

Fenol-Kloroform-İzoamilalkol DNA izolasyonunda kullanılacak solüsyonlar hazırlandı.

RBC Lizis Solüsyonu (Red Blood Cell Lysis Solution)

12 g NH₄Cl (Amonyum Klorit) (Sigma A9434)

1,2 g Na₂CO₃ (Sodyum Karbonat) (Sigma S7795)

200 µl 0,5 M pH:8,0 EDTA (Etilen -Diamin-Tetra Asetikasit) (Sigma E5134)

Az miktar suda çözülerek son hacim distile su ile 1000 ml'ye tamamlandı.

Sodyum Tris EDTA (STE)

0.1 M NaCl (S5886)

10 mM TrisHCl pH:8,0 (Sigma T5941)

1 mM EDTA pH:8,0 (Sigma E5134)

%10'luk Sodyum Dodecyl Sülfat (SDS)

100 g SDS (Scharlau SD0010) alınıp 900 ml distile suda 68°C'ye ayarlanmış su banyosunda (Lab-Line Aquabath) eritildi ve toplam hacim 1000 ml'ye tamamlandı. HCl (Riedel-de Haen 07102) ile pH metrede (CRISON pH meter Basic 20) pH:7,2'ye ayarlandı.

0,2 µm çaplı filtreden (Sartorius Minisart) geçirilerek steril edildi.

Sodyum Asetat

408,1 g Sodyum Asetat (Sigma S2889) 800 ml distile suya ilave edilerek manyetik karıştırıcı (Daihan Scientific) yardımıyla çözüldü. Asetik asitle pH:7,0'a ayarlanarak son hacim 1000 ml'ye tamamlanarak otoklavlandı (Hirayama Hiclave HVA-110). Kan hacminin 2,5 katı kadar RBC lizis solüsyonu eklenerek çalkalandı ve 20 dakika buzda bekletildi. 4°C'de 4000 rpm'de 20 dakika santrifüj (Hettich Mikro 22 R) edildi. Süpernatant alındı ve dipte lökosit tabakası elde edilene kadar işlem tekrarlandı. Pelet elde edilince üstteki sıvı yavaşça döküldü. 1000 µl RBC lizis solüsyonu eklenerek vortekslendi (Lms Vtx-3000L). 200 µl örnek tüpe alındı üzerine 500 µl STE, 30 µl %10'luk SDS ve 20 µl Proteinaz K (600 u/ml, 200 mg/ml, Fermentas EO0491) ilave edildi. Tüpün ağzı parafilm ile kapatılarak 56°C'ye ayarlanan su banyosunda bir gece inkübasyona bırakıldı. Ertesi gün su banyosundan alınan tüpüne 25:24:1 oranındaki Fenol:Kloroform:İzoamilalkol (Fluka 77617) karışımından 750 µl ilave edildi. Fenol:Kloroform:İzoamilalkol karışımı kullanılmadan önce çalkalandı. Tüpün ağzı kapatılıp parafilmleildi. Beyaz renk alıncaya kadar vorteks yardımıyla karıştırıldı. 20 dakika buzda bekletildi ve 4°C'de 4000 rpm'de 20 dakika santrifüj edildi. Pipet ile tüpün yüzeyinde oluşan faz alınarak başka bir tüpe konuldu. Temiz tüpe alınan süpernatant üzerine 750 µl kloroform eklendi ve vortekslendi. 20 dakika buzda bekletildikten sonra

4°C’de 4000 rpm’de 20 dakika santrifüj edildi. Tüpün yüzeyinde oluşan faz pipet yardımıyla alınarak temiz bir tüpe aktarıldı üzerine 1500 µl % 95’lik etanol (AppliChem A4230) ve 75 µl Sodyum Asetat eklendi. DNA beyaz yumak şeklinde görünür hale gelinceye kadar tüp yavaşça alt üst edildi. Tüpler -20°C’de bir gece bekletildi. Ertesi gün -20°C’den alınan tüp 4°C’de 4000 rpm’de 20 dakika santrifüj edilerek DNA çöktürüldü. Tüpteki alkol dikkatlice döküldü ve 500 µl %70’lik soğuk etanol eklendi. 4°C’de 4000 rpm’de 20 dakika santrifüj edildi. DNA kaybedilmeden alkol uzaklaştırıldı. Tüpün ağzı açık bırakılarak alkolün iyice uçması sağlandı. Kurutma işleminden sonra DNA üzerine 100 µl distile su veya TE tampon eklendi. Tüpün ağzı parafilmlelenerek 37°C’de su banyosunda bir gece bekletildi. Daha sonra kullanılmak üzere DNA 4°C’de saklandı.

3.2.2. DNA’nın Spektrofotometrik Analizi

DNA örneklerinin saflığı ve konsantrasyonunun belirlenmesi için spektrofotometre (NanoDrop ND 1000) kullanılarak ölçüm yapıldı. 260 nm DNA’nın, 280 proteinin ve 230 ise fenolün maksimum absorpsiyon verdiği dalga boyudur. DNA’nın saflığı değerlendirilirken A260/A280 ve A260/A230 oranlarına bakılmaktadır. 260/280 oranının ~1,8 ve A260/ A230 oranının ise 1,8-2,2 olması DNA’nın saf olduğunu göstermektedir. Protein, fenol veya diğer kontaminantların varlığı durumunda oranlar farklı çıkmaktadır.

3.2.3. DNA’nın Agaroz Jel Elektroforezi

DNA örneklerinin degrade olup olmadığının belirlenmesi için %1 lik agaroz jelde yürütüldü.

Kullanılan Solüsyonların Hazırlanması

5xTBE (Tris-HCl, Borik asit, EDTA)

54 g Trisma Base (Sigma T6066), 27,5 g Borik asit (Sigma B6768) ve 20 ml 0,5 M EDTA (Sigma E5134) deiyonize suyla (ddH₂O) 1000 ml’ye tamamlanarak hazırlandı.

Ethidium Bromide (10 mg/ml)

Etidyum bromid (Scharlau ET0109) son konsantrasyon 10 mg/ml olacak şekilde sulandırıldı.

DNA Ladder (Fermentas Gene Ruler 100 bp DNA Ladder Plus)

DNA Ladder'a 166 µl 6x Loading Dye (Fermentas 0611) ve 734 µl distile su eklenerek vortekslendi.

%1'lik Agaroz Jel

Agaroz (Sigma A 5093-500G) 1 g/100 ml (%1) olacak şekilde hassas terazide (Shimadzu BX320H) tartıldı. Üzerine 1x TBE (Trisma Base EDTA) ilave edildi. 1xTBE solüsyonu stok olarak 5xTBE solüsyonundan su ile 1/5 oranında seyreltilerek hazırlandı. Agaroz istenilen yüzdede hazırlandıktan sonra mikrodalga fırında (Vestel, Türkiye) çözülmesi sağlandı. Çeker ocak altında üzerine 5 µl etidyum bromid ilave edildi. İyice karıştırıldıktan sonra jel tablasına dökülerek donması için 23.30 dakika bekletildi. %1'lik agaroz jel elektroforez tankına yerleştirildi. 2 µl DNA, 4 µl 6x loading dye ile karıştırılarak jele yüklendi. Örnekler 100 voltta 30 dakika yürütüldü. UV görüntüleme sistemi (Syngene Gene Genius Bio Imaging System) kullanılarak görüntülendi ve termal yazıcı (SonyUP-D895) ile çıktı alındı.

3.3. 250K DNA Mikrodizin

3.3.1. Restriksiyon Enzim ile DNA'nın kesilmesi

Çizelge 3.1 izlenerek buz üzerinde DNA kesim karışımı hazırlandı, 2 saniye orta derecede vortekste karıştırıldı ve 2000 rpm'de 1 dakika santrifüj edildi.

Çizelge 3.1: Kesim karışımı hazırlanışı

Reaktif	Hacim
dH ₂ O	15 µl
Nsp (10U/µL)	1,25µl
BSA'lı 10X Tango Tamponu	2,5 µl
Genomik DNA (50ng/µl)	6,25 µl

Reaksiyon da hızlı kesen “Fast Digest” (Fermantas Nsp Fast Digest FD1474) enzim kullanıldı. 37°C’de 5 dakika inkübe edildi. İnkübasyon sonrası 65°C’de 5 dakika enzim inaktivasyonu gerçekleştirildi.

3.3.2. Ligasyon

Çizelge 3.2 izlenerek buz üzerinde ligasyon karışımı hazırlandı, 2 saniye orta derecede vortekste karıştırıldı ve 2000 rpm’de 1 dakika santrifüj edildi. 16°C’de 3 saat, 70°C’de 20 dakikada reaksiyon gerçekleştirildi. Ligasyonu yapılmış DNA’lar PCR reaksiyonuna geçmeden önce 75µl PCR’a uygun su ile dilüe edildi.

Çizelge 3.2: Ligasyon karışımı hazırlanışı

Reaktif	Hacim
Adaptor (50 µM)	0,75 µl
T4 DNA Ligaz Tamponu (10X)	2,5 µl
Kesilmiş DNA	19,75 µl
T4 DNA Ligaz (400U/ µl)	2 µl

3.3.3. PCR ve PCR Sonrası Jel Elektroforezi

PCR da kullanılacak solüsyonlar Çizelge 3.3 de görüldüğü gibi hazırlandı.

5M Betain

17,5725 gr betain (Fluca 61962) tartıldı son hacim 30 ml olacak şekilde su ile tamamlandı. 0,2 µm çaplı filtreden (Sartorius Minisart) geçirilerek steril edildi. Alikotlanıp +4 C de saklandı.

Örneğe bağlı olarak 3 ila 5 adet PCR ürünü, 1 arrayin hibridizasyonu için gerekli miktarı sağlamaktadır. Bu nedenle örnekler için 5 PCR reaksiyonu kuruldu. Kontaminasyon olup olmadığını anlamak için mutlaka negatif kontrol kullanıldı.

Tablo 4 gösterilen miktarlar çalışılan örnek için yeterli olacak şekilde hazırlandı ve 2 saniye orta derecede vortekste karıştırıldı. Karışım 2000 rpm'de 1 dakika santrifüj edildi. MJ DNA Engine Tetrad'da 94°C'de 3 dakika, 94°C'de 30 saniye, 60°C'de 30 saniye, 68°C'de 15 saniye, 68°C'de 7 dakika olmak üzere 30 döngü PCR reaksiyonu yapıldı. 3µl PCR ürünü 3 µl 2xJel yükleme boyası ile karıştırılarak 2% agaroz jele yüklendi ve 120 voltta yaklaşık 1 saat yürütüldü.

Çizelge 3.3: PCR karışımı hazırlanışı

Reaktif	Hacim
10X Dream Taq PCR Tamponu (Fermantas)	10µl
dNTP (2,5mM)(Fermantas)	14µl
MgCl ₂ (25mM)(Fermantas)	14µl
PCR Primer 002 (100pmol)	4,5µl
Dream Taq DNA Polimeraz (500U/ µl)(Fermantas)	4µl
Betain (5M)	20µl
dH ₂ O	23,5µl
Dilüe edilmiş adaptör takılmış DNA	10µl

3.3.4. PCR Purifikasyon ve Elüsyonu

PCR pürifikasyonunda kullanılacak solüsyonlar hazırlandı.

0,1M EDTA

37.22 gr EDTA (Sigma E5134) tartıldı ve son hacim 1 ml olacak şekilde su ile tamamlandı.

PCR purifikasyonu ve elüsyonu Clontech clean-up plate ile yapıldı. Her PCR reaksiyonunun üzerine 8 µl 0.1 M EDTA eklendi ve 2 saniye orta derecede vortekste karıştırıldı. Karışım 2000 rpm'de 1 dakika santrifüj edildi. Clean-up plate'nin her bir kuyusunun içine her bir örneğin 3 adet PCR reaksiyonu eklendi. Purifikasyon plate manifoldun üzerine yerleştirildi. Plate üzerinde kullanılmayan yerler alüminyum seal ile kapatıldı. 600 mbar vakumda purifikasyon gerçekleştirildi. Kuyular tamamen kuruyana kadar vakum uygulandı. Kuyular tamamen kuruduktan sonra 50 µl PCR suyu eklenerek yıkama işlemi gerçekleştirildi ve bu yıkama basamağı 2 kere daha tekrarlandı. Her yıkama arasında membranın tamamen kuruması beklendi. Son yıkamadan sonra, her kuyu tamamen kuruduktan sonra elüsyon tamponu eklenerek elüsyon yapıldı. Yıkama basamakları bittikten sonra vakum kapatıldı ve plate manifolddan çıkarıldı. Her kuyuya 45 µl RB buffer eklendi. Plate, karıştırıcıya yerleştirilerek oda sıcaklığında 10 dk karıştırıldı. Tüm kuyulardaki PCR ürünleri temiz tüplere alındı.

3.3.5. Purifiye PCR Ürünlerinin Miktar Ölçümü

Pürifiye PCR ürünlerinin konsantrasyonları spektrofotometrede (Nanodrop) ölçüldü.

3.3.6. Fragmentasyon ve Fragmentasyon Sonrası Jel Elektroforezi

Çizelge 3.4 izlenerek buz üzerinde fragmentasyon karışımı hazırlandı, 2 saniye orta derecede vortekste karıştırıldı. Karışım 2000 rpm'de 1 dakika santrifüj edildi.

Fragmentasyon tüpleri önceden 37 derecede ısıtılmış termal cycler'a yerleştirildi. 37°C'de 35 dakika ve 95°C'de 15 dakikada reaksiyon gerçekleştirildi. Reaksiyon bittikten sonra 4 µl fragmente edilmiş PCR ürünü ile 4 µl jel yükleme boyası karıştırılarak 4% agaroz jele yüklenerek 120 voltta 30 dakika yürütüldü.

Çizelge 3.4: Fragmentasyon karışımı hazırlanışı

Reaktif	Hacim
Purifiye PCR ürünü (90 µg in RB tamponu)	45 µl
10xFragmentasyon Tamponu	5 µl
Dilue fragmentasyon reaktifi (0.05U/µl)	5 µl

3.3.7. İşaretleme

Çizelge 3.5 izlenerek buz üzerinde işaretleme (labeling) reaksiyonu hazırlandı, 2 saniye orta derecede vortekste karıştırıldı. Karışım 2000 rpm'de 1 dakika santrifüj edildi. 37°C'de 4 saat ve 95°C'de 15 dakikada işaretleme reaksiyonu gerçekleştirildi.

Çizelge 3.5: İşaretleme karışımı hazırlanışı

Reaktif	Hacim
5X TdT Tamponu	14µl
GeneChip DNA Labeling Reaktifi (30 mM)	2µl
TdT (30 U/µl)	3.5µl
Fragmente edilmiş DNA	50.5µl

3.3.8. Hibridizasyon

Çizelge 3.6 izlenerek buz üzerinde hibridizasyon karışımı hazırlandı, 2 saniye orta derecede vortekste karıştırıldı. Karışım 2000 rpm’de 1 dakika santrifüj edildi.

Çizelge 3.6: Hibridizasyon karışımı hazırlanışı

Reaktif	Hacim
MES (12X; 1.22 M)	12µl
DMSO (100%)	13µl
Denhardt’s Solusyonu (50X)	13µl
EDTA (0.5 M)	3µl
HSDNA (10 mg/mL)	3µl
Oligonükleotid Kontrol	2µl
Human Cot-1 (1 mg/mL)	3µl
Tween-20 (3%)	1µl
TMACL (5M)	140µl
İşaretli DNA	70µl

260 µl hibridizasyon karışımı ve işaretli DNA 99°C ısı bloğunda 10 dakika denatüre edildi ve 10 saniye buzda bekletildi. 2000 rpm’de 1 dakika santrifüj edildi. Tüpler 49°C’ki ısı bloğunda 1 dakika tutuldu. Oda ısısına getirilen array içine 200 µl denatüre edilmiş hibridizasyon karışımı enjekte edildi. Denatürasyon sırasında tüpte çökelek oluşmuşsa, solüsyonu arraye enjekte etmeden önce bu çökelek pipet yardımıyla karıştırıldı. Hibridizasyon fırınında (GeneChip Hybridization Oven 640, 800139) 49°C’de 18 saat 60 rpm’de hibridizasyon reaksiyonu gerçekleştirildi.

3.3.9. Yıkama, Boyama ve Tarama

Yıkama ve boyamada kullanılacak solüsyonlar hazırlandı.

12 X MES Stok

(1.22 M MES, 0.89 M [Na+])

1000 ml için:

70.4 g MES Hydrate

193.3 g MES Sodium Salt

800 ml ultra saf su

Karıştırıldı ve son hacim 1,000 ml, pH 6.5 ve 6.7 arasında olacak şekilde ayarlandı.

0.2 µM filtreden geçirildi.

Wash A: Non-Stringent Wash Buffer

(6X SSPE, 0.01% Tween 20)

1000 ml için:

300 ml of 20X SSPE

1.0 ml of 10% Tween-20

699 ml ultra saf su

0.2 µm filtreden geçirildi. Oda sıcaklığında saklandı.

Wash B: Stringent Wash Buffer

(0.6X SSPE, 0.01% Tween 20)

1000 ml için:

30 mL of 20X SSPE

1.0 mL of 10% Tween-20

969 ml ultra saf su

0.2 µm filtreden geçirildi. Oda sıcaklığında saklandı.

0.5 mg/mL Anti-Streptavidin Antibody

1 ml suda çözüldü ve 4°C'de saklandı.

1X Array Holding Buffer

(1X 100 mM MES, 1M [Na+], 0.01% Tween-20)

100 ml için:

8.3 mL of 12X MES Stok Tamponu

18.5 mL of 5M NaCl

0.1 mL of 10% Tween-20

73.1 ml distile su karıştırıldı.

Işıktan koruyarak 2°C-8°C'de saklandı.

1X Stain Buffer

800,04 µl H₂O

360 µl SSPE (20X)

3.96 µl Tween-20 (3%)

24 µl Denhardt's (50X) karıştırıldı . Her çalışmada taze olarak hazırlandı.

18 saat hibridizasyondan sonra arrayden hibridizasyon karışımı çıkarıldı ve array 200 µl Array Holding Tamponu ile tamamen dolduruldu.

Aşağıdaki çizelgeler (3.7 ve 3.8) izlenerek boyama kimyasalları hazırlandı, 2 saniye orta derecede vortekste karıştırılır. Karışım 2000 rpm’de 1 dakika santrifüj edildi. SAPE içeren tüp, yıkama istasyonunun (GeneChip Fluidics Station 450, 03.0079) 1. modülüne yerleştirildi. Antibody içeren tüp istasyonunun 2. modülüne yerleştirildi. 3. amber tüp içine 820 µl array holding tamponu eklenerek istasyonun 3. modülüne yerleştirildi. Yıkama ve boyama bittikten sonra array tarandı (GeneChip Scanner 3000 7G , 03.0205). Tarama sonucunda ilgili (Affymetrix GeneChip Operating Software 1.4 690031 ve Affymetrix GeneChip DNA Analysis Software 4.0, 690051) programlar kullanılarak analiz dosyaları oluşturuldu.

Çizelge 3.7: Boyama karışımının hazırlanışı (1.tüp)

Reaktif	Hacim
Boyama Tamponu	594µl
1mg/mL SAPE	6µl

Çizelge 3.8: Boyama karışımı hazırlanışı (2.tüp)

Reaktif	Hacim
Boyama Tamponu	594µl
0.5 mg/mL biotinylated antibody	6µl

3.4. Biyoinformatik Analizler

3.4.1. Ardışık Hipotez Testleri (Sequential Probability Ratio Test, SPRT)

Bu tez çalışmasında elde edilen 51 trio'luk veri, ekibimiz tarafından önerilen ve az sayıda trio ile sonuç üretilebileceğini gösteren SPRT analizi (Ilk et al. 2011) ile analiz edilmiştir.

3.4.2. Birleşik Dengesiz Aktarım Testi (C-TDT) ve Permütasyon Testleri

Tez kapsamında trioların yanısıra toplanan duolarında analiz edilebilmesi için SPRT'nin yanısıra, trio ve duo haplotip verilerini bir arada inceleyebilen C-TDT ve permütasyon testleri de kullanılmıştır. Böylece, sadece trioları kullanan TDT veya sadece duoları kullanan 1-TDT testlerinden daha güçlü (daha fazla veri kullanması nedeniyle) bir test kullanılmakla birlikte, permütasyon testi sayesinde herhangi bir dağılım varsayımı yapılmadan ve küçük örneklemelere uygulanabilen bir yöntem ortaya çıkmıştır. Bu yöntemin, bu projede toplanan verilere uygulanabilmesi için MATLAB kodu yazılmış ve elimizdeki tüm verilerle (51 trio ve 33 duo) analiz yapılmıştır.

C-TDT istatistiğinin kullanımı şöyle bir örnek üzerinden anlatılabilir. Öncelikle, sadece bir tane hasta çocuğu olan anne ve/veya baba dan oluşan trio ya da duo aile grubundan oluşan bir aile temelli genom bazlı bir asosiyasyon çalışması ele alalım. Her bir SNP belirteç lokusu için her bir kişinin genotipi AA, AB veya BB ile gösterilebilsin.

Bu durumda C-TDT istatistiği aşağıdaki gibidir:

Bu formülde, anne-baba-çocuk üçlüleri için:

T = heterozigot anne ve babalardan taşıyıcı çocuğa A alelini aktarıp, fakat B alelini aktaramayan toplam anne ve baba sayısını,

NT = heterozigot anne ve babalardan taşıyıcı çocuğa A alelini aktaramayıp, fakat B alelini aktaran toplam anne ve baba sayısını göstermektedir.

Sadece bir ebeveynin bulunduğu aileler (anne-çocuk ikilisi ve baba-çocuk ikilisi) için,

n_m = toplam anne-çocuk ikilisi sayısını ve

n_p = toplam baba-çocuk ikilisi sayısını gösterirken, $\omega = n_p/(n_m+n_p)$ formülü ile hesaplanmaktadır.

Anne-çocuk ikilileri için,

$N_{M<C} = \sum I_{M<C}$ = annenin taşıyıcı çocuktan daha az A alel kopyası taşıdığı anne-çocuk ikilisi sayısını gösterirken,

$N_{M>C} = \sum I_{M>C}$ = annenin taşıyıcı çocuktan daha fazla A alel kopyası taşıdığı anne-çocuk ikilisi sayısını gösterir.

Baba-çocuk ikilileri için,

$N_{F<C} = \sum I_{F<C}$ ve $N_{F>C} = \sum I_{F>C}$ değerleri benzer şekilde bulunur. Son olarak da, $N_{F<C}$ ve $N_{F>C}$ olmak üzere, yukarıda bulunan değerlerden faydalanarak hesaplanır.

C-TDT istatistiği, malesef, TDT istatistiği gibi, boş hipotezi altında belirgin bir dağılıma sahip değildir. Belirgin bir dağılımın olmaması, bizleri test istatistiğinin önemliliğini test etmekten alıkoymaktadır. Bu nedenle, Hu ve Zhou (2010) C-TDT istatistiğinin önemlilik testi için, ampirik bir dağılım yaratma imkânı vermesi açısından permütasyon yöntemini kullanmaktadırlar. Permütasyon yöntemi orijinal verideki gözlemlerin yerlerini, önceden belirlenmiş kurallar çerçevesinde, ve belirtilen permütasyon sayısı kadar değiştirerek yeni veriler yaratıp, bu yeni yaratılmış olan her veri için de bahsi geçen test istatistiğinin tek tek hesaplanmasına dayanır.

Zhao ve arkadaşları (2000) anne-baba-çocuk üçlüsü genotip verisinin sırasını değiştirmek için bir permütasyon yöntemini önermiştir. Hu ve Zhou (2010) ise anne-çocuk ikilisi ve baba-çocuk ikilisi genotip verisinin sırasını değiştirmek için kendi permütasyon yöntemlerini önermektedir. Sonrasında, Hu ve Zhou (2010) kendi yöntemleri ile Zhao ve

arkadaşlarının (2000) yöntemini de birleştirerek, C-TDT istatistiği için permütasyonun nasıl yapılacağını anlatmaktadır (Zhao et al. 2000; Hu et al. 2010).

Bu bağlamda, Hu ve Zhou'nun (2010) makalelerinde de belirtildiği üzere C-TDT'nin permütasyonu için gerekli aşamalar aşağıdaki gibi sıralanabilir.

1. Öncelikle, herhangi bir belirteç için, anne-baba-çocuk üçlüsü, anne-çocuk ikilisi ve baba-çocuk ikilisinden gelen genotip bilgilerine bakarak, C-TDT test istatistiğinin değeri yukarıda verilen formülden hesaplanır ve kaydedilir.
2. Permütasyon sayısı farzedelim ki, $S=1000$ olarak alınır.
3. Her bir permütasyon altında, anne-baba-çocuk üçlüsü genotip verisinin ve bir ebeveyn-çocuk ikilisi genotip verisinin sıraları ayrı ayrı değiştirilmek üzere iki parçadan oluşur. Şöyle ki;
 - a) Anne-baba-çocuk üçlülerinden oluşan genotip verisi için, Zhao ve diğerlerinin (2000) önerdiği permütasyon yöntemi şu şekilde yapılır. Taşıyıcı çocuğa, anne ve babasından aktarılmış, kendi gerçek genotipi olsun veya olmasın, o belirteçde o anne ve babanın genotiplerinin eşleşmesinden ortaya çıkacak tüm olası genotiplerden bir tanesi, eşit olasılıkla rasgele olarak, taşıyıcı çocuğa, yeni genotip olarak atanır. Böylece anne-baba-çocuk üçlülerinde, taşıyıcı tüm çocukların genotipleri değiştirilmiş olur ve yeni bir anne-baba-çocuk üçlüsü genotip verisi yaratılmış olur.
 - b) Tek ebeveyn-çocuk ikililerinden oluşan genotip verisi için Hu ve Zhou'nun (2010) önerdiği permütasyon yöntemi, biraz daha karışık olmak üzere, aşağıda verilen basamaklardan oluşmaktadır.
 - i) $\{(C_i, M_i), 1 \leq i \leq n_m\}$ ve $\{(C_i, F_i), n_m + 1 \leq i \leq n_m + n_p\}$ dizileri sırasıyla n_m tane anne-çocuk ikilisinin genotiplerini ve n_p tane baba-çocuk ikilisinin genotiplerini gösterebilir. Burada, i .nci ebeveyn-çocuk ikilisinde, $C_i/M_i/F_i$ sırasıyla, çocuk/anne/ baba genotiplerini belirtir.
 - ii) $\{1, \dots, n_m, n_m+1, \dots, n_m+n_p\}$ dizisi permüte edilsin ve $\{j_1, \dots, \dots, \dots\}$ dizisine dönüşmüş olsun. Böylece, n_m tane anne-çocuk ikilisinin yeni genotipleri, aşağıda tanımlanan şekilde elde edilebilir.

uygulanmıştır. Başka bir deyişle, mevcut 262,264 SNP için, 1-5 basamaklarını hesaplayarak, aralarından hastalıkla istatistiksel olarak önemli ilişkili olanların bulunması hedeflenmiştir.

3.4.2.1. Birleşik Dengesiz Aktarım Testi (C-TDT) ve Permütasyon Testleri için Benzetim Çalışması

Hu ve Zhou makalelerinde yaptıkları benzetim çalışmalarında, çalışmanın anne-baba-çocuk üçlüsü ve çocuk-ebeveyn ikilisinden oluşması durumunda, C-TDT istatistiğinin, sadece anne-baba-çocuk üçlüsü kullanan TDT'ye göre daha güçlü sonuçlar verdiğini göstermektedir (Hu et al. 2010). Ne var ki, bu ve literatürdeki diğer çalışmalar, elimizde az sayıda veri olduğunda C-TDT'nin istatistiksel gücü hakkında bilgi vermemektedir. Literatürdeki bu açığı kapatmak ve elimizdeki verinin yeterli olup olmadığını anlamak amacıyla, bu proje kapsamında, C-TDT test istatistiğinin farklı örneklem büyüklükleri karşısındaki gücünü ölçmek için bir benzetim çalışması düzenlenmiştir. Bu bağlamda,

10 anne-baba-çocuk üçlüsü ve 4 anne veya baba-çocuk ikilisi,

30 anne-baba-çocuk üçlüsü ve 18 anne veya baba-çocuk ikilisi,

50 anne-baba-çocuk üçlüsü ve 30 anne veya baba-çocuk ikilisi,

70 anne-baba-çocuk üçlüsü ve 40 anne veya baba-çocuk ikilisi,

90 anne-baba-çocuk üçlüsü ve 55 anne veya baba-çocuk ikilisi,

110 anne-baba-çocuk üçlüsü ve 65 anne veya baba-çocuk ikilisi ve

130 anne-baba-çocuk üçlüsü ve 75 anne veya baba-çocuk ikilisi olmak üzere 7 farklı örneklem büyüklüğüne karar verilmiştir.

Her bir örneklem büyüklüğü altında simülasyon çalışması 100 kez tekrar edilmiştir. Herhangi bir örneklem büyüklüğü altında simülasyon çalışması yaparken, gerçek biyolojik veride olduğu üzere, bir kişinin toplam 262,264 SNP'nin gözlemlendiği varsayılmıştır.

İlk olarak anne-baba-çocuk üçlülerinin SNP'ler boyunca genotip değerlerinin üretiminden bahsetmek gerekirse, öncelikli olarak, belirtilen sayıda anne ve babanın genotipleri

262,264 SNP boyunca rastgele olarak üretilmiştir. Herhangi bir aile için, herhangi bir SNP üzerinde anne ve babanın genotipleri gözlemlendikten sonra, bu anne-baba çiftinden olacak bir çocuğun aynı SNP üzerinde alabileceği genotip değerlerini belirlemek için çizelge 3.9'dan faydalanılmıştır.

İstatistiksel güç analizi yapabilmek için, çocuğun genotipi için şu şekilde bir yol izlenmiştir. Herhangi bir çocuğa ait 262,264 SNP'in % 98'lik kısmını oluşturan 257,019 SNP, C-TDT test istatistiğinde hipotezi lehine katkıda bulunacağını düşündüğümüz şekilde üretilmiştir. Bu sebeple, aşağıdaki tabloda görüleceği üzere, C-TDT test istatistiğinde yer alan veya şartlarını yerine getiren genotiplere yüksek olasılıklar verilerek, diğer kalan genotiplere ise düşük olasılıklar verilerek üretim yapılmıştır. Örneğin, çizelge 3.9 da görüleceği gibi, annenin genotipinin AB, babanın genotipinin AB olarak gözlenmesi durumunda, çocuk 0.25 olasılıkla AA, 0.5 olasılıkla AB, 0.25 olasılıkla BB değeri alabilecekken, çocuğun AB olması yönünde destekleyici olacağından, çocuğun AB olma olasılığı 0.5'ten daha yüksek tutulmuştur.

Geriye kalan 5245 SNP (%2) ise aleyhinde, yani, istatistiksel olarak önemli asosiyasyon verecek şekilde üretilmiştir. Bu sebeple, aşağıdaki tabloda görüleceği üzere, C-TDT test istatistiğinde yer alan , şartlarını yerine getiren genotiplere daha yüksek olasılıklar verilerek, diğer kalan genotiplere ise daha düşük olasılıklar verilerek üretim yapılmıştır. Aslında durumu da 'i destekleyici veri olsa da, bu koşulda üretilen veri koşuluyla üretilen veriyi C-TDT formülünde iptal edeceğinden kullanılmamıştır. Örneğin, aşağıdaki tablo11 de görüleceği gibi, annenin genotipinin AA, babanın genotipinin AB olarak gözlenmesi durumunda, çocuk 0.5 olasılıkla AA, 0.5 olasılıkla AB değeri alabilecekken, çocuğun AA olması yönünde destekleyici olacağından, çocuğun AA olma olasılığı 0.5'ten daha yüksek tutulmuştur.

Çizelge 3.9: Herhangi bir SNP'te anne-baba-çocuk üçlüsü genotiplerinin, C-TDT test istatistiğinde, T ve NT terimlerine olan katkıları

	Anne	Baba	Çocuk	T	NT
1	AA	AA	AA	0	0

2	AA	AB	AA	1	0
3	AA	AB	AB	0	1
4	AA	BB	AB	0	0
5	AB	AA	AA	1	0
6	AB	AA	AB	0	1
7	AB	AB	AA	2	0
8	AB	AB	AB	1	1
9	AB	AB	BB	0	2
10	AB	BB	AB	1	0
11	AB	BB	BB	0	1
12	BB	AA	AB	0	0
13	BB	AB	AB	1	0
14	BB	AB	BB	0	1
15	BB	BB	BB	0	0

Anne veya baba çocuk ikilisinin SNP'ler boyunca genotip değerlerinin üretiminden bahsetmek gerekirse, öncelikle belirtilen sayıda anne ve baba için 262,264 SNP boyunca genotip veri rastgele olarak üretilmiştir. Herhangi bir aile için, herhangi bir SNP üzerinde annenin (ya da babanın) genotipi gözlemlendikten sonra, bu anneden (veya babadan) olacak bir çocuğun aynı SNP üzerinde alabileceği genotip değerleri için aşağıdaki çizelge 3.10' dan faydalanılmıştır.

Çizelge 3.10: Herhangi bir SNP’te anne veya babanın genotipinin gözlemlenmesi üzerine, aynı SNP’te çocuğun genotipinin alabileceği değerler ve bu genotiplerin C-TDT test istatistiğinde, $N_M > C$ veya $F > C$, $N_M < C$ veya $F < C$, $N_M = C$ veya $F = C$ terimlerine katkıları

	Anne veya Baba	Çocuk	Frekans			
1	AA	AA	2	0	0	1
2	AA	AB	2	1	0	0
3	AB	AA	2	0	1	0
4	AB	AB	4	0	0	1
5	AB	BB	2	1	0	0
6	BB	AB	2	0	1	0
7	BB	BB	2	0	0	1

İstatistiksel güç analizi yapabilmek için, çocuğun genotipi için ortaya çıkan değerlerin olasılıkları için benzer bir yol izlenmiştir: Herhangi bir çocuk için 262,264 SNP’in % 98’lik kısmını oluşturan 257,019 SNP, C-TDT test istatistiğinde hipotezi lehine katkıda bulunacağını düşündüğümüz şekilde üretilmiştir. Bu sebeple, yukarıdaki tabloda görüleceği üzere, C-TDT test istatistiğinde yer alan (şartlarını yerine getiren genotiplere daha yüksek olasılıklar verilerek, diğer kalan genotiplere ise daha düşük olasılıklar verilerek üretim yapılmıştır. Örneğin, yukarıdaki tabloda da görüleceği gibi, annenin (veya babanın) genotipinin AA olarak gözlenmesi durumunda, çocuk 0.5 olasılıkla AA, 0.5 olasılıkla AB olabileceken, çocuğun AA olması yönünde destekleyici olacağından, çocuğun AA olma olasılığı 0.5’ten daha yüksek tutulmuştur.

Geriye kalan 5245 SNP (%2) ise aleyhinde, yani, istatistiksel olarak önemli asosiyasyon verecek şekilde düzenlenmiştir. Bu sebeple, çizelge 3.10 da görüleceği üzere, C-TDT test istatistiğinde yer alan (şartlarını yerine getiren genotiplere daha yüksek olasılıklar verilerek, diğer kalan genotiplere ise daha düşük

hesaplanmış,buradan yola çıkarak da,

hesaplanmıştır.

Oranlar her simülasyon verisi için hesaplandığından, örneğin 100 adet test gücü elde edilmiştir. Bunları özetlemek için, 100 oranın ortalaması alınmıştır. Permütasyon yöntemlerinin yoğun hesaplama gerektirmesi, hesaplama süresini uzatmaktadır. Bu sebeple çalışmamızın bu aşamasında, programlama için C programlama dili kullanılmıştır.

3.5. Biyoinformatik Analiz Sonuçlarının Doğrulanması için Yapılan Çalışmalar

3.5.1. SNP Genotiplendirme Çalışmaları

Yapılan analiz sonucunda anlamlı çıkan ($p < 0,001$) kromozom bölgelerindeki 5 gen için SNP genotiplendirme çalışması yapılmaya karar verilmiştir. TIB MOLBIOL tarafından sentezlenen simple probler kullanılarak hedef değişikliğin genotiplendirilmesi Roche 480 Light Cycler Cihazı kullanılarak yapılmıştır.

3.5.2. Kopya Sayısı Değişiklikleri Analizi (Copy Number Variation, CNV)

Çalışmada kullanılan tüm hasta bireyler (84 sporadik kolorektal hastası) için CNV analizi yapılmıştır. CNV analizi Affymetrix CNAT4.0 programı ile yapılmıştır. CNV analizinde kullanılan kontrol örnek referans listesi, hastalardan bağımsız olarak seçilen sağlıklı hasta anne ve/veya babalar kullanılarak hazırlanmıştır.

Tüm hastaların CNV profilini karşılaştırarak ortak bölgelerin gösterilmesi için “multi data analysis” adlı program kullanılmıştır (www.hgmbiyoinformatik.com.tr, <http://cnat.probecat.com/>).

3.5.3. Haplotip Analizi

Biyoinformatik analiz sonucunda aynı kromozom lokusunda birbirini izleyen SNP'ler öncelikli olarak değerlendirilmiştir. Haplotip analizi ile bazı ailelerde bu bölgelerde haplotip konsensusu gözlenmiş ve kurucu etkisi (founder effect) saptanmıştır.

Biyoinformatik analiz sonrasında hastalıkla ilişkili olarak saptanan 3p22.3 bölgesinde lokalize olan SNP'lerin geniş ölçekte analiz edilebilmesi için haplotip analizi yapılmıştır. Haplotip analizi 3p22.3 (35013470 bç- 35978493 bç arasında) bölgesinde yer alan rs6550313 ile rs4538325 arasındaki SNP lerin anne, baba ve hasta çocukta nasıl kalıtıldığı incelenmiştir.

3.5.3.1. DNA Dizi Analizi

Haplotip analizi ile de sporadik kolorektal kanserle ilişkisi kesinleştirilen 3p22.3 bölgesinde *ARPP21* ve *MIR 128-2* olarak iki gen bulunmaktadır. Her iki genin de aday olabileceği düşünülmüş ve bu iki aday gen için primer dizayn edilmiş (çizelge 3.11 ve 3.12) , sekans PCR' ları yapılmış ve sonunda da DNA dizi analizi (Beckman Coulter CEQ 8000) gerçekleştirilmiştir.

Çizelge 3.11: MİR128-2 geni için dizayn edilen PCR ve sekans primerleri

Primer adı	Primer dizisi 5'-3'
MİR128-2 F	GAGTGCTTCCTCTGTTCT
MİR128-2 R	GAGCAGGCATCTATGAGT

Çizelge 3.12: ARPP21 geni için dizayn edilen PCR ve sekans primerleri

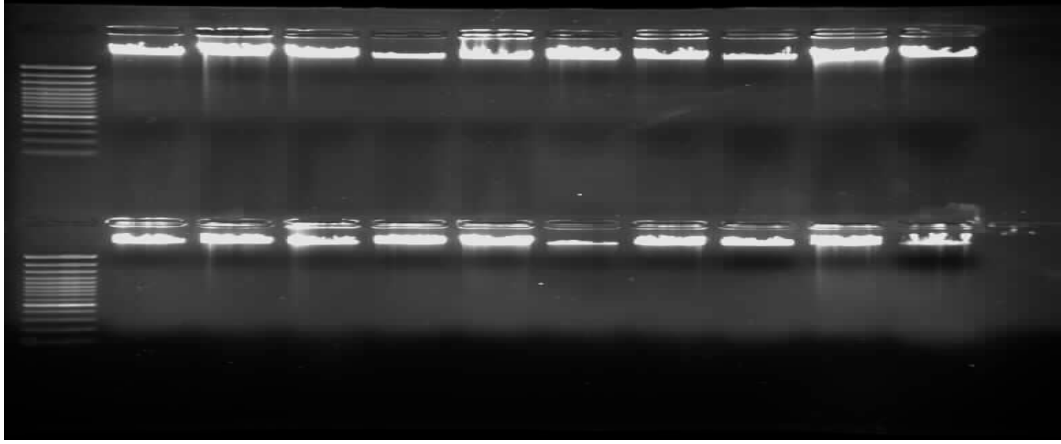
Primer adı	Primer dizisi 5'-3'
EKZON 1 F1	TATCCTCTCCAGGCATGATT
EKZON 1 R1	CATATCCAATGGTCAGCCTT
EKZON2 F	GGTGATCCTGATCCTCATTG
EKZON2 R	TCAGTGAAGTCTGGATGC
EKZON3 F	AACTGCATCTGCCTCTTGTG
EKZON3 R	ACCTTGAGGTGTTCTGGA
EKZON4 F	GACTTCTCTACTGTAGCCGGT

EKZON4 R	AGCACCAGTACAGCCGAGAT
EKZON5 F	TGTGGAGATGATTTGGCAGT
EKZON5 R	TCATGCACACAAAGGTGGAT
EKZON6 F1	GGAGCCAGACAGAGATGAT
EKZON6 R1	CTGTTAGAGGCGGTTGTCAT
EKZON6 R2	TCTGTGACTATGTGTGTGC
EKZON6 R3	CACACACACAGATACGTACAC
EKZON7 F1	CCACTTCCCCTCCTATGACA
EKZON7 R1	GAACAAAACGAAGGCCAAGT
EKZON7 F2	ACCATCATTCTGACAGCTC
EKZON8 F	CAGTCAATACTTGGCCTTCG
EKZON8 R	AGAAGCACACATGGTCCAAG
EKZON9 F1	ATGCAATAGGCTTAGTGGTT
EKZON9 R1	TGCCTCTAGGTTCCGAGCTT
EKZON9 R2	CATAGAGTACTGATGGAAG
EKZON9 R3	TGCCACACTACTTCTGATA
EKZON10 F	TTCAGCCATGATGAGCAGCA
EKZON10 R	AGGTTGGAGATAGGACAGTGG
EKZON 11 F	ACCTCTCACCATTGCATGTC
EKZON 11 R	CTGAACCAGATCCTCTTCC
EKZON 12 F	CTTCTGACCTCCATGCCGAAC
EKZON 12 R	GCAAGTTCCTGTGATGTCTCC
EKZON 13 F	TCTTGGATATGTGCCTCTGG
EKZON 13 R	GACTTCCACGCTAGCTAT
EKZON 14 F	GTCTTGTGTGGATCACCTTC
EKZON 14 R	CTTGACCTAGGCACATAACT
EKZON 15 F	CTCATCCTACAGCATGCGTC
EKZON 15 R	CTCGTCTTAGCTCTTAAGCA
EKZON 15 R2	GTTCATGCAACTAGCTT
EKZON 16 F	CAGTAAGGTCTTCACATTGG
EKZON 16 R	AGAACTCTATGCTGCCAGGT
EKZON 17 F1	ACAACCTCCAAGAATGTGTCC
EKZON 17 R1	CTGAGCTGGATATCACATGG
EKZON 18 F	CATCTTCACAACCTACCACA
EKZON 18 R	GCTTCACCTTCTTGTACCT
EKZON 19 F1	CCAGGAGAAGAGGATAGTCAG
EKZON 19 F2	GAGATGTCTGAAGGCTGC
EKZON 19 R1	ATGGACCAAGCACGCATTAC
EKZON 20 F1	GCAGAGCTGGAATTGATGACC
EKZON 20 R1	GGCATACTAGCAGACATGAC
EKZON 20 F2	ATGGAAGAGGAACAAGGTGG
EKZON 20 R2	TTGCTACGGCATATTCC

4. ARAŞTIRMA BULGULARI

4.1. DNA Bütünlük, Kalite ve Miktar Tayini

İzole edilen DNA'lar bütünlüklerinin kontrolü için %1 lik agaroz jelde yürütülmüş ve örneklerin 250K mikrodizin çalışmaları için uygun olduğu saptanmıştır (Şekil 4.1). Çalışmada kullanılan DNA'ları temsil eden uygun jel görüntüleri ve spektrofotometrik okuma sonuçları aşağıda birkaç örnekle gösterilmiştir (Çizelge 4.1).



Şekil 4.1: Genomik DNA'ların agaroz jel elektroforezindeki görüntüleri

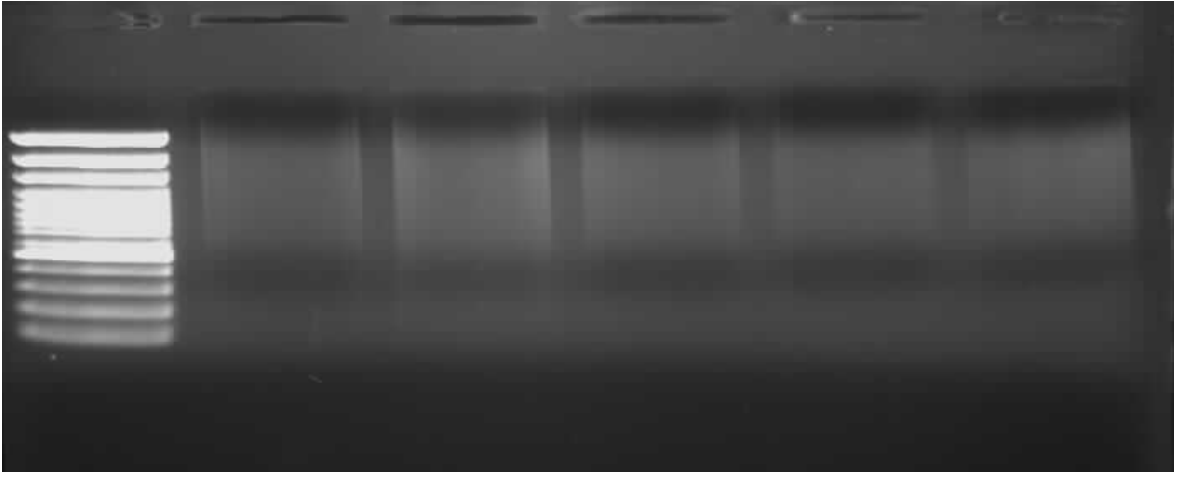
Çizelge 4.1: Genomik DNA'ların spektrofotometrik okumaları

Örnek Kodu	ng/ul	A260	A280	260/280	260/230
08/36	651,33	13,027	6,893	1,89	2,34
08/65	858,7	17,714	9,117	1,88	2,13
09/138	166,63	3,333	1,789	1,86	2,22
09/308	469,3	9,386	4,461	2,1	2,19
10/45	852,03	17,041	9,243	1,84	2,04
10/59	913,61	18,272	9,718	1,88	2,16

4.2. 250K DNA Mikrodizin Analizi

4.2.1. Restriksiyon Endonükleaz Kesim Sonuçları

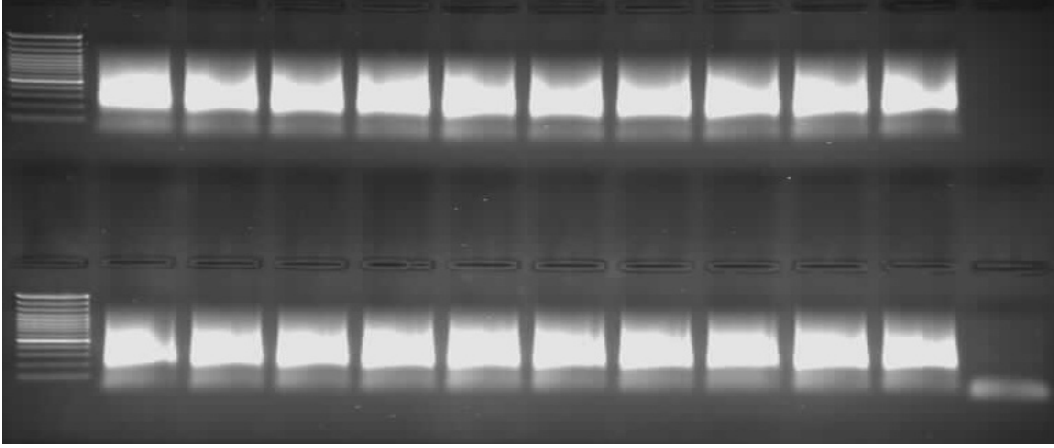
SNP array protokolleri restriksiyon enzimleri ile kesim sonrası kesim sonucunun teyidini öngörmemektedir. Ancak tecrübelerimiz bu aşamanın başarı ile gerçekleşip gerçekleşmediğinin de kontrol edilmesi gerektiğini ortaya koymuştur. Yalnızca 312.5 ng DNA kesildiği için %1'lik jelde restriksiyon endonükleaz kesimi belli belirsiz gözükmemektedir. Burada önemli olan kesilmemiş DNA'nın gözlenmemesidir (Şekil 4.2).



Şekil 4.2: DNA örneklerinin restriksiyon endonükleaz ile kesimi sonrası %1'lik agaroz jel elektroforezi görüntüsü

4.2.2. PCR Elektroforez Sonuçları

250K Array analizlerinde restriksiyon endonükleaz analizi sonrası kesilmiş DNA fragmentlerinin uçlarına kitle bulunan adaptör DNA ligaz enzimi ile bağlanmakta ve böylece başı ve sonundaki DNA dizisi bilinen bu fragmentler adaptör diziyeye komplementer tek bir primer kullanılarak PCR'a alınmaktadır. PCR sonrası her bir örneğin miktarının 90ug'a ulaşması gerektiğinden her örnek PCR'ı 3 ila 5 tüpte gerçekleştirilmektedir. PCR ürünleri 200 ve 1100 bp arasında olmaktadır (Şekil 4.3).



Şekil 4.3: PCR jel görüntüleri

4.2.3. PCR Pürifikasyon Sonrası Miktar Tayini

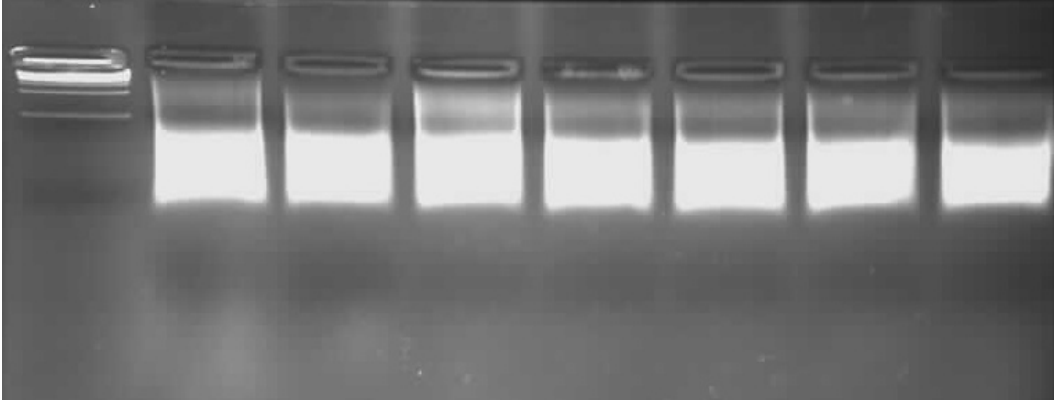
Herbir örnek için kurulan 5 PCR reaksiyonu, pürifikasyon sırasında biraraya getirilmekte ve reaksiyonların toplamından elde edilen DNA miktarının hibridizasyon için 90µl içinde en az 90ug'a ulaşması beklenmektedir. Çalışmada kullanılan PCR ürünleri temsil eden uygun spektrofotometrik okuma sonuçları aşağıda birkaç örnekle gösterilmiştir (Çizelge 4.2).

Çizelge 4.2: PCR ürünlerinden bazılarının spektrofotometrik okumaları

Örnek Adı	ng/ul konsantrasyon	A260	A280	260/280	260/230
08/47	1236,46	24,729	12,238	2,02	2,21
08/65	1110	22,2	11,163	1,99	2,21
08/38	1306,18	26,124	13,123	1,99	2,2
08/39	1122,56	22,451	11,26	1,99	2,19

4.2.4. Fragmentasyon ve İşaretleme

PCR sonrası 90 ug'a ulaşan örneklerin hacmi, vakumlu santrifüjde 45ul'ye indirildikten sonra fragmentasyon reaktifi ile efektif işaretleme ve hibridizasyonu sağlamak için fragmente edilmektedir. Başarı ile fragmente olan örneklerin fragment boylarının 200 bç'nin altında yoğunlaşması beklenmektedir (Şekil 4.4).



Şekil 4.4: Fragmentasyon görüntüleri

4.2.5. Hibridizasyon, Yıkama, Boyama ve Tarama

Fragmentasyon sonucu gerçekleştirilen işaretleme reaksiyonu ve bunu takip eden hibridizasyon, yıkama, boyama ve tarama işlemlerinin sonucunda herbir array'e ait genotipleme raporu alınan sonucun başarısını ortaya koymaktadır. %85'in üzerindeki sinyal oranı (call rate) bu aşamaya dek yapılan aşamaları başarılı kabul etmek için gerekli alt sınır olarak kabul edilmiştir.

Bu raporların sadece bir tanesi örnek teşkil etmek üzere aşağıda bulunmaktadır.

4.3. Biyoinformatik Bulgular

4.3.1. SPRT Analizi Sonuçları

Çalışmamız kapsamında toplanan 51 trio ile SPRT analizi yapılmıştır. Analiz sonuçlarına göre 75 SNP ($\alpha=0.1\%$ ve $\beta=20\%$) hastalıkla ilişkili bulunmuştur. Bu 75 SNP, toplam 20 kromozomda 67 farklı kromozom lokusunda bulunmaktadır. Ayrıca 75 SNP'in toplam 119 genin intronunda, üst bölgesinde (upstream) veya alt bölgesinde (downstream) bulunduğu belirlenmiştir (Çizelge 4.3).

Çizelge 4.3: SPRT analizi hastalıkla ilişkili bulunan SNP'ler

Kromozom	Fiziksel Pozisyonu	Kromozom lokalizasyonu	Genin sembolü	SNP numarası
1	21840709	p36.12	ALPL	SNP_A-2236414
1	21840886	p36.12	ALPL	SNP_A-4221240
1	57377810	p32.2	C8A	SNP_A-1895216
1	71010207	p31.1	CTH , PTGER3	SNP_A-1882816
1	118908373	p12	SPAG17 , TBX15	SNP_A-2268795
1	155080090	q22	EFNA3 , EFNA1	SNP_A-1797826
1	167234166	q24.2	POU2F1	SNP_A-2000449
1	207409288	q32.2	C4BPA , CD55	SNP_A-4197863
1	241720559	q43	KMO	SNP_A-4221145
2	18648687	p24.2	NT5C1B , KCNS3 , RDH14	SNP_A-1946248
2	50947512	p16.3	NRXN1	SNP_A-2043314
2	137679501	q22.1	CXCR4 , THSD7B	SNP_A-2044723
2	172283513	q31.1	METTL8	SNP_A-2139004
3	14377469	p25.1	LSM3 , SLC6A6	SNP_A-4202186

3	35616208	p22.3	PDCD6IP , ARPP21	SNP_A-1972343
3	36189216	p22.3	STAC , ARPP21	SNP_A-2076419
3	158776311	q25.32	MFSD1 , IQCJ	SNP_A-2037548
3	168467339	q26.2	GOLIM4 , MECOM , C3orf50	SNP_A-2294629
3	187473372	q27.3	BCL6 , LPP	SNP_A-2089111
4	27202594	p15.2	STIM2 , PCDH7	SNP_A-2183762
4	35471869	p15.1	PCDH7 , ARAP2	SNP_A-2039589
5	41506080	p13.1	PLCXD3	SNP_A-4200449
5	53874176	q11.2	SNX18 , ESM1	SNP_A-2270763
5	54256550	q11.2	SNX18 , ESM1	SNP_A-2162113
5	102710525	q21.1	PPIP5K2 , NUDT12	SNP_A-2261676
5	148971305	q32	ARHGEF37	SNP_A-4200252
6	3824924	p25.2	C6orf145 , FAM50B	SNP_A-2272871
6	11040183	p24.2	ELOVL2 ,	SNP_A-2219269
6	14771957	p23	CD83 , JARID2	SNP_A-2257626
6	41495339	p21.1	NCR2 , FOXP4	SNP_A-2188623
6	52374823	p12.2	TRAM2	SNP_A-2303765
6	54763018	p12.1	FAM83B	SNP_A-1986112
6	123229277	q22.31	SMPDL3A , CLVS2	SNP_A-2254772
7	43051925	p14.1	MRPL32 , HECW1	SNP_A-2164167
8	81303433	q21.13	TPD52 , ZBTB10	SNP_A-1952692
8	105169450	q22.3	RIMS2	SNP_A-1882853
8	123599769	q24.13	HAS2 , ZHX2 , HAS2AS	SNP_A-2205361
8	130225474	q24.21	GSDMC , MYC , MIR1208	SNP_A-2127975

8	137102995	q24.23	KHDRBS3 FAM135B	SNP_A-2305116
9	8494221	p24.1	PTPRD	SNP_A-2167105
9	13568063	p23	MPDZ , NFIB	SNP_A-1930351
9	34396346	p13.3	C9orf24	SNP_A-1934576
9	82116305	q21.31	PSAT1 , TLE4	SNP_A-4205164
10	9778028	p14	CELF2 , GATA3 , LOC254312	SNP_A-2109711
10	95098494	q23.33	MYOF	SNP_A-2204654
10	110503788	q25.1	SORCS1 , XPNPEP1	SNP_A-1913097
11	18866730	p15.1	PTPN5 , MRGPRX1	SNP_A-4197674
11	56248136	q12.1	OR5M3 , OR5M8	SNP_A-2234370
11	97987244	q22.1	JRKL , CNTN5	SNP_A-2247154
11	98011781	q22.1	JRKL , CNTN5	SNP_A-2090144
11	98025525	q22.1	JRKL , CNTN5	SNP_A-1798471
11	98025817	q22.1	JRKL , CNTN5	SNP_A-2035542
11	98043380	q22.1	JRKL , CNTN5	SNP_A-4233313
11	114702663	q23.3	CADM1 , FAM55B	SNP_A-2122717
12	68049075	q15	DYRK2	SNP_A-2136483
12	96196642	q22	NTN4 , SNRPF	SNP_A-2277914
13	20732487	q12.11	GJA3	SNP_A-2003156
13	98453492	q32.2	RAP2A , IPO5	SNP_A-4228019
14	68690442	q24.1	RAD51L1	SNP_A-1867874
14	79886827	q31.1	NRXN3	SNP_A-4192710
14	93186618	q32.12	LGMN	SNP_A-1862567
15	26933734	q12	GABRB3	SNP_A-2262915

15	40840557	q15.1	CHST14 , C15orf57 , MRPL42P5	SNP_A-2273484
15	46212662	q21.1	SQRDL , SEMA6D	SNP_A-2079910
17	10885174	p12	PIRT , SHISA6	SNP_A-4224546
18	64095565	q22.1	CDH7 , CDH19	SNP_A-1925229
20	1039416	p13	RSPO4 , PSMF1	SNP_A-1942899
20	17077510	p12.1	OTOR , PCSK2	SNP_A-1887582
20	17086061	p12.1	OTOR , PCSK2	SNP_A-2134216
20	54722875	q13.2	CBLN4 , MC3R	SNP_A-1860033
21	28030838	q21.3	CYYR1 , ADAMTS1	SNP_A-4215750
21	32060613	q22.11	KRTAP21-2 , KRTAP20-3 , KRTAP21-3	SNP_A-2308352
21	33275860	q22.11	HUNK	SNP_A-1970018
X	5066057	p22.32	PRKX , NLGN4X	SNP_A-1998124
X	121440080	q25	GLUD2 , GRIA3	SNP_A-4235371

4.3.3. C-TDT Analizi İçin Yapılan Benzetim Çalışmasının Sonuçları

Benzetim çalışmasının sonucunda, beklenildiği gibi, örneklem büyüklüğü arttıkça, testin gücünün arttığı ve hesaplanmış anlamlılık düzeyinin ise düşerek, gerçek değer olan %5'e yaklaştığı gözlemlenmiştir. Ayrıca, 50 anne-baba-çocuk-üçlüsü ve 30 anne veya baba-çocuk ikilisine ait verinin toplanmasıyla % 83'e yakın bir güç elde edildiği gözlenmiştir (Çizelge 4.4). İstatistiksel analizlerde genellikle %80'in üstünde bir güç hedeflendiğinden, bu proje kapsamında toplanan veriye C-TDT uygulayarak tatmin edici sonuçlar aldığımız söylenebilir. Örnek sayısı 70 trio ve 40 duo olduğu takdirde güç %92'ye çıkarırken, bu artışın hesaplanmış anlamlılık düzeyini pek de değiştirmedeği görülmektedir (Çizelge 4.4). (yayınlanmamış veri, makale dergiye gönderme aşamasında (Rajabli F., Inan G., İlk O, (2011), Power analysis of small sample size genome-wide association studies by the joint use of case-parent trios and pairs.))

Çizelge 4.4: Örneklem büyüklüklerine göre güç ve hesaplanmış anlamlılık düzeyler

Aile türü		Güç (%)	Hesaplanmış anlamlılık düzeyi (%)
Anne-baba-çocuk üçlüsü	Anne veya baba- çocuk		
10	4	41.3	14.4
30	18	66.1	8.9
50	30	83	7.8
70	40	92	7.3
90	55	96.5	6.9
110	65	98.5	6.7
130	75	99.4	6.6

4.3.4. C-TDT Analizi Sonuçları

Bu tez çalışması kapsamında çalışılan 51 trio ve 33 duo ile yaptığımız C-TDT analizi sonucunda 15,093 SNP anlamlı ($p < 0.001$) çıkmıştır, yani bu 15,093 SNP sporadik kolorektal kanser ile asosiyasyon göstermektedir. Fakat bu 15,093 SNP içinden biyolojik olarak da en anlamlı olanları doğrulamak ve sahte olanları elimine edebilmek için ileri analizlere ihtiyaç duyulmaktadır.

Anlamlı çıkan tüm SNP'lerin doğrulanması, binlerce hasta ve kontrol bireyin anlamlı SNP'ler için tekrar genotiplendirilmesini gerektirmektedir. Bu konfirmasyon (doğrulama) yaklaşımı, projeye ekstra maliyet getireceğinden ve binli sayıda yeni hastaya ulaşmanın zorluğu gibi nedenlerle tercih edilmemiştir.

Eğer böyle bir doğrulama çalışması yapılamayacaksa elimizdeki verinin anlamlılık kazanabilmesi için başka analiz yaklaşımları yapılabilmektedir. Bunlar;

- Asosiyasyon saptanan kromozom bölgelerinde fonksiyonel olarak anlamlı genler bulup, bu genlerin içindeki SNP lerin anlamlılığını test etmek,
- Asosiyasyon saptanan kromozom bölgelerinde haplotip yapıp ortak segrege olan bölge yakalamak,

c. Çalışılan örneklerde kopya sayısı değişiklerini (copy number variation, CNV) değerlendirip anlamlı bir değışiklik saptamak olabilmektedir.

d. SNP'lerin yoğunlaştığı bölgeleri/genleri bu uzun listede saptamak. Aynı bölgede bulunan SNP'lerin aynı bilgiyi verdiği aşıkardır.

Yukarıda adı geçen analiz yöntemleri hem ayrı ayrı hem de hep birlikte yapıldıklarında ekstra maliyete gerek duymaksızın elimizdeki verinin güçlendirilmesi adına önemli yararlar sağlamaktadırlar. Bu projede yukarıda belirtilmiş olan 4 yöntem ile hem SPRT hem de C-TDT sonucu elde ettiğimiz veriler analiz edilmiştir.

Bu dört yönteme örnek olarak yukarıdaki kapsamın d. maddesinden hareketle saptanan SNP'ler arasında bir öncelik listesi belirlerken SNP'lerin kromozomlar üzerindeki lokuslarına bakıldığında aynı bölgede ≥ 30 ve < 35 SNP'in bulunması esas alındığında 17 lokus (Çizelge 4.5), ≥ 35 ve < 40 SNP esas alındığında 7 lokus (Çizelge 4.6), ≥ 40 SNP esas alındığında ise 11 lokus (Çizelge 4.7) saptanmıştır. Aynı kromozom bölgesinde ≥ 30 ve < 35 SNP'in bulunması durumunda saptanan 17 lokus değerlendirildiğinde 2., 3., 4., 5., 6., 8., 9., 11., ve 14. kromozomlardaki bazı lokuslar olduğu belirlenmiştir (Şekil 4.5).

Çizelge 4.5: Aynı lokusta ≥ 30 SNP < 35 bulunması esas alındığında belirlenen lokuslar

Kromozom lokusları	Görülen SNP sayısı
2p16.3	34
2p22.3	33
2q14.1	32
2q22.1-2q22.2	30
3p26.1	34
4q28.3	30
4q34.3	34
5p15.2	34

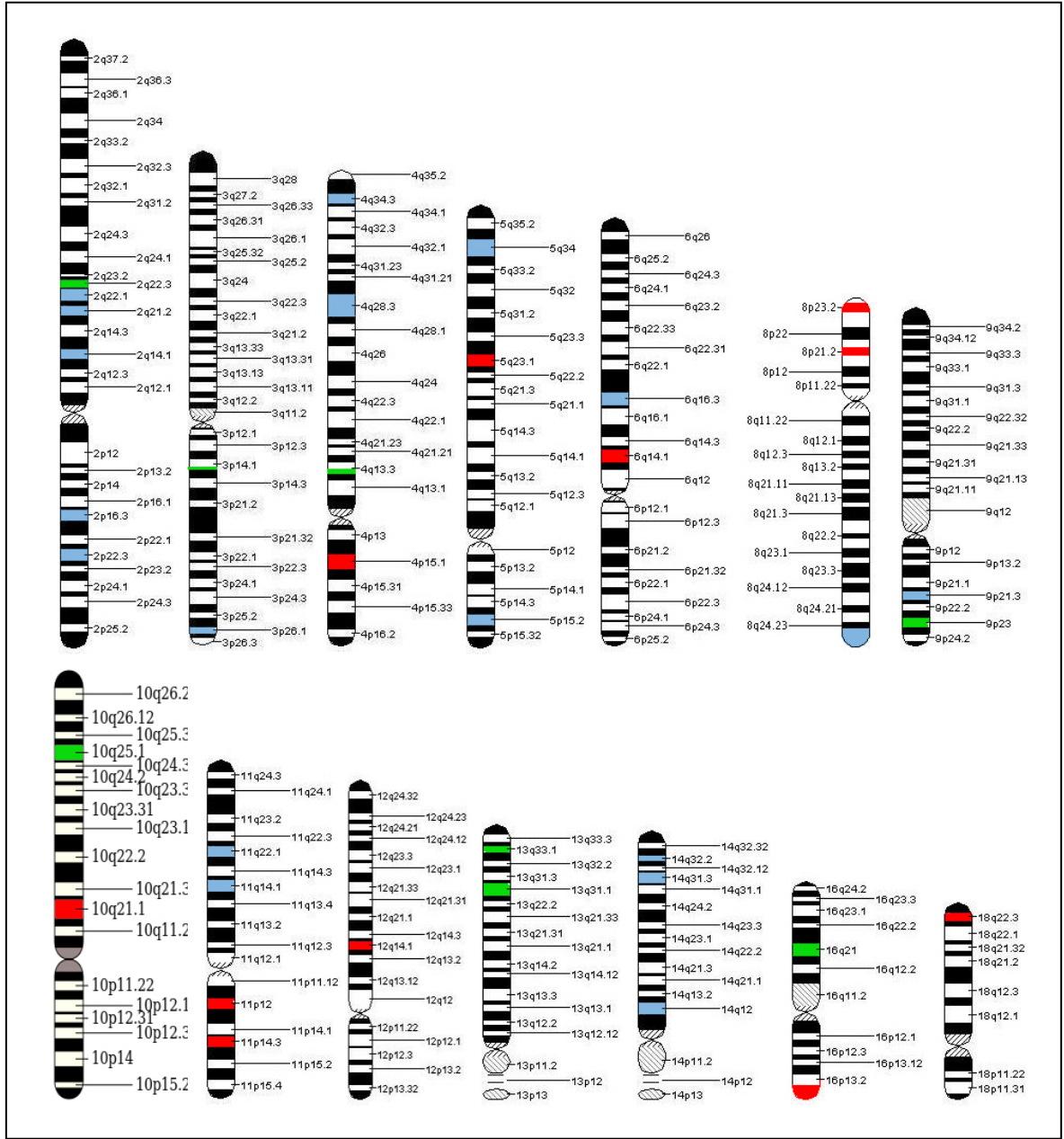
5q34	31
6q16.3	31
8q24.3	32
9p21.3	33
11q14.1	30
11q22.1	33
14q12	31
14q31.3	31
14q32.2	33

Çizelge 4.6: Aynı lokusta ≥ 35 SNP < 40 bulunması esas alındığında belirlenen lokuslar

Kromozom lokusları	Görülen SNP sayısı
2q32.3	35
3p14.2	39
4q13.2	39
9p23	37
10q25.1	38
13q31.1	39
16q21	38

Çizelge 4.7: Aynı lokusta < 40 SNP esas alındığında belirlenen lokuslar

Kromozom lokusları	Görülen SNP sayısı
4p15.1	55
5q23.1	49
6q14.1	40
8p12	48
8p23.2	57
10q21.1	44
11p12	51
11p14.3	43
12q14.1	45
16p13.3	42
18q22.3	52



Şekil 4.5: SNP'lerin yoğun bulunduğu kromozom bölgeleri

(**Mavi** ile boyalı bölgeleri ≥ 30 ve < 35 SNP, **yeşil** ile boyalı bölgeleri ≥ 35 ve < 40 SNP, **kırmızı** ile boyalı bölgeleri ise ≥ 40 SNP işaret etmektedir.)

4.4. Biyoinformatik Analiz Sonuçlarının Doğrulanması için Yapılan Çalışmaların Sonuçları

4.4.1. SNP Genotiplendirme Çalışmalarının Sonuçları

Yapılan analiz sonucunda anlamlı çıkan ($p < 0,001$) kromozom bölgelerindeki 5 gen için SNP genotiplendirme çalışması yapılmaya karar verilmiştir. Bu 5 genin seçimi sırasında literatür araştırması ve yolak analizleri de yapılmıştır. Bu çalışmalar sonucunda kolorektal kanser yolağında yer alan DCC geni, kolorektal karsinogenezle daha önceden ilişkilendirilmiş COL11A1 ve CXCR4 genleri ve karsinogenez ile ilişkili olabileceği düşünülen TMBIM4 ve PDCD6IP genleri aday genler olarak seçilmişlerdir. Bu 5 aday gen içinden 1 SNP genotiplendirilmek üzere seçilmiştir. Genler içindeki bu SNP'ler, protein dizisindeki amino asitte farklılık yaratan SNP (nonsynonomous SNP) olması ve daha önce çalışılan popülasyonlarda frekans bilgisi var ise frekansının çok yüksek olmamasına dikkat edilerek seçilmişlerdir.

84 sporadik kolorektal kanser hastası ve 96 kontrol birey kullanılarak (3 hasta örneğinin deneylerinin çalışmaması nedeni ile hasta grubu rs8793 no'lu SNP'de 81 olarak kullanılmıştır), verilerin doğrulanması için yapılan genotip çalışmasında belli bir genotip sıklığının hasta ve kontrol grubunda dağılımını incelemek amacıyla Fisher exact test kullanılmıştır. İstatistiksel testler için anlamlılık düzeyi 0.05 olarak alınmıştır. Bu beş SNP (Çizelge 4.8) için P-değerleri sırasıyla 0.13, 0.98, 0.96, 0.90 ve 1 olarak elde edilmiştir. 5 SNP de elde edilen P-değerlerinin verilen anlamlılık düzeyi olan 0.05'ten büyük olması sonucu, 5 SNP için de hasta ve kontrol grupları arasında anlamlı bir fark olmadığı saptanmıştır. Herbir lokustaki genotiplerin hasta ve kontrol gruplarındaki dağılımları aşağıdaki tablolarda (Çizelge 4.9, 4.10, 4.11, 4.12, 4.13) verilmiş, testlere ait P-değerleri belirtilmiştir.

Çizelge 4.8: Doğrulama analizleri için genotiplendirilmek üzere seçilen SNP'ler

Gen	SNP ID	Değişim	Kromozom	Atasal alel	Frekans	Fonksiyon
DCC	rs2229080	C-G	18:50432602	C	0.500	Tümör baskılayıcı, adezyon, migrasyon
TMBIM4	rs8793	C-T	12:66546100	C	0.300	Apoptozis
CXCR4	rs56400844	T-G	2:136873341	T	bilgi yok	Anjiogenez, metastaz
COL11A1	rs3753841	C-T	1:103379918	C	0.500	Kolorektal kanserde over eksprese
PDCD6IP	rs3792594	T-C	3:33877626	T	0.400	Apoptozis

Çizelge 4.9: rs8793 (TMBIM4 geni) lokusunda genotiplerin hasta ve kontrol gruplarında dağılımı

Genotip	Hasta Grubu (%)	Kontrol Grubu (%)	P-değeri
	(n=81)	(n=96)	
CC	11	24	0.13
CT	40	37	
TT	30	35	

Çizelge 4.10: rs2229080 (DCC geni) lokusunda genotiplerin hasta ve kontrol gruplarında dağılımı

Genotip	Hasta Grubu (%)	Kontrol Grubu (%)	P-değeri
	(n=84)	(n=96)	
CC	21	24	0.98
CG	45	53	
GG	18	19	

Çizelge 4.11: rs3753841(COL11A1) lokusunda genotiplerin hasta ve kontrol gruplarında dağılımı

Genotip	Hasta Grubu (%)	Kontrol Grubu (%)	P-değeri
	(n=84)	(n=96)	
CC	13	15	0.96
CT	37	40	
TT	34	41	

Çizelge 4.12: rs3792594 (PDCD6IP) lokusunda genotiplerin hasta ve kontrol gruplarında dağılımı

Genotip	Hasta Grubu (%)	Kontrol Grubu (%)	P-değeri
	(n=84)	(n=96)	
CC	37	46	0.90
CT	40	42	
TT	7	8	

Çizelge 4.13: rs56400844(CXCR4) lokusunda genotiplerin hasta ve kontrol gruplarında dağılımı

Genotip	Hasta Grubu (%)	Kontrol Grubu (%)	P-değeri
	(n=84)	(n=96)	
TT	84	96	1
TG	0	0	
GG	0	0	

4.4.2. Kopya Sayısı Değişiklikleri Analizi Bulguları

Proje kapsamında çalışılan hastalardaki CNV'lerin belirlenmesi amacı ile çalışılan herbir hastanın 250K verisi, sağlıklı bireyin (25 tane hasta ile akraba olmayan) 250K verisi ile kıyaslanmıştır. Bu analizlerin sonucunda bu hastalarda ortak olarak bulunan delesyon ve/veya amplifikasyonları saptamak için yapılan analiz sonucunda 22 hastada 4q13.2 de 105,066 bç lik, 24 hastada 9p12 bölgesinde 23,055 bçlik, 30 hastada 15q11.1 de 449,582 bç'lik, 22 hastada 15q11.2 de 267,962 bç'lik delesyonlar belirlenmiştir (Çizelge 4.14).

Delete olduğu gösterilen bu 4 bölgenin, kopya sayısı varyasyonları veri tabanında, kopya sayısı varyasyonu olarak bildirildiği belirlenmiştir.

24 hastada 2q11.1 bölgesinde 130,798 bç lik, 20 hastada 17q21.31 lokusunda 186,953 bç lik, 27 hastada 19p13.2 29,762 bç lik, 23 hastada Xp11.23 de 251,771 bç lik, 20 hastada Xq21.1 de 104,917 bç lik, 29 hastada Xq28 de 92 bç lik, amplifikasyonlar saptanmıştır (Çizelge 4.15). Amplifiye olduğu gösterilen bu bölgelerden 17q21.31 ve Xp11.23 bölgesinin, kopya sayısı varyasyonları veri tabanında, kopya sayısı varyasyonu olarak bildirildiği belirlenmiş olup diğer bölgelerin ise veri tabanında bulunmadığı saptanmıştır.

Çizelge 4.14: Kopya sayısı değişiklikleri analizinde delete olduğu saptanan bölgeler

Delete olan bölge	Bu lokusta bulunan gen sayısı	Delete bölgesinin büyüklüğü (baz çifti, bç)	Bölgenin delete olduğu hasta sayısı
4q13.2	3 gen	105,066 bç	22
9p12	gen yok	23,055 bç	24
15q11.1	9 gen	449,582 bç	30
15q11.2	13 gen	267,962 bç	22

Çizelge 4.15: Kopya sayısı değişiklikleri analizinde amplifiye olduğu saptanan bölgeler

Amplifiye olan bölge	Bu lokusta bulunan gen sayısı	Amplifiye bölgesinin büyüklüğü (baz çifti, bç)	Bölgenin amplifiye olduğu hasta sayısı
2q11.1	4 gen	130,798 bç	24
17q21.31	5 gen	186,953 bç	20
19p13.2	4 gen	29,762 bç	27
Xp11.23	10 gen	251,771 bç	27
Xq21.1	1 gen	104,917 bç	20
Xq28	1 gen	92 bç	29

4.4.3. Haplotip Analizi

Yapılan C-TDT analizi sonucunda anlamlı çıkan SNP'ler incelendiğinde 3p22.3 bölgesinde toplanan 15 SNP gözlenmiştir. Bu bölge için tüm trioların haplotip analizi yapılmış ve sonuçta 3p22.3 lokusunun kolorektal kanser ile ilişkili olabileceği saptanmıştır.

3p22.3 bölgesinde bulunan rs6801949 (35565537 bp) ve rs10865865 (35941124 bp) arasındaki yaklaşık 0.5 Mb lik alan kritik bölge olarak saptanmış ve bu bölgede iki gen bulunduğu tespit edilmiştir. Bu iki genin de aday olabileceği düşünülmüş ve bu genlerin sekanslanmasına karar verilmiştir (Çizelge 4.16).

Özellikle bazı hastaların (6 hastanın) aynı haplotipi gösterdikleri ve bu bölgeyi homozigot olarak aldıkları belirlenmiştir. Aynı hastaların anne ve babalarının genotipleri incelendiğinde, anne ya da babadan birinin ilgili bölge için heterozigot durumda olduğu da tespit edilmiştir (Çizelge 4.16). Bu hastaların memleketleri araştırıldığında 5 hastanın Zonguldak ve çevresinden olduğu bilgisine ulaşılmış, 1hastanın ise memleket bilgisine ulaşılamamıştır. Aynı coğrafi bölgede görülen bu durum, kurucu etkisi (founder effect) ile açıklanabilmektedir.

Cizelge 4.16: Haplotip Analizi (Homozigosite Haritalaması)

Kromozom	Fiziki pozisyon	BARA	ANNE	HASTA	BARA	ANNE	HASTA	BARA	ANNE	HASTA	BARA	ANNE	HASTA	BARA	ANNE	HASTA	BARA	ANNE	HASTA
SNP A-1972337	35540241	rs6801949	AB	AB	AA	AA	AA	AB	AB	AA	AA	AA	AA	AB	AB	AA	AA	AA	AA
SNP A-2038792	35540249	rs4678281	AA	AB	AA	AA	AA	AB	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA
SNP A-4209403	35540850	rs1303333	AA	NoCall	BB	BB	NoCall	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB
SNP A-1972339	35568508	rs1464585	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB
SNP A-1972340	35569126	rs2037358	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB
SNP A-197244	35571626	rs207339	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA
SNP A-197242	35572939	rs4323088	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB
SNP A-2109391	35591085	rs1171827	AB	NoCall	AA	AA	NoCall	NoCall	AA	AA	AA	AA	AA	AB	AA	AA	NoCall	AA	AA
SNP A-1972343	35591212	rs1688314	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB
SNP A-1972345	35613272	rs2197228	BB	AB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB
SNP A-1972346	35613656	rs9311101	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA
SNP A-1972347	35614122	rs1094087	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB
SNP A-4209404	35650059	rs2399774	BB	AB	BB	BB	NoCall	BB	BB	BB	BB	BB	BB	AB	BB	BB	BB	BB	BB
SNP A-2052330	35658108	rs860326	BB	AB	BB	BB	NoCall	BB	BB	BB	BB	BB	BB	AB	BB	BB	BB	BB	BB
SNP A-2062227	35658446	rs1703361	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB
SNP A-2165662	35670516	rs4678289	BB	AB	BB	BB	BB	BB	BB	BB	BB	BB	BB	AB	BB	BB	BB	BB	BB
SNP A-2051597	35670956	rs112366	AA	AB	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA
SNP A-4208338	35671103	rs112366	AA	AB	AA	AA	NoCall	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA
SNP A-2297060	35688763	rs1313472	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA
SNP A-4209405	35692279	rs1303367	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA
SNP A-2282041	35720823	rs927664	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA
SNP A-1969421	35730396	rs1171888	AB	AB	BB	BB	BB	BB	BB	BB	BB	BB	BB	AB	BB	BB	BB	BB	BB
SNP A-2189357	35750253	rs4678822	AB	AB	BB	BB	NoCall	BB	BB	BB	BB	BB	BB	AB	BB	BB	BB	BB	BB
SNP A-2297623	35750931	rs4678823	NoCall	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	AB	BB	BB	BB	BB	BB
SNP A-1945158	35751109		AB	AB	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA
SNP A-2092114	35760001	rs6503367	BB	AB	AB	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA
SNP A-4222891	35762465	rs6801990	AA	AB	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA
SNP A-2198339	35762893	rs1303374	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA
SNP A-1972346	35771281	rs1333643	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA
SNP A-4209406	35775717	rs9311103	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB
SNP A-2021323	35795649	rs2016955	BB	BB	BB	BB	NoCall	AB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB
SNP A-2197988	35798612	rs2016973	AA	AB	AA	AA	AA	AA	AA	AA	AA	AA	AA	AB	AA	AA	AA	AA	AA
SNP A-1819409	35807987	rs1308394	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB
SNP A-1886672	35819665	rs1192171	AA	AB	AA	AA	NoCall	AB	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA
SNP A-2209487	35820264	rs1329481	BB	AB	BB	BB	AA	NoCall	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB
SNP A-4231215	35820210	rs1781488	AA	NoCall	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA
SNP A-1972354	35829121	rs4678509	BB	AB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB
SNP A-4209407	35834447	rs1515106	BB	AB	BB	BB	NoCall	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB
SNP A-1897032	35865656	rs6799649	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA
SNP A-1951892	35866737	rs6799649	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA
SNP A-2044376	35869926	rs7630333	BB	BB	BB	BB	NoCall	AB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB
SNP A-2188995	35872778	rs5847358	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB
SNP A-1846579	35916126	rs1308384	BB	BB	BB	BB	AB	AB	AB	AB	AB	AB	AB	AB	AB	AB	AB	AB	AB

4.4.3.1. DNA Dizi Analizi Sonuları

MİR 128-2 geni ve 20 ekzondan oluřan ARPP 21 geninin tm ekzonları ile DNA dizi analizi gerekleřtirilmiřtir. Her iki gende de DNA dizi analizi sonucunda hasta ve saėlıklı bireylerin DNA dizileri arasında herhangi bir farklılık saptanmamıřtır.

5. TARTIŞMA ve SONUÇ

Kanser ve Kolorektal Kanser

Kanser tüm dünyada en önemli insan sağlığı problemlerinden birisidir. Kompleks ve multifaktöryel bir hastalık olan kanserin ortaya çıkışında hem çevresel faktörler hem de genetik faktörler önemli rol oynamaktadır.

Kanser nedenli ölümlerin başında gelen ve dünya çapında yılda yaklaşık 639,000 kişinin ölümüne neden olan kolorektal kanserin prevalansı sabit bir biçimde artmaktadır (Naccarati et al. 2007; Altekruse et al. 2010).

T.C Sağlık Bakanlığı Kanserle Savaş Daire Başkanlığının 2004-2006 yılları kanser istatistiğine göre, kadınlarda görülen kanserler içinde 2. sırada yer alan kolorektal kanser erkeklerde ise 4. sırada yer almaktadır.

Ülkemiz için de önemli sağlık problemlerinin başında gelen kolorektal kanserin moleküler mekanizmasının aydınlatılması büyük önem taşımaktadır.

Asosiyasyon Çalışmaları

Birçok genetik ve çevresel faktörün etkisi ile gelişen bir hastalık olan kanserde ailesel kümelenme görülmekle beraber çoğunlukla herhangi bir mendelyan kalıtım paterni izlenmemektedir. Ailesel kümelenme gösteren bazı kanserlerde, yüksek penetrans gösteren kanser yatkınlık genleri bağlantı analizleri ile tanımlanabilirken, düşük penetranslı allellerin tanımlanmasında bağlantı analizleri yetersiz kalmaktadır. Bu noktada devreye giren asosiyasyon çalışmaları geniş popülasyonlardaki bireylerde, genetik varyantlar ve hastalık arasındaki asosiyasyonu belirleyebilmektedir.

Geçmişte yapılan asosiyasyon çalışmaları daha çok aday gen asosiyasyon çalışmaları şeklinde tasarlanarak kanser gelişiminde önemli rol oynayabileceği düşünülen aday genlerle çalışılması yaklaşımına dayanmaktaydı. Günümüzde ise artan genom bilgisi ve gelişen teknoloji ile paralel olarak kanser araştırmalarında en aktif olarak kullanılan metodoloji genom boyu asosiyasyon çalışmalarıdır.

Genom boyu asosiyasyon alıřmaları poplasyon temelli ya da aile temelli yaklařım ile tasarlanabilmektedir. Poplasyon temelli asosiyasyon alıřmalarının nndeki en nemli sorun olarak bilinen poplasyon katmanlařması sorununu ařmak iin binlerce hatta onbinlerce rnek ve kontrol ile alıřılırken aynı aileden gelen hasta ve kontrol rnekleri ile alıřıldıđı iin poplasyon katmanlařması gibi bir sorunun olmadığı aile temelli asosiyasyon alıřmalarında yzly sayılarda anne, baba, hasta ocuk lleri, hasta ocuk, anne veya baba ikilileri, kardeř iftleri gibi rneklerle alıřılabilmektedir. Yukarıdaki bilgilerden anlařılacađı zere aile temelli asosiyasyon alıřmaları hem poplasyon katmanlařması probleminin olmaması hem de alıřılması gereken rnek sayısının daha az olması nedeni ile maddi olarak poplasyon temelli asosiyasyon alıřmalarına kıyasla avantajlıdır.

Farklı varsayımlar altında tekrarlanan simlasyonlar, aile bazlı asosiyasyon alıřmalarında makul bir gce ulařmak iin yzlerce trioya ihtiya olduđunu raporlamaktadır (Chen et al. 2001). Az sayıda anne-baba-hasta triolarına ulařılabilmesi istatistiksel testlerin gcn dřrmektedir. TDT testinin ilk kez nerildiđi Spielman ve arkadařları tarafından yazılmıř olan 1993 tarihli makale bugne kadar Web of Science'a gre 2600'den fazla atıf almıřtır (Spielman et al. 1993). Literatrde genel kabul grmř olduđu ařıkarak olan bu test de aile temelli asosiyasyon analizlerinde kabul edilebilir bir gte olabilmesi iin en az 200 trioya ihtiya gstermektedir.

Ancak aile temelli asosiyasyon yaklařımında yeterli sayıda trio bulma sıkıntısı yařanmaktadır. Bu sıkıntı zellikle ge yařta ortaya ıkan hastalıklarda anne, baba, ocuk llerinin eksiksiz bir řekilde toplanmasındaki lojistik ve fiziki glklerden kaynaklanmaktadır.

Kanser arařtırmalarında yapılan genom boyu asosiyasyon alıřmalarında ve/veya bu alıřmaların dođrulama basamaklarında ise rnek materyali olarak poplasyondan toplanan olgu ve kontrol rnekleri ve/veya bunlara ek olarak trio'lar kullanılmaktadır.

Bugne kadar yayınlamıř genom boyu asosiyasyon alıřmalarında yaygın olarak kullanılan alıřma dizaynı poplasyon temelli yaklařımdır. Aile temelli genom boyu asosiyasyon yaklařımı genellikle dođrulama basamaklarında kullanılmasına (Hinney et al. 2007; Graham et al. 2008; Wallace et al. 2008) rađmen hem ilk tarama basamađında hem de

doğrulama basamağında aile temelli yaklaşım kullanan çalışmalar da literatürde yer almaktadır (Scuteri et al. 2007; Concannon et al. 2008; Lasky-Su et al. 2008)

İlk tarama basamağında populasyon temelli yaklaşımı kullanan çalışmalarda, doğrulama, tekrarlama basamağında aile temelli yaklaşımın kullanılması popülasyon stratifikasyonundan kaynaklanacak potansiyel anlamlı asosiyasyonları ortadan kaldırmaktadır. Bu nedenle bu tür çalışmalardaki bu yaklaşım oldukça cazip bir dizayn olarak karşımıza çıkmaktadır.

Genom-boyu asosiyasyon çalışmalarında tekrarlama basamağında devamlı kullanılan aile temelli yaklaşım gelecekte daha yaygın hale gelecektir.

SPRT

Gerçekleştirilmiş bu tez projesinin araştırma konusunun geç yaşta ortaya çıkan kolorektal kanser olması nedeniyle TDT testinin ihtiyaç duyduğu 200 trioya çıkılması hedefine ulaşıp ulaşılamayacağı tez projesinin başından itibaren sorulan ve üzerinde titizlikle durulan bir konu olmuştur. Bu nedenle bir yandan ulaşılan hasta sayısını arttırmaya çalışırken diğer yandan da mevcut trio verileri ile hangi sonuçlara ulaşılabılır sorusunun cevabı aranmıştır. Bu kapsamda istatistik biliminin araçlarından biri olan ardışık hipotez testlerinin (SPRT) aile temelli asosiyasyon analizlerine uygulanabilirliği sorgulanmış ve bu yaklaşımın tez projesi kapsamında sorulan soruya cevap olabildiği gösterilmiştir. Yeni geliştirilen bu analiz ile aile temelli asosiyasyon çalışmalarındaki kısıtlı örnek sayısı sorununa oldukça önemli bir çözüm sunulmuştur.

Yukarıda da ifade edildiği üzere geliştirilen bu yeni analizde ardışık hipotez testleri (sequential probability ratio test, SPRT) prensibi kullanılmıştır. Buna göre SPRT testleri, çalışmanın başında örneklem sayısına karar vermez. Aksine, örneklemeleri teker teker alır, testi uygular ve üç karardan birine ulaşır: Hipotezi reddeder, kabul eder veya yeterince bilgi olmadığını ve örneklem alınmasına devam edilmesi gerektiğini belirtir. Daha önce yapılan simülasyon çalışmaları, ortalamada, ardışık hipotez testlerinin çalışmanın başında örneklem sayısını sabitlemeyi gerektiren klasik testlerden daha az örneklem gerektirdiğini göstermiştir (van der Tweel et al. 2003). Ayrıca, gerçek veriyle yapılan çalışmalar da, SPRT'nin örneklem boyutu açısından avantajını desteklemiştir. Örneğin, van der Tweel ve van Noord makalelerinde SPRT'nin biyolojik örneklem, zaman, bütçe vb. kaynaklardan

%35-65 arasında kazanç sağlayabileceğini belirtmiştir. Özellikle kalite-kontrol gibi bir çok alanda kullanılan bu testler *bugüne dek aile bazlı asosiyasyon çalışmalarında önerilmemiştir.*

Tez verisine uygulanan yeni analiz yaklaşımının aile temelli asosiyasyon çalışmalarına uygulanabilirliği ve sonuçlarının güvenilirliğini sorgulayabilmek üzere öncelikle bir benzetim (simülasyon) çalışması gerçekleştirilmiş ve SPRT adını verdiğimiz bu analiz yaklaşımı TDT ile karşılaştırıldığında özellikle az sayıda trio olması durumunda, birçok kriterde SPRT'nin TDT'ye üstünlüğü gözlenmiştir. SPRT, 50 trio ile bile yüksek düzeyde asosiyasyon gösterdiği varsayılan SNP'lerin % 60'ı hakkında karar verilebilmiştir. Geriye kalan %40'ı içinse 'daha fazla trio toplamaya devam et' kararı çıkmıştır. 100'den ufak trio sayısı için, SPRT %87.9 ila %93.8 arasında doğruluk değeri alırken, TDT sadece %39.9 ila %74.8 arasında değerler almıştır. 50 trio için doğruluk değerlerine bakıldığında SPRT %92.7'lerde iken TDT'nin ise %57.9'larda kaldığı görülmektedir (Çizelge 2.3).

Proje için toplanan 51 trio gerçekleştirilen SPRT analizi sonuçlarına göre 75 SNP anlamlı bulunmuştur. Bu 75 SNP, toplam 20 kromozomda 67 farklı kromozom lokusunda bulunmaktadır (Çizelge 4.7). Ayrıca 75 SNP'in toplam 119 genin intronunda, üst bölgesinde (upstream) veya alt bölgesinde (downstream) bulunduğu belirlenmiştir. Bu 119 gen ile yolak analizleri yapıldığında kolorektal kanser yolağında, hücre döngüsü yolağında ve birçok kanser yolağında da bulunan ve önemli bir onkogen olduğu bilinen MYC geni, genel kanser yolağında MYC genine ek olarak MECOM (MDS1 ve EVI1 kompleks lokusu) geni, MAPK sinyal yolağında MECOM, PRKX, PTPN5 ve yine MYC geni bulunmaktadır. Wnt sinyal yolağında ise MAPK yolağı ile ortak olarak PRKX ve MYC geni bulunmaktadır. Hücre adezyon moleküllerinden CADM1 de kolorektal kanserle ilişkili olabileceği düşünülen 119 genden biridir. Ayrıca apoptozis ve Hedgehog sinyal yolaklarında ise PRKX geni bulunmaktadır. Yukarıda adı geçen tüm sinyal yolaklarının kolorektal kanser gelişimi üzerinde etkili olduğu bilinmektedir. Böylece, SPRT analizi sonucunda ilişkili olabileceği belirlenmiş genlerin ne derecede önemli olduğu açıkça görülmektedir.

C-TDT

Tezin materyel kısmında ifade edildiği üzere bu proje kapsamında 51 trionun yanısıra 33 duo toplanmıştır. Yöntem bölümünde ayrıntılı olarak açıklanan C-TDT ve permütasyon testleri trio verilerinin duo verileri ile analizini mümkün kılmaktadır. Böylece, sadece trioları kullanan TDT veya sadece duoları kullanan 1-TDT testlerinden daha güçlü (daha fazla veri kullanması nedeniyle) bir test kullanılmakla birlikte, permütasyon testi sayesinde herhangi bir dağılım varsayımı yapılmadan ve küçük örneklemelere uygulanabilen bir yöntem ortaya çıkmıştır (Hu et al. 2010).

C-TDT istatistiği hakkında şunlar söylenebilir: Daha önce de belirtildiği üzere, çalışmanın anne-baba-çocuk üçlüsü ve çocuk-ebeveyn ikilisinden oluşması durumunda, C-TDT istatistiğinin, sadece anne-baba-çocuk üçlüsü kullanan TDT'ye göre daha güçlü sonuçlar verdiği gösterilmiştir (Hu et al. 2010). Diğer bir taraftan, bu test istatistiği asimptotik bir dağılım yerine, permütasyon yöntemine dayandığı için, küçük örneklem büyüklüklerinde de etkili bir şekilde kullanılabilirdiği Good tarafından bildirilmektedir (Good 2001). Ancak bu yöntem birden fazla taşıyıcı çocuktan oluşan ailelere veya sadece anne-çocuk ikililerinden ya da sadece baba-çocuk ikililerinden oluşan çalışmalara uygulanamamaktadır (Hu et al. 2010). Ayrıca, bu test istatistiği, bir ebeveynin kayıp olma sebebinin genotipinin olmadığı şartını arayan, tamamıyla rastlantısal kayıp varsayımı altında çalışmaktadır (Hu et al 2010).

C-TDT'nin aile temelli asosiyasyon çalışmalarına uygulanabilirliği ve sonuçlarının güvenilirliğini sorgulayabilmek üzere öncelikle bir benzetim (simülasyon) çalışması gerçekleştirilmiş ve bu çalışmanın sonuçlarına göre 50 trio ve 30 duo ile yapılan C-TDT analizinde % 83'e yakın bir güç elde edildiği gösterilmiştir.

Tez projesi kapsamında toplanan 51 trio ve 33 duo ile gerçekleştirdiğimiz C-TDT sonuçlarına göre ise 15,093 SNP ($p < 0.001$) hastalıkla ilişkili olarak tanımlanmıştır. C-TDT analizi sonucunda, daha önce kolorektal kanserle ilişkili olduğu bildirilen 10 lokustan (18q21.1, 8q23.3, 8q24.21, 10p14, 11q23.1, 14q22.2, 15q13.3, 16q22.1, 19q13.11, 20p12.3) 4'ü (8q23.3, 8q24.21, 10p14, 14q22.2) hastalık ile ilişkili olarak tanımlanmıştır.

Doğrulama Çalışmaları

Aday 5 SNP

Yapılan biyoinformatik analizler (SPRT ve C-TDT) sonucunda anlamlı çıkan kromozom bölgelerinde bulunan bazı genler aday olarak seçilmişlerdir.

Kolorektal karsinogenezde aday tümör baskılayıcı gen olarak tanımlanmış ve kolon kanseri hücrelerinde hücre adezyonu ve migrasyonunda rol oynadığı gösterilmiş DCC (rs2229080) (kolorektal kanserde delete olmuş, deleted in colorectal cancer) geni ilk aday gen olarak seçilmiştir (Martin et al. 2006).

Kolorektal karsinogenezle daha önceden ilişkilendirilmiş aday genlerimizden biri, normal kolon dokusunda ekspresyonu olmayan fakat kolorektal karsinoma dokularında ekspresyonu saptanmış ve dolayısıyla kolorektal kanserde malignite ile ilişkilendirilmiş olan COL11A1(rs3753841) genidir (Fischer et al. 2001). Diğeri ise stromal hücre kökenli faktör 1 (stromal cell-derived factor 1, SDF1)' in reseptörü olan ve birçok farklı tümörde olduğu gibi kolorektal kanserde de anjiogenez ve metastaz süreçlerinde etkin rol oynadığı gösterilen ve bu nedenle özellikle anti anjiogenik terapide terapötik potansiyel hedef olabileceği düşünülen CXCR4 (rs56400844) genidir (Ingold et al. 2009; Kollmar et al. 2010).

Ayrıca karsinogenezde önemli bir yolak olan apoptoz yolağında yer alan genler arasında bulunan TMBIM4 (rs8793) ve overekspresyonunun apoptozisi blokladığı gösterilen PDCD6IP (rs3792594) genleri kolorektal karsinogenezdeki olası rolleri nedeni ile aday genler olarak seçilmişlerdir.

Bu aday 5 gen içinden 1 SNP genotiplendirilmek üzere seçilmiş ve genotiplendirilmiştir. Genler içindeki bu SNP'ler, protein dizisindeki amino asitte farklılık yaratan SNP (nonsynonymous SNP) olması ve daha önce çalışılan populasyonlarda frekans bilgisi var ise frekansının çok yüksek olmamasına dikkat edilerek seçilmişlerdir.

5 lokusta da elde edilen P-değerlerinin verilen anlamlılık düzeyi olan 0.05'ten büyük olması sonucu, 5 lokus için de hasta ve kontrol grupları arasında anlamlı bir fark olmadığı saptanmıştır.

Haplotip Analizi

Biyoinformatik analizler (SPRT ve C-TDT) sonucunda elde edilen anlamlı SNP'ler, kromozomdaki fiziksel pozisyonlarına göre düzenlendiğinde çocuğa (hasta) aktarılan allellerin durumu hakkında net bir bilgi sahibi olunabilmektedir.

Özellikle akraba evliliklerinin olduğu durumlarda resesif özellik gösteren genetik hastalıkların nedeni olan genlerin tespitinde yararlanılan homozigosite haritalaması yaklaşımının aile içi evliliğin bulunmadığı durumlarda da kullanılabilirliği gösterilmiştir (Hildebrandt et al. 2009). Burada hareket noktası hastalıkla ilgili aday genlerin homozigot allel blokları içinde yer alma durumudur.

SPRT ve C-TDT'nin ortaya çıkardığı istatistiksel olarak anlamlı SNP verileri üzerinden gerçekleştirilen homozigosite haritalaması, 6 trio'da 3p22.3 lokusunda böyle bir bölgenin varlığını işaret etmiştir. Bu durum anılan lokusun kolorektal kanser için aday yatkınlık lokusu olabileceğini işaret etmektedir. Bu nedenle bu lokusun daha ayrıntılı incelenmesine karar verilmiştir. Bu lokusta sadece iki gen (ARPP21 ve MIR128-2) bulunduğu belirlenmiş ve bu iki gen de DNA dizi analizi ile ayrıntılı olarak incelenmiştir. Her iki genin DNA dizisinde de bir mutasyon saptanmamıştır.

Kopya Sayısı Değişiklikleri

84 hastanın ayrı ayrı CNV analizleri yapıldıktan sonra bu hastalarda ortak olarak bulunan delesyon ve/veya amplifikasyonları saptamak için yapılan analiz sonucunda delete olan ilgili kromozom lokuslarında bulunan genler incelendiğinde:

4q13.2 lokusunda iki tanesi pseudogen ve daha önce endometriyal kanserle ilişkilendirilmiş UGT2B17 geni olmak üzere toplam 3 gen bulunmakta olduğu, 9p12 lokusunun ilgili bölgesinde hiç gen bulunmadığı ve 15q11.1bölgesinde 9 genin ve 15q11.2 bölgesinde 13 genin bulunmakta olduğu görülmüştür (Çizelge 4.18).

Amplifiye olan kromozom lokuslarında bulunan genler incelendiğinde ise, 2q11.1 bölgesinde daha önce kanserle ilişkilendirilmemiş iki gen ve iki de pseudogen bulunduğu, 17q21.31 bölgesinde 4'ü pseudogen olmak üzere toplam 5 gen olduğu, 19p13.2 bölgesinde

daha önce kanserle ilişkilendirilmiş RNASEH2A geni, kanserle ilişkilendirilmemiş 2gen ve 1 pseudogen bulunduğu, Xp11.23 lokusunda ise 8'i miRNA genleri olmak üzere toplam 10 gen bulunduğu, Xq21.1'de 1 pseudogen bulunduğu ve Xq28 de sadece daha önce mide kanseri ile ilişkilendirilmiş TKTL1 geninin bulunduğu görülmüştür (Çizelge 4.19).

CNV veri tabanı incelendiğinde delesyon saptanan tüm bölgelerin daha önce bildirildiği 2q11.1, 19p13.2, Xq21.1 ve Xq28 lokuslarındaki amplifikasyonların daha önce bildirilmediği, 17q21.31 ve Xp11.23 lokuslarındaki kopya sayısı değişikliğinin daha önce bildirildiği saptanmıştır.

SONUÇ ve ÖNERİLER

Kolorektal kanser arařtırmalarında yaygın olarak kullanılan genom boyu asosiyasyon çalışmalarında, hem sporadik hem de aile öyküsü olan kolorektal kanser olguları kullanılmıştır. Bugüne kadar yapılan asosiyasyon çalışmalarının ortak özellikleri ise popülasyon temelli asosiyasyon çalışmaları olmalarıdır. Sunulan bu tez çalışması aile bazlı asosiyasyon yaklaşımı ile sporadik kolorektal kanserlerde yeni kanser yatkınlık genlerinin saptanmasında *bir ilk olmuştur*.

Bu tez çalışması kapsamında, geç yaşlarda gözlenen bir hastalık incelendiği için, klasik istatistiksel yöntemlerin gerektirdiği sayılarda anne-baba-çocuk üçlüsüne ulaşılmakta problem yaşanmıştır. Bu problemi aşmak için, iki istatistiksel yöntem üzerinde çalışılmıştır. Öncelikle, sadece üçlü veri kullanılarak ardışık hipotez testlerini ('sequential hypothesis tests') önerip, bir benzetim çalışması yürütülmüş ve bu çalışma European Journal of Human Genetic dergisinde yayınlanmıştır. Ekibimizce geliştirilip, bu tez çalışmasında elde ettiğimiz 51 trio' dan oluşan kolorektal kanser verisine uyguladığımız bu yeni analiz (SPRT) yaklaşımı ile aile temelli genom boyu asosiyasyon çalışmalarının önündeki en büyük sorun olan örnek sayısının yetersizliği problemi çözülmüş olup, az sayıda trio ile çalışılmasına imkan verilerek, bilim dünyasına önemli bir katkıda bulunulmuştur.

Üzerinde çalışılan ikinci yöntem olarak, Hu ve Zhou (2010) çalışması takip edilerek hem üçlü hem de ikili veri (trio ve duo) bir arada değerlendirilerek C-TDT uygulanmıştır.

Bugüne kadar İngiltere, İskoçya, Japonya, İsrail ve Kanada’da kolorektal kanser üzerinde yapılan asosiyasyon çalışmaları ile eşik değeri $p < 10^{-5}$ olan 10 farklı (8q24.21, 8q23.3, 10p14, 11q23.1, 14q22.2, 15q13.3, 16q22.1, 18q21.1, 19q13.11 ve 20p12.3) yatkınlık lokusu tanımlanmıştır (Le Marchand 2009; Galvan et al. 2010).

Bu tez çalışmasında SPRT analizleri 75, C-TDT analizleri 15,093 SNP’yi kolorektal kanser ile ilintilendirmiştir. Bu analizlerin sonuçları bir anlamda genom üzerinde kolorektal kanser ile ilintili olabilecek lokusları işaret eden bir pusula niteliğindedir. Bu verilerin belirlediği ve işaret ettiği bilgileri maddeler halinde kristalize etmek gerekirse:

SPRT analizinin belirlediği 75 SNP, C-TDT’in tespit ettiği 15,093 SNP içinde yer almaktadır.

Tespit edilen SNP’lerin kromozomlar üzerindeki lokuslarına bakıldığında SPRT de aralarında en fazla 500 000 bç olan SNP ler seçildiğinde, C-TDT de ise aynı bölgede 30 veya daha fazla SNP’in bulunması esas kriteri olarak alındığında:

SPRT ile 1p36.12, 3p22.3, 5q11.2, 11q22.1, 20p12.1 ve C-TDT ile 2p16.3, 2p22.3, 2q14.1, 2q22.1-2q22.2, 3p26.1, 4q28.3, 4q34.3, 5p15.2, 5q34, 6q16.3, 8q24.3, 9p21.3, 11q14.1, 11q22.1, 14q12, 14q31.3, 14q32.2 lokuslarını asosiyasyona aday bölgeler olarak ortaya koymuştur.

SPRT’nin ortaya koyduğu lokusların tümü (1p36.12, 3p22.3, 5q11.2 11q22.1, 20p12.1) daha önce kolorektal kanser ve diğer kanserlerle ilintilendirilmiş bölgeleri içermektedir.

C-TDT’nin ortaya koyduğu lokuslardan 6q16.3, 9p21.3, 11q22.1, 14q31.3, 14q32.2 daha önce kanser ile ve 2p16.3, 2p22.3, 2q14.1, 2q22.1-2q22.2, 3p26.1, 4q28.3, 4q34.3, 5p15.2, 5q34, 8q24.3, 11q14.1, 14q12 lokusları ise daha önce kolorektal kanser ile ilintilendirilmiş bölgeleri içermektedir.

C-TDT analizi sonucunda, daha önce kolorektal kanserle ilişkili olduğu bildirilen 10 lokustan (18q21.1, 8q23.3, 8q24.21, 10p14, 11q23.1, 14q22.2, 15q13.3, 16q22.1, 19q13.11, 20p12.3) 4’ü (8q23.3, 8q24.21, 10p14, 14q22.2) hastalık ile ilişkili olarak tanımlanmıştır.

SPRT analizlerinin ortaya koyduğu 75 SNP esas alınarak yürütülen haplotip analiz ve homozigosite haritalaması 3p22.3 bölgesini 6 trio için aday olarak göstermektedir. C-TDT

analizinin de belirlemiř olduđu bu lokustaki genlerin DNA dizi analizleri yapılmıř fakat DNA dizisinde herhangi bir deęiřiklik saptanmamıřtır.

CNV analizi ile 4q13.2, 8p23.1, 9p12 ve 15q11.1-15q11.2 kromozom lokuslarında çeřitli baz çifti uzunluklarındaki bölgelerde delesyonlar ve 2q11.1, 17q21.31, 19p13.2, Xq21.2, Xq28, Xp11.23 lokuslarında ise yine farklı uzunluklardaki bölgelerde amplifikasyonlar saptanmıřtır. Delesyon saptanan tüm bölgelerin ve amplifikasyon saptanan bazı bölgelerin daha önce kopya sayısı varyasyonu olarak tanımlandığı tespit edilmiřtir.

Tez projesinde kolorektal kanser ile ilintili bulunmuř olan SPRT'nin ve C-TDT'nin tespit ettiđi SNP verisine bađlı lokus ve genlerin analizleri devam edecektir.

KAYNAKLAR

- Allen, A.S., Rathouz, P.J., Satten, G.A. 2003. Informative missingness in genetic association studies: case-parent designs. *Am J Hum Genet*, 72(3): 671-80.
- Altekruse S.F, Kosary, C.L, Krapcho, M., Neyman, N., Aminou, R., Waldron, W., Ruhl, J., Howlander, N., Tatalovich, Z., Cho, H., Mariotto, A., Eisner, M.P., Lewis, D.R., Cronin, K., Chen, H.S., Feuer, E.J., Stinchcomb, D.G., Edwards, B.K. 2010. SEER Cancer Statistics Review, 1975-2007. Bethesda.
- Ashktorab, H., Schaffer, A.A., Daremipouran, M., Smoot, D.T., Lee, E., Brim, H. 2010. Distinct genetic alterations in colorectal cancer. *PLoS One*, 5(1): e8879.
- Benjamini, Y., Drai, D., Elmer, G., Kafkafi, N., Golani, I. 2001. Controlling the false discovery rate in behavior genetics research. *Behav Brain Res*, 125(1-2): 279-84.
- Benke, K., Fallin, M.D. 2010. Methods: genetic epidemiology. *Clin Lab Med*, 304: 795-814.
- Bertram, J. S. 2001. The molecular biology of cancer. *Molecular Aspects of Medicine*, 21: 167-223.
- Bowcock, A.M. 2007. Genomics: Guilt by association. *Nature*, (447): 645-646.
- Campisi, J. 2005. Senescent cells, tumor suppression, and organismal aging: good citizens, bad neighbors. *Cell*, 120(4): 513-22.
- Cardon, L.R., Palmer, L.J. 2003. Population stratification and spurious allelic association. *Lancet*, 361(9357): 598-604.
- Cervino, A.C.L., Lakiss, S., Sow, O., Hill, A.V.S. 2000. Allelic association between the NRAMP1 gene and susceptibility to tuberculosis in Guinea-Conakry. *Ann Hum Genet*, 64(6): 507-12.
- Cheah, P.Y., 2009. Recent advances in colorectal cancer genetics and diagnostics. *Crit Rev Oncol Hematol* 69(1):45-55. Epub 2008 Sep 6.
- Chen, W.M., Deng, H.W. 2001. A general and accurate approach for computing the statistical power of the transmission disequilibrium test for complex disease genes. *Genet Epidemiol*, 21(1): 53-67.
- Cichon, S., Craddock, N., Daly, M., Faraone, S.V., Gejman, P.V., Kelsoe, J., Lehner, T., Levinson, D.F, Moran, A., Sklar, P., Sullivan, P.F. 2009. Genomewide association studies: history, rationale, and prospects for psychiatric disorders. *Am J Psychiatry*, 166(5): 540-56.

- Collins, A. R. 2007. Linkage disequilibrium and association mapping: an introduction. *Methods Mol Biol*, 376: 1-15.
- Concannon, P., Onengut-Gumuscu, S., Todd, J.A., Smyth, D.J., Pociot, F., Bergholdt, R., Akolkar, B., Erlich, H.A., Hilner, J.E., Julier, C., Morahan, G., Nerup, J., Nierras, C.R., Chen, W.M., Rich, S.S., Type 1 Diabetes Genetics Consortium. 2008. A human type 1 diabetes susceptibility locus maps to chromosome 21q22.3. *Diabetes*, 57(10): 2858-61.
- Dracopoli NC., H. J., Korf BR., Moir DT., Morton CC., Seidman CE., Seidman JG., Smith DR. 1994. Genetic Mapping : Current Protocols in Human Genetics.
- Duncan, D.C., Thomas, R.W. Duggan, D. 2005. Recent Developments in Genomewide Association Scans: A Workshop Summary and Review. *Am. J. Hum. Genet*, (77): 337-345.
- Engel, L.S., Taioli, E., Pfeiffer, R., Garcia-Closas, M., Marcus, P.M., Lan, Q., Boffetta, P., Vineis, P., Autrup, H., Bell, D.A., Branch, R.A., Brockmüller, J., Daly, A.K., Heckbert, S.R., Kalina, I., Kang, D., Katoh, T., Lafuente, A., Lin, H.J., Romkes, M., Taylor, J.A., Rothman, N.2002. Pooled analysis and meta-analysis of glutathione S-transferase M1 and bladder cancer: a HuGE review. *Am J Epidemiol*, 156(2): 95-109.
- Ewens, W. J., Spielman, R.S. 1995. The transmission/disequilibrium test: history, subdivision, and admixture. *Am J Hum Genet*, 57(2): 455-64.
- Fischer, H., Salahshor, S., Stenling, R., Björk, J., Lindmark, G., Iselius, L., Rubio, C., Lindblom, A. 2001. COL11A1 in FAP polyps and in sporadic colorectal tumors. *BMC Cancer*, 1: 17.
- Freedman, M.L., Monteiro, A.N., Gayther, S.A., Coetzee, G.A., Risch, A., Plass, C., Casey, G., De Biasi, M., Carlson, C., Duggan, D., James, M., Liu, P., Tichelaar, J.W., Vikis, H.G., You, M., Mills, I.G. 2011. Principles for the post-GWAS functional characterization of cancer risk loci. *Nat Genet.*, 43(6):513-8.

- Galvan, A., Ioannidis, J.P.A., Dragani, T.A. 2010. Beyond genome-wide association studies: genetic heterogeneity and individual predisposition to cancer. *Trends Genet*, 26(3): 132-41.
- Garcea, G., A., Sharma, R.A., Dennison, A., Steward, W.P., Gescher, A., Berry, D.P. 2003. Molecular biomarkers of colorectal carcinogenesis and their role in surveillance and early intervention. *Eur J Cancer*, 39(8): 1041-52.
- Good, P. 2001. *Permutation Tests: A Practical Guide to Resampling Methods for Testing Hypotheses*. *The Indian Journal of Statistics*, 63(1):134–137.
- Graham, R.R., Cotsapas, C., Davies, L., Hackett, R., Lessard, C.J., Leon, J.M., Burtt, N.P., Guiducci, C., Parkin, M., Gates, C., Plenge, R.M., Behrens, T.W., Wither, J.E., Rioux, J.D., Fortin, P.R., Graham, D.C., Wong, A.K., Vyse, T.J., Daly, M.J., Altshuler, D., Moser, K.L., Gaffney, P.M. 2008. Genetic variants near TNFAIP3 on 6q23 are associated with systemic lupus erythematosus. *Nat Genet*, 40(9): 1059-61.
- Gupta, P.K., Rustgi, S., Mir, R.R. 2008. Array-based high-throughput DNA markers for crop improvement. *Heredity*, 101(1);5-18.
- Hardy, J., Singleton, A. 2009. Genomewide association studies and human disease. *N Engl J Med*, 360(17):1759-68.
- Hildebrandt, F., Heeringa, S.F., Rüschenhoff, F., Attanasio, M., Nürnberg, G., Becker, C., Seelow, D., Huebner, N., Chernin, G., Vlangos, C.N., Zhou, W., O'Toole, J.F., Hoskins, B.E., Wolf, M.T., Hinkes, B.G., Chaib, H., Ashraf, S., Schoeb, D.S., Ovunc, B., Allen, S.J., Vega-Warner, V., Wise, E., Harville, H.M., Lyons, R.H., Washburn, J., Macdonald, J., Nürnberg, P., Otto, E.A. 2009. A systematic approach to mapping recessive disease genes in individuals from outbred populations. *PLoS Genet*, 5(1): e1000353.
- Hildesheim, A., Apple, R.J., Chen, C.J., Wang, S.S., Cheng, Y.J., Klitz, W., Mack, S.J., Chen, I.H., Hsu, M.M., Yang, C.S., Brinton, L.A., Levine, P.H., Erlich, H.A. 2002. Association of HLA class I and II alleles and extended haplotypes with nasopharyngeal carcinoma in Taiwan. *J Natl Cancer Inst*, 94(23): 1780-9.
- Hinney, A., Nguyen, T.T., Scherag, A., Friedel, S., Brönner, G., Müller, T.D., Grallert, H., Illig, T., Wichmann, H.E., Rief, W., Schäfer, H., Hebebrand, J. 2007. Genome wide association (GWA) study for early onset extreme obesity supports the role of fat mass and obesity associated gene (FTO) variants. *PLoS One*, 2(12): e1361.
- Horvath, S., Laird, N.M. 1998. A discordant-sibship test for disequilibrium and linkage: no need for parental data. *Am J Hum Genet*, 63(6): 1886-97.

- Hu, Y. Q., Zhou, J.Y. 2010. Inferring haplotype/disease association by joint use of case-parents trios and case-parent pairs. *Annals of Human Genetics*, (74): 263-274.
- Ilk O, R. F., Dungul, DC., Ozdag, H., Ilk, HG. 2011. A novel approach for small sample size family-based association studies: sequential tests. *European Journal of Human Genetics*, 2011 Mar 23, Epub ahead of print.
- Ingold, B., Schulz, S., Budczies, J., Neumann, U., Ebert, M.P., Weichert, W., Röcken, C. 2009. The role of vascular CXCR4 expression in colorectal carcinoma. *Histopathology*, 55(5): 576-86.
- Iyengar, S.K., Elston, R.C. 2007. The genetic basis of complex traits: rare variants or "common gene, common disease"? *Methods Mol Biol*, 376: 71-84.
- Kinzler, K.W., Vogelstein, B. 1997. Cancer-susceptibility genes. Gatekeepers and caretakers. *Nature*, 386(6627): 761, 763.
- Kleinsmith, L.J. 2006. *Principles of Cancer Biology*. Benjamin Cummings, 352, Michigan.
- Knudson, A.G. 1971. Mutation and cancer: statistical study of retinoblastoma. *Proc Natl Acad Sci USA*, 68(4): 820-3.
- Kollmar, O., Rupertus, K., Scheuer, C., Nickels, R.M., Haberl, G.C., Tilton, B., Menger, M.D., Schilling, M.K. 2010. CXCR4 and CXCR7 regulate angiogenesis and CT26.WT tumor growth independent from SDF-1. *Int J Cancer*, 126(6): 1302-15.
- Kruglyak, L. 2008. The road to genome-wide association studies. *Nat Rev Genet* 9(4): 314-8.
- Rodriguez-Revenga, L., Mila, M., Rosenberg, C., Lamb, A., Lee, C. 2007. Structural variation in the human genome: the impact of copy number variants on clinical diagnosis. *Genetics in Medicine*, 9(9): 600-606.
- Landi, S., Moreno, V., Gioia-Patricola, L., Guino, E., Navarro, M., de Oca, J., Capella. G., Canzian, F. 2003. Association of common polymorphisms in inflammatory genes interleukin (IL)6, IL8, tumor necrosis factor alpha, NFkB1, and peroxisome proliferator-activated receptor gamma with colorectal cancer. *Cancer Res*, 63(13): 3560-6.
- Lasky-Su, J., Anney, R.J., Neale, B.M., Franke, B., Zhou, K., Maller, J.B., Vasquez, A.A., Chen, W., Asherson, P., Buitelaar, J., Banaschewski, T., Ebstein, R., Gill, M., Miranda, A., Mulas, F., Oades, R.D., Roeyers, H., Rothenberger, A., Sergeant, J.,

- Sonuga-Barke, E., Steinhausen, H.C., Taylor, E., Daly, M., Laird, N., Lange, C., Faraone, S.V. 2008. Genome-wide association scan of the time to onset of attention deficit hyperactivity disorder. *Am J Med Genet B Neuropsychiatr Genet*, 147B(8): 1355-8.
- Le Marchand, L. 2009. Genome-wide association studies and colorectal cancer. *Surg Oncol Clin N Am*, 18(4): 663-8.
- Lichtenstein, P., Holm, N.V., Verkasalo, P.K., Iliadou, A., Kaprio, J., Koskenvuo, M., Pukkala, E., Skytthe, A., Hemminki, K. 2000. Environmental and heritable factors in the causation of cancer--analyses of cohorts of twins from Sweden, Denmark, and Finland. *N Engl J Med*, 343(2): 78-85.
- Lynch, H. T., Boland, C.R., Gong, G., Shaw, T.G., Lynch, P.M., Fodde, R., Lynch, J.F., de la Chapelle, A. 2006. Phenotypic and genotypic heterogeneity in the Lynch syndrome: diagnostic, surveillance and management implications. *Eur J Hum Genet*, 14(4): 390-402.
- Lynch, H. T., Lynch, J.F., Attard, T.A. 2009. Diagnosis and management of hereditary colorectal cancer syndromes: Lynch syndrome as a model. *CMAJ*, 181(5): 273-80.
- Manolio, T.A., Brooks, L.D., Collins, F.S. 2008. A HapMap harvest of insights into the genetics of common disease. *J Clin Invest.*, 118(5):1590-605.
- Martin, E. R., Monks, S.A, Warren, L.L., Kaplan, N.L. 2000. A test for linkage and association in general pedigrees: the pedigree disequilibrium test. *Am J Hum Genet*, 67(1): 146-54.
- Martin, M., Simon-Assmann, P., Kedinger, M., Martin, M., Mangeat, P., Real, F.X., Fabre, M. 2006. DCC regulates cell adhesion in human colon cancer derived HT-29 cells and associates with ezrin. *Eur J Cell Biol*, 85(8): 769-83.
- Martinez et al., J. D., Parker, M.D., Fultz, K.E., Ignatenko, N.A., Gerner, E.W. 2003. *Molecular Biology of Cancer*. John Wiley&Sons, Inc, Chapter One, 1-50, Arizona
- McCarthy, M.I., Abecasis, G.R., Cardon, L.R., Goldstein, D.B., Little, J., Ioannidis, J.P., Hirschhorn, J.N. 2008. Genome-wide association studies for complex traits: consensus, uncertainty and challenges. *Nat Rev Genet*, 9(5): 356-69.

- Miller, R.D., Kwok, P.Y. 2001. The birth and death of human single-nucleotide polymorphisms: new experimental evidence and implications for human history and medicine. *Hum Mol Genet*, 10(20):2195-8.
- Moore, J. H., Asselbergs, F.W., Williams, S.M. 2010. Bioinformatics challenges for genome-wide association studies. *Bioinformatics*, 26(4): 445-55.
- Mutch, M. G. 2007. Molecular profiling and risk stratification of adenocarcinoma of the colon. *J Surg Oncol*, 96(8): 693-703.
- Naccarati, A., Pardini, B., Hemminki, K., Vodicka, P. 2007. Sporadic colorectal cancer and individual susceptibility: a review of the association studies investigating the role of DNA repair genetic polymorphisms. *Mutat Res*, 635(2-3): 118-45.
- Oláh, E. 2005. Basic Concepts of Cancer: Genomic Determination." *The Journal of The International Federation of Clinical Chemistry and Laboratory Medicine* ,16(2).
- Pharoah, P. D., Dunning, A.M., Ponder, B.A., Easton, D.F. 2004. Association studies for finding cancer-susceptibility genetic variants. *Nat Rev Cancer*, 4(11): 850-60.
- Pitot, H. C. 1993. The molecular biology of carcinogenesis. *Cancer*, 72(3): 962-70.
- Burt, R. W., Bishop, D.T., Cannon-Albright, L., Sarnowitz, W.S., Lee, R. L., DiSario, J. A., Skolnick, M.H. 2006. Hereditary aspects of colorectal adenomas. *Cancer*, 70(S3): 1296-1299.
- Sachse, C., Smith, G., Wilkie, M.J., Barrett, J.H., Waxman, R., Sullivan, F., Forman, D., Bishop, D.T., Wolf, C.R. 2002. A pharmacogenetic study to investigate the role of dietary carcinogens in the etiology of colorectal cancer. *Carcinogenesis* 23(11): 1839-49.
- Schork, N.J. 1997. Genetics of complex disease: approaches, problems, and solutions. *Am J Respir Crit Care Med*, 156(4 Pt 2):S103-9.
- Scuteri, A., Sanna, S., Chen, W.M., Uda, M., Albai, G., Strait, J., Najjar, S., Nagarajah, R., Orru, M., Usala, G., Dei, M., Lai, S., Maschio, A., Busonero, F., Mulas, A., Ehret, G.B., Fink, A.A., Weder, A., Cooper, R., Galan, P., Chakravarti A., Schlessinger, D., Cao, A., Lakatta, E., Abecasis, G.R. 2007. Genome-wide association scan shows genetic variants in the FTO gene are associated with obesity-related traits. *PLoS Genet*, 3(7): e115.

- Slattery, M. L., Herrick, J., Curtin, K., Samowitz, W., Wolff, R.K., Caan, B.J., Duggan, D., Potter, J.D., Peters, U. 2010. Increased risk of colon cancer associated with a genetic polymorphism of SMAD7. *Cancer Res*, 15(70): 1479-85.
- Spielman, R. S., Ewens, W.J. 1998. A sibship test for linkage in the presence of association: the sib transmission/disequilibrium test. *Am J Hum Genet*,62(2): 450-8.
- Spielman, R.S., McGinnis, R.E., Ewens, W.J. 1993. Transmission test for linkage disequilibrium: the insulin gene region and insulin-dependent diabetes mellitus (IDDM). *Am J Hum Genet*, 52(3): 506-16.
- Stadler, Z.K., Vijai, J., Thom, P., Kirchhoff, T., Hansen, N.A., Kauff, N.D., Robson, M., Offit, K. 2010. Genome-wide association studies of cancer predisposition. *Hematol Oncol Clin North Am*, 24(5): 973-96.
- Sun, F., Flanders, W. D., Yang, Q., Khoury, M.J. 1999. Transmission disequilibrium test (TDT) when only one parent is available: the 1-TDT. *Am J Epidemiol*, 150(1): 97-104.
- Takayama, T., Miyamishi, K., Hayashi, T., Sato, Y., Nitsu, Y. 2006. Colorectal cancer: genetics of development and metastasis. *J Gastroenterol*, 41(3): 185-92.
- Tomlinson, I. P., Dunlop, M., Campbell, H., Zanke, B., Gallinger, S., Hudson, T., Koessler, T., Pharoah, P.D., Niittymäki, I., Tuupainen, S., Aaltonen, L.A., Hemminki, K., Lindblom, A., Försti, A., Sieber, O., Lipton, L., van Wezel, T., Morreau, H., Wijnen, J.T., Devilee, P., Matsuda, K., Nakamura, Y., Castellví-Bel, S., Ruiz-Ponte, C., Castells, A., Carracedo, A., Ho, J.W., Sham, P., Hofstra, R.M., Vodicka, P., Brenner, H., Hampe, J., Schafmayer, C., Tepel, J., Schreiber, S., Völzke, H., Lerch, M.M., Schmidt, C.A., Buch, S., Moreno, V., Villanueva, C.M., Peterlongo, P., Radice, P., Echeverry, M.M, Velez, A., Carvajal-Carmona, L., Scott, R., Penegar, S., Broderick, P., Tenesa, A., Houlston, R.S. 2010. COGENT (COlorectal cancer GENEtics): an international consortium to study the role of polymorphic variation on the risk of colorectal cancer. *Br J Cancer*, 102(2): 447-54.
- Van der Tweel, I., van Noord, P. A. 2003. Early stopping in clinical trials and epidemiologic studies for "futility": conditional power versus sequential analysis. *J Clin Epidemiol* 56(7): 610-7.
- Varghese, J. S., Easton, D. F. 2010. Genome-wide association studies in common cancers-- what have we learnt? *Curr Opin Genet Dev*, 20(3): 201-9.
- Vogelstein, B., Fearon E. R., Hamilton, S.R., Kern, S.E., Preisinger, A.C., Leppert, M., Nakamura, Y., White, R., Smits, A.M., Bos, J.L. 1988. Genetic alterations during colorectal-tumor development. *N Engl J Med*, 319(9): 525-32.

- Wald, A. 1947. *Sequential Analysis*. John Wiley and Sons, New York.
- Wallace, C., Newhouse, S. J., Braund, P., Zhang, F., Tobin, M., Falchi, M., Ahmadi, K., Dobson, R.J., Marçano, A.C., Hajat, C., Burton, P., Deloukas, P., Brown, M., Connell, J.M., Dominiczak, A., Lathrop, G.M., Webster, J., Farrall, M., Spector, T., Samani, N.J., Caulfield, M.J., Munroe, P.B. 2008. Genome-wide association study identifies genes for biomarkers of cardiovascular disease: serum urate and dyslipidemia. *Am J Hum Genet*, 82(1): 139-49.
- Webb, E.L., Houlston, R. S. 2007. Association studies using familial cases: an efficient strategy for identifying low-penetrance disease alleles. *Methods Mol Biol*, 376: 151-9.
- Weinberg, C.R. 1999. Allowing for missing parents in genetic studies of case-parent triads. *Am J Hum Genet*, 64(4): 1186-93.
- Worthley, D.L., Whitehall, V. L., Spring, K.J., Leggett, B.A. 2007. Colorectal carcinogenesis: road maps to cancer. *World J Gastroenterol*, 13(28): 3784-91.
- Yang, Q., Xu, X., Laird, N. 2003. Power evaluations for family-based tests of association with incomplete parental genotypes. *Genetics*, 164(1): 399-406.
- Yende, S., Kammerer, C.M., Angus, D.C. 2006. Genetics and proteomics: deciphering gene association studies in critical illness. *Crit Care*, 10(4):227.
- Zhang, J., Powell, S.N. 2005. The role of the BRCA1 tumor suppressor in DNA double-strand break repair. *Mol Cancer Res*, 3(10): 531-9.
- Zhao, H., Zhang, S., Merikangas, .KR., Trixler, M., Wildenauer, D.B., Sun, F., Kidd, K.K. 2000. Transmission/disequilibrium tests using multiple tightly linked markers. *Am J Hum Genet*, 67(4): 936-46.

ÖZGEÇMİŞ

Adı Soyadı: DİLAY ÇIĞLIDAĞ DÜNGÜL

Doğum Yeri :Ankara

Doğum Tarihi: 09/03/1979

Medeni Hali: Evli

Yabancı Dili: İngilizce

Eğitim Durumu

Lise: Yıldırım Bayezıt Anadolu Lisesi

Lisans: Hacettepe Üniversitesi Fen Fakültesi, Biyoloji Bölümü

Yüksek Lisans: Dokuz Eylül Üniversitesi, Tıp Fakültesi, Tıbbi Biyoloji ve Genetik

Çalıştığı Kurum/Kurumlar ve Yıl: Ankara Üniversitesi Biyoteknoloji Enstitüsü/2006-2009
(sözleşmeli araştırmacı)

Yayımları (SCI ve diğer):

Dilay Ciglidag Dungul, Hilal Ozdag, Nejat Akar, The 3'UTR (+861 A→G) Polymorphism Detected In Hemoglobin Alpha 2 Gene In Turkish Population, The Egyptian Journal of Medical Human Genetics, 12:1, May 2011,

Ozlem Ilk, Farid Rajabli, **Dilay Ciglidag Dungul**, Hilal Ozdag and Hakki Gokhan Ilk, A novel approach for small sample size family-based association studies: sequential tests, European Journal of Human Genetics 23 March 2011,doi:10.1038/ejhg.2011.51

Yurt İçi Bildirileri:

Ç.Eresen, O.Altungöz, F.Paralı, E.Ay, Ş.Çakmakoglu, O.Çetin, Y.Kılıç, **D.Çığlıdağ**, G.Bulut, Ö.Altıok, M.Sakızlı.: DEÜTF Tıbbi Biyoloji ve Genetik ABD Genetik ve DNA Diagnostik Laboratuvarında, 1999-2004 Yılları Arasında Saptanan Dengeli Translokasyon Olguları,VI.Ulusal Prenatal Tanı ve Tıbbi Genetik Kongresi özet kitabı , S:140, KLP02, Nisan 2004, Antalya

Evin Özen, **Dilay ÇIĞLIDAĞ**, Esra Erdal, Hilal Özdağ, Neşe Atabey .: Hepatosit Büyüme Faktörü ve Heparan Sulfat Proteoglikanların Hepatosellüler Karsinoma Hücrelerinde Motilite ve İnvazyon Üzerine Etkisi, 2008, Çanakkale

Dilay Çığlıdağ Düngül, Hilal Özdağ, Nejat Akar.: Türk Populasyonunda Hemoglobin Alfa 2 Geninde Saptanan 3'UTR (+861 A→G) Polimorfizmi, 2009, Antalya

Dilay Çığlıdağ Düngül, Nevin Belder, Mehmet Korkmaz, Arzu Ensari, Berna Savaş, Ayhan Kuzu, Ethem Geçim, Özlem İlk, H.Gökhan İlk, Hilal Özdağ, Sporadik Kolorektal Kanser Vakalarında Genom Ebadında TNP Profilinin Belirlenmesi ile Yeni Genetik Yatkınlık Genlerinin ve Kanser Gelişmesinde Etken Olan Genlerin Belirlenmesi, 2009, Antalya

Yurt Dışı Bildirileri

Dilay Cıglıdag, Murat Cokakli, Nese Atabey.: Cell-Extracellular Matrix Interactions Modulate HGF Induced Adhesion, Proliferation, Motility and Invasion of HCC Cells, 31. FEBS Kongresi "Molecules in Health & Disease" kongre özet kitabı, S:104, PP104, Haziran 2006, İstanbul

Dilay Çığlıdağ Düngül, Farid Rajabli, Gül İnan, Nevin Belder, Ethem Geçim, Ayhan Kuzu, Berna Savaş, Arzu Ensari, H.Gökhan İlk, Özlem İlk, Hilal Özdağ, A Family Based Genome Wide Association Study in Turkish Sporadic Colorectal Cancer Cases, Aralık 2010, İngiltere

