

ANKARA ÜNİVERSİTESİ
BİYOTEKNOLOJİ ENSTİTÜSÜ

BİYOİNFORMATİK
YÜKSEK LİSANS TEZİ

**KOLOREKTAL KANSERDE TRANSKRİPTOMİK META ANALİZ VE YENİ
BİYOBELİRTEÇLERİN KEŞFİ**

Hakiimu KAWALYA

Danışman Öğretim Üyesi

Prof. Dr. Hilal ÖZDAĞ

ANKARA

Ekim

2016

ETİK BEYAN

Bu tez çalışmasının; akademik kural ve etik ilkelere baęlı kalınarak hazırlandığını, çalışmada yararlanılan ve bu çalışma ürünü olmayan bütün bilgiler için kaynak yayınlara atıfta bulunulmuş olduğunu beyan ederim.

Hakiimu KAWALYA

İmzası

ONAY

Prof. Dr. Hilal Özdağ danışmanlığında Hakiimu Kawalya tarafından hazırlanan bu çalışma 25/10/2016 tarihinde aşağıdaki jüri tarafından Biyoteknoloji Anabilim Dalı'nda yüksek lisans tezi olarak kabul edilmiştir.

Başkan: Prof. Dr. Can AKÇALI

İmza:

Üye: Prof. Dr. Hilal ÖZDAĞ

İmza:

Üye: Yrd. Doç. Dr. Bala GÜR DEDEOĞLU

İmza:

Yukarıdaki sonucu onaylarım.

Prof. Dr. Aykut ÖZKUL

Enstitü Müdürü

ÖZET

Yüksek Lisans Tezi

KOLOREKTAL KANSERDE TRANSKRİPTOMİK META ANALİZ VE YENİ BİYOBELİRTEÇLERİN KEŞFİ

Hakiimu KAWALYA

Ankara Üniversitesi Biyoteknoloji Enstitüsü

Danışman: Prof. Dr. Hilal Özdağ

Kolorektal kanser (KRK) günümüzde dünyada en yaygın görülen kanser tiplerinden biridir ve gelişmiş ülkelerde mortalite sıralamalarında 3. sırada yer almaktadır. Erken teşhis, KRK'nın tedavi ve önlenmesinde önemlidir. Bugüne kadar bu tür malignitelerin taramasında kullanılan en makul tanı aracı biyobelirteçlerin kullanımı olmuştur. Ancak mevcut biyobelirteçlerin yetersizliği nedeniyle yeni biyobelirteçlerin belirlenmesine veya geliştirilmesine ihtiyaç duyulmaktadır. Muhtemel kolorektal kanser biyobelirteçlerinin tanımlanması için seçilebilecek olan yollardan birinde, ilk olarak gen ekspresyon profillerinin belirlenerek, kanserli ve normal dokunun ifade profillerini karşılaştırılması gerekmektedir. Bu amaçla yaptığımız çalışmada keşif seti olarak tanımladığımız araştırma ekibimizin daha önce yürütmüş olduğu projeden elde edilen gen ekspresyon profillerini doğrulamak ve kesin sonuçlara ulaşmak amacıyla, keşif seti örnekleri gen ekspresyon veri tabanlarından (GEO_Gene Expression Omnibus) alınan ve validasyon seti olarak adlandırılan örneklerle karşılaştırılmıştır.

Kullanılan keşif setinde, 50 normal örnek farklı sınıflaması olan ve eşli olduğu 50 tümör örneğiyle karşılaştırılmıştır. Doğrulama seti 1260 örneklerden oluşmaktadır ve bu örnekler GEO veri tabanından belirlenmiştir. Her bir set RMA ön işleme algoritması ile normalize edilmiştir. Bu meta-analiz metodolojisinde, genlerin filtre edilmesinde $p \leq 0.001$ anlamlılık düzeyinde ve 2 kat değişim eşliğinde ortak olarak farklı ifade edilip, setler arasında örtüşen genlerinin, gen anotasyon gruplarında $\geq 1,3$ zenginleşme skoruna sahip olma kriterleri kullanılmıştır.

ANOVA analizi sonucunda tümör-normal kontrol karşılaştırmasında keşif ile validasyon setlerinin gen listeleri arasında %59'luk uyum tespit edilmiştir. DAVID biyoinformatik aracı ile gerçekleştirilen ontoloji ve yolak analizleri, bu genlerin bulunduğu önemli zenginleştirilmiş kümelerin arasında kanserdeki yolaklar, fokal adezyon ve aldesteron bağımlı sodyum geri emilimi gibi yolaklar olduğunu göstermiştir. 28 genin KRK hastalarında biyobelirteç potansiyeline sahip olduğu görülmüştür. Tümör derecelerine (grade) göre yapılan karşılaştırmalarda gen liste örtüşmeleri, grade 2 normal kontrol karşılaştırmasında %66 ve grade 3 normal kontrol karşılaştırmasında ise keşif seti ile %68 örtüştüğü görülmüştür. Grade 2 hastalarda hücre döngüsü ve p53 sinyal yolağı gibi pre-invasif genlere bağlantılı yolaklar belirlenirken, grade 3 hastalarında PPARD, CLCA1, CLCA4 ve NR3C1'i içeren özel gen imzası ortaya çıkmıştır. Ayrıca, keşif ve validasyon setlerinin arasında çakışan evre (stage) genlere bakıldığında, normal örnekler ile evre 2 karşılaştırması sonucunda hastalara ait genin % 61'i valide edilirken evre 3 hastalarına ait genin % 68'i doğrulanabilmiştir. Hippo sinyal yolağın yanısıra, MUC2, MUC5B ve ATP2A3 genlerin evre 3 hastalarına özgü olduğu belirlenmiştir.

Bu çalışma sonucunda KRK tanısı, evrelendirme ve sınıflanmasında kullanılacak biyobelirteç adayı olabilecek genler belirlenmiştir.

ABSTRACT

MSc Thesis

TRANSCRIPTOMIC META-ANALYSIS AND DISCOVERY OF NEW BIOMARKERS IN COLORECTAL CANCER

Hakiimu KAWALYA

Ankara University Biotechnology Institute

Supervisor: Prof. Hilal Özdağ

Colorectal cancer (CRC) is one of the most prevalent tumors worldwide with the third highest mortality rate in developed countries. Early diagnosis is crucial to the treatment and prevention of colorectal cancer and currently, the most plausible diagnosis tool for CRC screening is the use of biomarkers. Identification of probable colorectal cancer biomarkers often commences with comparing the expression profiles of cancerous and normal tissue by performing gene expression profiling. This study compared our samples against samples from the gene expression database so as to validate our findings and reach more robust conclusions.

Design: In our discovery set 50 normal samples were matched with 50 tumor samples of different grading and staging. A validation set of 1260 samples was identified from GEO database and each set was normalized using RMA tool embedded in the Partek software. A meta-analysis methodology that derived overlaps from differentially expressed genes ($p \leq 0.001$, fold change ≥ 2) of discovery and validation sets and used an enrichment score threshold (e.s ≥ 1.3) for biological process and molecular functional annotation clustering of overlap gene lists was developed.

Results indicated a 59% validation of the ANOVA tumor vs normal gene-list of the discovery set. Analysis using DAVID bioinformatics tool related significant enriched clusters to pathways such as P13K-Akt pathway, pathways in cancer, focal adhesion, ECM receptor interaction, metabolic pathways and aldosterone regulated sodium reabsorption. Consequently a 28-gene signature was highlighted as possessing diagnostic potential in Turkish CRC patients. Meanwhile, grade gene-list overlaps validated 66% of grade 2 vs

normal and 68% of grade 3 vs normal of our cohort. Cell cycle and p53 signaling pathways were found to be linked to only grade 2 samples while genes such as PPARD, CLCA1,CLCA4 and NR3C1 were unique to the grade 3 gen signature. Meanwhile, stage gene-list overlaps validated 61% of stage 2 and 52% of stage 3 samples of our patients in comparison to normal samples. MUC2, MUC5B and ATP2A3 genes were specific to stage 3 patients. Hippo pathway was also linked to only the stage 3 gen signature.

Therefore this study may be useful in developing kits for CRC diagnosis and providing more insight into scanty studied genes believed to have a role in CRC grading and staging.



2016, 106 pages

Keywords: Meta-analysis, Microarray, Biomarker, Colorectal cancer.

TEŞEKKÜR

Tez çalışmam ve yüksek lisans eğitimim boyunca benimle hem akademik bilgi hem de kendisinin hayat tecrübesini paylaşan değerli hocam Prof. Dr. Hilal Özdağ'a,

Kendisinin verisi benimle paylaşmakla kalmayıp süreç içinde hep yanımda olan ve desteklerini esirgemeyen Dr. Nevin Belder'e,

Ankara Üniversitesi'nin Biyoteknoloji Enstitüsü'ndeki eğitim hayatımın sürecinde birlikte çalışmaktan hep mutluluk duyduğum, bana her zaman yardımcı olan, Blm. Uzm. Ali Emre Akpınar, Dr. Reza Dastouri, Dr Aynur Karadağ, Blm. Uzm. Seda Taşır Yılmaz, Dr. Günseli Deniz, Blm. Uzm. Özge Cumaoğulları ve Dr. Semih Dalkılıç'a

Yüksek lisans boyunca beni finansal olarak destekleyen Türkiye Cumhuriyeti Yurtdışı Türkler ve Akraba Topluluklar Başkanlığı'na ve

En önemlisi de dünyanın en güzel hazineleri olan canım ailem, babam ve anneme

Sonsuz teşekkürlerimi sunarım

HAKIIMU KAWALYA

Ekim, 2016

İÇİNDEKİLER

ETİK BEYAN	i
ÖZET	iii
ABSTRACT	v
TEŞEKKÜR	vii
1. GİRİŞ	1
2. KURAMSAL TEMELLER	4
2.1. GENOM VE GENOMİK	4
2.2. TRANSKRİPTOM VE TRANSKRİPTOMİK	4
2.3. MİKRODİZİN VERİ ANALİZİ VE GEN İFADE ÇALIŞMALARINDAKİ UYGULANMASI	5
2.3.1. MİKRODİZİN VERİTABANLARI	5
2.3.2. PROBLAR VE MİKRODİZİN PLATFORMU (AFFYMETRIX)	7
2.3.3. ÖN İŞLEME	9
2.3.3.1. Görüntü Kalitesi	9
2.3.3.2. Saçılım/Dağılım Ve MA Grafikleri	10
2.3.3.3. Normalizasyon	11
2.3.3.4. Özetleme	12
2.3.3. TEMEL BİLEŞENLER ANALİZİ (PRINCIPLE COMPONENT ANALYSIS)	12
2.3.4. FARKLI OLARAK EKSPRESE EDİLEN GENLER	13
2.3.5. KÜMELEME (CLUSTERING)	13
2.3.6. GEN ONTOLOJİ, ZENGİNLEŞME VE YOLAK ANALİZLERİ	14
2.4. META ANALİZ	15
2.4.1. MİKRODİZİN ÇALIŞMALARINDA META-ANALİZ	15
2.4.2. KOLOREKTAL KANSERDE META-ANALİZ ÖRNEK ÇALIŞMASI	17
2.5. KANSER	19
2.5.1. KOLOREKTAL KANSER	20
2.5.2. KOLOREKTAL KANSERİN ETYOLOJİSİ	21
2.5.3. KOLOREKTAL KANSERİN RİSK FAKTÖRLERİ	22
2.5.4. KOLOREKTAL KANSERDE MOLEKÜLER GENETİK DEĞİŞİKLİKLER	22
2.5.4.1. Tümör Supresör Genlerdeki Değişimler	23
2.5.4.2. Protoonkogenlerin Aktivasyonu	24
2.5.4.3. DNA Onarım Genlerindeki Değişiklikler (Mismatch Repair)	24
2.5.5. KOLOREKTAL KANSER SINIFLANDIRMA	24
2.5.5.1. Kolorektal Karsinomlarda Histolojik Dereceleme (Grade)	24
2.5.5.2. Kolorektal Karsinomlarda Evreleme (Stage)	25

2.5.6. KOLOREKTAL KANSERİN GEN İFADE PROFİLİ	27
3. GEREKÇE VE AMAC	29
4. MATERYEL VE METOD	33
4.1. SEÇİLEN ÇALIŞMALARINDAN VERİ TOPLANMA.	33
4.2. VERİNİN ALT GRUPLARA KOYUNMASI	34
4.3. HAVUZLANMIŞ VERİLERİNİN (POOLED DATA) ÖN İŞLEMESİ	37
4.4. GEN LİSTESİNİN ÇIKARTILMASI	37
4.5. ÖRTÜŞME SAPTANMASI	38
4.6. ÇAKIŞAN GEN LİSTELERİN FONKSİYONEL YOLAK VE ZENGİNLEŞME ANALİZİ	39
5. BULGULAR	40
5.1. TÜMÖR VE KONTROL KARŞILAŞTIRILMALARININ BULGULARI	40
5.1.1. KEŞİF SETİNİN BULGULARI	40
5.1.2. VALİDASYON SETLERİN ARASINDAKİ KIYAS BULGULARI	42
5.1.3. KEŞİF SETİ İLE VALİDASYON SETLERİN KIYAS BULGULARI	47
5.2. GRADE BULGULARI	52
5.2.1. KEŞİF SETİN GRADE BULGULARI	52
5.2.2. VALİDASYON SETİNİN GRADE BULGULARI	53
5.2.3. KEŞİF VE VALİDASYON SETLERİNDEKİ GRADELER ARASINDAKİ ÇAKIŞAN GENLER	56
5.3. EVRE BULGULARI	61
5.3.1. KEŞİF SETİNİN EVRE BULGULARI	61
5.3.2. VALİDASYON SETİN EVRE BULGULARI	62
5.3.3. KEŞİF VE VALİDASYON SETLERİNDEKİ EVRELER ARASINDAKİ ÇAKIŞAN GENLER	65
6. TARTIŞMA	70
6.1. TÜMÖR–KONTROL ÇAKIŞMALARI	71
6.2. GRADELER- KONTROL ÇAKIŞMALARI	73
6.3. EVRELER- KONTROL ÇAKIŞMALARI	74
7. KAYNAKLAR	78
8. EKLER	86

ŞEKİLLER DİZİNİ

Şekil 2.1: Affymetrix platformu (22)	8
Şekil 2.2: Meta-analizin şematik gösterimi (38).	17
Şekil 2.3: Kalın bağırsağın anatomisi.....	21
Şekil 2.4: Tumor evrelendirmesindeki Astra-Coller sistemi(50).	27
Şekil 3.1: Keşif setin normalleri ile tümör, grade ve evre karşılaştırılmasında elde edilen bulguları.....	30
Şekil 4.1: Validasyon setinin çalışmalarından alınan örneklerin bölgelere ve etkin yapısına göre demografik dağılımları	33
Şekil 4.2: Mikrodizin veri analiz süreci.....	38
Şekil 4.3: Meta-analiz sonucunda önemli genlerin belirtme şeması	39
Şekil 5.1: Keşif setindeki tümör ve normalin aralarında çıkan PCA ve heatmap sonuçları	40
Şekil 5.2: Keşif setinin tümör ve normallerin karşılaştırmasından elde edilen ifadesi artan ve azalan genlere ait yolak analizinin bulguları	41
Şekil 5.3: Tüm validasyon setinin örneklerin Spearman korelasyon çizimi	43
Şekil 5.4: Validasyon setlerin PCA bulguları.....	44
Şekil 5.5: Validasyon setlerin tümör ve normallerin karşılaştırmasından elde edilen heatmap bulguları.	44
Şekil 5.6: Validasyon setlerin tümör ve normallerin karşılaştırmasından elde edilen ifadesi azalan genlere ait yolak analizinin bulguları	45
Şekil 5.7. Validasyon setlerin tümör ve normallerin karşılaştırmasından elde edilen ifadesi artan genlere ait yolak analizinin bulguları.	46

Şekil 5.8: Keşif ve Validasyon setlerinin tümör- normal örtüşmesi	47
Şekil 5.9: Keşif set ile tüm, eşleşmiş ve eşleşmemiş validasyon setlerinin karşılaştırılmasında çakışan ve ifadesi azalan gen listesinin ontoloji sonuçları.....	49
Şekil 5.10: Keşif set ile tüm, eşleşmiş ve eşleşmemiş validasyon setlerinin karşılaştırılmasında çakışan ve ifadesi azalan gen listesinin yolak analizinin sonuçları.....	49
Şekil 5.11: Keşif set ile eşleşmiş ve eşleşmemiş validasyon setlerinin karşılaştırılmasında çakışan ve ifadesi artan gen listesinin yolak analizinin sonuçları.	50
Şekil 5.12: Keşif set ile tüm, eşleşmiş ve eşleşmemiş validasyon setlerinin karşılaştırılmasında çakışan ve ifadesi artan gen listesinin yolak analizinin sonuçları.....	50
Şekil 5.13: Keşif setinin grad 2 vs normallerinin PCA ve heatmap sonuçları.	52
Şekil 5.14: Keşif setinin grad 2 vs normallerinin PCA ve heatmap sonuçları.	53
Şekil 5.15: Validasyon setinin grade ve normal örneklerin Spearman korelasyon çizimleri	54
Şekil 5.16: Validasyon setinin grade 2 vs normallerinin PCA ve heatmap sonuçları.....	55
Şekil 5.17: Validasyon setinin grade 3 vs normallerinin PCA ve heatmap sonuçları.	55
Şekil 5.18: Keşif ve Validasyon setinin grade örtüşmesi	56
Şekil 5.19: Grade2vK ve grade3vK için keşif ve validasyon setlerin arasındaki çakışan ve ifadesi azalan gen listesinin ontoloji sonuçları.....	57
Şekil 5.20: Grade2vK ve grade3vK için keşif ve validasyon setlerin arasındaki çakışan ve ifadesi azalan gen listesinin yolak analizinin sonuçları.....	57
Şekil 5.21: Grade2vK için keşif ve validasyon setlerin arasındaki çakışan ve ifadesi artan gen listesinin ontoloji sonuçları.....	58
Şekil 5.22: Grade2vK ve grade3vK için keşif ve validasyon setlerin arasındaki çakışan ve ifadesi artan gen listesinin yolak analizinin sonuçları	58
Şekil 5.23: p53 sinyal yolağı üzerindeki FİG'ler	60

Şekil 5.24: Keşif setinin evre 2 vs normallerinin PCA ve heatmap sonuçları.	61
Şekil 5.25: Keşif setinin evre 3 vs normallerinin PCA ve heatmap sonuçları.	62
Şekil 5.26: Validasyon setinin evre ve normal örneklerin Spearman korelasyon çizimleri	63
Şekil 5.27: Validasyon setinin evre 2 vs normallerinin PCA ve heatmap sonuçları.....	64
Şekil 5.28: Validasyon setinin evre 3 vs normallerinin PCA ve heatmap sonuçları.	64
Şekil 5.29: Keşif ve Validasyon setlerinin Evre örtüşmesi	65
Şekil 5.30: Evre2vK ve evre3vK için keşif ve validasyon setlerin arasındaki çakışan ve ifadesi azalan gen listesinin ontoloji sonuçları.	66
Şekil 5.31: Evre2vK ve evre3vK için keşif ve validasyon setlerin arasındaki çakışan ve ifadesi azalan gen listesinin yolak analiz sonuçları.....	66
Şekil 5.32: Evre2vK ve evre3vK için keşif ve validasyon setlerin arasındaki çakışan ve ifadesi azalan gen listesinin ontoloji sonuçları.....	67
Şekil 5.33: Evre2vK ve evre3vK için keşif ve validasyon setlerin arasındaki çakışan ve ifadesi artan gen listesinin yolak analiz sonuçları	67
Şekil 5.34: Hippo sinyal yolağının üzerindeki FiG'leri	69

ÇİZELGELER DİZİNİ

Çizelge 2.1: Chan ve ark çalışmada kullandığı çalışmaları ve platıformları.....	19
Çizelge 2.2: TNM ve Duker sistemleri	26
Çizelge 4.1: Eşleşmiş örneklerini içeren çalışmalar	34
Çizelge 4.2 Eşleşmiş olmayan örneklerini içeren çalışmalar	34
Çizelge 4.3: Derecelendirilmiş ve Evrelendirilmiş örneklerini içeren çalışmalar.....	35
Çizelge 4.4: Seçilen Çalışmaların Toplamı	36
Çizelge 5.1: Tümör ve normallerin kıyaslamasının özeti.....	51
Çizelge 5.2: Grade kıyaslamasının özeti	59
Çizelge 5.3: Evre kıyaslamasının özeti.....	68

SİMGELER DİZİNİ

AMSTRA	A measurement tool for the 'assessment of multiple systematic reviews
APC	Adenomatöz polipozis koli
CASP	Critical Appraisal Skills Programme
cDNA	Tamamlayıcı deoksiribonukleik asit
Cy3	Cyanine dye 3 (Siyanin boyası 3)
Cy5	Cyanine dye 5 (Siyanin boyası 5)
DAVID	Database for Annotation, Visualization and Integrated Discovery
DNA	Deoksiribonükleik asit
DSÖ	Dünya Sağlık Örgütü
FAK	Fokal adezyon
FAP	Ailesel adenomatöz polip
GO	Gen ontoloji
GWAS	Genom boyu assosiyasyon çalışmaları
KRK	Kolorektal kanser
MOOSE	Meta-analysis of Observational Studies in Epidemiology
MM	Mismatch
PM	Perfect match
PZR	Polimeraz zincir reaksiyonu
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
TBA	Temel Bileşenler Analizi

1. GİRİŞ

Dünya Sağlık Örgütü'nün (DSÖ) değerlendirmesine göre, 2012 yılında kanser nedeniyle ölümlerin sayısı 8,2 milyona ulaşmıştır. En yaygın görülen kanser tiplerinden biri olan kolorektal kanser (KRK), yılda 694.000 kişinin ölümüne sebep olmasıyla birlikte ölümlerle sonuçlanan kanserler sıralamasında 4. sırada yer almaktadır. ABD'de ise kolorektal kanser, üçüncü en yaygın kanserdir ve 2014 yılındaki yeni kolorektal kanser hasta sayısı 136.830 kişi olarak tespit edilmiştir ve bu hastaların 50.310'unun ölümlerle sonuçlandığı belirtilmiştir (1). Türkiye'de ise, 2009 yılının verilerine göre, ülkede en yaygın kanserlerden biri olan kolorektal kanserin, erkeklerde %21,0 oranla dördüncü, kadınlarda ise % 13,4 oranla üçüncü en yaygın kanser tipi olduğu belirtilmiştir.

Moleküler biyoloji kapsamında en çok üzerinde çalışılan tümörlerinden biri kolorektal kanserdir. Vogelstein modeli, kolorektal karsinogenezinin aşamalarını anlatan ilk model olduğu bilinmektedir (2). Bu modelin ardından karsinogenik mekanizmaları açıklayan çeşitli literatürler yayınlanmıştır. Kanser ile ilgili mevcut bilimsel bilgiye göre kanser tedavi sürecinde, kanserin erken tanısı çok önemlidir. Ayrıca kanser başlangıcında belirli ve özgün semptomların olmaması sebebiyle geleneksel yöntemleri kullanarak hastalığın erken tanısı oldukça zordur. Kolonoskopi başta olmak üzere, diğer tanı yöntemleri invazif ve zahmetli olduğundan, genellikle hastalar tarafından kabul edilmesi zordur. Bu nedenle günümüzde bu sahadaki araştırmaların önceliği tümör taraması için en uygun ve etkin tanı aracı biyobelirteçlerin tanımlanması ve kullanılmasıdır. 2007 yılında Srivastava ve arkadaşları tarafından yapılan bir çalışma (3) ve benzeri birçok diğer makalede biyobelirteçlerin, kolorektal kanserin taramasında en etkin araç olduğunu göstermektedir. Bununla yanısıra bazı araştırmacılar, adjuvan tedaviden yararlanacak uygun hastalarının belirlenmesi için prognostik biyobelirteçlerin geliştirilmesi planlanlama çalışmaları yaparken (4), bazıları ise hastalık riski taşıyan hastaları izole etmeye çalışmaktadırlar (5).

Türkiye'de şimdilik kolorektal taramaların kapsamı %20-30 arasındadır ve teşhis edilen olguların yarısından çoğu ancak ileri evrede yakalanabilmektedir. Bu veri, kanserin erken tanısında sorun yaşandığını göstermektedir. Bunun yanında var olan karsinoembriyojenik Antijen (CEA), karbonhidrat/kanser antijeni 19-9 (CA19-9) vb. belirteçlerin düşük özgünlük ve duyarlılığa sahip olması nedeniyle, daha etkin kanser teşhis ve tedavi

yöntemlerinin geliştirilmesi için yeni biyobelirteçlerin keşfedilmesine ihtiyacı duyulmaktadır.

Yeni biyobelirteçleri tespit etmek amacıyla kullanılan yaklaşımlardan biri normal (kontrol) ve patolojik hücrelerde, ifade düzeyi değişen gen profillerinin mikrodizin analizleri ile belirlenmesidir. DNA mikrodizin teknolojisinin temelleri 1990'lı yıllara dayanmaktadır. Eskiden “bir gen bir deney” kavramı varken mikrodizin teknolojisi ile birlikte, birden çok genin transkripsiyon seviyelerini aynı anda belirleme fırsatı bulunmuştur (6).

Petersen ve arkadaşları (2005) (7), farklı platformlarda yapılmış olan mikrodizin deneyleri arasında gen ifade düzeylerinde uyumsuzluk olmasının sık karşılaşılan bir durum olduğunu belirtmişlerdir. Bu farklılıklar, sadece platformların varyasyonu ile sınırlı değildir; farklı prob setleri veya mikrodizin deneylerinde kullanılan karakteristik olarak farklı örneklerle bağlantılıdır. Bu yüzden aynı istatistiksel analiz yöntemi uygulanmasına rağmen, tutarsız sonuçların çıkma olasılığı yüksek olmaktadır. Bu teknolojinin devamında araştırmacılar tarafından meta-analiz olarak adlandırılan bir analitik yöntem geliştirmişlerdir. Bu metodoloji, mikrodizin verilerinde yer alan önemli parametreleri, karakteristik olarak farklı olsalar bile, ortak bir formulasyon sağlayarak, tüm verilerin entegre edildiği bir süreçtir.

Meta-analiz, daha sağlam ve tutarlı istatistiksel sonuçları üretmek için geliştirilen bir araçtır. Aslında çoklu veri setlerini birleştirerek, her gen veya her prob için bir genel skor (meta skoru) çıkartan bir metottur (8). Meta-analiz, daha anlamlı yolaklar belirlemek için, direk olarak herbir veri setine ait olan işlevsel kategorilerine veya bütün ekspresyon veri setine uygulanabilir (9). Genom boyu ilişkilendirme çalışmalarında, meta-analiz teknikleri yeni gen setleri keşfetmek için kullanılmaktadır. Ayrıca önemli gen fonksiyonlarının keşfedilmesinde ve öngörülmesinde bu tekniklerin yardımcı olabileceği düşünülmektedir (10).

2013 yılında, Ankara Üniversitesi Biyoteknoloji Enstitüsü'nde kolorektal kanserde gen ifadesi profillemesi üzerine bir çalışma yapılmıştır. Bu çalışmaya çoğunluğunun evre 2 ve evre 3 olduğu 50 adet tümör örneğini ile bu tümör örneklerinin cerrahi sınırda yer alan normal epitelinden alınmış 50 adet eşleştirilmiş normal (kontrol) örneği dâhil edilmiştir. Bu çalışmadan elde edilen verileri kullanarak tezimizde genom veritabanında yer alan benzer çalışmaların incelenmesiyle, bir doğrulama (validasyon) seti oluşturulmuştur. Meta-

analiz yöntemini kullanarak, yukarıda bahsedilmiş çalışmadan elde edilen sonuçları (keşif seti) ve validasyon setindeki sonuçlarını karşılaştırıp, gen listelerinden yeni genlerin bulunması için ayrıntılı inceleme yapılmıştır. Kısaca, bu tezin amacı söz konusu olan çalışmadan çıkarılmış gen ekspresyon bulguları, literatür arşivlerinde yer alan benzer çalışmalarla karşılaştırarak doğrulama yapılmasıdır.

Tezimizden elde edilecek anlamlı seviyede ifade edilen ve doğrulanan genler, yeni model oluşturulması ve özellikle Türkiye’de erken kolorektal kanser tarama kitlerinin geliştirilmesinde kullanılabilir aday biyobelirteçler olabilecektir.



2. KURAMSAL TEMELLER

2.1. GENOM VE GENOMİK

Genom bir organizma tarafından taşınan tüm kalıtsal bilgiyi içermektedir (11) ve organizmanın çekirdeğinde bulunan DNA'daki bilginin yanı sıra mitokondri veya kloroplast gibi bağımsız genetik materyali de kapsamaktadır. Organizmalar, genom yapısına göre ökaryotik ve prokaryotik olarak iki gruba ayrılmaktadırlar. Bunun yanında, germ hattı genomu, ebeveynlerden gelen genetik materyalin tümü anlamına gelir. Somatik genom ise hayatımız boyunca değişen soylarımıza aktaramadığımız genetik materyalin tümü olarak adlandırılır.

Genomik kavram ise genetik materyal analizi ile ilgili tüm çalışmaları içermektedir. Bu çerçevede yapısal ve işlevsel fonksiyonları kodlayan tüm genlerin tanımlanmasının yanı sıra, aralarında ve çevre ile etkileşimleridir. Genomik çalışmalar üç gruba ayrılmaktadır. Yapısal genomik, genetik dizinin oluşturulması, gen keşfi ve gen haritalaması gibi konuları içeren çalışmalardır. İşlevsel genomik, genlerin biyolojik işlevleri, düzenlenmeleri ve ürünleri içerirken, karşılaştırmalı genomik ise işlevsel ve evrimsel akrabalığı belirlemek için farklı gen dizilerini karşılaştırılmasıdır.

Gen terimi ilk kez 1909 yılında botanikçi Wilhem Johannsen tarafından kullanılmıştır ve 1972 yılında bir genin dizisi ilk kez Walter Fiers ve ekibi tarafından belirlenirken (11), Frederick Sanger ve arkadaşı, 1977 yılında bakteriyofaj üzerinde ilk tüm genom dizisinin çıkarılmasını başarmışlardır (12). Bu çalışmalar ve daha sonraki benzer moleküler biyolojik buluşlar ciddi bir şekilde genomik alanının omurgasını oluşturmuştur. Son olarak 2001 yılında insan genom projesinin tamamlanması sonucunda, insan genomunun genler arasında DNA dizileri ile ayrılmış yaklaşık 25,000-30,000 gen içerdiği belirlenmiştir.

2.2. TRANSKRİPTOM VE TRANSKRİPTOMİK

Transkriptom; bir organizmanın transkriptlerinin toplamıdır ve hücrede bulunan tüm RNA moleküllerini içermektedir. Bu RNA'lar içinde mRNA, rRNA, tRNA ve kodlanmayan RNA'lar yer almaktadır. Transkriptomik ise, hücre genomundan ifade edilmiş tüm mRNA

transkriptlerinin incelenme sürecini kapsayan bir kavramdır. Aslında transkriptomik, özel bir dokuda belli bir zamanda aktif olmuş gen ürünlerini kapsamaktadır. Bu gen ürünleri çoğu zaman dokunun kendisine spesifik imza taşırlar. Bu spesifik moleküler imzalar organizmaya göre değişiklik göstermektedir ve bu nedenle hastalıkların tanısında veya belirtilmesinde kullanılmaktadır. Bu tez kapsamında en çok araştırılan kanserlerden biri olan kolorektal kanserin transkriptomik profilleri değerlendirilmiştir.

2.3. MİKRODİZİN VERİ ANALİZİ VE GEN İFADE ÇALIŞMALARINDAKİ UYGULANMASI

Northern blot ve RT-PCR, genetik teknolojilerin mRNA analizine yönelik ilk teknikleri arasındadır. Bu tekniklerde tek gen incelenirken daha sonra ortaya çıkan mikrodizin tekniği ile genom boyunca gen analizleri yapılmaya başlanmıştır.

Mikrodizin analizi, tek seferde tek gen ifadesinin analizinin yerine bütün tanımlanmış transkriptlerin aynı anda analiz edebildiği geniş kapsamlı bir tekniktir (13). Mikrodizin tekniği otomatikleştirilmiş süreçleri kullanarak DNA'yı cama bağlamasıyla southern blot tekniğinden farklıdır. Böylece daha fazla spot yoğunluklarına izin verip sonuç olarak bir defada araştırılan örneklerin sayısını artırmış olur. DNA mikrodizin deneyleri transkriptomdaki değişikliklerin tespit edilmelerinde, bir genin ekspresyonu veya ifade düzeylerini belirlemek amacıyla bu deneyler belirli bir zamanında bir deneysel dokuda transkripte edilen genlerin kopya miktarını ölçmek için immobilize DNA dizilerini kullanmaktadır. Ancak bu teknik gen ekspresyon profillemeye ile sınırlı değil, hastalıkların tanısı, SNP tespiti ve farmakogenetik çalışmalar için de uygulanabilir. Mikrodizin ayrıca karışık popülasyondaki bakteri türlerinin belirleme aracı olarak kullanılabilir. Bunun bir örneği gıdalardaki mikroorganizmaların belirlemesidir.

2.3.1. MİKRODİZİN VERİTABANLARI

Mikrodizinlerin kullanıldığı bilimsel deneylerin ham verilerinin (klinik veya patolojik özellikleri dâhil) yanı sıra elde edilen gen ifade verileri büyük veritabanlarına yüklenmekte ve buralarda muhafaza edilmektedir. Son yıllarda mikrodizin verilerinin paylaşılması ile deney sonuçlarının yorumlanmasını sağlamak için standartlaşmış bir formatın

geliştirilmesine ihtiyaç duyulmuştur. Mikrodizin deneylerinde MIAME (Minimum information about a microarray experiment) olarak adlandırılan bir standart format uygulanarak veri aktarılması önerilmiştir (14). MAGE-TAB projesinde gösterildiği gibi deneylerin MIAME-uyumlu tanımlarının üretilmesi için bilim adamlara yardımcı olacak çeşitli yazılım araçları ve veri alışveriş formatları mevcuttur.

Ayrıca MIAME'nin oluşturulmasında katkıda bulunan en kritik altı tane unsur aşağıda verilmiştir.

1. Her hibridizasyon için ham veri (örneğin; CEL veya GPR dosyaları),
2. İşlenmiş (normalize) veri,
3. Deney faktörleri ve değerleri dâhil olmak üzere esas örnek anotasyon (doz gibi temel bilgiler),
4. Örnek-veri ilişkileri de dâhil olmak üzere deneysel tasarım,
5. Dizinin yeterli anotasyon bilgileri (örneğin; gen tanımlayıcıları, genomik koordinatlar, prob oligonükleotid dizileri veya referans ticari dizi katalog numarası gibi bilgiler),
6. Gerekli laboratuvar ve veri işleme protokolleri (örneğin: normalizasyon yöntemi).

Tüm araştırmacıların kullanımına açık veritabanlarından en önemlileri Gene Expression Omnibus (15) ve ArrayExpress'tir (16). Ayrıca, genel kullanım amaçlı veritabanları haricinde spesifik organizmalara veya miRNA gibi moleküllere ait özel veritabanları da mevcuttur (17), (18).

Gene Expression Omnibus (GEO) Amerika kökenli ve dünyanın genelinde en kapsamlı biyolojik veritabanı olan NCBI'nin (National Center for Biotechnology Information) altındaki mikrodizin veritabanıdır. Şimdilik bu veritabanındaki 73.399 çalışmanın içinde toplam 1.934.448 örnek ve 16.392 çeşit platform bulunmaktadır. ArrayExpress ise Avrupa'nın biyoinformatik merkezi olan EBI'nin (The European Bioinformatics Institute) altındaki mikrodizin veritabanıdır. Şu an bu veritabanında yaklaşık 68.0916 çalışma mevcuttur ve bunların içinde toplam 2.090.592 örnek bulunmaktadır.

2.3.2. PROBLAR VE MİKRODİZİN PLATFORMU (AFFYMETRIX)

Bugünlerde çok çeşitli firmalar mikrodizin teknolojisi için kullanılan çipleri üretmektedirler. Bunların örneği Affymetrix, Illumina, Agilent ve Exiqon gibi firmalardır ayrıca bu tezin çerçevesinde Affymetrix platformu kullanılmıştır.

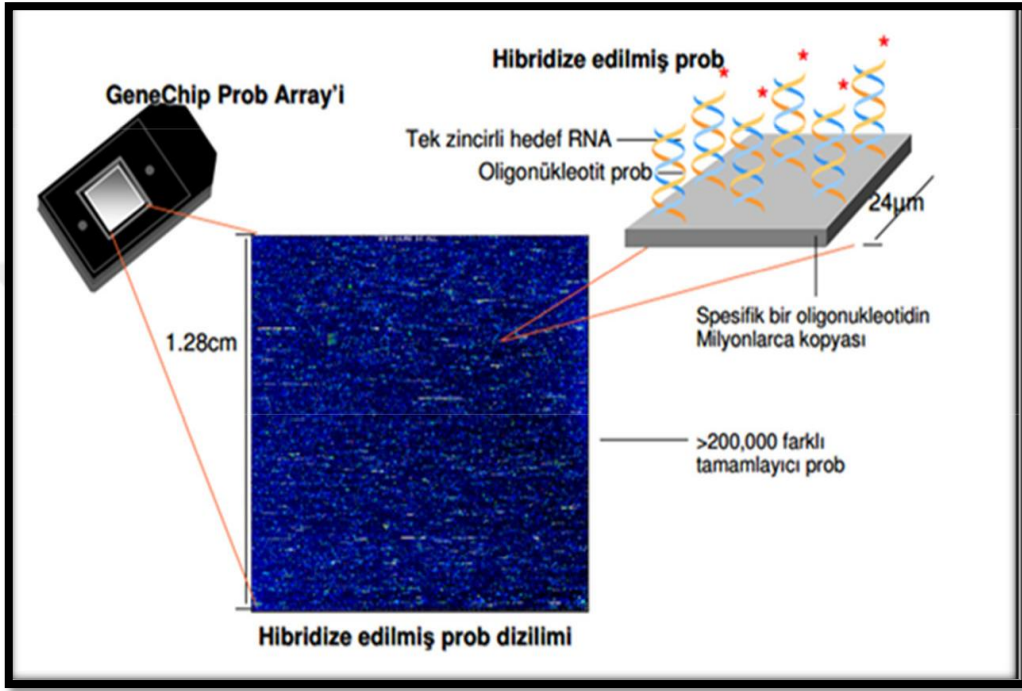
Mikrodizinler hareketsizleştirilmiş veya devinimsiz kılınmış genomik sekanslar olarak bilinen prob tiplerine göre sınıflandırılabilir. Mikrodizindeki prob sekansları analiz edilen örneğin transkriptomu içindeki ilgili dizinin varlığını ve miktarını hibridizasyon yolu ile tanımlamaktadır. Problar, örnekteki nükleik asit miktarı ile gen ifadesi hakkında yorumlar yapmak için kullanılan dijital okumalara dönüştürmektedir.

İlk nesil mikrodizinlerde problar cDNA kütüphane klonlarının PCR ile çoğaltılmasından elde edilmekteydi. Bunlar PCR amplifiye cDNA 100-3000 bazı çifti uzunluğunda olabiliyordu. Bir probun tercihi araştırılan organizmaya bağlıdır. Oligonükleotid temelli mikrodizin platformları genom dizisi bilinen ve anote edilmiş organizmalar için tasarlanabilmekte ve bu durumda oligonükleotidler muhtelif metodlarla sentezlenmektedir. Aşağıda bahsi geçecek olan Affymetrix platformu buna bir örnektir.

Duyarlılık, seçicilik/özgüllük, gürültü ve bias/yanlılık gen ifade deneyinde kullanılacak probun seçilmesinde dikkat edilmesi gereken önemli hususlardır. Seçicilik bir probun sadece kendi hedef mRNA'ları belirleme kabiliyetidir. Duyarlılık ise bir prob ve hedefi arasında ne kadar az hibridizasyon sinyali kaybedildiği ölçüsüdür. Gürültü belirli bir prob-hedef hibridizasyon sinyalinin rastgele varyasyonu olarak tanımlanmaktadır. Bias ise bir ölçümün prob-özgün etkileşiminden veya teknik sınırlılıklarına bağlı olarak gerçek hibridizasyon sinyalinden sistematik sapmasıdır (19).

Mikrodizin teknolojisinde çeşitli platformlar mevcuttur. Bu tezin kapsamında 1990'ların başında üretilen Affymetrix platformu kullanılmıştır. Affymetrix, fotolitografi ve katı faz DNA sentezi (20) tekniklerinden faydalanarak ışımaya yönelik sentez esasına dayanan teknolojiyi kullanmaktadır. Wafer olarak adlandırılan katı destek substratın bölgelerine selektif bir şekilde ultraviyole ışığı düşürülerek nükleik asitler veya diğer moleküller fotokimyasal yönlendirilmiş tepkimelerden sentezlenmektedir. Affymetrix 25-mer oligonükleotid problar sentezler ve Affymetrix dizisindeki her gen tipik olarak yaklaşık 11-12 multipl prob çiftinden ibarettir (şekil 2.1). Her çift, araştırılan transkripte tamamlayıcı olan "perfect-match" (PM) probu ve prob sekansında santral uyumsuzluğu sahip olan

“mismatch” (MM) probu olmak üzere iki oligonükleotid dizisinden oluşmaktadır. Bu mismatch probu spesifik olmayan hibridizasyonu tespit etmek/modellemek için geliştirilmiştir (21). Bir ekspresyon değeri ise belirli bir genin tüm problemleri kullanılarak hesaplanmaktadır.



Şekil 2.1: Affymetrix platformu (22)

Mikrodizin teknolojisinin sürecini mRNA örneğinin ters transkripsiyon tekniği ile cDNA'ya çevrilmesi ile başlayıp in vitro transkripsiyon ile cRNA (complementary RNA) sentezlenmesi ve biyotin işaretli nükleotidlerin inkorporasyonu ile devam etmektedir. Elde edilen aRNA'ların pürifikasyonu ve fragmentasyon reaksiyonlarının ardından GeneChip array'e hibridize edilen örnekler taranarak, farklı olarak ifade edilen genlerin tespit edilmesi ile sona ermektedir (22). Ayrıca, verideki biyolojik anlamlılığın çıkarılması ve analiz edilmesi için biyoinformatik uygulamalar bulunmaktadır.

2.3.3. ÖN İŞLEME

Herhangi bir mikrodizin deneyi sırasında tespit edilen gen ifade değişiklikleri genellikle transkripsiyondaki selektif ve biyolojik olarak anlamlı değişimleri yansıtmaktadır. Ancak bu değişiklikler deneysel faktörlerin sonucundan ortaya çıkmış olabilirler. Elbette, gen ifade profillerinde gözlenen değişiklikler gerçekten ifade farklılıkları olabileceği gibi aşağıda belirtilen işlem hataları yüzünden de kaynaklanabilir;

- Floresanla işaretlenmiş nükleotidlerin farklı etiketleme verimleri
- Floresan tarayıcının performans varyasyonları
- Biyolojik örnekler arasında RNA miktar varyasyonları
- RNA'nın saflaştırma, etiketleme ve mikrodizine hibridasyon yollarındaki varyasyonlar
- Spesifik olmayan bağlanmaları uzaklaştırmak için mikrodizin yıkama farklılıkları
- Sinyal ölçme yollarındaki farklılıklar

Veri ön işlemenin ana fikri, biyolojik olarak önemli transkripsiyon değişiklikler sonucunda ortaya çıkan gen ifade varyasyonları korunurken mümkün olduğunca sistematik önyargının kaldırılmasıdır (23). Ön işleme yönteminin seçimi daha sonrasındaki (downstream) analizleri ciddi bir şekilde etkilemektedir (24). Böylece ön işleme dikkatli bir şekilde yürütülmesi gereken önemli bir basamaktır.

Ön işleme basamakları gen ekspresyon ölçümlerindeki önyargıyı düşürerek doğruluğu ve varyasyonu azaltarak hassasiyetini iyileştirmek için tasarlanmıştır. Bu arada çalışmalar, hem RMA hem de GCRMA algoritmaları doğruluk ve hassasiyet açısından iyi performans göstermişlerdir. Ön işleme birçok önemli anahtar basamakları içermektedir. Bu basamaklar; görüntü ölçümü, arka plan çıkarımı, normalizasyon, özetleme ve kalite değerlendirmesidir.

2.3.3.1. Görüntü Kalitesi

Herhangi bir mikrodizin deneyinde görüntü ölçümü, “ıslak laboratuvar” prosedürlerinden hesaplamalı olanlara geçişini göstermektedir. Bu nedenle verinin kalitesi ve olasılığını incelenme aşamasında yeterli zaman harcanmalıdır. Genel olarak görüntü analizi negatif ve

pozitif kontroller beklenen doğrulamayı içermektedir, birbirine yakın olan verim ölçümlerini replike eden doğrulama ve artefaktlar, önyargılar ya da hataların varlığını kontrolü içermektedir.

Diğer bir deyişle, görüntü analizinin ilk adımı spot tespitidir ve daha sonra spot sınırlarının tahmini takip etmektedir. Bunu da daha sonra her spot için tek ortak bir yoğunluk elde etmek için önemli olan görüntülerin işlenmesi basamağı takip etmektedir. Kırmızı-yeşil atmosferinde bir spotun renk değişiklik yoğunluğu da saptanabilir ve son olarak spot kalite değerlendirmesi yapılmıştır. Unutulmaması gereken önemli bir husus, kırmızı ve yeşil boyaları farklı etiketleme verimliliğinden ya da görüntü kalitesindeki iki örnekten alınan total RNA'nın farklı miktarından kaynaklanan sistemik hata yalnızca normalleştirme aşamasında aşılabilmektedir.

Görüntü analizi için çeşitli yazılım paketleri mevcut ve sağlamlıkları açısından birbirinden farklılık göstermektedir. Farklı mikrodizin teknoloji tiplerinde farklı düzensizlik tipleri ortaya çıkabilmektedir ve bir platform için iyi bir segmentasyon veya ölçüm algoritması diğer bir platform için uygun olmayabilir. Yang et al (2002) tarafından yapılmış noktalı/benekli cDNA dizilerinin görüntü analiz metotlarını değerlendiren çalışma (25) dahil olmak üzere birçok çalışma görüntü analizini değerlendirmiştir.

2.3.3.2. Saçılım/Dağılım Ve MA Grafikleri

Mikrodizin verilerinin görselleştirme yöntemleri arasında en fazla kullanılan yöntemi saçılma (scatter) plotudur. İki örneğin gen ifade değerinin karşılaştırmasını göstermektedir. Çoğu veri noktaları tipik olarak 458 hattı üzerinde düşmektedir, ancak yukarı veya aşağıya doğru regüle edilmiş transkriptler çizginin dışında yerleşmiştir. Saçılma plotu, deneyde çarpıcı/dramatik ve farklı olarak regüle edilmiş transkriptleri göstermektedir. MA plotları genellikle bir verinin normalleştirilmesine ihtiyacı olup olmadığını göstermektedir.

2.3.3.3. Normalizasyon

Sıklıkla mikrodizin deneme sürecinin karmaşıklığının neden olduğu sistematik hataların düzeltilmesidir. Ölçümler örneklerin tekrarları kullanılarak elde edilir ve bu amaç busüreçte sistematik hatayı rastgele hata komponentinden ayrılmaktadır.

Cy3 ve Cy5 boyaarı farklı verimliliklerle cDNA'ya dahil edildikleri için bu boyaarla işaretlenmiş örneklerinin göreceli ifadelerinin hatasız bir şekilde değerlendirmesini olanaksız hale getirdiği için mikrodizin verilerinin normalleştirme gerekliliği şüphesizdir. Normalleştirme yine de bir gen ekspresyonunun multipl mikrodizin deneylerinde karşılaştırılma izni verdiği için esansiyeldir.

Normalizasyon bazal yoğunluk sinyalinin ölçüp ve daha sonra her genin sinyalinden çıkarılmasını içermektedir (26). Dizideki boş spotları yine bazalı tespit etmek amacıyla kullanılabilir. Normalizasyonun başka bir yaklaşımı, ekspresyon değeri sabit olduğu varsayılan ve dizideki diğer ekspresyon değerlerinin normalleştirilmesinde bazal olarak görev yapan house-keeping genleri kullanılmaktadır. Affymetrix yazılımı bu yaklaşımı uygulamaktadır ve önemli house keeping genler arasında β -aktin ve gliseraldehit-3- fosfat dehidrojenazdır (GAPDH). Her bir dizi deneyindeki her gen ekspresyon değerini bu house-keeping genlerin ortalama ekspresyon değerleri ile belirlenir

Bu tezin metodoloji bölümünde kullanılan yöntemde, sağlam (robust) çoklu dizi analizi (RMA) dilim normalizasyonu ile birleştirilmiştir. Dilim normalizasyon bir deneydeki tüm diziler için aynı genel dağılımı üreten bir yaklaşımdır (27). Eğer iki örnekten elde edilen ve işaretlenmiş RNA miktarı aynı ve teorik olarak benzeyen iki diziye uygulanmışsa, aynı genel yoğunluk dağılımlarına sahip olmaları gerektiği varsayımı ile çalışan parametrik olmayan bir metottur.

Bu arada, küresel ve yerel normalleşme, mutlak sinyal yoğunluğu boyunca heterojen olarak dağıtılmış önyargı ve varyansı düzeltmektedir. Standardizasyon ve mikrodizin veri normalleşme (SNOMAD) gibi yazılım paketleri bu tür normalleştirme tiplerinde kullanılabilir.

2.3.3.4. Özetleme

Görüldüğü gibi genlerin ifadesinde her prob seti için değişen sayıdaki proplar söz konusu olduğundan, aynı prob setindeki propların intensite değerleri “median-polish” algoritması ile tek bir ifade değeri verecek şekilde özetlemek gerekmektedir. Özetleme platform-spesifiktir ve başka bir deyimle Affymetrix için bir transkriptin ekspresyon düzeyi için tek bir ölçümü elde etmek için multipl proplar ile ilgili bilgilerin birleştirilip entegre edildiği bir süreçtir. MAS4, MAS5, dChip, PLIER, RMA ve GCRMA gibi çeşitli algoritmalar bu aşamada kullanılabilirlerdir.

2.3.3. TEMEL BİLEŞENLER ANALİZİ (PRINCIPLE COMPONENT ANALYSIS)

Temel Bileşenler Analizi (TBA/PCA) mikrodizin deneylerinden gen ifade verilerinde yer alan desenleri/ patenlerini bulmak için kullanılan bir keşif tekniğidir. Mikrodizin sonuçlarını analiz ederken PCA, aynı örneklerin birbirlerine benzer olup olmadığı sorusuna cevap vermek amacıyla kullanılır. Bu araç mikrodizin veri boyutluluklarını azalttığı için örnekler arasındaki ilişkinin görülmesini kolaylaştırır (28), ve söz konusu olan verinin orijinal hali hakkında maksimum bilgi vermektedir. PCA'nın temel prensibi bir takım değişkenleri, temel bileşenler olarak adlandırılan, daha az sayıda bağılantısız değişkenlere dönüştürmektedir. PCA verideki varyasyonu muhafaza eden, orijinal değişkenlerin temel eksen rotasyonudur. Sonuç olarak orijinal değişkenlerin toplam varyansı, temel bileşenlerin total varyansına eşittir.

Bir mikrodizin deneyinde PCA, ekspresyon değeri değişmeyen ve böylece örnek davranış farklılıkları hakkında bilgi vermeyen genler gibi gereksiz bilgileri tespit eder ve çıkarır. PCA yine tek değer ayrışımı olarak ta adlandırılmaktadır (29).

İki ya da üç boyutlu PCA plotunda herhangi iki nokta arasındaki yakın mesafe, genel gen ekspresyonu pateni benzer olduğunu göstermektedir. Diğer taraftan PCA grafiğinde uzak mesafede duran noktalar, farklı genel profillere sahiptir. Bu nedenle bu teknik, verideki aykırı ve majör etkilerin belirlenmesinde önemlidir. PCA yine otolog genler veri tabanı kümelerindeki tüm dizilmiş genomlar arasındaki ilişkilerini ifade etmek için kullanılmaktadır.

2.3.4. FARKLI OLARAK EKSPRESE EDİLEN GENLER

Gen ekspresyonu deneylerinin temel amacı, farklı şartlar altında farklı davranış gösteren genleri belirlemek ve bu farklı olarak davranan genler için bir güven ölçümünü saptamaktır. Farklı olarak eksprese edilen genlerin analiz basamakları T-testi, varyans filtrasyonu, çoklu test düzeltmesi ve kat değişim testi kapsamaktadır.

Mikrodizin analizleri yüksek rezolüsyonlu deneylerdir. Dolayısıyla karşılaştırma için bir takım t-test gerekmektedir ve karşılaştırmalar arasındaki bağımsızlığın yokluğu yanı sıra bu da yalancı pozitif olasılığını artırmaktadır (30). Varyans analizi (ANOVA) daha mantıklı bir alternatiftir. ANOVA çok gruplu verilerin analizini yapar ve tek bir p-değeri ile sonuçlanır. Bu da önem düzeyinin bağımsızlığını göstermektedir. Gen ekspresyonundaki varyansının geldiği kaynak sayısına göre, ANOVA tek yön, çift yönlü hatta çok faktörlü olabilir.

Çoklu test yapma, aynı anada birden fazla hipotezi değerlendiren bir süreçtir. Çoklu test yapmadaki zorluk fazla sayıda yalancı pozitiflerin olması ve bu yalancı pozitifler Yanlış Keşif Oranı (FDR) sayesinde düzeltilebilmektir. FDR pozitif bulguları arasındaki beklenen yalancı pozitiflerin oranıdır (31). Bir gen için kat değişimi, bir koşuldaki numunelerin gen ekspresyonunun ortalaması bölü başka bir koşuldaki numunelerin gen ekspresyonunun ortalaması olarak hesaplanır. Bu yaklaşım biyolojik çiftleri olmayan veri setleri için uygundur. Genellikle bir eşik değeri belirlenir ve bu eşik değerinin üstündeki herhangi bir fark (log formunda) anlamlı düşünülmemektedir.

2.3.5. KÜMELEME (CLUSTERING)

Kümeleme, farklı alt gruplardaki desenler ve benzerliği gösterme yöntemlerinden biridir. Kümelemede, gruplar genellikle ilgili birimlere ayrılır ve birimler arasındaki ilişkinin yüksek olması durumunda örneklerin aynı gruba ait olabileceği uzak veya düşük ilişki isen bileşenlerin aynı grupta olmadığı anlaşılmaktadır.

Kümelenme, aynı zamanda belirli bir veri setinin genel özelliklerinin anlaşılmasında yardımcı olan bir araştırma veri analizi aracı ve bir veri setinin daha anlamlı bir şekilde değerlendirilmesinde yararlıdır. Çoğu durumlarda kümeleme aracı çalıştırıldığında,

kümelerin oluşması beklenir ancak her oluşan küme biyolojik anlam taşıdığı anlamına gelmemektedir.

Kümeler denetimli veya denetimsiz olarak kategorize edilebilirler. Denetimli kümeleme metodları, kümeleme algoritmasını tanımlamak için özel işlevsel olarak ilişkili genlerdeki biyolojik bilgileri kullanmaktadır. Diğer bir deyişle, eğer kümeleşmesi beklenen genler hakkında ön bilgi varsa ve sonuçlar daha önceden edinilmiş bu bilgilerin doğrultusunda şekillenmişse denetimli kümeleme ortaya çıkar.

Hiyerarşik kümeleme aglomeratif bir yaklaşımdan oluşmaktadır ve gen ekspresyon veri analizinde dünya çapında en çok kullanılan kümeleme yöntemlerinden biridir. Bu tezde yapılan çalışmada hiyerarşik kümeleme yöntemi kullanılmıştır.

2.3.6. GEN ONTOLOJİ, ZENGİNLEŞME VE YOLAK ANALİZLERİ

Ontoloji, kavramların açıklamasını ifade etmektedir. Gen Ontolojisi (GO) Konsorsiyumu dinamik, genlerin farklı yönleri ile ilgili kontrollü ifadeninmesi ve gen ürünleri veya proteinleri toplayan bir projedir (32). Bu ifadelerin en önemli kullanım amacı RNA transkriptlerinin mikrodizin deneylerinin sonuçlarını açıklamak ve yorumlamaktır.

GO veritabanı merkezileştirilmiş değildir ancak, her gen ya da gen ürünü diğer veritabanlarında açıklanmış GO terimlerine dayanmaktadır. Birçok web tarayısı GO terimlerini aramak için temel ağ geçitleri olarak hizmet vermektedir. Ayrıca NCBI'deki Entrez Gen ve Entrez Protein girişleri GO terimleri içermektedir.

GO'nun organize düzenleyici ana prensipleri moleküler fonksiyon, biyolojik süreç ve hücresel bileşen olmak üzere üçe ayrılmaktadır. Moleküler fonksiyon, bireysel gen ürünleri tarafından gerçekleştirilen işlemleri tanımlamaktadır. Biyolojik süreç ise bir gen ürünü (protein) ile ilişkili olan geniş bir biyolojik hedeflere değinmektedir. Hücresel bileşen, proteinin subsellüler lokalizasyonunu tarif etmektedir. Herhangi bir protein birden fazla molekül fonksiyona, biyolojik sürece veya hücresel bileşene katılabilir. Bazı durumlarda alınan sonuç, farklı GO terim düzeyleri arasındaki ilişkiyi gösteren bir grafik ağaç görünümünü içermektedir.

GO terimleri, zenginleşme ve KEGG yolak analizleri “DAVID” yazılımı üzerinden “Fisher’s exact” testi uygulanarak gerçekleştirilmektedir. DAVID bütünleştirilmiş bir veritabanı olmakla birlikte geniş gen ve protein listelerindeki biyolojik anlamın çıkarılması amacıyla kullanılan sistematik bir analitik araçtır (33). DAVID yazılımı ile listede yer alan genler, majör gen aileleri, gen listelerinde hangi anotasyon terimleri ve grupların zenginleşmiş olduğunu, listelerdeki genler için hangi yolaklar ön planda olup tek tek genlerin hangi yolaklarda yer aldıklarını ve listelerdeki genlerin hangi hastalıklarla ilişkilendirilmiş olduğunu belirlemek amacıyla kullanılır.

2.4. META ANALİZ

Meta-analiz daha önceden yapılmış araştırmaların sonuçlarına bakarak bir takım yeni sonuçlara varma yöntemidir. Meta analiz, ilişkili fakat bağımsız çalışma sonuçlarının niceliksel olarak gözden geçirilmesi ve sentezi olarak tanımlanabilmektedir. Çoklu çalışmaların nicel sentez tekniği olarak, meta analizinin kullanımı hızla yaygınlaşmıştır. Aynı konudaki farklı çalışmalardan bilgi toplayarak yapılan birleştirme analizi, sadece tek bir çalışmaya dayanan analizden daha fazla istatistiksel güce sahip olacaktır (34).

2.4.1. MİKRODİZİN ÇALIŞMALARINDA META-ANALİZ

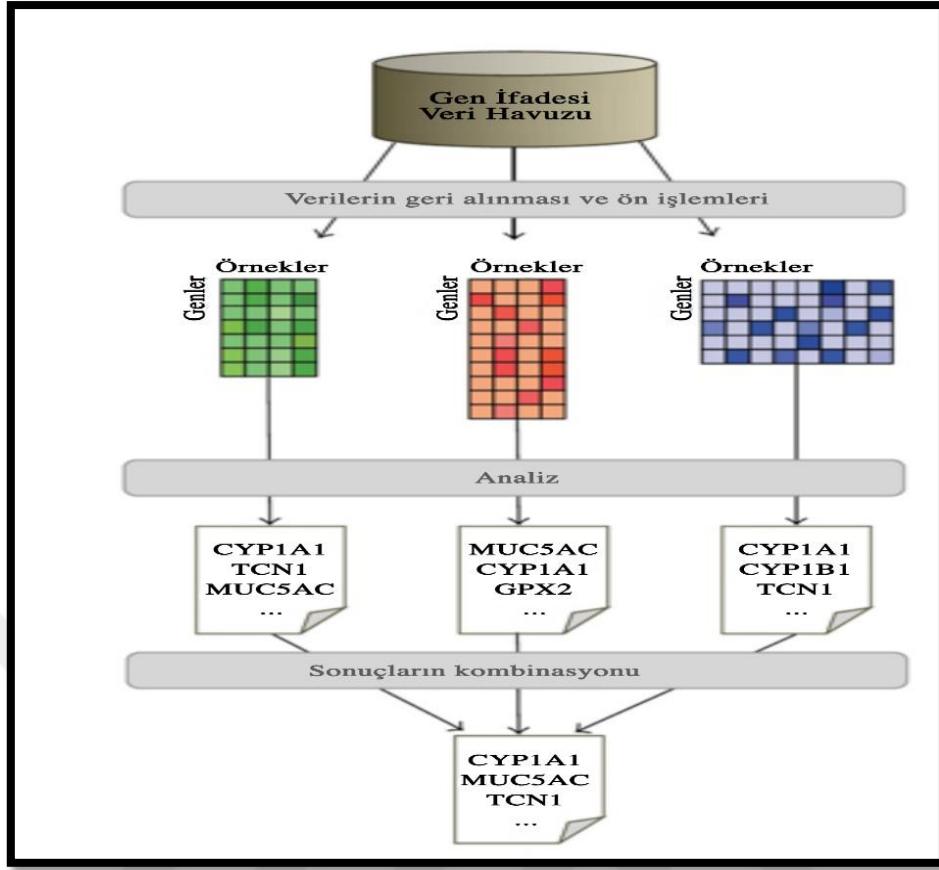
Meta-analiz yaklaşımı, gen ifade çalışmalarında hastalığın alt tiplerinin ortaya çıkarılması için (35), sağkalım tahmininde (36) ve biyobelirteçler ve tedavi hedeflerinin keşfinde (37) kullanılmaktadır.

Mikrodizin çalışmalarında meta-analiz genellikle GEO ve ArrayExpress gibi veritabanlarından ham verilerinin toplanması ile başlamaktadır. Ayrıca bu veritabanı tarama yapılırken AMSTRA, CASP, MOOSE ve PRISMA gibi kriterlere uygun olması şarttır. Meta-analiz yaklaşımında farklı platformlara ait verilerin birbirleriyle karşılaştırılabilmesi için ortak bir normalizasyon yöntemini gerektiğinden dolayı veritabanı arama yaparken dikkat edilmesi gereken husus verilerin farklı platformlardan geliyor olmasıdır. O yüzden bu tezin çerçevesinde tek platformdan gelen veriler kullanılmıştır.

Genel anlamda meta-analiz metotları iki ana gruba ayrılabilir. Bunlardan birincisi çalışmalardaki her gen ya da her prob için genel bir skor (meta-skoru) kullanan yaklaşımlardır (ör: GeneMeta, RankProduct vb). Başka bir deyişle çalışmalar önce bireysel olarak analiz edilip elde edilen sonuçları istatistiksel yöntemler kullanılarak bir araya getirilmektedir (şekil 2.2).

İkinci ana metot ise Naive meta-analiz yöntemidir. Burada çalışmaların ham verilerinin hiçbir işlemle geçirilmeden bir araya getirilmesi ve tek bir çalışma gibi analiz edilmesidir. Bu tez kapsamında naive meta-analiz metodu kullanılmıştır. Ayrıca bu yöntemin en büyük sorunu farklı parametrelere (kullanılan mikrodizin platformu gibi) sahip olan çalışmalar bir araya getirilmesi durumunda yaşanan istatistiksel güç kaybıdır. Bu nedenle bu tezin çerçevesinde bu farkı azaltmak için sadece Affymetrix platformu içeren çalışmaları kullanılmıştır.

Kabaca, meta-analiz yöntemlerini kullandığı çalışmalarda her bir çalışmanın kendine özgü amacına bağlı olarak analize dahil edilen münferit araştırma çalışmalarının bazı parametreleri esas alınarak meta-analiz çalışması tasarlanır.



Şekil 2.2: Meta-analizin şematik gösterimi (38).

2.4.2. KOLOREKTAL KANSERDE META-ANALİZ ÖRNEK ÇALIŞMASI

Chan ve ark. (2008) tarafından yapılan bir çalışmada (39), kolorektal kanser meta-analizini oldukça ayrıntılı bir şekilde araştırılmıştır. Yapılan bu çalışmada, kanser ve sağlıklı kontrol dokularının ifade profillerinin karşılaştırıldığı 23 bağımsız çalışma incelenmiştir. Bu çalışmalardan elde edilen sonuçlara göre örtüşen ve kolorektal kanserde önemli rol oynayan genleri saptamak için bir meta-analiz metodu geliştirmişlerdir. Araştırmaya dahil edilen çalışmalarda farklı mikrodizin platformları kullanılmış olduğu için her bir çalışmadan elde edilen farklı gen ifade profili gösteren prob verileri Genbank gen numaraları ile eşleştirilerek platformlara özgül gen listeleri elde edilmiştir. Chan ve ark, bu çalışmada simülasyon tekniğini kullanarak çalışmalarına dahil ettikleri 23 bağımsız araştırmayı 2'li gruplardan başlayarak artan sayıda gruplarda incelemiş ve bu oluşturulan bu gruplar içinde KRK tümör ve normal dokuları arasında farklı ifade profili sergileyen örtüşen genleri tespit etmiştir. Tespit edilen bu genler arasında ilgilenilen öncelikle gen

listesini oluşturmak için bu genleri önceliğe göre sıralamak (ranking) için aşağıdaki 3 kriter belirlenmiştir;

- (a) Çakışma seviyesi (bir genin karşılaştırılan listelerde bulunması ve ifade seviyesinin aynı yönde değişim göstermesi)
- (b) Örtüşen / çakışan genlerin tespit edildiği toplam çalışma sayısı
- (c) Uyum içinde olan çalışmalarından hesaplanan ortalama kat değişimi.

Sonuç olarak $p < 0,0001$ anlamlılık seviyede en azından 2 ortak çalışmada 573 genin farklı ifadeldendiği belirlenirken 3 ortak çalışmada bulunan gen sayısı ise 95 olarak kaydedilmiştir. Bunun yanı sıra 4 çalışmada çakışan gen sayısı 30 iken, 9 çalışmada iki çakışan gen bulunmuştur. Böylece Chan ve arkadaşları, sadece üç veya daha fazla çalışmalardaki örtüşen genleri qPCR validasyon olmak üzere ilerideki analizlere dahil etmişlerdir.

Kısaca, Chan ve ark (2008) tarafından yapılan bu çalışmada (39), Kolorektal kanserin genetik profilleri ile ilişkili önemli sonuçlar elde etmek amacıyla meta-analizi yöntemi başarılı bir şekilde kullanılmıştır. Üstelik birçok platformun (Çizelge 2.1) kullanımına bağlı olarak örtüşmeler tespit edilirken, örtüşmeyen genler analiz dışı bırakılması söz konusu olduğu için Meta-analizin Rank metodun uygulanması uygun bulunmuştur.

Çizelge 2.1: Chan ve ark çalışmada kullandığı çalışmaları ve platformları

Çalışma	Platform	Gen sayısı
Habermann et al.	Hs-UniGEM2 human cDNA microarray	9,128
Lin et al.	Custom cDNA microarray	23,040
Buckhaults et al.	SAGE	N/A
Notterman et al.	Affymetrix Human 6500 GeneChip Set	7,457
Galamb et al.	Human Atlas Glass 1.0 cDNA microarray	1,090
Wang et al.	TGS s-4k cDNA microarray	3,800
Croner et al.	Affymetrix HG-U133A	22,283
Kwon et al.	Macrogen MAGIC cDNA microarray	4,608
Bertucci et al.	Custom nylon cDNA microarray	8,074
Ohmachi et al.	Agilent cDNA microarray	12,814
Mori et al.	Human Atlas Glass 1.0 cDNA microarray	1,090
Kim et al.	Oligonucleotide microarray from Compugen/Sigma-Genosys	18,861
Zou et al.	Custom cDNA microarray	8,000
Koehler et al.	Atlas Human Cancer 1.2 Array	1,185
Ichikawa et al.	Custom cDNA microarray	20,784
Jansova et al.	Human 19K microarrays (Clinical Genomic Centre)	19,201
Grade et al.	National Cancer Institute oligonucleotide arrays (Operon V2 oligo set)	21,543
Bianchini et al.	Human 19K microarrays (Clinical Genomic Centre)	19,201
Agrawal et al.	Affymetrix Human 6800 GeneChip Set	7,129
Sugiyama et al.	Human Cancer Pathway Finder Gene Arrays (Superarray Bioscience)	96
Kitahara et al.	Custom cDNA microarray	9,216
Williams et al.	Custom cDNA microarray	9,592
Takemasa et al.	Custom cDNA microarray	4,608

2.5. KANSER

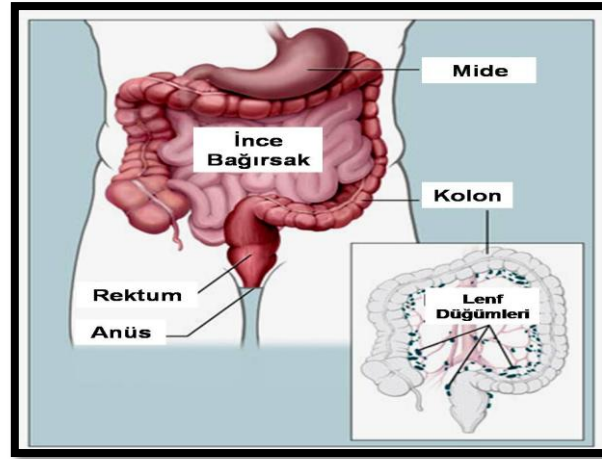
Karsinos ya da karsinoma kelimeleri, tıbbın atalarından biri olan yunan fizikçi Hippocrates (MÖ 460-370) tarafından ülser oluşturan ve ülser oluşturmeyen tümörler için kullanılan ilk kanser terimi olmuştur. Kanser, hücrenin proliferasyonu, hücre döngüsünü, hücre ölümünü kontrol eden genler ile hücrenin stabilitesini belirleyen genlerdeki mutasyon ve anormal aktivasyon sonucu olarak gerçekleşen bir biyolojik bir durumdur (40). Kanser, tek bir normal olmayan hücreden köken alsa da metastaza uğrayınca, mikroevrimsel bir süreç dönüşmeye başlamaktadır.

Kanserler köken aldıkları doku ve hücre türüne göre kategorize edilmektedir. Örneğin epitel hücrelerinden köken alan kanserler karsinom, bağ dokusu ya da kas hücrelerinden köken alanlar sarkom olarak isimlendirilmektedir. Genel olarak tek bir mutasyon, kansere neden olmak için yeterli değildir. Kanserde rol oynayan mutasyonlar genetik değişikliğinin yanı sıra epigenetik değişikliklerden de kaynaklanmaktadır ancak kanserlerin çoğu genetik değişikliklerden oluşmaktadır (41).

Kanser tek bir hastalık gibi görünmesine rağmen, aslında hücre ve dokuları etkileyen karmaşık ve çeşitli şekilleri mevcuttur. Bunlar, Meme Kanseri, Testis Kanseri, Kalınbağırsak ve Rektum Kanseri, Akciğer Kanseri, Rahim Boynu Kanseri, Rahim İç Zarı Kanseri, İdrar Yolu Kanseri, Mesane Kanseri, Ağız Kanseri, Gırtlak Kanseri, Prostat Kanseri ve Cilt Kanseri. Ayrıca Hoadley ve ark tarafından yapılan bir araştırmada (2014); Meme kanseri, Akciğer kanseri, Ovaryen kanseri, Mesane kanseri, Kalın Bağırsak Kanseri, Endometriyal kanser, Baş ve boyun karsinoma, Renal hücre karsinomu ve Akut miyelojen lösemi gibi en sık görülen kanser tiplerinin farklı genetik oluşumun içerildikleri gösterilmiştir (42). Bu tezin çerçevesinde ise kolorektal kanser üzerinde odaklanılmıştır.

2.5.1. KOLOREKTAL KANSER

Kolorektal kanser, dünya genelinde en yaygın görülen ve kanser sebepli ölümlerde ilk sıralarda yer almaktadır (43). Bu kanser türü, özellikle mide ve rektum kısmını ilgilendirmektedir ve kolon kısmı kalın bağırsağın büyük bir bölümüne denk gelirken rektum kısmı ise anüsten önceki yaklaşık 10 cm'lik kısımdır (şekil 2.3). Dünya çapında en sık görülen ve ölümlerle sonuçlanan bir kanser olduğu için kolorektal karsinoma hem etiyolojisi, hem de moleküler mekanizması, tanı ve tedavisi açısından yoğun bir şekilde pek çok araştırmacının ilgisini kazanmıştır. Kolorektal kanserin gelişimi bakımından çeşitli tipleri mevcuttur ayrıca bu tez, Sporadik kolorektal kanser ile ilgilendirmektedir. Sporadik kavramı ailede kanser öyküsü olmayan ve ailesinde ilk kez kolorektal kansere yakalanan bireyler için kullanılmaktadır. Sporadik KKR genellikle 60 yaşından sonra görülmektedir ve tüm kolorektal kanser olguların yaklaşık %70'ini oluşturmaktadır (22)



Şekil 2.3: Kalın bağırsağın anatomisi

2.5.2. KOLOREKTAL KANSERİN ETYOLOJİSİ

2012 yılındaki Globacan verilere göre, dünyada 14.1 milyon yeni kanser vakasının ortaya çıktığını, kanser ile yaşayan kişi sayının ise 32.6 milyon olup, kanser sebepli ölümlerin ise 8.2 milyona ulaştığı belirtilmiştir. Bunlar içinde Kolorektal kansere ait olan rakamlar ise erkekler için 746,000 yeni vaka (toplamın % 10'u) ile üçüncü sırada yer alıp kadınlarda ise 614,000 yeni vaka, (toplamın % 9.2'si) ile birlikte ikinci sırada yer almaktadır. Yeni vakaların neredeyse % 55'i gelişmiş ülkelerde meydana gelmekte ayrıca cinsiyet açısından, dünya genelinde kolorektal kanserin insidansı yaklaşık 10 kat bir değişim göstermektedir. En sık Avusturalya, Yeni Zelanda ve Kuzey Amerika'da görülürken, en az Batı Afrika ve Asya'da görülmektedir. Son yıllarda farklı coğrafyalarda oran değişsede de kolorektal kanser kadın ve erkekler arasında neredeyse eşit oranda görülen tek kanser türüdür (44). Ölüm oranı biraz düşük (694,000 ölümü, toplamdan % 8.5%) olarak kaydedilirken, en sık görülen ölüm sayısı ise %52 civarındaki bir rakam ile az gelişmiş ülkelere aittir.

ABD 2013 Kanser İstatistik raporunda, Amerika'da 2013 yılında 1.660.290 yeni kanser vakası olduğu ve 580,350 kanser sebepli ölüm gerçekleştiği bildirilmiştir. Bu raporda kolon kanseri, tahmini yeni vaka sayısında hem erkek hem de kadınlarda üçüncü sırada yer almaktadır. Yine kanser sebepli ölümlerde kolon kanseri hem erkeklerde (26.300) hem de kadınlarda (24.530) üçüncü sırada yer almaktadır (1). Türkiye'deki Sağlık Bakanlığı'nın 2013 yılındaki verilerine göre, kolorektal kanserin görülme sıklığı açısından tüm kanserler içinde % 8,3 ile kadınlarda üçüncü ve % 9,1 ile erkeklerde de üçüncü sırada yer almaktadır (45).

2.5.3. KOLOREKTAL KANSERİN RİSK FAKTÖRLERİ

Kolorektal kanser gelişiminde karsinojenler, mutajenler, vitaminler, diyet tipi, alkol, çevre etkiler ve benzer diğer risk faktörlerini önemli rol oynamaktadır.

Radyasyon; Pelvik radyasyona maruz kalma rektal kanserin bir risk faktör olduğu bilinmektedir. Bazen başka tür kanserlerin tedavi süresinde uygulanan bu pelvik radyoterapi, yaklaşık 15 yıl sonra hastalarda rektum kanseri gelişme riski artmaktadır.

Çevresel faktörler ve beslenme; Çevresel faktörler genetik yapıyı etkileyerek mutasyona sonuçlanıp kanserin gelişmesine yol açabilmektedir. Bunun yanı sıra, çeşitli çalışmalarda görüldüğü gibi, yağ ve kalori oranı düşük ve fiber içeriği yüksek diyetle beslenenlerde kolorektal kanserin daha az tespit edilip yüksek miktar yağ tüketimi ise kolon kanserin gelişmesinde kolaylaştırıcı bir faktör olduğu gösterilmiştir. Diyetteki yağ safra asitleri prostaglandin sentezini artırmak suretiyle kolon epitelinde hücre çoğalmasını hızlandırmaktadır. Diyetteki Fiber ise kolon lümenindeki pH'yı düşürerek karsinojen maddelerin aktivasyonu azalmaktadır.

D vitamini metabolitleri ise önemli hücresel aktiviteler üzerinde düzenleyici rolü oynayan bir kolorektal kanserin risk faktörüdür. Kolorektal kanser ile D vitamini ve kalsiyum tüketimi arasında ters ilişkiyi gösteren çeşitli çalışmalar olmasına rağmen bu tamamen açığa kavuşulan bir konu değildir.

Diğer kolorektal kanserin riskini artıran bir faktörü ise yıllarca sigara ve alkol (kronik) kullanımınıdır. Steroidal olmayan antiinflamatuvar ilaçlar, hormon replasman tedavisi ve fiziksel faaliyeti bu kanser türü için kolaylaştırıcı faktörler olarak gösterilmiştir (46).

2.5.4. KOLOREKTAL KANSERDE MOLEKÜLER GENETİK DEĞİŞİKLİKLER

Normal bir hücre çoğalma, farklılaşma, yaşlanma ve apoptoz gibi koordine ve düzenlenen süreçlerden geçmektedir. Tümör oluşumu ise genellikle hücrenin bu önemli dengesini bozan kontrolsüzlükten dolayı ortaya çıkmaktadır. Tümörler, kanser riskini taşımayıp vücudun diğer organlarına yayılmayan iyi huylu (benign) ve hızlı bir şekilde büyüyüp diğer organlara yayılma potansiyeline sahip olan kötü huylu (malign) olmak üzere iki ana gruba ayırmak mümkündür. Karsinogenezin ilk genetik değişiklikleri histopatolojik

incelemeler ile tespit edilemeyen oldukça hassas olayların olduğu da hatırlanmasını gerekmektedir.

Sporadik kanser, çevresel faktörlerin yanısıra tümör baskılayıcı gen ve onkogenlerdeki değişiklikler dâhil genetik değişikliklerden meydana gelir (47). Genel olarak kolorektal kanserin gelişimine sebep olan genetik değişiklikler; Tümör supresör gen aktivitesinin azalması veya kaybolması, protoonkogenlerde oluşan değişiklikler ve DNA onarımı (mismatch repair) ile ilgili genlerdeki değişiklikler olarak üç temel gruba ayrılabilir. Bu kategorilere göre sporadik kolorektal kanser vakalarının % 85’inde APC gen mutasyonu, %15’inde ise tamir genlerindeki yetersizlik (mikrosatellit instabilite) izlenmektedir (22).

2.5.4.1. Tümör Supresör Genlerdeki Değişimler

Tümör baskılayıcı genler ilk defa 1986 yılında Alfred Knudson tarafından keşfedilmiştir ve hücre çoğalmasında negatif yönde rol oynayan genlerdir. Bu genlerde, kansere neden olan mutasyonlar genellikle çekiniktir. En sık görülen mutasyonlar; nokta mutasyonları, inversiyon, mikrolezyon veya makro delesyonlar ve translokasyonlardır. Genellikle, tümör supresör genler iki kategori olarak sınıflandırılmaktadır. Birinci kategoride, proliferasyonda hücre apoptozisini etkileyen Bekçi (gatekeeper) genler bulunur. Bunlar TP53, APC, CDKN2A gibi genlerdir. İkinci kategoride ise Bakıcı (caretaker) genler yer alır. Genomun bütünlüğünden sorumlu DNA tamir genleri olan bakıcıların kendileri işlev kaybettirici bir mutasyona uğradığı zaman, genom boyunca mutasyonlar ortaya çıkmaya başlamaktadır. Başka bir deyimle bakıcı genlerde bir mutasyon geliştiğinde genomik instabilite gelişmektedir (48).

Tümör supresör genler genellikle proliferasyonu engellenmektedir. APC geninin (Adenomatous polyposis coli) inaktivasyonu ile başlayan bu süreç heterozigote kaybı olarak adlandırılır (Loss of heterozygosity-LOH). APC geni mutasyonunun kolorektal karsinogenezin erken dönemlerinde görünmektedir. Hücre çoğalması ve farklılaşmasının düzenlenmesinde rol oynayan p53 geni mutasyonu ise kromozom 17p’nin delesyon ile oluşan bir mutasyondur (49).

2.5.4.2. Protoonkogenlerin Aktivasyonu

Protoonkogenler hücrede uyarı iletiminde ve hücre büyümesinin kontrolünde rol oynayan genlerdir. Aslında protoonkogenler, hücre çoğalmasında itici rol oynamaktadırlar. Dolayısıyla bu genler normaldir ama herhangi bir nedenle mutasyonu edinirlerse, aktif edilmiş şekillere (onkogenler) dönüştürüp tümör oluşmasına yol açmaktadırlar. Onkogenler, kanser hücrelerindeki kontrolsüz proliferasyonu başlatan faktörlerdir. Protoonkogenlerin onkogenlere dönüşme sürecinde ise viral transdüksiyonla uyarımı, nokta mutasyonları, translokasyonlar ve gen amplifikasyonları gibi çeşitli mekanizmaları söz konusudur.

Önemli onkogenlerden biri olup p21 adlı hücre proliferasyonunda görev alan bir proteini kodlayan Ras geninin mutasyonlarının çoğu adenom gelişmesinin orta aşamalarında oluşmaktadır.

2.5.4.3. DNA Onarım Genlerindeki Değişiklikler (Mismatch Repair)

DNA onarım genlerindeki mutasyonlar ve kayıplar kolorektal karsinogenezin mühtemelen birden çok basmağında etkilidir. Kromozomlardaki nükleotid şifresi hücre bölünmesi sırasında yeni oluşan hücrelere aktarılmaktadır ve her 10 milyar nükleotid çiftinin kopyalanma sırasında bir hata oluşması normal olarak kabul edilebilir bir mutasyon oranıdır. Daha fazla miktarlardaki kopyalanma hatalarının düzeltilmesini sağlayan onarım genleri mevcuttur (Miss match repair genes – MMR). Genomik kararsızlık (genomic instability), yeterince mutasyona uğramış bir hücrenin kanser (malign) hücreğine dönüşümüne uygun bir durum oluşturur ve kolon karsinogenezinde rol oynayan önemli bir mekanizmadır (50).

2.5.5. KOLOREKTAL KANSER SINIFLANDIRMA

2.5.5.1. Kolorektal Karsinomlarda Histolojik Dereceleme (Grade)

Gradler, tümörün davranışının değerlendirilmesi, prognoz ve tedavi seçiminde önem taşımaktadır ve tümörlerin invaziv karakteristikleri ileri gradelerde artmaktadır.

International Union Against Cancer (UICC)' ye göre dört tip derece (grade) mevcuttur ve patolojik ölçüm genellikle hücrel diferansiyasyon seviyesi ile uyumludur.

G1: İyi diferansiye,

G2: Orta diferansiye,

G3: Kötü-az diferansiye,

G4: İndiferansiye.

WHO ise iki dereceyi kabul edilir. Bunlar; Düşük derece (G1 - G2 içerir) ve Yüksek derece (G3 - G4 içerir) ve bu sınıflama yöntemi klinikte kullanım açısından yeterli olarak görünmektedir. Kanselerde düşük ve yüksek dereceli alanlar birlikte bulunduğu zaman yüksek dereceli olarak rapor edilmesi gerekmektedir (51).

2.5.5.2. Kolorektal Karsinomlarda Evreleme (Stage)

Kolorektal kanser evrelemesi için değişik yöntemler kullanılmaktadır. Bu evreleme yöntemleri içinde 3 en sık kullanılan sistemleri; Dukes, Astra-Coller ve TNM'dır. TNM ve Dukes evreleme sistemleri çizelge 2.2'de ayrıntılı olarak verilmiştir. TNM evrelendirme ilk 2002 yılında yapıldıktan 8 yıl sonra son baskı (7. Basım) yeni TNM 2010 yılının başından itibaren yürürlüğe girmiştir, fakat bazı yayınlarda 2009 TNM olarak da adlandırıldığı görülmektedir. Astra-Coller sistemi ise şekil 2'de izah edilmiştir. Diğer önemli evreleme sistemi ise AJCC (American Joint Committee on Cancer) evrelemesidir (52).

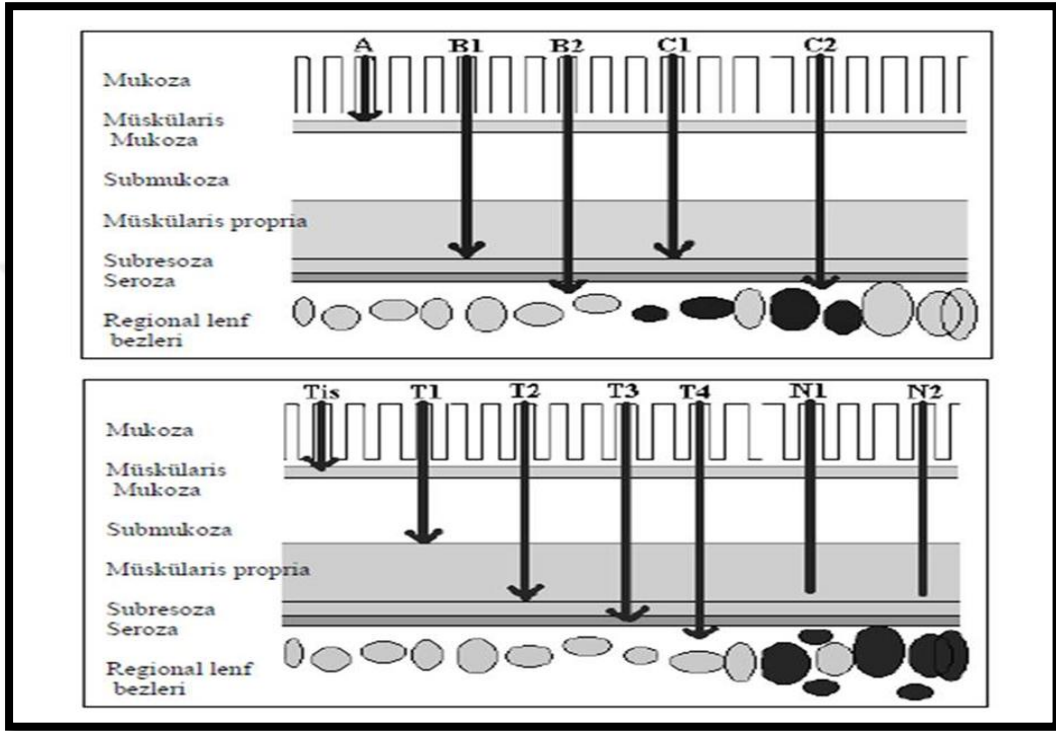
Çizelge 2.2: TNM ve Dukes sistemleri

AJCC Evre	TNM Evre	TNM Evre kriterleri	Dukes sınıflaması	Tümör invazyon derecesi
Evre 0	Tis N0 M0	Tis: Tümör mukozaya sınırlıdır; Karsinoma insitu		Gözle görülür tümör yok
Evre 1	T1 N0 M0	T1: Tümör submukozaya invaze	A	Tümör mskularis mukoza ve submukozaya sınırlı, mskularis propriayı aşmamış
Evre 1	T2 N0 M0	T2: Tümör mskularis propriaya invaze		
Evre II-A	T3 N0 M0	T3: Tümör subseroza ya da nonperitonealize perikolik/perirektal dokuya invaze	B1	Tümör mskularis propriaya sınırlı
Evre II-B	T4 N0 M0	T4: Tümör komşu organ/yapılara invazyon göstermekte ve/veya visseral peritonu perforer eder	B2	
EvreII I-A	T1-2, N1 M0	N1: 1-3 lenf bezi tutulumu mevcuttur. T1 veya T2.	C1	Serozaya ulaşmış, komşu organ tutulumu olabilir
EvreII I-B	T3-4, N1 M0	N1: 1-3 lenf bezi tutulumu mevcut. T3 veya T4.		
EvreII I-C	Herhangi bir T, N2 M0	N2: 4 veya daha fazla lenf bezi tutulumu mevcut. Herhangi bir T.		
Evre IV	Herhangi bir T, herhangi bir N, M1	M1: Uzak metastazlar mevcut. Herhangi bir T, Herhangi bir N.	D	Uzak metastaz var

T = Primer tümör, N = Bölgesel lenf düğümü tutulumu, M = Metastaz için kullanılır. TNM sistem ile ilgili diğer terimleri ise şöyle kullanılmaktadır. TX: Primer tümörü bilinmeyen, T0: Primer tümör yok, NX: Bölgesel lenf bezleri değerlendirilememekte, N0: Lenf bezi metastaz yok, MX: Uzak metastaz değerlendirilememekte ve M0: Uzak metastaz yok, anlamları taşımaktadır.

Astra-coller sistemi

Evre A: Tümör mukozada sınırlı, Evre B: Tümör submukozaya sınırlı, lenf bezi invazyonu yok, Evre B2: Tümör kas tabakasına sınırlı, lenf bezi invazyonu yok, Evre C1: Tümör barsak duvarını aşmadığı halde lenf bezi metastazı mevcut. Evre C2: Tümör barsak duvarını aşmış ve lenf bezi metastazı mevcut.



Şekil 2.4: Tumor evrelendirmesindeki Astra-Coller sistemi(50).

2.5.6. KOLOREKTAL KANSERİN GEN İFADE PROFİLİ

Mikrodizin çalışmalarındaki veri klinik çerçevesinde Kolorektal kansere uygulanır hale gelmiştir. Mikrodizin analizi, karsinogenezis prosesi olmak üzere moleküler sınıflandırmasını ve tedaviye yönelik çalışmalarını aydınlatır. Bugüne kadar binlerce tümör-normal karşılaştıran analizleri yapılmıştır ve bu analizlerde 2'den fazla kat değişimi ile diferansiyel olarak ifade edilen genin belirlenmesine sonuçlanır.

Kolorektal kanser, özellikle henüz tam anlaşılmayan moleküler alt yapı, moleküler alt tiplere olan ihtiyaç, tanı ve tedavide kullanılabilecek biyobelirteçler gibi pek çok başlık altında moleküler biyoloji ve genetik alanında sık çalışılan bir kanser türüdür. Bu

noktaların çözümünde; tümör oluşması ve ilerlemesi, prognostik ve prediktif değeri olan gen imzası ve tabiki yeni tümör hedeflerinin tespiti gibi önemli sonuçlar içeren küresel gen ifade çalışma örnekleri mevcuttur. Ancak klinikte değerlendirilebilmeleri için zamana ihtiyaç olduğu ve daha ileri çalışmaların yapılması gerektiği belirtilmektedir (53).



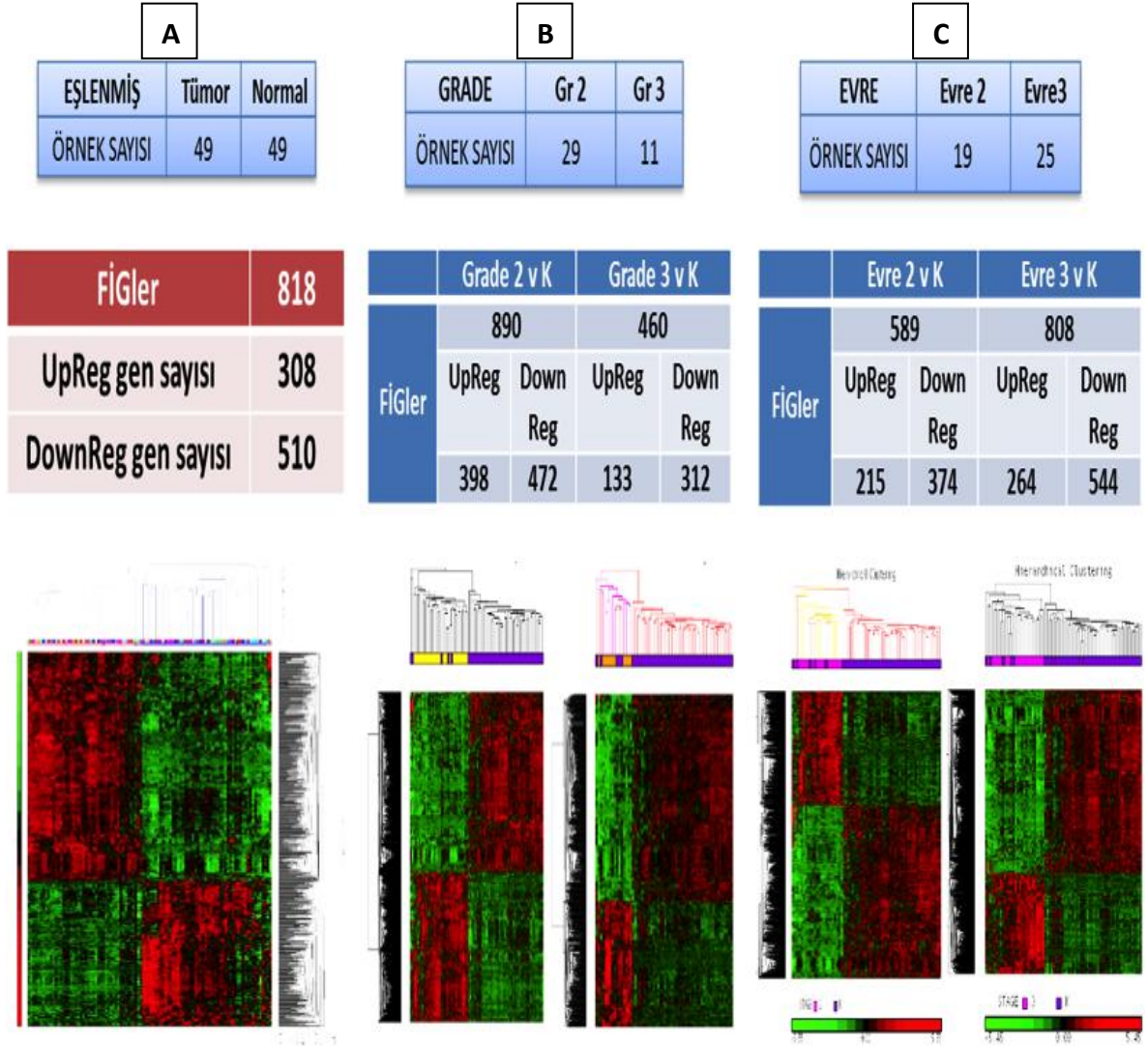
3. GEREKÇE ve AMAÇ

Kanser, belirli bir zaman süresinde, hücre çekirdeğinde, genomda ard arda mutasyonlar veya epigenetik kaynaklardan gelen değişikliklerin birikimi sonucunda ortaya çıkmaktadır. Karsinogenez karmaşık bir süreçtir yani kanserin gelişmesi, birçok faktörün etkisi altında, çeşitli genetik değişikliklerden etkilenmiş ayrıntılı bir süreci içermektedir (54). Gen ekspresyon profillemeye pek çok diğer omik teknik gibi bu sürecinin izah edilmesi amacıyla kullanılmaktadır. Gen ekspresyonu profillemeye yönteminin, geleneksel klinik ve patolojik parametrelere göre daha geniş kapsamlı klinik tahmin sonuçları sağlayabileceği öngörülmektedir (55). 2000’li yılların başlarından bugüne dek kanser sürecini aydınlatan ve kansere özgül moleküler imzalarını belirleyen binlerce gen ifade profillemeye çalışmaları yapılmıştır. Tümör-normal kıyasının yanı sıra karsinogenezin daha kolayca incelenip anlaşılabilmesi için kanserlerin sınıf ve evrelere göre çeşitli TNM (Tumor, Node, Metastasis), AJCC (American Joint Committee on Cancer) veya UICC (Union for International Cancer Control) sınıflandırma sistemleri geliştirilmiştir.

Bugüne kadar kolorektal kanser, meme kanseri, akciğer kanseri ve prostat kanseri gibi çeşitli kanser tiplerinin moleküler imzalarını belirlemeye yönelik çalışmalar yürütülmüştür. Bu tezin kapsamında ise kolorektal kanseri sağlıklı kolorektal dokudan ve KRK’nın farklı grade ve evrelerini birbirinden ayırabilen moleküler imzaların belirlenmesine yönelik olarak çalışılmıştır.

Diğer bütün kanserler gibi heterojen bir yapıya sahip olan kolorektal kanser, kalın bağırsak ve rektumda görülen kötü huylu (malign) oluşumlardır ve diğer kanser tipleri gibi hücresel düzeyde ilgili genlerdeki mutasyonlar veya gen ifadesindeki değişikliklerinden kaynaklanan genetik bir bozukluktur (56). Kolorektal kanserin erken teşhis ve doğru prognostik tayini tedavi edilebilirliği açısından son derece önemlidir. Erken teşhis ve prognozun öngörülebilmesi için biyobelirteçler etkindir ve kullanılan biyobelirteçler DNA, RNA, protein veya metabolit temelli olabilir (57). Bu bağlamda, gen ifadesinin, kolorektal kanser terapisinin karmaşık ağındaki boşlukları doldurmak için önemli bir araç olduğunu kanıtlanmıştır.

Araştırma ekibimiz 2009-2012 yılları arasında sporadik kolorektal kanser hastalarının parafine gömülü tümör ve eşli normal doku (n=49) materyelleri üzerinden yürütmüş olduğu gen ifade profillemesi çalışması sonucunda tümör/normal, grade ve evre temelli olarak yapılan karşılaştırmalarda Şekil 3.1’de görüleceği üzere farklı moleküler imzalar saptamıştır.



Şekil 3.1: Keşif setin normalleri ile tümör, grade ve evre karşılaştırılmasında elde edilen bulguları. Tümör ve normalin karşılaştırılmasından FİGler ve heatmap içeren elde edilen bulguları (A). B kısmı keşif setin grade bulgularını gösterirken C kısmı ise keşif setin evre bulgularını temsil etmektedir.

Bu tezin kapsamında bu bulgulara dayanarak Őu soru ortaya çıkmıŐtır. 49 KRK tımör ve eŐli kontrol dokudan elde edilmiŐ olan moleküler imzaları daha önce yapılan benzer kolorektal kanser alıŐmalarında saptamak mımkmün olabilecek midir? Diđer bir ifade ile saptamıŐ olduđumuz moleküler imzalar sporadik KRK iin evrensel nitelik taŐımakta, KRK sımrecinin anlaŐılması iin bir yol haritası sunmakta mıdır? Bu soru erevesinde aŐađıdaki hipotez bu tez kapsamında irdelenmiŐtir.

Hipotez

H₀: KeŐif setinin tımör, normal, evre ve sınıf karŐılaŐtırmalarından elde edilen KRK moleküler imzaları yalnızca bu hasta setine özđül deđil evrenseldir/btımün sporadik KRK dokuları iin geerlidir.

AraŐtırma ekibimizin Tıbitak desteđi ile yımrtmıŐ olduđu 100S477 nolu proje kapsamında elde edilen 49 KRK tımör ve eŐleŐmiŐ kontrol dokusuna ait gen ifade profil verisi bu tez kapsamında “KeŐif seti” olarak isimlendirilmiŐtir.

Bu tezin kapsamında aynı platformun (Affymetrix) üzerinde yımrtılmıŐ, sporadik kolorektal kanser alıŐmalarının tımör, evre ve sınıf bilgisi olan gen ifade verileri genomik veritabanlarında yıklenmiŐ olan alıŐmalar saptanıp indirilmiŐ ve “Validasyon seti” olarak adlandırılmıŐtır.

Tez alıŐması kapsamında keŐif ve validasyon setinden elde edilen moleküler imzalar sistematik bir Őekilde analiz edilerek karŐılaŐtırılarak incelenmiŐtir. Bu veri setlerinin arasındaki kıyaslanmanın ana hedefi, hasta örneklerimizden ıkan moleküler imzalarının valide edilmesidir. Bu tez alıŐmasına tek bir platform ile yapılan alıŐmalar dahil edilmiŐtir, zira birden fazla platformun kullanılması durumunda, kullanılan farklı mikrodizin platformlarının gen kapsamaları farklı olduđu iin, ortak olan/olmayan genlerin ifade dımeylerinin belirlenmesinde sonuların dođruluđunu azaltabilmektedir. Bu nedenle apraz platformların uyumsuzluklarını azaltmak iin tarafımızca yapılan bu meta-analiz alıŐmasında sadece Affymetrix platformu dikkate alınmıŐtır.

Kısaca, gımümüze kadar kolorektal kanser üzerine yapılan yımlerce alıŐmaya rađmen, kanserin oluŐumu, ilerlemesi ve yayılması ile ilgilenen hımresel, genetik ve moleküler mekanizmaların henüz btımünüyle anlaŐılamamıŐtır (58). Bu tez alıŐması sonucunda valide

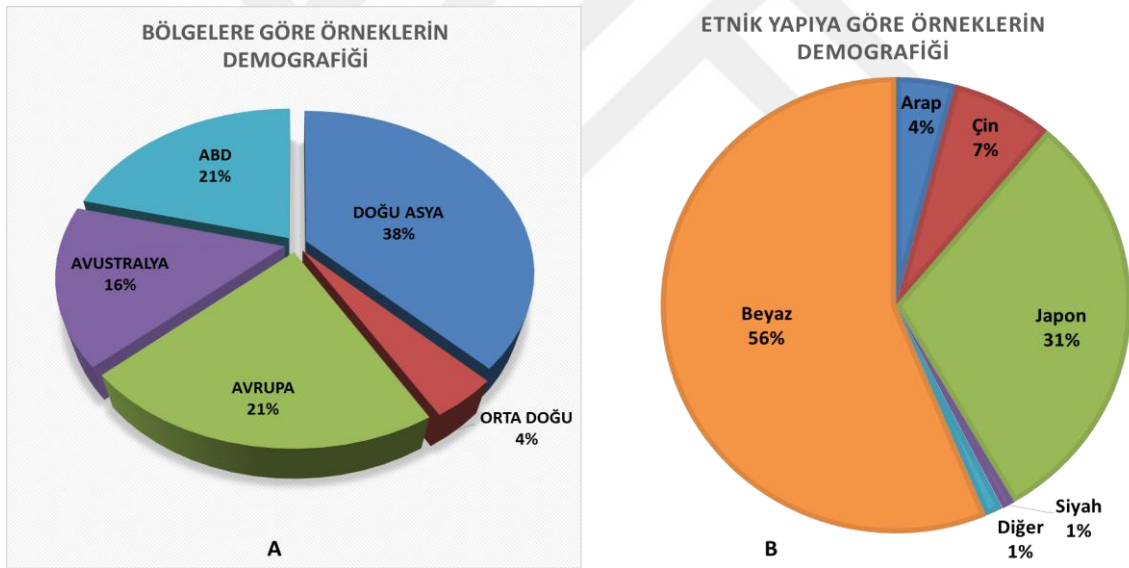
edilen Kolorektal kanser moleküler imzalar, bu kanser tipinin moleküler mekanizmasının aydınlatılmasına katkıda bulabilecektir.



4. MATERYEL ve METOD

4.1. SEÇİLEN ÇALIŞMALARINDAN VERİ TOPLANMA.

GEO ve Array Express'i Express olmak üzere başka genomik veritabanlarından Affymetrix platformun üzerinde hayatta geçiren tüm kolorektal kanserdeki gen ifade profilleri içeren çalışmalarını sorgulayarak bizim kohortumuza benzeyen çalışmaları içeren bir validasyon seti elde edilmiştir. Büyük örnek sayısına sahip olan çalışmaları tercih edilmiştir. Bu örnekler cerrahi rezeksiyondan elde edilen kanserli tümör, normal kolorektal epitel ve adenomatöz dokuları içermektedir (Çizelge 4.4). Validasyon setinde kullanılmak üzere diğer çalışmalardan alınan örneklerin demografik dağılımları ise şekil 4.1' de verilmiştir.



Şekil 4.1: Validasyon setinin çalışmalarından alınan örneklerin bölgelere göre (A) ve etnik yapısına göre (B) demografik dağılımları

4.2. VERİNİN ALT GRUPLARA AYRILMASI

Daha net inceleme yapabilmesi için toplam elde edilen çalışmaların alt gruplanması yapılmıştır. Örneğin tümör grupları (tüm tümör, eşleşmiş ve eşleşmiş olmayan), normal grubu, evre ve derecelendirmiş grupları ayrılmıştır. Bu alt gruplarının birbiriyle kıyaslaması oldukça kolaydır. Çeşitli alt grupların incelenmesi için, ayrıntılar 4.1, 4.2 ve 4.3 numaralı çizelgelerde bulunmaktadır

Çizelge 4.1: Eşleşmiş örneklerini içeren çalışmalar

Çalışma No	Yazar	Örnek Sayısı	Platform	Dergi	Yayın Tarihi
GSE32323	Galamb <i>et al</i>	34	Affymetrix	PLoS ONE	1/1/2013
GSE15960	Galamb <i>et al</i>	24	Affymetrix	BJC	3/3/2010
GSE22598	Okazaki <i>et al</i>	34	Affymetrix	IJO	4/12/2011

Çizelge 4.2 Eşleşmiş olmayan örneklerini içeren çalışmalar

Çalışma No	Yazar	Örnek sayısı	Platform	Dergi	Yayın Tarihi
GSE4107	Hong <i>et al.</i>	22	Affymetrix	Clin Cancer Res	20-Mar-07
GSE23878	Uddin <i>et al.</i>	59	Affymetrix	Am J path	31-Aug-10
GSE4183	Galamb <i>et al.</i>	53	Affymetrix	Dis Markers	25-Jan-08
GSE10714	Galamb <i>et al</i>	33	Affymetrix	Cancer Epid. Biomarkers	29-Jan-09
GSE9348	Hong Y	82	Affymetrix	Clinical & exptal metastasis	15-Feb-10
GSE37364	Galamb <i>et al.</i>	94	Affymetrix	PLoS ONE	1-Jan-13
GSE21510	Tsukamoto <i>etal</i>	148	Affymetrix	ClinCancer Res	28-Feb-11

Çizelge 4.3: Derecelendirilmiş (Grade) ve Evrelendirilmiş(Stage) örneklerini içeren çalışmalar

Çalışma No	Yazar	Örnek sayısı	Platform	Dergi	Yayın Tarihi
*GSE32323	Galamb <i>et al</i>	34	Affymetrix	PLoS ONE	1-1-2013
*GSE21510	Tsukamoto et al.	148	Affymetrix	Clin Cancer Res	28-Feb-11
*GSE23878	Uddin et al.	59	Affymetrix	Am J path	31-Aug-10
**GSE18088	Gröne et al	53	Affymetrix	Int J Colectal Dis	10-Apr-11
**GSE17538	Smith et al	244	Affymetrix	Gastroenterology	14-Nov-09

** Grade örnekleri içeren çalışmalarını işaret ederken * ise Evreleri gösterir.

Çizelge 4.4: Seçilen Çalışmaların Toplamı

Çalışma No	Örnek sayısı	Yazar	Dergi	Yayın Tarihi
GSE4107	22	Hong Y, Ho KS, Eu KW, Cheah PY	Clin Cancer Res	20-Mar-07
GSE23878	59	Uddin S, Ahmed M, Hussain A, Abubaker J, Al-Sanea N, Abdul Jabbar A, Ashari LH, Alhomoud S, Al-Dayel F, Jehan Z, Bavi P, Siraj AK, Al-Kuraya KS	Am J path	31-Aug-10
GSE37364	94	Galamb O, Wichmann B, Molnar B	PLoS ONE	1-Jan-13
GSE32323	34	Ahmed K, Toshiaki I, Kazuro S, Kaoru M, Satoru I, Megumi I, Hiroshi M, Hiroshi T, Hiroyuki U, Kenichi S	Cancer genomics proteomics	20-Mar-12
GSE22598	34	Okazaki S, Ishikawa T, Iida S, Ishiguro M, Kobayashi H, Higuchi T, Enomoto M, Mogushi K, Mizushima H, Tanaka H, Uetake H, Sugihara K.	International Journal of oncology	4-Dec-11
GSE9348	82	Hong Y	Clinical & experimental metastasis	15-Feb-10
GSE15960	18	Galamb O, Molnar B, Spisak S, Sipos F	BJC	3-Mar-10
GSE4183	53	Galamb O, Györfy B, Sipos F, Spisák S et al.	Dis Markers	25-Jan-08
GSE10714	33	Galamb O, Molnar B, Spisak S, Sipos F	Cancer Epid Biomarkers	29-Jan-09
GSE18105	111	Matsuyama T, Ishikawa T, Mogushi K, Yoshida T, Iida S, Uetake H, Mizushima H, Tanaka H, Sugihara K	Int J Cancer	4-Feb-10
GSE27854	115	Kikuchi A, Ishikawa T, Iida S, Ishiguro M, Mogushi K, Mizushima H, Uetake H, Tanaka H, Sugihara K	Int J Cancer	16-Oct-12
GSE21510	148	Tsukamoto S, Ishikawa T, Iida S, Ishiguro M, Mogushi K, Mizushima H, Uetake H, Tanaka H, Sugihara K	Clin Cancer Res	28-Feb-11
GSE35896	62	Schlicker A, Beran G, Chresta CM, McWalter G, Pritchard A, Weston S, Runswick S, Davenport S, Jackson K, Castro DA, Orphanides G, French T, Wessels LF	BMC Med Genomics	14-Jan-13
GSE18088	53	Gröne J, Lenze D, Jurinovic V, Hummel M, Seidel H, Leder G, Beckmann G, Sommer A, Grützmann R, Pilarsky C, Mansmann U, Buhr H, Stein H, Hummel M	Int J Colectal Dis	10-Apr-11
GSE14333	290	Sieber OM	Clin Cancer Res	8-Jan-10
GSE17538	244	Smith JJ, Deane NG, Wu F, Merchant NB, Zhang B, Jiang A, Lu P, Johnson JC et al.	Gastroenterology	14-Nov-09

4.3. HAVUZLANMIŞ VERİLERİNİN (POOLED DATA) ÖN İŞLEMESİ

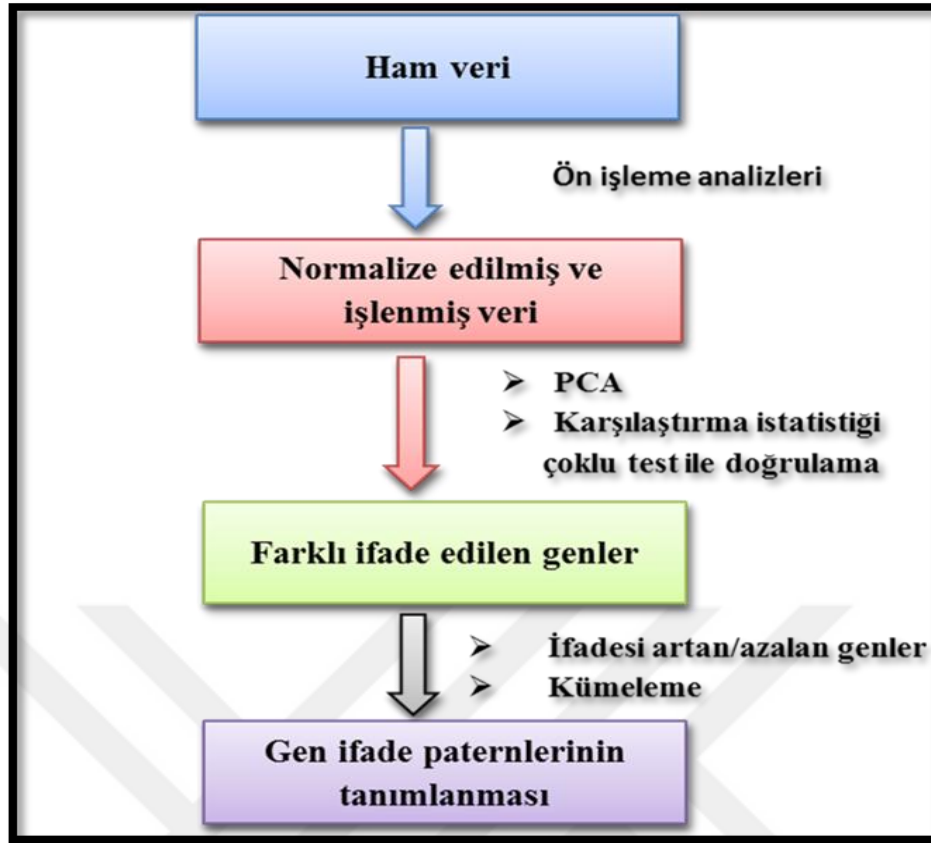
Doğruluk ve hassasiyet açısından iyi performans gösteren Robust Multiarray Ortalama (RMA) algoritması ve Partek yazılımı mikrodizin gen ekspresyon düzeyleri ile ilişkili çip arka plan çıkarmak için kullanılmıştır. Normalizasyon, çalışmaların arasındaki sistematik farklılıkları ortadan kaldırmaktadır (13) dolayısıyla gen listelerini almadan normalizasyon işlemleri yapılmıştır. Özellikle bu çerçevede verilerin farklı laboratuvarlardan elde edilmiş olması nedeniyle olası teknik farklılıklardan kaynaklı veri kalitesi farklılıklarının tespitine yönelik olarak BRB dizi araçlar (BRB-ArrayTools) yazılımı kullanılarak Spearman korelasyon analizleri yapılmıştır.

Kısaca, her setin örnekleri birleştirilerek Partek'e taşınıp farklı setlerin tüm gen ekspresyon değerlerinin bir araya getirilmesinden önce ön işleme yapılmıştır.

4.4. GEN LİSTESİNİN ÇIKARTILMASI

Partek® Genomic Suite™ 6.6 yazılımını kullanarak öncelikle keşif setin (yani kohortumuzun) diferansiyel olarak anlamlı genleri belirlenmiştir. Aynı şekilde validasyon setinin tüm alt grupların temelinde gen listeleri de elde edilmiştir. Gene listesinin çıkarılmasında çok gruplu verilerin analizini yapıp tek bir p-değeri ile sonuçlayan ANOVA metodu kullanılırken, aynı anada birden fazla hipotezi değerlendirilmesi söz konusu olduğu için Çoklu test'e ait olan fazla sayıda yalancı pozitifleri Yanlış Keşif Oranı (FDR) sayesinde düzeltilmiştir (şekil 4.1).

Bu tezin çerçevesinde, farklı gen listelerine ait desenler ve benzerliğinin gösterilmesi için aglomeratif bir yaklaşımından oluşan hiyerarşik kümeleme yöntemleri kullanılmıştır. Böylece her setin örnekleri ile ifade azalan/artan genlerin ilişkisi incelenmekle birlikte setler arasındaki kümeleme benzerliği de değerlendirilmiştir. Aynı şekilde çıkan gen listelerin kıyaslanmasına yardım etmek amacıyla her set için tümör, evre veya grade örnekleri ile normal örnekleri ilişkisinin saptanması için PCA yöntemi kullanılmıştır.



Şekil 4.2: Mikrodizin veri analiz süreci

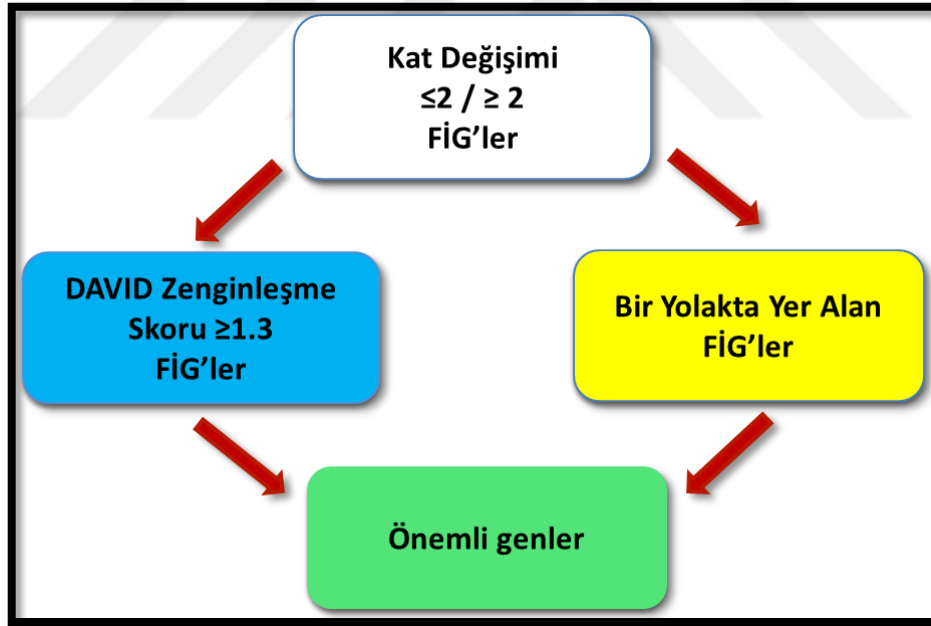
4.5. ÖRTÜŞME SAPTANMASI

Venny biyoinformatik yazılımını kullanarak Keşif ve Validasyon setlerinin gen listelerinin arasındaki örtüşen genler saptanmıştır. Yani, keşif setindeki n tane gen varsa bunlardan validasyon setinin gen listesinde ne kadar olduğu belirlenmiştir. Evre ve gradeler gibi çeşitli alt gruplarındaki çakışma oranı da belirlenmiştir. Örneğin keşif setinin eşleşmiş alt grubu ne kadar validasyon setinin eşleşmiş alt grubuna benziyor diye sorgulanarak bu iki alt grubun üstü üste gelen genleri belirlenmiştir. DEG'ler tek platform ve aynı süreç ile üretilmiş oldukları için çakışma öncesinde sıralama (ranking) veya bir takım dengeleme işlemlerinin ihtiyacı duyulmamıştır.

4.6. ÇAKIŞAN GEN LİSTELERİN FONKSİYONEL YOLAK VE ZENGİNLEŞME ANALİZİ

Ontoloji, zenginleşme ve KEGG yolak analiz kullanarak alt gruplar dâhil çeşitli validasyon setlerin keşif setin arasındaki gen listelerinde bulunan önemli genleri (artan ya da azalan) içeren özet bir değerlendirme yapılmıştır. Bu tez çalışmasının aşmasında ifadesi azalan ve artan gen listeler özellikle “DAVID” yazılımı üzerinden ayrı ayrı incelenmiştir. Kabul edilebilir belli istatistiksel anlamlılıkta farklı olarak ifade edilen genler:

1. Kat değişimi $\leq 2 / \geq 2$ mi? (Çakışan genlerin hepsi bu kriteri yerine getirmektedir)
2. Bir yolağın üzerinde yer alıyor mu?
3. Zenginleşme skoru ≥ 1.3 olan bir DAVID fonksiyonel anotasyon kümesinde yer alıyor mu?



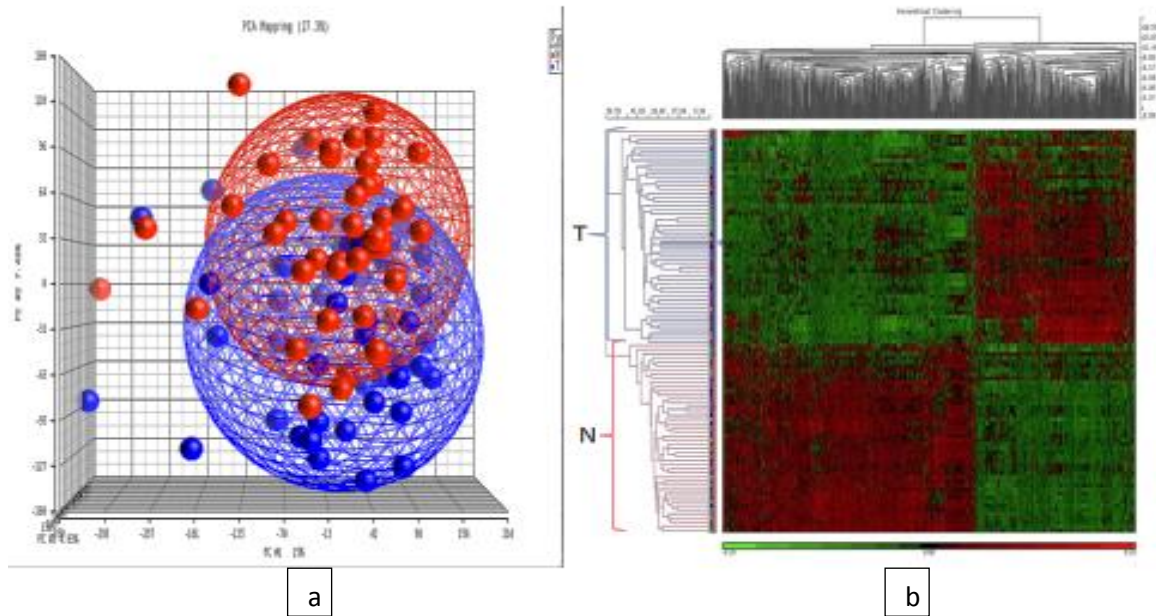
Şekil 4.3: Meta-analiz sonucunda önemli genlerin belirtme şeması

5. BULGULAR

5.1. TÜMÖR VE KONTROL KARŞILAŞTIRILMALARININ BULGULARI

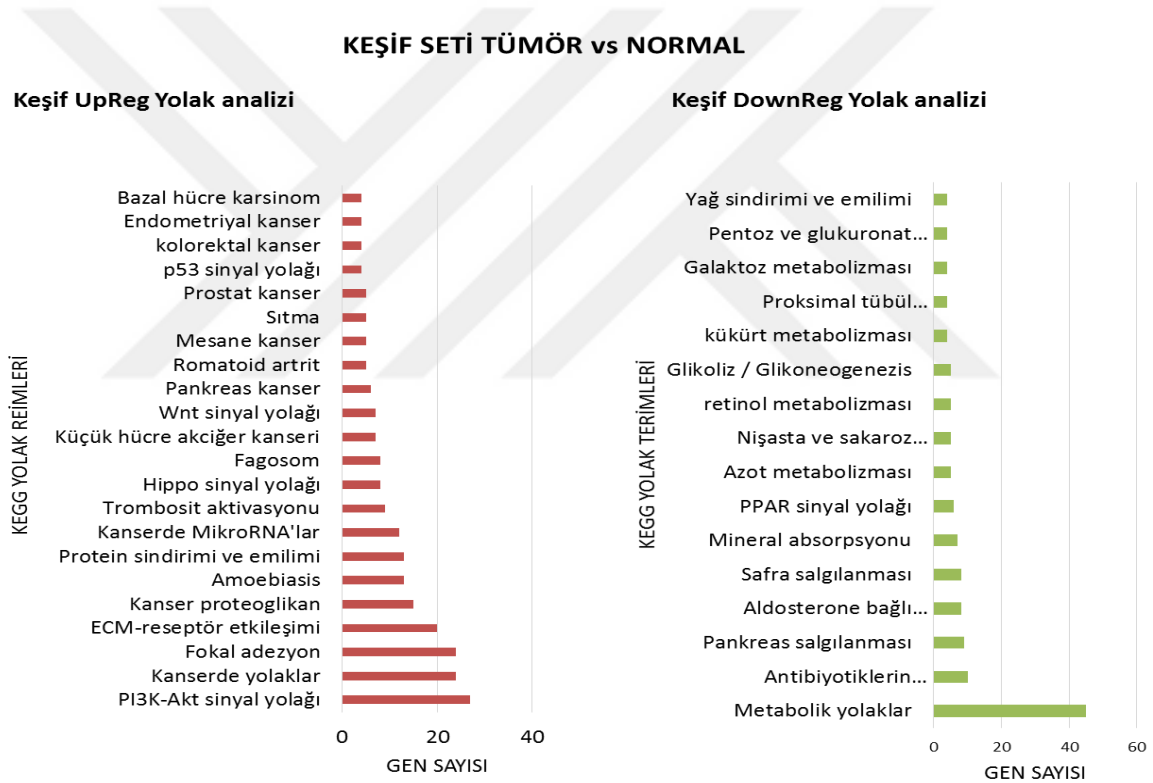
5.1.1. KEŞİF SETİNİN (49 TÜMÖR VE 49 EŞLEŞMİŞ KONTROL) BULGULARI

Keşif setindeki bulgulara göre en az 2 kat değişimi ile 1333 gen ($p < 0,001$) farklı olarak ifade edilmiş olup; 816 genin ifadesinde artma gözlenirken, 517 genin ifadesinde azalma olduğu tespit edilmiştir. Toplam 1333 genden ise 1186 gen (%89) Entrez Gene ID'si ile tanımlanabilmiştir. PCA (Temel Bileşenler Analizi-Principal Component Analysis) sonucu 49 normal ve 49 kanser örneğinin dağılımları incelendiğinde, bazı normal ve kanser örneklerinde kendi aralarında gruplanırken, bazı kanser ve normal örnekler ortak olarak toplanmıştır. Hiyerarşik kümeleme analizi sonucunda ise örneklerin gen ifade paternlerine göre tümör ve kontrol olmak üzere 2 gruba oldukça net bir şekilde ayrıldığı görülmektedir (şekil 5.1).



Şekil 5.1: Keşif setindeki tümör ve normalin aralarında çıkan PCA ve heatmap sonuçları; **a.** kırmızı normalleri ve mavi tümör örneklerini içeren PCA **b.** Hiyerarşik kümeleme analizi sonucu elde edilen heatmap .

Yapılan bu tez çalışması kapsamında DAVID yolak analizi sonuçlarına göre metabolik yolak, PPAR sinyal yolağı, aldosterone bağı sodyum reabsorbsiyonu, azot metabolizması, pankreas sekresyonu, mineral absorbsiyonu, kükürt metabolizması, nişasta ve sakaroz metabolizması, retinol metabolizması ve galaktoz metabolizması gibi yolakların keşif setinde yer alan ifadesi azalan genler için önemli yolak oldukları gösterilmiştir. İfadesinde artış gösteren genlere bakıldığında, P13K-Akt-sinyal yolağı, kanserde yolaklar, fokal adzeyon, ECM-reseptör etkileşimi, p53 sinyal yolağı, hippo sinyal yolağı, kanser proteoglikan, kanserde rol oynayan mikro RNA'lar, amoebiasis, trombosit aktivasyonu, Wnt sinyal yolağı, protein sindirimi ve absorbsiyonu, bazal hücre karsinoma, kolorektal kanser ve endometriyal kanser gibi KEGG yolaklarının önemli olduğu gösterilmiştir (Şekil 5.2).



Şekil 5.2: Keşif setinin tümör ve normallerin karşılaştırmasından elde edilen ifadesi artan ve azalan genlere ait yolak analizinin bulguları

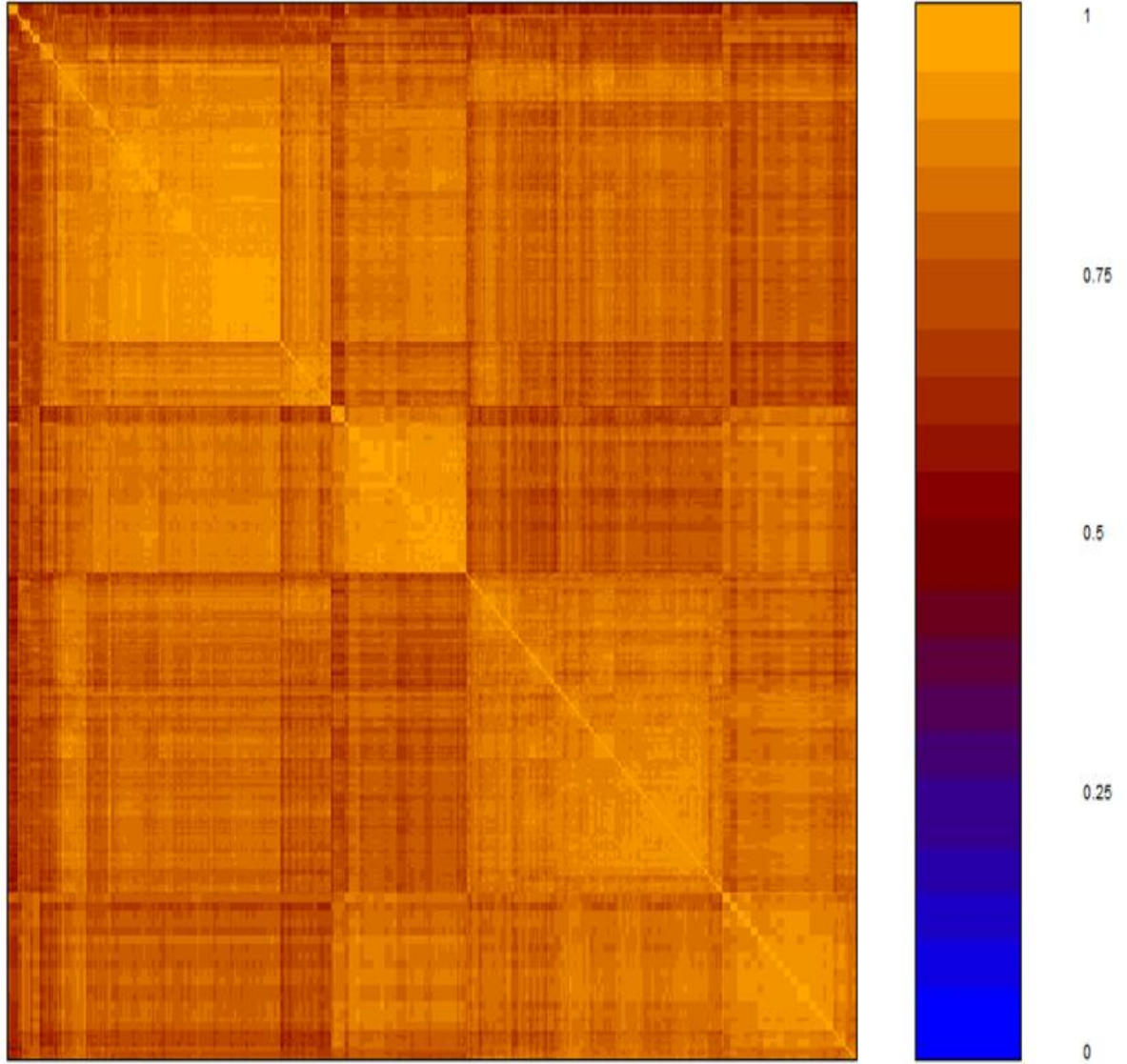
5.1.2. VALİDASYON SETLERİN ARASINDAKİ KIYAS BULGULARI

Tüm validasyon setinin (225 tümör - 160 karışık kontrol) incelenmesinde 2008 genin ifade seviyeleri değişiklik göstermiştir. Bu genlerden 1176 genin ifade seviyelerinde artış ve 832 genin ifade seviyelerinde düşüş görülmüştür. Toplam 2008 genden 1920 gen (%89) Entrez Gene ID'si ile tanımlanabilmiştir. Normalizasyon hemen ardından örneklerin yaklaşık %85-95 oranı ile tekniksel olarak uyum halinde olduğunu gösterilirken (Şekil 5.3), PCA analizi sonucu ise bazı normal ve kanser örneklerinin bir arada toplandığı, diğer bazı kanser ve normal örneklerinin ise birbirlerinden ayrıldığı gözlenmiştir (Şekil 5.4). Hiyerarşik kümeleme analizi sonucunda ise örneklerin gen ifade profillerine göre tümör ve kontrol olmak üzere 2 gruba çok net bir şekilde ayrıldığı görülmektedir (Şekil 5.5). DAVID biyoinformatik yazılım tarafından yapılan analize göre, bu setteki ifadesini artan genlerinin KEGG analizleri sonucunda 28 yolda yer aldığı, 47 yolağın ise ifadesi azalan genleri içerdiği belirlenmiştir (Şekil 5.6, 5.7).

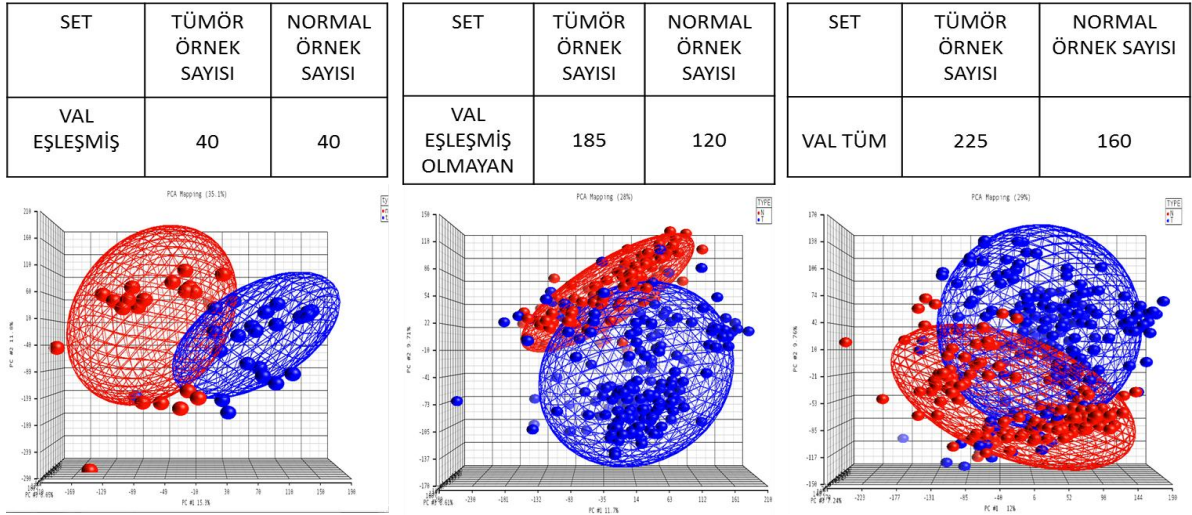
Eşleşmiş validasyon seti (40 tümör - 40 eşleşmiş kontrol) analiz edildiğinde ise toplam 3113 genin ifade seviyeleri değişiklik göstermiştir. Bu genlerden 1506 genin ifade seviyelerinde artış ve 1607 genin ifade seviyelerinde düşüş görülmüştür. Farklı ifade seviyesi görülen toplam 3113 genin 2950'sini (% 94.8) Entrez Gene ID'si ile tanımlanmak mümkün olmuştur. Yine, PCA yöntemi sonuçlarına göre normal ve kanser örneklerinin neredeyse tamamen ayrılmıştır (Şekil 5.3). Bununla birlikte, hiyerarşik kümeleme analizi ve örnek-gen ifade seviyesinin ilişkisi oldukça belirgin bir şekilde ayrıldığı görülmüştür (Şekil 5.5). DAVID analize göre, bu setin ifadesini artan genlerinin KEGG analizi sonucunda 24 yolağın gösterilirken, ifadesi azalan genlerin 48 yolağın içinde yer aldığı belirlenmiştir (Şekil 5.6, 5.7).

Eşleşmiş olmayan validasyon setinin (185 tümör - 120 eşleşmemiş kontrol) analizinden çıkan sonuçlara göre, 1976 genin ifade seviyeleri değişiklik göstermiştir. Bu genlerden 1213 genin ifade seviyelerinde artış ve 763 genin ifade seviyelerinde düşüş olduğu belirlenmiştir. Toplam 1976 genden, 1868 gen (% 94.5) Entrez Gene ID'si ile tanımlanabilmesi mümkün olmuştur. PCA ve hiyerarşik kümeleme tümör-normal örneklerin ve örnek-ifadesi artan/azalan genlerin sırasıyla açık bir şekilde ayrıldığı görülmüştür (Şekil 5.4, 5.5). DAVID sonrasındaki sonuçlarına göre, bu set içerisinde yer alan ifadesini artan genlerinin KEGG analizinin sonucunda 21 yolağın gösterilirken, ifadesi azalan genler 39 yolağın içerisinde dağılım göstermiştir (Şekil 5.6, 5.7).

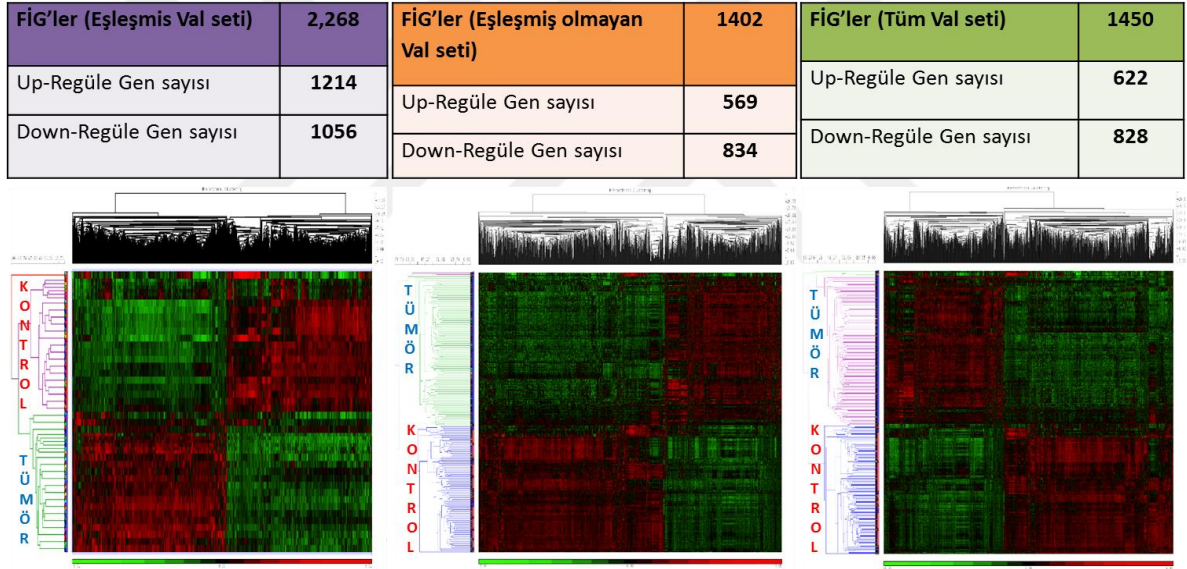
Spearman correlation plot of log(intensity)



Şekil 5.3: Tüm validasyon setinin örneklerin Spearman korelasyon çizimi.



Şekil 5.4: Validasyon setlerin PCA bulguları



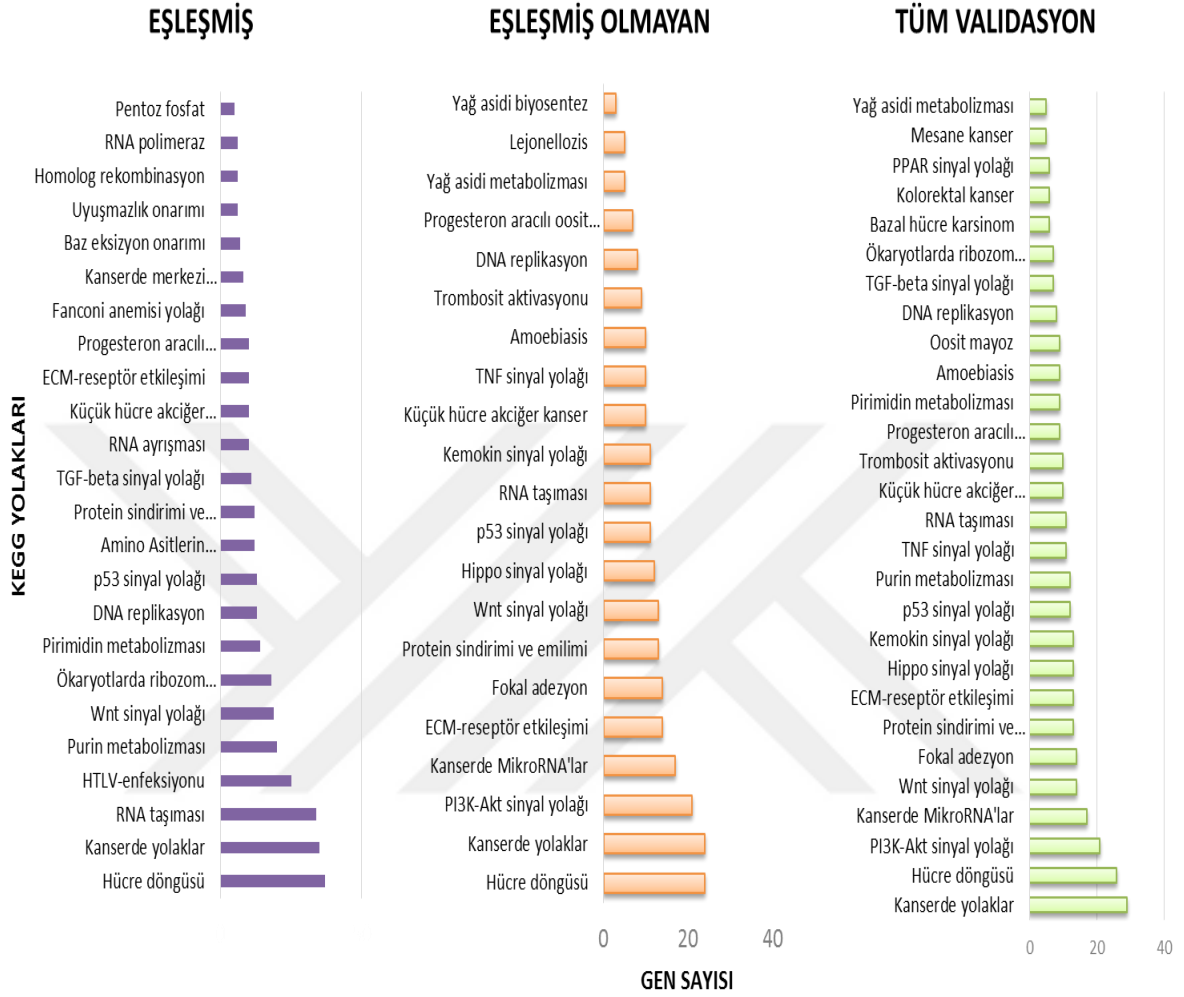
Şekil 5.5: Validasyon setlerin tümör ve normallerin karşılaştırmasından elde edilen heatmap bulguları.

Validasyon setlerin Tümör vs Kontrol DownReg Yolak Analizleri



Şekil 5.6: Validasyon setlerin tümör ve normallerin karşılaştırmasından elde edilen ifadesi azalan genlere ait yolak analizinin bulguları

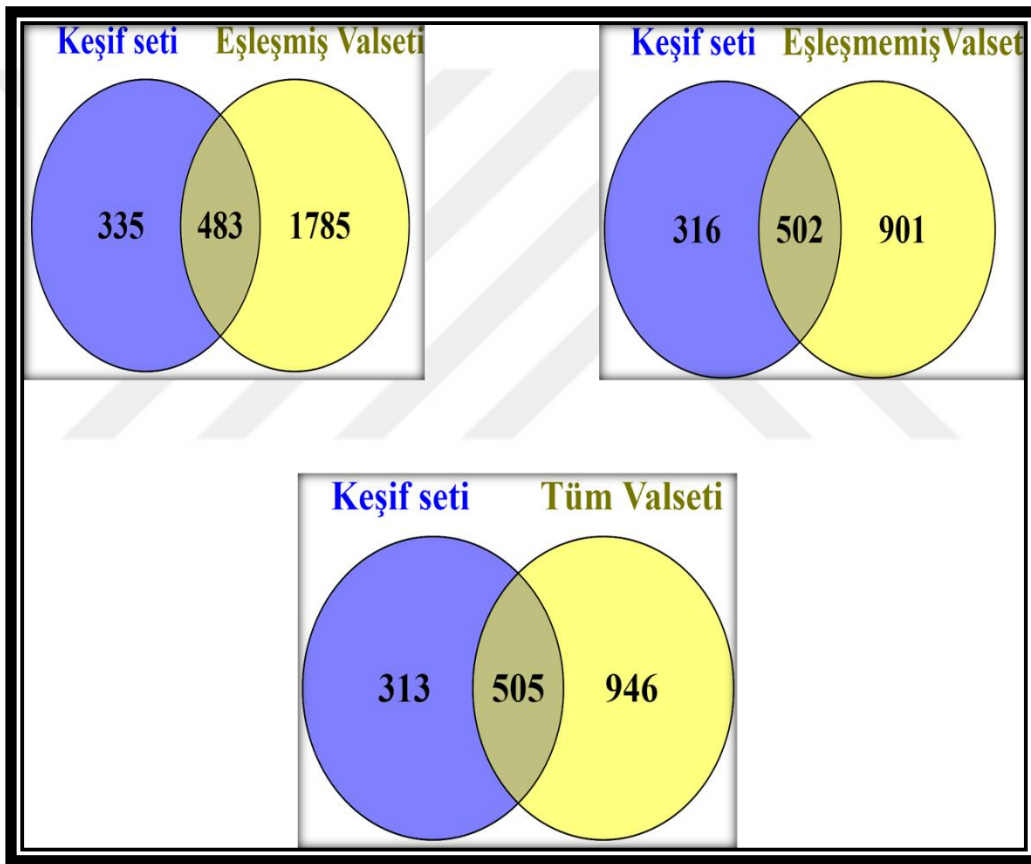
Validasyon setlerin Tümör vs Kontrol UpReg Yolak analizleri



Şekil 5.7. Validasyon setlerin tümör ve normallerin karşılaştırmasından elde edilen ifadesi artan genlere ait yolak analizinin bulguları.

5.1.3. KEŞİF SETİ İLE VALIDASYON SETLERİN KİYAS BULGULARI (ÇAKIŞMA BULGULARI)

Venny yazılımı kullanarak çakışan genleri içeren A (Keşif ve Tüm validasyon seti), B (keşif ve eşleşmiş validasyon seti) ve C (keşif ve eşleşmiş olmayan validasyon seti) listelerinin analizleri sonucu; çakışan A gen listesinin 505 gen içerdiği, yani referans set olarak kullanılan keşif setinde yer alan 851 genin %59'unu ifade etmektedir. 483 gen içeren B gen listesi keşif setinin %57'i valide ederken C gen listesinde ise 502 gen ile keşif seti %59 oranla benzerlik göstermiştir. Çakışma özeti şekil 5.8' te verilmiştir.



Şekil 5.8: Keşif ve Validasyon setlerinin tümör- normal örtüşmesi

DAVID biyoinformatik yazılım ile yapılan analizler sonucunda çakışan A, B ve C gen listelerinde ifadesi azalan genler; sodyum iyonu transmembran taşıması, kadmiyum, karbonik anhidraz (alfa-sınıf, korunmuş sitesi), P-tipi ATPaz, sitoplazmik alan azot ve sülfür oksidasyonu olmak üzere beş ayrı küme ile bağlantısı gözlenmiştir (Şekil 5.9). Ayrıca bu genlerin, metabolik yolak, aldosterone bağlı sodyum reabsorpsiyonu, azot metabolizm, pankreas sekresyonu, mineral absorpsiyonu, kükürt metabolizması, nişasta ve sakaroz metabolizması ve tat iletişim gibi KEGG yolaklarında yer aldığı saptanmıştır (Şekil 5.10).

İfadesini artan çakışan genler ise propeptid: C-terminali propeptidi, interferon-beta tepkisi ve metal iyonu bağlayıcı sitesi (çinko 2; inhibe formda) gibi işlevsel yollara ayrılmıştır (şekil 5.11). Ayrıca ifadesi artan bu genlerin, P13K-Akt-sinyal yolağı, kanser yolakları, fokal adzeyon, ECM-reseptör etkileşimi, kanserde mikroRNA'lar, amoebiasis, trombosit aktivasyonu, Wnt sinyal yolağı, protein sindirimi ve absorpsiyonu, bazal hücre karsinoma ve endometriyal kanser gibi KEGG yolaklarında yer aldığı belirlenmiştir (Şekil 5.12).

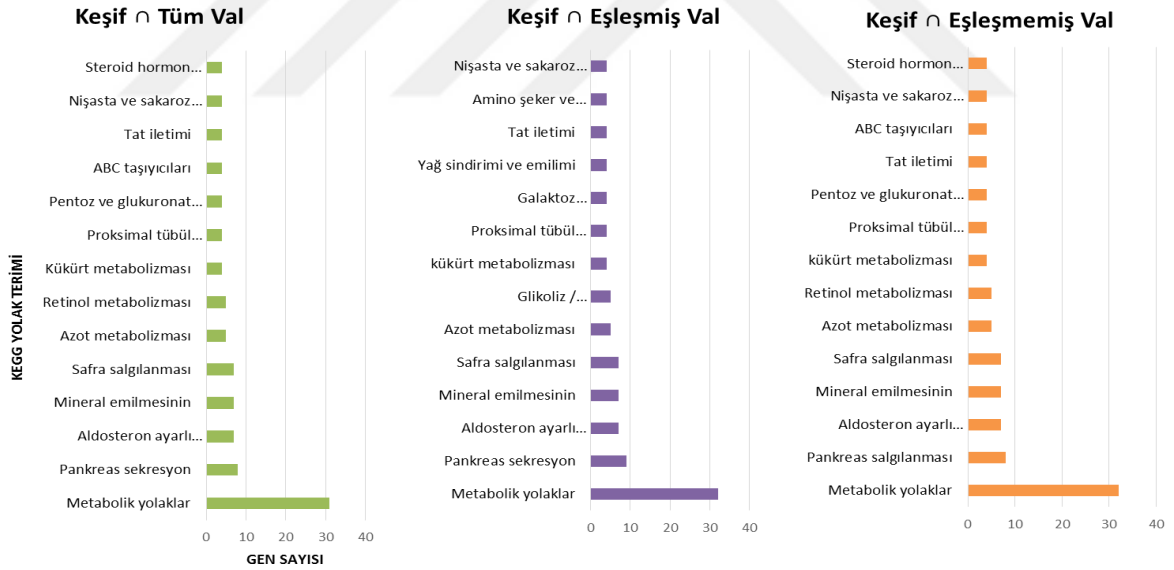
Keşif ve validasyon setlerinde çakışan gen listelerinin DAVID analizinden çıkan sonuçlarına bakıldığında, ifadesinde artış görülen COL11A1, COL12A1, COL1A1, COL1A2, COL27A1, COL5A1, COL5A2, MMP1 ve MMP3 genleri 1,3 veya daha fazla olan zenginleşme skorlarına ait bir fonksiyonel kümede yer alırken, ifadesi azalan genlerin ATP2A3, ATP2B1, CA1, CA12, CA2, CA4, CA7, ETHE1, MT1E, MT1F, MT1G, MT1H, MT1M, MUC2, MUC5B, SCN7A, SCNN1A, SCNN1B, SCNN1G, SLC4A4, SQRDL ve TST olduğu belirlenmiştir (Çizelge 5.1).

Çakışan DownReg genlerin Ontoloji analizleri

Keşif ∩ Tüm Val			Keşif ∩ Eşleşmiş Val			Keşif ∩ Eşleşmemiş Val		
Zenginleşme skoru	Anotasyon Adı	Gen Sayısı	Zenginleşme skoru	Anotasyon adı	Gen sayısı	Zenginleşme skoru	Anotasyon Adı	Gen Sayısı
2,26	Sodyum iyonu transmembran taşıması	7	2,26	Sodyum iyonu transmembran taşıma	7	2,26	Sodyum iyonu transmembran taşıması	7
4,42	Kadmiyum	5	4,23	Kadmiyum	5	4,42	Kadmiyum	5
4,02	Karbonik anhidraz, alfa-sınıf, korunmuş sitesi	5	4,01	Karbonik anhidraz, alfa-sınıf, korunmuş sitesi	5	4,01	Karbonik anhidraz, alfa-sınıf, korunmuş sitesi	5
2,23	B hücresi aktivasyonunun pozitif düzenlemesi	5	1,79	Plexin	4	2,23	B hücresi aktivasyonunun pozitif düzenlemesi	5
1,8	Plexin	4	1,56	P-tipi ATPaz, sitoplazmik alan N	4	1,56	P-tipi ATPaz, sitoplazmik alan N	4
1,57	P-tipi ATPaz, sitoplazmik alan N	4	2,1	Psödohipoaldosteronizm, tip I	3	1,97	P-tipi ATPaz, sitoplazmik alan N	4
1,97	Sülfür oksidasyonu, sülfid kullanılarak : kinon oksidoredüktazı	3	1,97	Sülfür oksidasyonu kullanılarak sülfid: kinon oksidoredüktazı	3	1,97	sülfür oksidasyonu sülfid kullanılarak: kinon oksidoredüktazı	3
1,89	Psödohipoaldosteronizm, tip I	3	1,92	Alan:VWFD 4	3	1,89	Psödohipoaldosteronizm, tip I	3
1,69	Alan: VWFD 4	3						

Şekil 5.9: Keşif set ile tüm, eşleşmiş ve eşleşmemiş validasyon setlerinin karşılaştırılmasında çakışan ve ifadesi azalan gen listesinin ontoloji sonuçları.

Çakışan DownReg genlerin Yolak analizleri



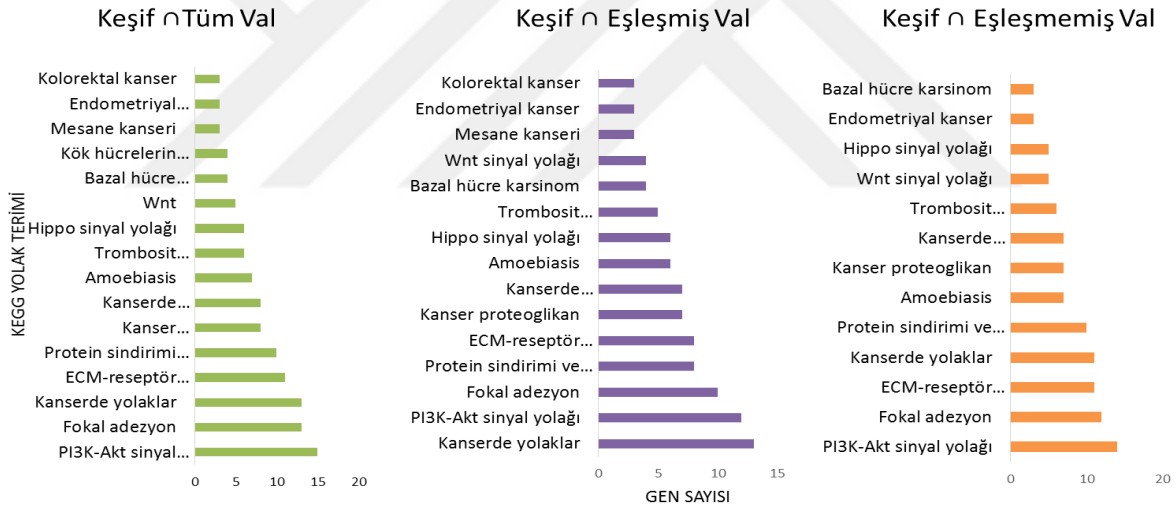
Şekil 5.10: Keşif set ile tüm, eşleşmiş ve eşleşmemiş validasyon setlerinin karşılaştırılmasında çakışan ve ifadesi azalan gen listesinin yolak analizinin sonuçları.

Çakışan Up-regüle genlerin Ontoloji analizi

Keşif ∩ Eşleşmiş Val			Keşif ∩ Eşleşmemiş Val		
Zenginleşme skoru	Anotasyon Adı	Gen sayısı	Zenginleşme skoru	Anotasyon Adı	Gen sayısı
5,43	Propeptid: C-terminali propeptidi	5	7,05	Propeptid: C-terminali propeptidi	6
1,56	Tip I İnterferon sinyal yolağını	4	3,28	alan: TSP N-terminal	5
2,4	interferon-beta tepkisi	3	2,79	alan: VWFC	4
1,86	Metal iyonu bağlayıcı sitesi: Çinko 2; inhibe formda	3	1,93	alan: NTR	3
1,73	Alan: VWFC	3	1,91	Metal iyon-bağlayıcı sitesi: Çinko 2; inhibe formda	3
1,45	Dönüşen büyüme faktör-beta-İlgili	3	1,6	interferon-beta tepkisi	3
1,36	alan: TSP N-terminal	3	1,49	alan: TSP tip-1 3	3

Şekil 5.11: Keşif set ile eşleşmiş ve eşleşmemiş validasyon setlerinin karşılaştırılmasında çakışan ve ifadesi artan gen listesinin yolak analizinin sonuçları.

Çakışan UpReg genlerin Yolak analizi



Şekil 5.12: Keşif set ile tüm, eşleşmiş ve eşleşmemiş validasyon setlerinin karşılaştırılmasında çakışan ve ifadesi artan gen listesinin yolak analizinin sonuçları.

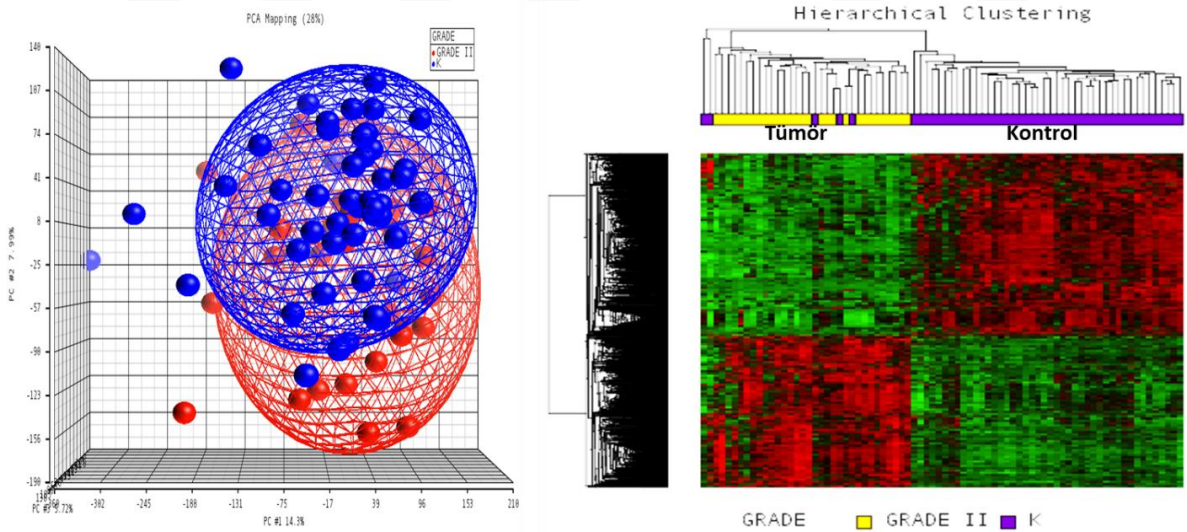
Çizelge 5.1: Tümör ve normallerin kıyaslamasının özeti

ÇAKIŞMA		GEN SAYISI	ÖNEMLİ GEN SAYISI (Z-Skor \geq Yolak)	ÖNEMLİ GENLERİ
Keşif & Tüm Val seti	UpReg	157	-	
	DownReg	348	26	ATP2A3, ATP2B1, CA1, CA12, CA2, CA4, CA7, ETHE1, MT1E, MT1F, MT1G, MT1H, MT1M, MUC2, MUC5B, PLXNA2, SCN7A, SCNN1A, SCNN1B, SCNN1G, SEMA4G, SEMA6A, SEMA6D, SLC4A4, SQRDL, TST
Keşif & Eşli Val seti	UpReg	140	10	BMP4, COL11A1, COL12A1, COL1A1, COL1A2, COL27A1, COL5A2, INHBA, MMP1, MMP3
	DownReg	341	26	ATP2A3, ATP2B1, CA1, CA12, CA2, CA4, CA7, ETHE1, MT1E, MT1F, MT1G, MT1H, MT1M, MUC2, MUC5B, PLXNA2, SCN7A, SCNN1A, SCNN1B, SCNN1G, SEMA4G, SEMA6A, SEMA6D, SLC4A4, SQRDL, TST
Keşif & Eşli olmayan val seti	UpReg	153	13	COL11A1, COL12A1, COL1A1, COL1A2, COL27A1, COL5A1, COL5A2, MMP1, MMP3, SFRP4, THBS2, TIMP1, TIMP3
	DownReg	349	20	ATP2A3, ATP2B1, CA1, CA12, CA2, CA4, CA7, ETHE1, MT1E, MT1F, MT1G, MT1H, MT1M, SCN7A, SCNN1A, SCNN1B, SCNN1G, SLC4A4, SQRDL, TST

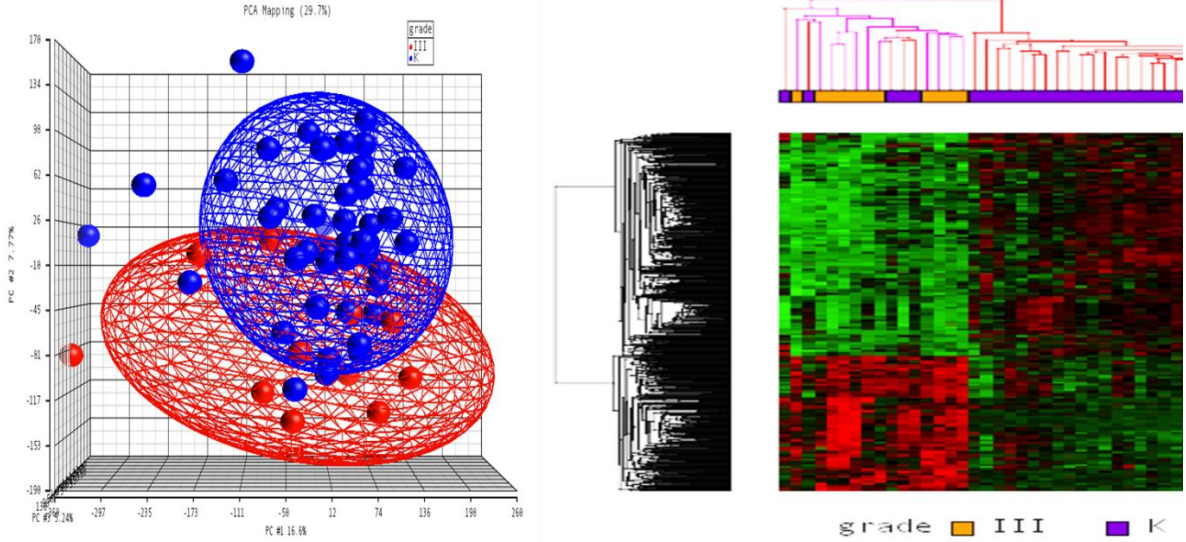
5.2. GRADE BULGULARI

5.2.1. KEŞİF SETİN GRADE BULGULARI (GRADE VE KONTROL KARŞILAŞTIRMALARI)

Keşif çalışmasının 2 kat değişimi ve $p < 0.001$ istatistiksel anlamlılıkta Grade 2 ve Grade 3 örnekleri arasındaki karşılaştırmasında, hiçbir genin farklı olarak ifade edilmediği görüldüğünden, sadece 40 farklı evreye ait olan hasta grubu, 40 eşli normal örnekler ile karşılaştırılmıştır. Ayrıca, grade ve normallerin kıyasında grade 1'e ait örneklerinin sayısı çok yetersiz olması sebebiyle bu grade ile ilgili karşılaştırmadan hiç bir sonuç elde edilememiştir. Grade 2 örnekleri, normalleri ile kıyasladığında, 398 gen ifade artışı gösterirken, 472 genin ifadesinde azalma olduğu görülmüştür ($p < 0.001$). Grade 3 ve normallerinin karşılaştırılmasında ise, 133 ifade artışı gösteren gen ve 312 ifadesinde azalma gösteren gen belirlenmiştir. Grade2 - normal ve Grade 3 - normal kıyaslarından elde edilen gen listelerinin kümeleme analizleri sonucunda, örnekler ile gen ifade profilleri ilişkisine göre, tümör-normal ve ifadeleri azalan/artan genler ayrılmıştır (Şekil 5.13, 5.14).



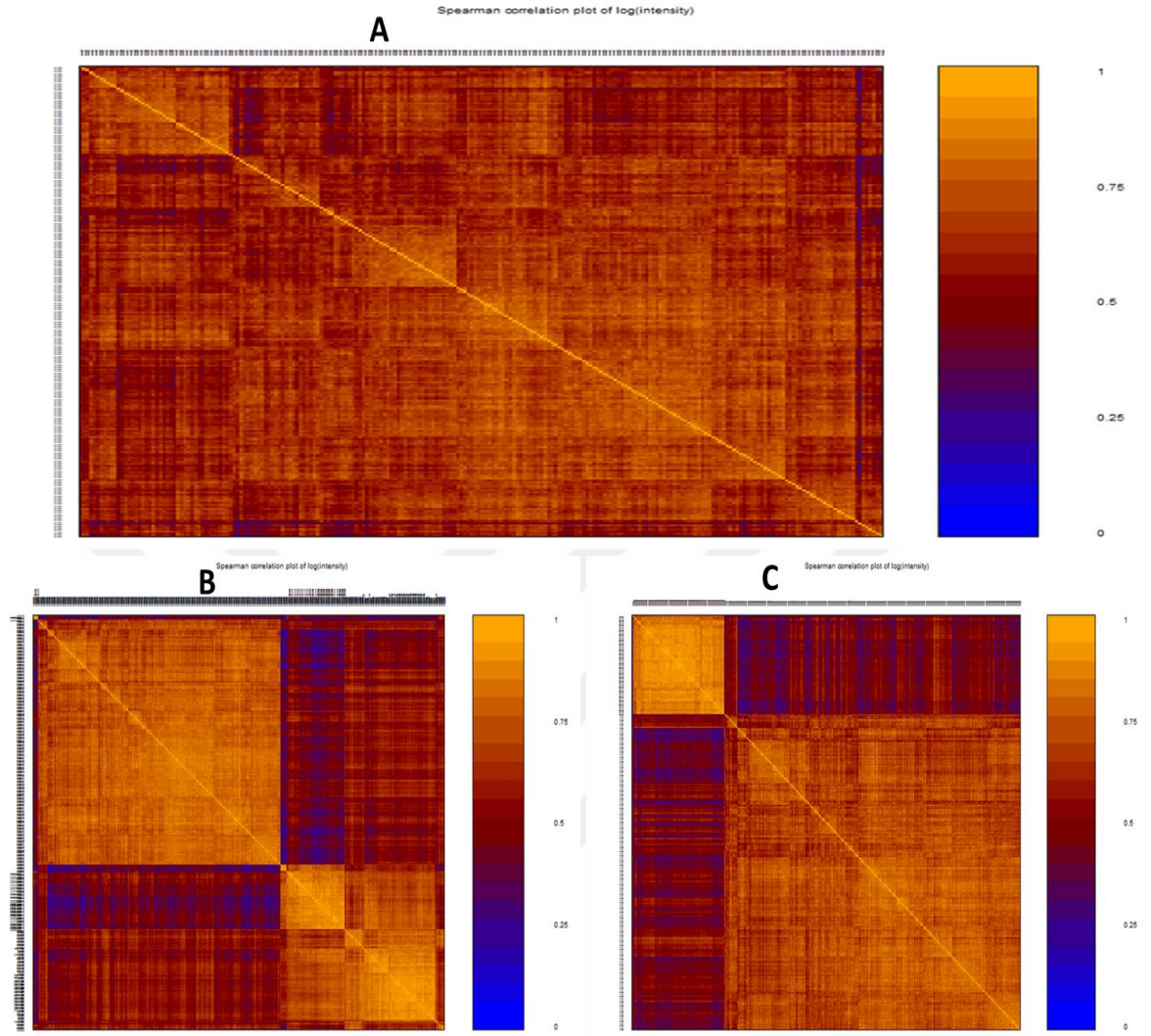
Şekil 5.13: Keşif setinin grad 2 vs normallerinin PCA (solda) ve heatmap (sağda) sonuçları. PCA'nın mavısı normalleri, kırmızısı ise tümörlerini temsil etmektedir



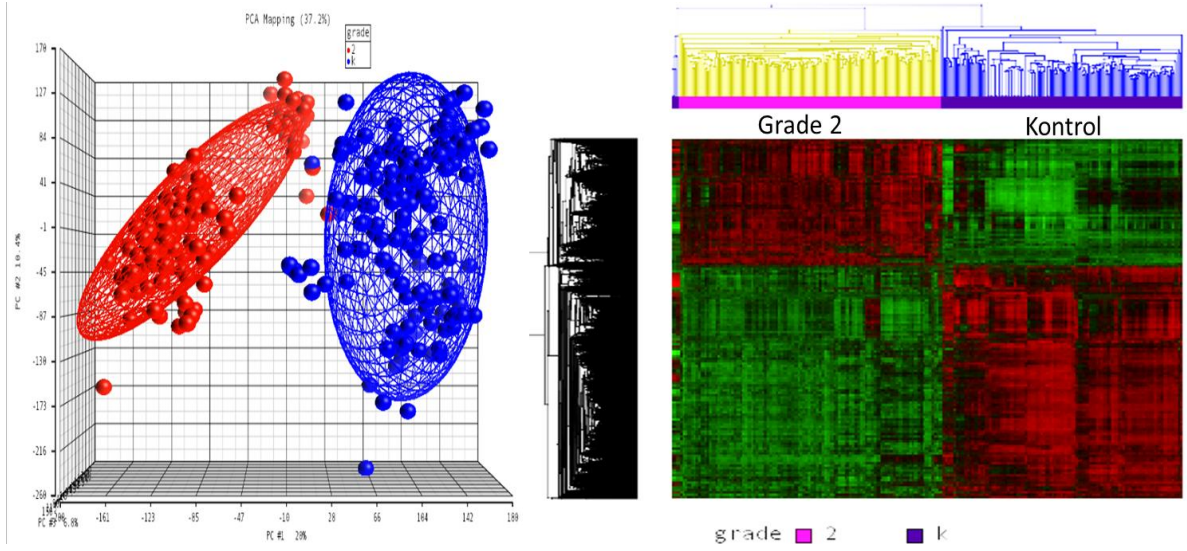
Şekil 5.14: Keşif setinin grad 2 vs normallerinin PCA (solda) ve heatmap (sağda) sonuçları. PCA'nın mavısı normalleri, kırmızısı ise tümörlerini temsil etmektedir

5.2.2. VALİDASYON SETİNİN GRADE BULGULARI (GRADE KONTROL KARŞILAŞTIRMALARI)

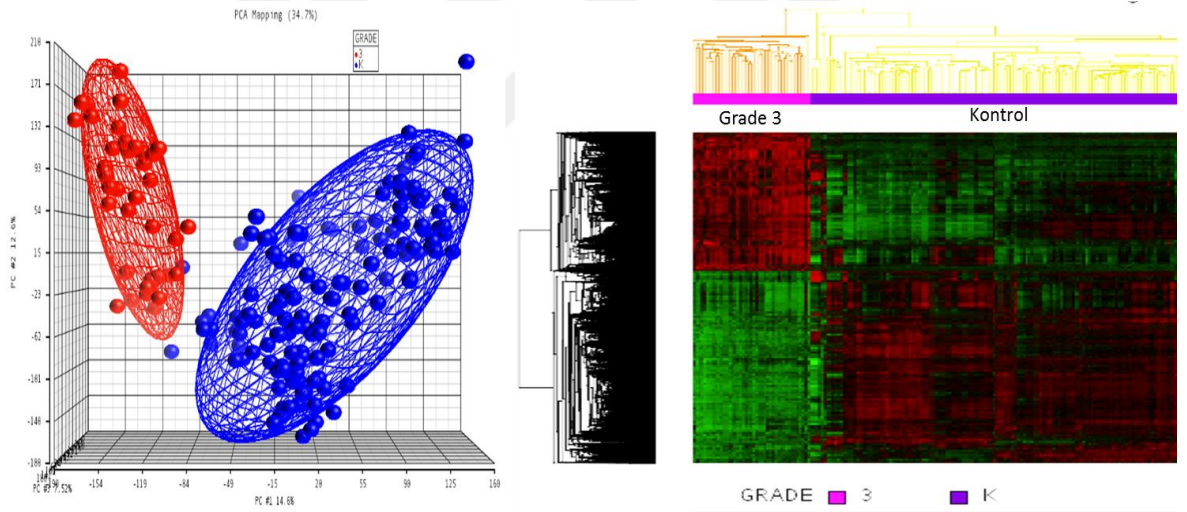
Keşif setinde Grade 1 örnekler yetersiz olması nedeniyle Validasyon setindeki 18 Grade 1 örneği de normaller ile kıyaslanmamıştır. Validasyon setinin oluşturduğu grade ve kontrol örnekleri tekniksel olarak GSE18088 veri seti ile GSE17538 veri seti arasında yaklaşık %75-85 oranında uyum tespit edilirken bu veri setlerinden alınan grade örnekleri ile 160 normal örneğin bütünüyle kıyaslandığında, uyumluluk çok yüksek olmamakla birlikte yalnızca eşleşmiş olmayan normaller (n=96) ile daha fazla uyum göstermiştir. (Şekil 5.15). Grade 2 (n=169) ve 160 örneğe sahip olan normallerin kıyaslanmasında 2 kat değişim eşiğinde artan gen sayısı 880 olup azalış gösteren gen sayısı ise 1413 olarak çıkmıştır. Tümör Grad 3 (n=43) örneklerinin normaller (n=160) ile kıyaslanmasında 1221 artış gösteren ve 1462 azalış gösteren farklı olarak ifade edilen genler tespit edilmiştir ($p < 0.001$ ve kat değişimi=2). PCA ve Hiyerarşik kümeleme analizlerinde, validasyon setindeki Grade 2 ve 3 örnekler normallerden ayrılmıştır (Şekil 5.16, 5.17).



Şekil 5.15: Validasyon setinin grade ve normal örneklerin Spearman korelasyon çizimleri. A grade validasyon alt grubunu oluşturan veri setlerinin korelasyonunu göstermektedir. B gradeler ve tüm normalleri gösterirken, C gradeler ve yalnız tüm validasyon setinin eşleşmiş olmayan normallerini göstermektedir.



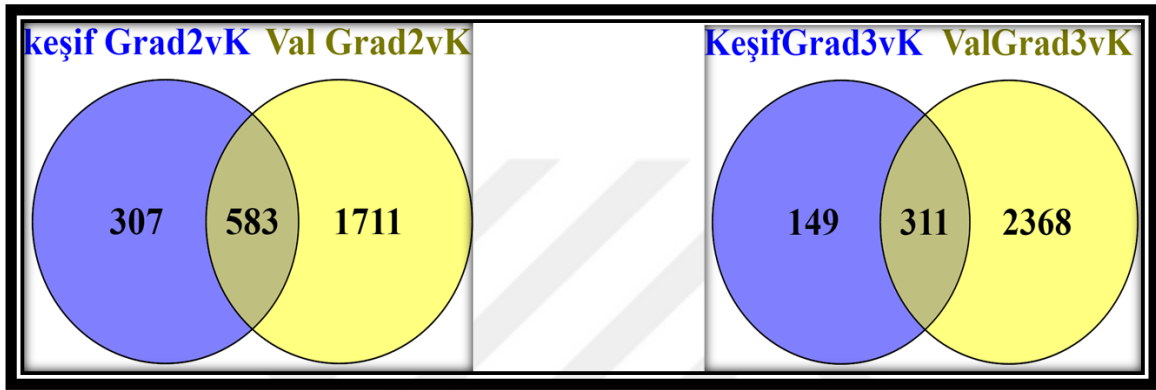
Şekil 5.16: Validasyon setinin grade 2 vs normallerinin PCA (solda) ve heatmap (sağda) sonuçları. PCA'nın mavisi normalleri, kırmızısı ise tümörlerini temsil etmektedir.



Şekil 5.17: Validasyon setinin grade 3 vs normallerinin PCA (solda) ve heatmap (sağda) sonuçları. PCA'nın mavisi normalleri, kırmızısı ise tümörlerini temsil etmektedir.

5.2.3. KEŞİF VE VALIDASYON SETLERİNDEKİ GRADELER ARASINDAKİ ÇAKIŞAN GENLER

Keşif ve validasyon setindeki grade 2 tümör örnekleri ile normalleri arasındaki karşılaştırma sonucu, üstü üste geldiği belirlenen 583 gen keşif setindeki tüm genlerin % 66'sını oluşturmaktadır. Grade 3 tümör örnekleri ve normallerinden ortaya çıkan 311 çakışan gen ise keşif setinin % 68'ini sembolize etmektedir. Venny yazılımından elde edilen bu sonuçlar aşağıda verilmektedir. Çakışma özeti Şekil 5.18' da verilmiştir.



Şekil 5.18: Keşif ve Validasyon setinin grade örtüşmesi

DAVID analizden sonra, keşif ve validasyon setleri birleştirilerek oluşturulan grade 2 tümör ile normal karşılaştırmasında, çakışan listesinde yer alan ve ifadesi artan genleri 13 fonksiyonel küme ile bağlantısı kurulurken (Şekil 5.21), keşif set ve grade 3 karşılaştırması sonucunda, ifadesi artan genler ise herhangi 1.3 veya daha fazla zenginleşme skoruna sahip olan bir işlevsel küme içerisinde yer almamıştır. Ayrıca, keşif seti ile grade 2 ve 3 ün karşılaştırılma analizleri sonucu ifadesi azalan genler, karbonik anhidraz (alfa-sınıf), P-tipi ATPaz sitoplazmik alan ve semaforin olmak üzere üç ortak fonksiyonel küme içinde yer almıştır (Şekil 5.19).

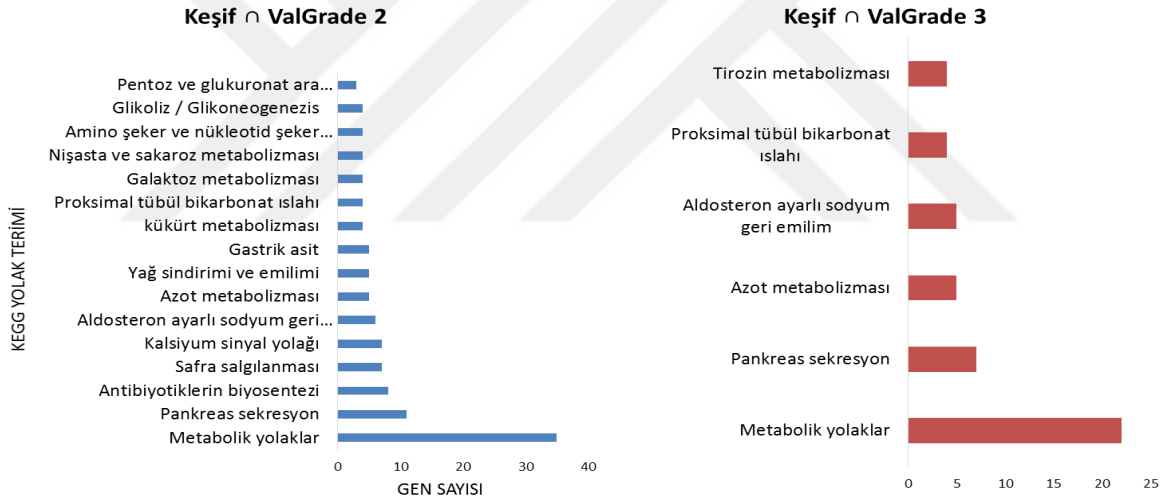
Her iki setin grade ile normal karşılaştırmaları sonucu, çakışık ifadesi artan genlerinin KEGG yolak sonuçlarına göre, keşif - grade 2 için 24 yolak belirlenirken, keşif – grade 3 için 13 yolağın mevcut olduğu görülmüştür (Şekil 5.22). Çakışık ifadesi azalan genlerinin yolak bulguları ise keşif – grade 2 için 16 yolak saptanırken, keşif – grade 3 için 6 yolak belirlenmiştir (Şekil 5.20).

Çakışan DownReg genlerin Ontoloji bulguları

Keşif \cap ValGrade 2			Keşif \cap ValGrade 3		
Zenginleşme skoru	Anotasyon Adı	Gen Sayısı	Zenginleşme skoru	Anotasyon Adı	Gen Sayısı
2,75	Karbonik anhidraz, alfa-sınıf	4	4,82	Karbonik anhidraz, alfa-sınıf, korunmuş sitesi	5
2,08	P-tipi ATPaz, sitoplazmik alan	4	3,15	P-tipi ATPaz, sitoplazmik alan N	5
1,59	Semaforin	3	2,1	Alan: VWFA	5
1,57	Galectin, karbonhidrat tanıma al	3	1,97	Çinko parmak bölgesi: NR C4 tipi	4
			1,95	Fosfolipid taşıyan P-tipi ATPaz, alt familyası IV	3
			1,37	Semaforin	3

Şekil 5.19: Grade2vK ve grade3vK için keşif ve validasyon setlerin arasındaki çakışan ve ifadesi azalan gen listesinin ontoloji sonuçları.

Çakışan DownReg genlerin Yolak analiz bulguları

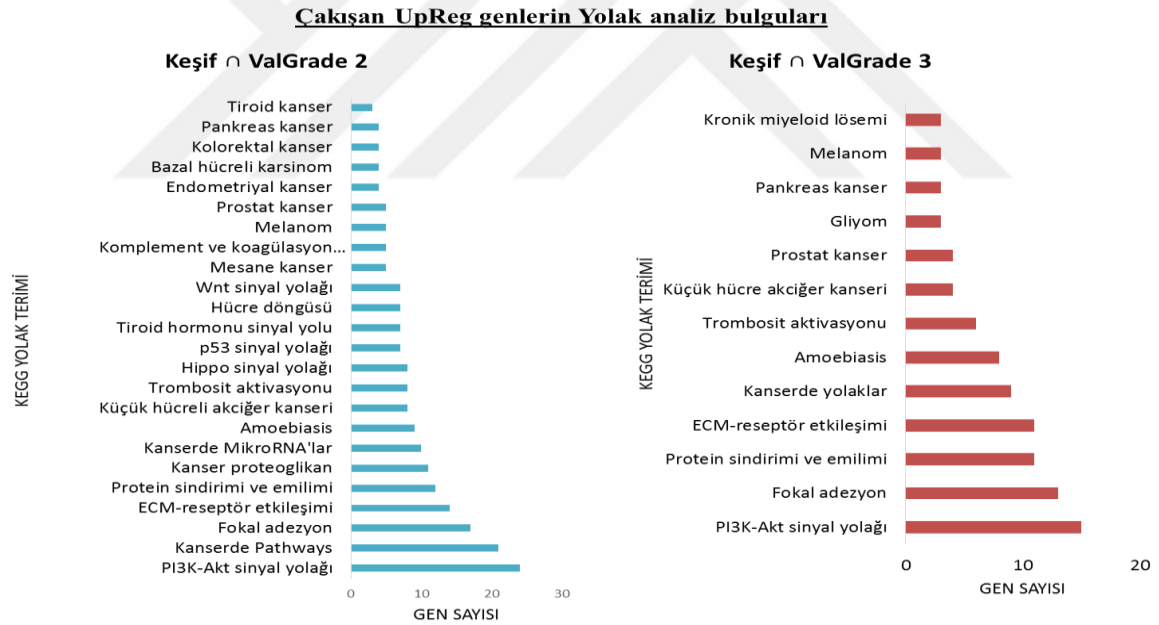


Şekil 5.20: Grade2vK ve grade3vK için keşif ve validasyon setlerin arasındaki çakışan ve ifadesi azalan gen listesinin yolak analiz sonuçları.

Çakışan Up-regüle genlerin Gen Ontoloji Analiz bulguları (Keşif \cap ValGrade 2)

Zenginleşme skoru	Anotasyon Adı	Gen sayısı
1,48	Tekrarlanma: LRR 6	8
1,78	Lösün açısından zengin tekrar içeren N-terminali	7
7,1	Propeptid: C-terminali propeptidi	6
2	Tekrarlanma: LRR 12	6
3,17	domain (alan): TSP N-terminal	5
2,53	domain: Kazal gibi	5
3	domain: tiroglobulin tip-1	4
2,72	Metal iyon-bağlayıcı sitesi: Çinko 2; inhibe formda	4
2,72	domain: IGFBP N-terminal	4
2,59	domain: NTR	4
1,93	domain: EGF-benzeri 17; Kalsiyum bağlanması	3
1,77	İnterferon-alfa tepkisi	3
1,45	Metal iyonu-bağlayıcı sitesine: Kalsiyum 4; karbonil oksijen üzerinden	3

Şekil 5.21: Grade2vK için keşif ve validasyon setlerin arasındaki çakışan ve ifadesi artan gen listesinin ontoloji sonuçları.

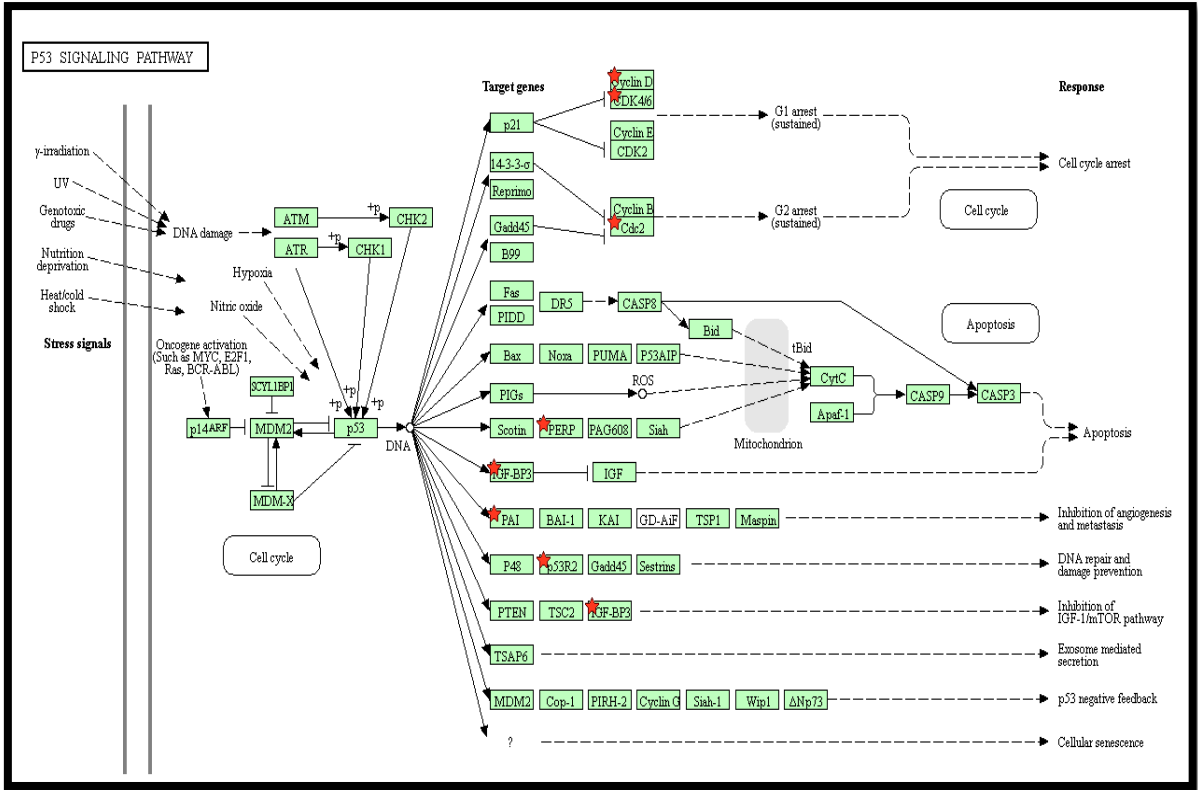


Şekil 5.22: Grade2vK ve grade3vK için keşif ve validasyon setlerin arasındaki çakışan ve ifadesi artan gen listesinin yolak analiz sonuçları

Çizelge 5.2: Grade kıyaslamasının özeti

ÇAKIŞMA		GEN SAYISI	ÖNEMLİ GEN SAYISI (Z-skoru \cap Yolak)	GENLER
Keşif & Val set Grade2vK	UpReg	338	16	COL11A1, COL12A1, COL15A1, COL1A1, COL1A2, COL3A1, COL5A1, COL6A3, IGFBP3, LUM, MMP1, MMP2, MMP3, NOTCH3, SFRP4, THBS2
	DownReg	198	9	ATP2A3, ATP2B1, CA1, CA2, CA4, CA7, SEMA4G, SEMA6A, SEMA6D
Keşif & Val seti Grade3vK	UpReg	248	0	
	DownReg	113	16	ATP2A3, ATP2B1, CA1, CA12, CA2, CA4, CA7, CLCA4, NR3C1, NR3C2, NR5A2, PPARD, SEMA4G, SEMA6A, SEMA6D

Bu tez çalışmasında göz önüne alınan kat değişimi, yüksek zenginleşme skoru ve yolak üzerinde yer alması gibi üç kriteri yerine getiren Grade 2 çakışan gen listesinde hiçbir gen bulunmazken Grade 3’de CLCA1, CLCA4, NR3C1, NR3C2, NR5A2 ve PPARD genlerinin çakışan gen listesinde yer aldıkları görülmüştür (Çizelge 5.2). Öte yandan, sadece Grade 2’de tespit edilen p53 sinyal yolağı, Wnt sinyal yolağı, kalsiyum sinyal yolağı ve hücre döngüsü yolaklarından p53 sinyal yolak şeması örnek olarak sunulmuştur (şekil 5.23).

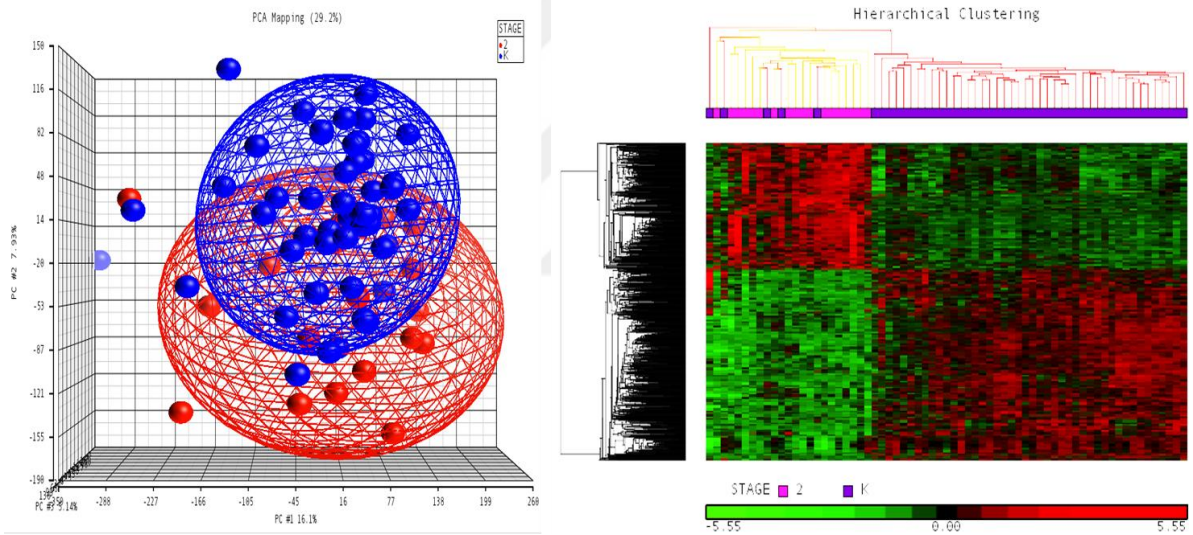


Şekil 5.23: p53 sinyal yoluğđ üzerindeki FİG'ler

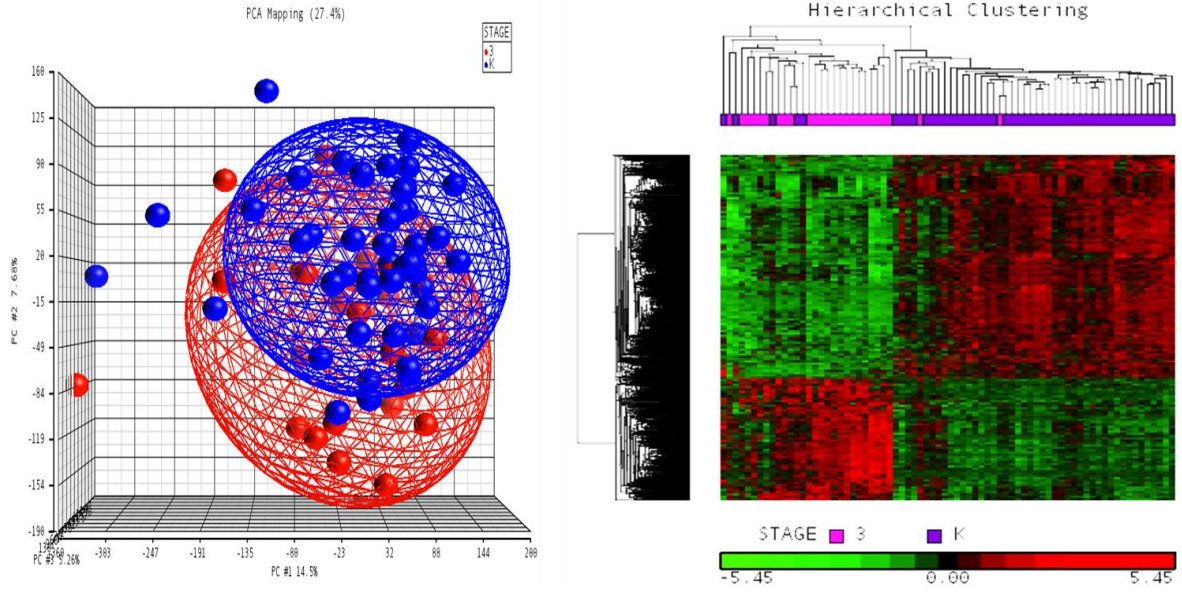
5.3. EVRE BULGULARI

5.3.1. KEŞİF SETİNİN EVRE BULGULARI (EVRE VE KONTROL KARŞILAŞTIRMALARI)

Keşif çalışmanın 2 kat değişimi ve $p < 0.001$ istatistiksel anlamlılıkta Evre 2 ve Evre 3 örnekleri arasındaki karşılaştırmada hiçbir genin farklı olarak ifade edilmemesi sebebiyle sadece 44 farklı evreye ait olan hasta grubu ile 44 eşli normal örnekler karşılaştırılmıştır. Evre 2 örnekleri, normalleri ile kıyaslandığında, 215 gende ifade artışı gözlenirken, 374 genin ifade azalışı tespit edilmiştir ($p < 0.001$). Evre 3 ve normallerinin karşılaştırılmasında ise, 264 genin ifadesinde artış gözlenirken, 544 genin ifadesinde azalma olduğu belirlenmiştir. Kümeleme analizleri sonucunda tümör Evre 3 örnekleri ve tümör evre 3 normal örneklerinden belirgin bir şekilde ayrılmıştır (Şekil 5.24, 5.25).



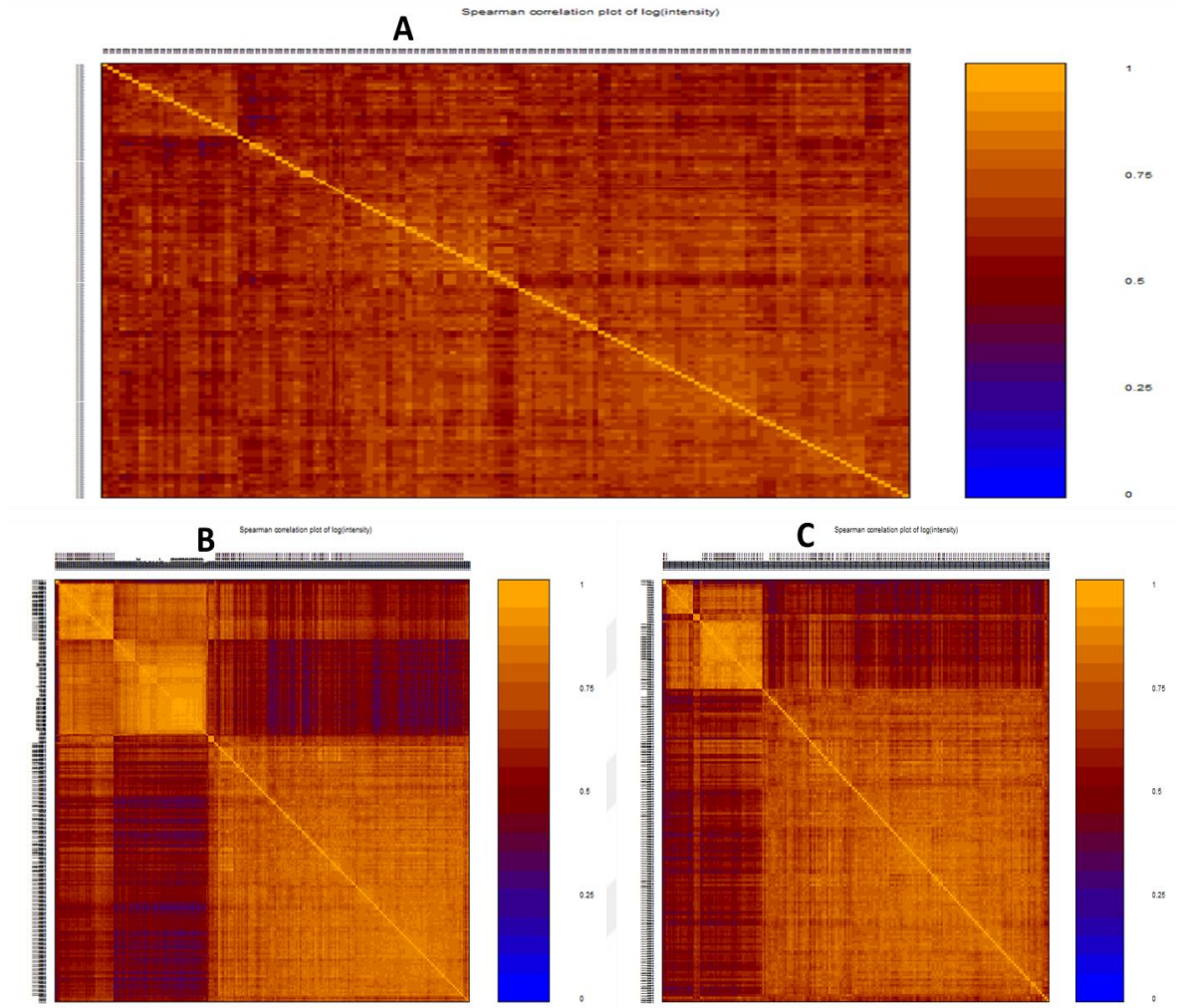
Şekil 5.24: Keşif setinin evre 2 vs normallerinin PCA (solda) ve heatmap (sağda) sonuçları. PCA'nın mavisi normalleri, kırmızısı ise tümörlerini temsil etmektedir.



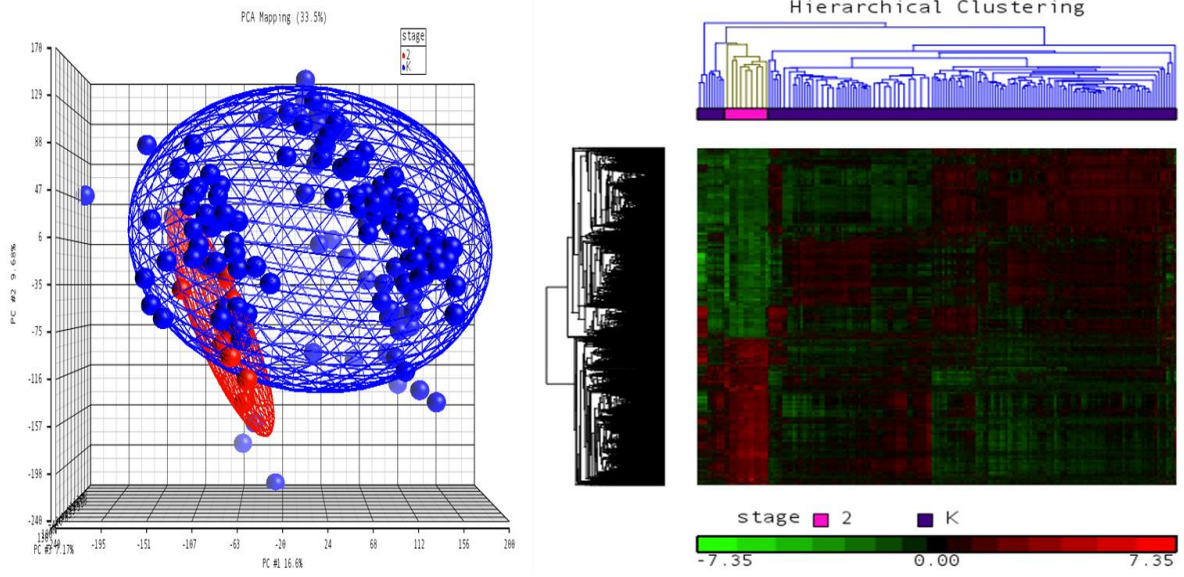
Şekil 5.25: Keşif setinin evre 3 vs normallerinin PCA (solda) ve heatmap (sağda) sonuçları. PCA'nın mavısı normalleri, kırmızısı ise tümörlerini temsil etmektedir.

5.3.2. VALİDASYON SETİN EVRE BULGULARI (EVRE VE KONTROL KARŞILAŞTIRMALARI)

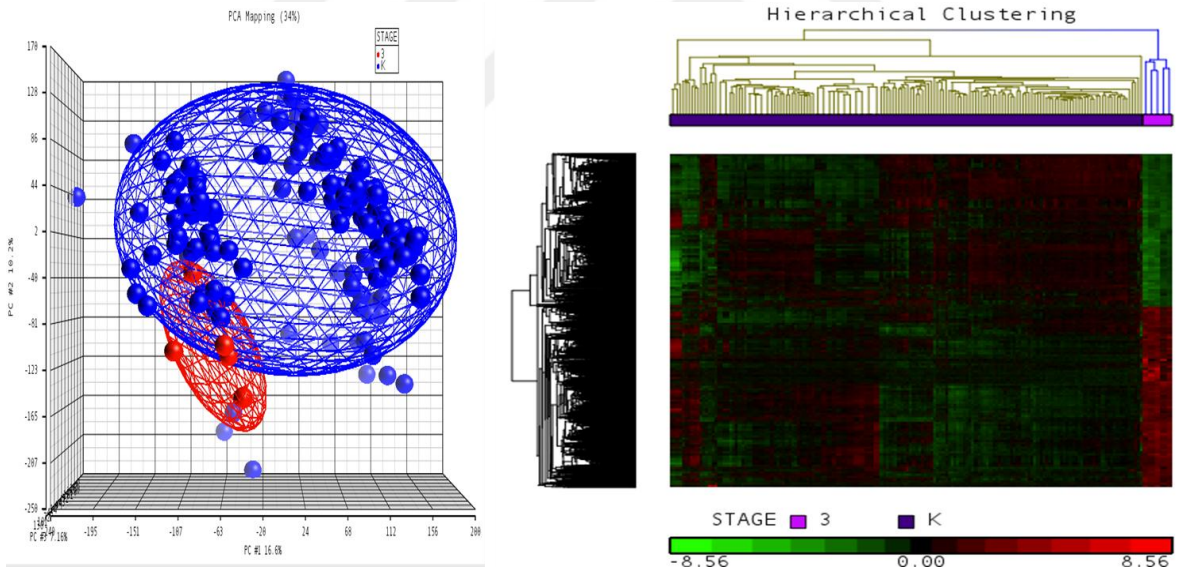
Validasyon setinde 33 tane tümör evre 1 örneği olmasına rağmen, bu örnekler analiz sürecine dahil edilmemiştir. Bunun sebebi, keşif setinde evre 1 örneklerinin yetersiz olmasıdır. Validasyon setinin oluşturduğu evre ve kontrol örnekleri tekniksel olarak GSE21510, GSE32323 ve GSE27854 veri setleri arasında yaklaşık %75-85 oranında uyum tespit edilirken bu veri setlerinden alınan evre örnekleri ile 160 normal örneğin kıyaslandığında, uyumluluk çok yüksek olmamakla birlikte eşleşmiş normaller ile daha fazla uyum göstermiştir (Şekil 5.26). Evre 2'nin 94 örneği 160 normal örnek ile kıyaslandığında, 1170 genin ifadesinde artış görülürken, 1376 genin ifadesinde azalma olduğu tespit edilmiştir ($p < 0.001$). Evre 3 ($n=79$) ve 160 normal örneğin karşılaştırılmasında ise, 1479 artış gösteren ve 1164 azalış gösteren farklı gen belirlenmiştir. Kümeleme analizleri sonucunda evre 2 tümör örnekleri ve evre 3 tümör örneklerinin kendi normalleri ile karşılaştırılması sonucunda gen ifade seviyelerinde net bir ayırım olduğu görülmüştür. Ancak evre 2 ve evre 3'ün normal örneklerin kendi içlerinde tam olarak kesin ayrımlar görülmemiştir (Şekil 5.27, 5.28).



Şekil 5.26: Validasyon setinin evre ve normal örneklerin Spearman korelasyon çizimleri. A evre validasyon alt grubunu oluşturan verisetlerinin korelasyonunu göstermektedir. B evreler ve tüm normalleri gösterirken C evreler ve yalnız tüm validasyon setinin eşleşmiş normallerini göstermektedir.



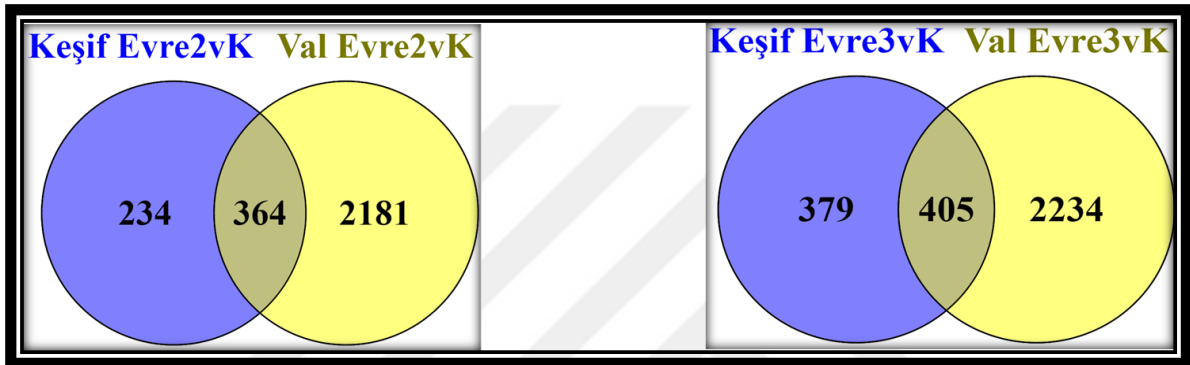
Şekil 5.27: Validasyon setinin evre 2 vs normallerinin PCA (solda) ve heatmap (sağda) sonuçları. PCA'nın mavısı normalleri, kırmızısı ise tümörlerini temsil etmektedir.



Şekil 5.28: Validasyon setinin evre 3 vs normallerinin PCA (solda) ve heatmap (sağda) sonuçları. PCA'nın mavısı normalleri, kırmızısı ise tümörlerini temsil etmektedir.

5.3.3. KEŞİF VE VALİDASYON SETLERİNDEKİ EVRELER ARASINDAKİ ÇAKIŞAN GENLER

Keşif ve validasyon setindeki grade 2 tümör örneklerinin normalleri ile karşılaştırmasında 583 genin üstü üste geldiği 583 olduğu ve bu genler keşif setindeki tüm genlerin % 66'sını oluşturduğu tespit edilmiştir. Grade 3 tümör örnekleri ve normallerinin karşılaştırılması sonucu ortaya çıkan 311 çakışan gen ise keşif setinin %68'ini temsil etmektedir. Venny yazılımından sonuçlar ve bu sonuçları özetleyen Şekil 5.29 aşağıda verilmektedir.



Şekil 5.29: Keşif ve Validasyon setlerinin Evre örtüşmesi

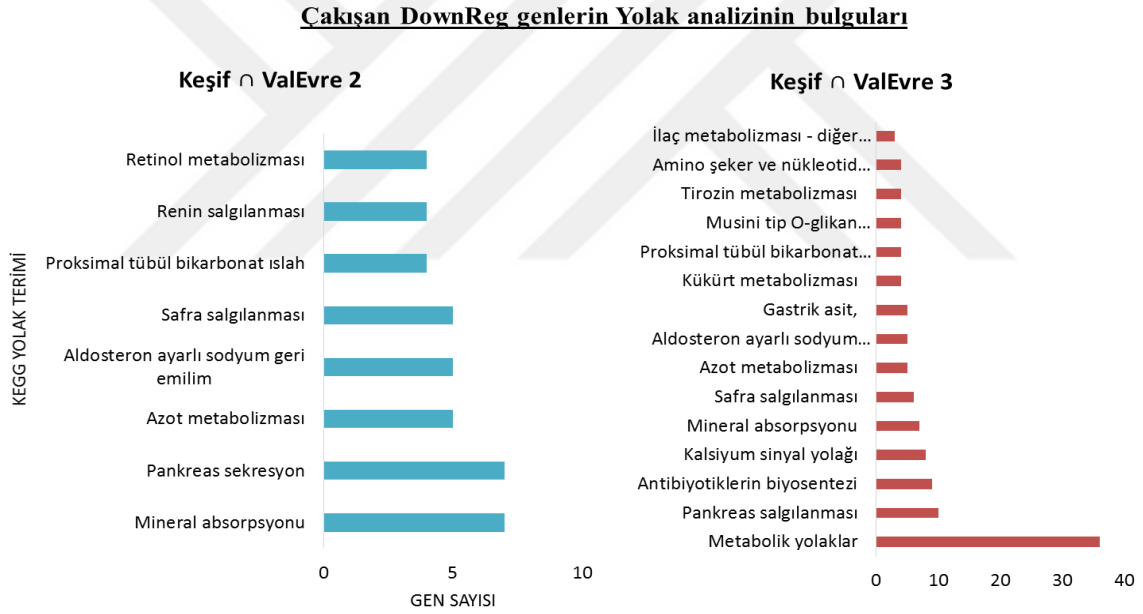
DAVID sonrasında, iki setin evre ile normal örnekleri karşılaştırıldığında, çakışıp ifadesi artan genlerinin fonksiyonel kümeleme sonuçlarına göre; keşif-evre 2 ve keşif-evre 3 karşılaştırmaları için 1,3 veya fazla zenginleşme skoruna sahip olan, kollajen üçlü sarmal tekrarlanması ve propeptid N-terminal propeptid olmak üzere iki fonksiyonel kümeye ayrıldığı belirlenmiştir (Şekil 5.32). Ayrıca, keşif-evre 2 ve 3 de yer alan ve ifadesinde azalma görülen genler, karbonik anhidraz (alfa-sınıf) ve P-tipi ATPaz sitoplazmik alan fonksiyonel kümelere ayrıldığı tespit edilmiştir (Şekil 5.30).

İki setin evre ile normal karşılaştırmalarında çakışan ve ifadesi artan genlerinin KEGG yolak sonuçlarına göre, keşif-evre 2 için 24 görülürken, keşif-evre 3 için ise 13 yolakta yer aldığı belirlenmiştir (Şekil 5.33). Çakışıp ifadesi azalan genlerinin yolak bulgularına bakıldığında keşif-grade 2 için 16 yolak, keşif-grade 3 için 6 yolak bulunmuştur (Şekil 5.31).

Çakışan DownReg genlerin Ontoloji bulguları

Keşif \cap ValEvre 2			Keşif \cap ValEvre 3		
Zenginleşme skoru	Anotasyon Adı	Gen sayısı	Zenginleşme skoru	Anotasyon Adı	Gen sayısı
4,85	Kadmiyum	5	4,46	Kadmiyum	5
4,41	Karbonik anhidraz, alfa-sınıf, korunmuş sitesi	5	4,04	Karbonik anhidraz, alfa-sınıf, korunmuş sitesi	5
1,45	3'5'-siklik nükleotid fosfodiesteraz korunmuş alanı	3	1,59	P-tipi ATPaz, sitoplazmik alan N	4
			1,54	B hücresi aktivasyonunun pozitif düzenlemesi	4
			1,71	alan: VWFD 4	3

Şekil 5.30: Evre2vK ve evre3vK için keşif ve validasyon setlerin arasındaki çakışan ve ifadesi azalan gen listesinin ontoloji sonuçları.

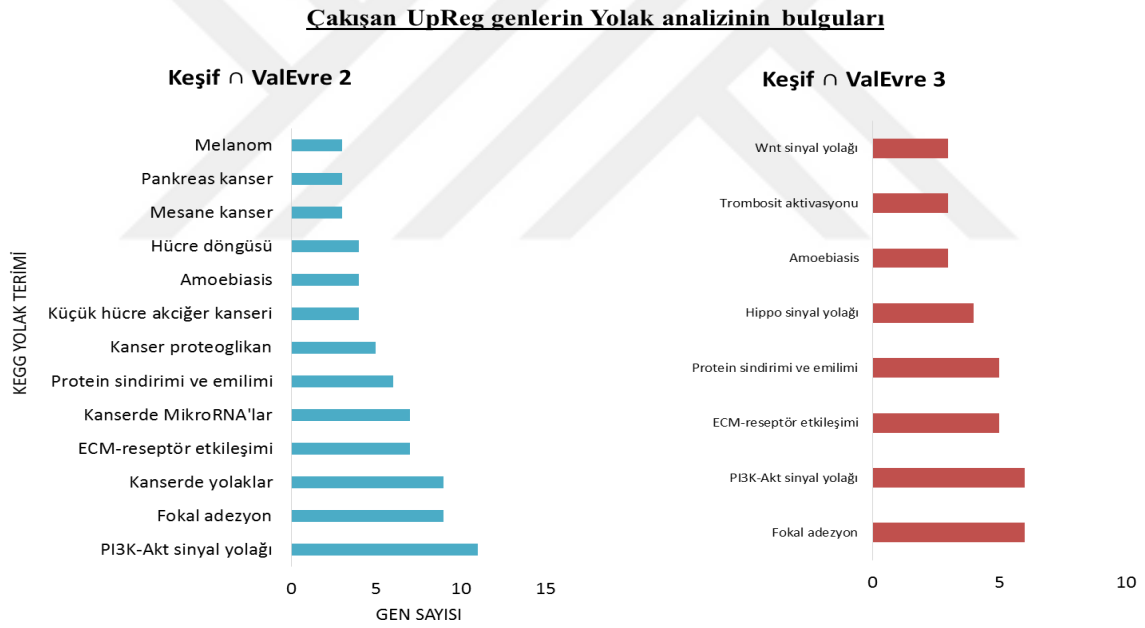


Şekil 5.31: Evre2vK ve evre3vK için keşif ve validasyon setlerin arasındaki çakışan ve ifadesi azalan gen listesinin yolak analizinin sonuçları

Çakışan UpReg genlerin Yolak analizinin bulguları

Keşif ∩ ValEvre 2			Keşif ∩ ValEvre 3		
Zenginleşme skoru	Anotasyon Adı	Gen sayısı	Zenginleşme skoru	Anotasyon Adı	Gen sayısı
4,67	Kollajen üçlü sarmal tekrarlanması	7	3,99	Kollajen üçlü sarmal tekrarlanması	6
2,63	Alan: TSP N-terminal	4	2,4	Propeptid N-terminal propeptid	3
2,39	Propeptid N-terminal propeptid	3	2,08	alan: VWFC	3
			1,68	alan: TSP N-terminal	3

Şekil 5.32: Evre2vK ve evre3vK için keşif ve validasyon setlerin arasındaki çakışan ve ifadesi azalan gen listesinin ontoloji sonuçları.



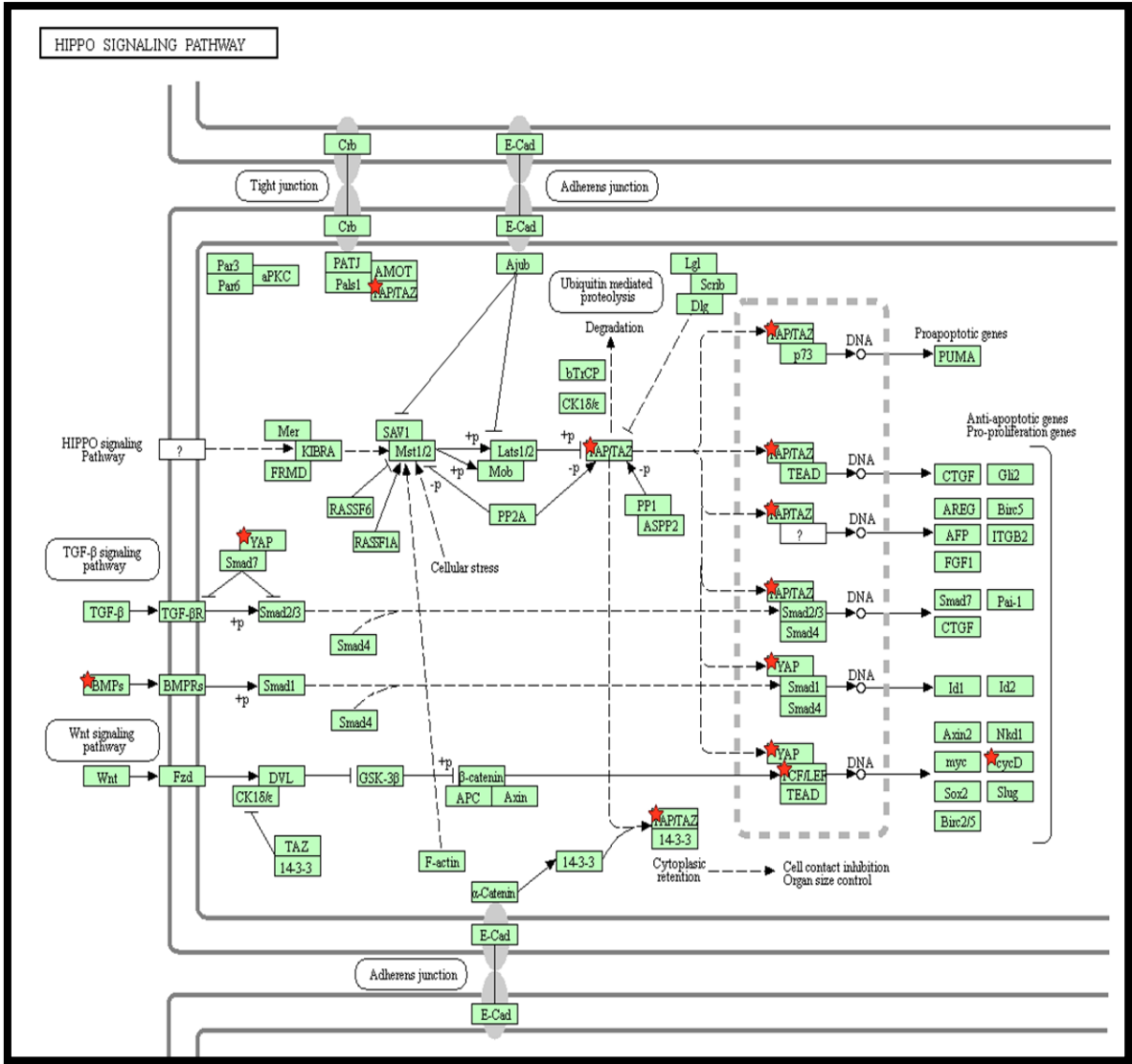
Şekil 5.33: Evre2vK ve evre3vK için keşif ve validasyon setlerin arasındaki çakışan ve ifadesi artan gen listesinin yolak analizinin sonuçları

Çizelge 5.3: Evre kıyaslamasının özeti

ÇAKIŞMA		GEN SAYISI	ÖNEMLİ GEN SAYISI (Z-skoru Yolak)	GENLER
Keşif & Val seti Evre2vK	UpReg	105	7	COL10A1, COL11A1, COL12A1, COL1A2, COL4A1, COL27A1, THBS2
	DownReg	257	13	CA1, CA2, CA12, CA4, CA7, MT1E, MT1G, MT1H, MT1F, MT1M, PDE3A, PDE4D, PDE9A
Keşif & Val seti Evre3vK	UpReg	86	6	COL10A1, COL11A1, COL12A1, COL1A2, COL4A1, COL27A1, THBS2
	DownReg	317	14	ATP2A3, ATP2B1, CA1, CA2, CA12, CA4, CA7, MT1E, MT1G, MT1H, MT1F, MT1M, MUC2, MUC5B

Bu tez çalışmasında göz önüne alınan kat değişimi, yüksek zenginleşme skoru ve yolak üzerinde yer alması gibi üç kriteri yerine getiren evre 2 çakışan gen listesine bakıldığında COL4A1, COL27A1 genlerini ifadesinde artış görüldüğü. PDE3A, PDE4D, PDE9A genlerinde ifade azalması gözlenmiştir. ATP2A3, ATP2B1, MUC2 ve MUC5B genlerinin ise yalnızca evre 3 çıkan gen listesinde yer aldığı tespit edilmiştir (Çizelge 5.3).

Öte yandan, sadece evre 2 de tespit edilen ve ifadesi azalan genler için kalsiyum sinyal yolağı ve ifadesi artan genler için ise hücre döngüsü yolağı saptanmıştır. Aşağıda bu yollara örnek olarak kalsiyum sinyal yolağı şeması sunulmuştur (Şekil 5.34).



Şekil 5.34: Hippo sinyal yolağının üzerindeki FİG'leri

6. TARTIŞMA

Meta-analiz, aynı konudaki benzer araştırma sonuçlarını niteliksel ve niceliksel olarak birleştirmeye yardımcı olan istatistiksel bir metottur (59). Keşif setinin doğrulanması için ham verilerinin hiçbir işlemlemeden bir araya getirilmesi ve tek bir çalışma gibi analiz edilmesine yönelik naive meta-analiz metodu kullanılmaktadır. Bu tez çalışmasında Chan ve ark (2008) tarafından yapılan çalışmada (39) olduğu gibi, farklı platformlardan elde edilen verilerin biraraya getirilmesi nedeniyle pek çok örtüşmeyen genin analiz dışı bırakılması sonucunun ortaya çıkmaması için aynı platformda (Affymetrix) yürütülen KRK gen ifade profillemesi araştırmaları kapsam dahiline alınmıştır.

Kanser ile ilgili DNA mikrodizin çalışmalarındaki en önemli çıktı differansiyel olarak ifade edilmiş genlerdir (60). Açıkça bir gen imzasının toplanmasındaki en önemli basamak, araştırmada yer alan iki grubun (normal ve hasta örnekler vb.) gen ekspresyon düzeylerinin karşılaştırılmasıdır. Bu da farklı olarak ifade edilmiş gen listesi ile sonuçlanır (FİG'ler). Ancak, mikrodizin çalışmalarından elde edilen verilerin meta-analizde olduğu gibi, birden fazla çalışmanın birleştirilmesi durumunda, farklı önem düzeylerine sahip genlerin yanı sıra, farklı regülasyon yönleri (bir çalışmada ifade artışı gösteren gen diğer bir çalışmada ifade azalışı gösteriyor çıkabilmektedir) sebebiyle bağımsız FİG listelerinin uyumsuzluğu oluşabilir. Meta-analiz gerçekleştirilirken öncelikle analize dahil edilmek üzere seçilen verisetlerin (bağımsız çalışmaların) mümkün olduğunca homojen olmasına gayret edilmelidir. Homojenite, bağımsız verisetlerindeki örneklerin klinik/patolojik/demografik özellikleri, çalışmanın yürütüldüğü platform, deney parametreleri (başlangıç RNA miktarı) gibi parametreler temelinde değerlendirilmelidir. Homojen verisetleri analize alındıktan sonra karşılaştırmalar yapılmak üzere keşif ve doğrulama grupları oluşturulur. Keşif grubunda yapılan karşılaştırmalar sonucu çıkan moleküler imza ile doğrulama grubunda yapılan karşılaştırmalarda elde edilen moleküler imza birebir örtüşmeyecektir. Araştırılan biyolojik sorunun mümkün olan en temiz moleküler analizini yapabilmek için elde edilen bu moleküler imza listelerinin kesişimlerine bakılır. Esas itibarıyla keşif genler, sorulan biyolojik soru karşısında tekrar edecek şekilde karşımıza çıkan genlerdir. Diğer bir deyişle tespit edilen moleküler imza valide edilmiş olur. Bu tez kapsamında elde edilen bulguların aşağıda ilgili başlıklar altında tartışılmıştır.

6.1. TÜMÖR-KONTROL ÇAKIŞMALARI

Keşif setindeki tümör ve normal karşılaştırmasından çıkan imzalarının valide edilmesi için tümör ve normal karşılaştırma sonuçları ile validasyon setlerinin alt grupları kıyaslanmıştır. Çıkan bu çakışma bulgularına göre keşif ve eşleşmiş validasyon setlerinden 483 ortak FİG elde edilirken, keşif ve eşleşmiş olmayan setlerden çıkan FİG sayısı 502 olarak belirlenmiştir. Keşif ve tüm validasyon setlerinin örtüşmesi sonucu ise 505 FİG elde edilmiştir. Böylece keşif setinde yer alan 818 FİG' in % 57, % 59 ve %59' u sırasıyla eşlemiş, eşlenmemiş ve tüm validasyon setleri tarafından valide edilmiştir.

Bireyler arasındaki temel farklılıklar göz önünde bulundurulduğunda, tümör ve komşu normal dokuların örneklerinin eşleştirilmesi, analizde netlik sağlamaktadır (61). Bu çerçevede keşif seti ile eşleşmiş validasyon alt grubu kıyası sonucu elde edilen gen listesinin birbirini en fazla doğrulaması beklenirken, çıkan çakışma oranlarının yanı sıra, PCA grafikleri ve ısı haritaları kıyaslamalı olarak incelendiğinde üç validasyon seti arasında çok büyük bir fark olmadığı görülmektedir.

Valide edilmiş bu çakışan gen listeleri, biyolojik ve moleküler fonksiyon düzeyinde incelenmiş ve gen listelerinde öncelikli genler filtrelenebilmiştir. Bu bağlamda, KEGG yolak analizinden çıkan sonuçlara göre, keşif ve validasyon setlerinde çakışan gen listelerinin oluşturduğu, ifadesi azalan gen listelerine bakıldığında, metabolik yolak, aldosteron bağlı sodyum reabsorpsiyonu, azot metabolizması, pankreas sekresyonu, mineral absorpsiyonu, kükürt metabolizması, nişasta ve sakaroz metabolizması ve tat iletişim gibi ortak yolaklar doğrulanmıştır. İfadesi artan çakışan gen listelerine bakıldığında ise, P13K-Akt-sinyal yolağı, kanserde rol oynayan yolaklar, fokal adezyon, ECM-reseptör etkileşimi, kanserde rol oynayan mikro RNA'lar, amoebiasis, trombosit aktivasyonu, Wnt sinyali yolağı, protein sindirimi ve absorpsiyonu, bazal hücre karsinoma ve endometriyal kanser gibi KEGG yolaklarının ön plana çıktığı valide edilmiştir.

Genellikle sinyal ve metabolik yolaklardaki aksamalar, KRK'deki tümör oluşumunda gözlenmektedir (62). Bu tez çerçevesinde metabolik yolaklar, içerisinde en fazla geni taşıyan yolak olarak belirlenirken, günümüzde yapılan çalışmalarda, metabolik profillemenin KRK'de önemli konuma geldiği görülmektedir. Yüzlerce çalışmada metabolik teknikleri uygulanmak suretiyle normal mukozayı tümörden ayırarak metabolik imzaları belirlenip KRK' in anatomik ve klinikopatolojik özellikleri ile

ilişkilendirilmektedir. Holmes ve arkadaşları tarafından 2008 yılında yapılan bir çalışmada, tümör mukozasında laktat ve glisin yüksek miktarda bulunmuş ve lipidler ve glukoz/galaktoz eşleşmiş normal mukozada miktar olarak arttığı gözlenmiştir (63). Bu tez çalışmasında metabolik yolların yanı sıra, galaktoz metabolizmasının ifadesi azalan genler ile bağlantısı tespit edilmiştir, ancak bu yollarda yer alan her gen ve bu genin kodlayan her metabolite ait bilginin elde edilmesi için ileri çalışmaların yapılması gerekmektedir.

Nam ve Park (2012) tarafından yapılan çalışmada Nam (64), bu tez çalışmasında ön plana çıkan yollardan biri olan fokal adezyon yolağının, kolorektal kanserin başlangıcı olarak tahmin edilen genleri (terapötik hedefler) içerdiği iddia edilmiştir. Bu yolak özellikle hücre çoğalması, hücre göçü ve metastazın direkt etkilenmesi veya tümör çevresindeki ekstraselüler matriksin birikmesiyle beraber tümör hücre büyümesinde önemli bir rol oynamaktadır Coulson (65). Ekstraselüler matriks (ECM) etkileşimine bakıldığında, tümör mikro çevresinin malign hücrelerin davranışında önemli bir rol oynadığını belirlenmiştir. Malignite sürecini takiben, tümör ve çevresindeki dokuların fonksiyonel ilişkileri değişmekte ve ECM'i yeniden şekillendirmektedir (66). PI3K-Akt sinyal yolağının kolon kanserde önemli olup hastaların neredeyse % 20'sinde PIK3CA mutasyonlar olduğu yapılan başka bir çalışmada belirlenmiştir (67).

Metot kısmında ifade edildiği üzere KEGG yolağında yer alan 1.3 veya daha fazla zenginleşme skoruna sahip olan anotasyon kümesinde bulunan genler, elde edilen sonuç gen listesine dahil olmuşlardır. Keşif setinde bulunup sonuç listesinde yer alan ve ifadesi artan genler, *COL11A1*, *COL12A1*, *COL1A1*, *COL1A2*, *COL27A1*, *COL5A2*, *MMP1* ve *MMP3* olarak belirlenirken, ifadesi azalan genlerin ise *ATP2A3*, *ATP2B1*, *CA1*, *CA12*, *CA2*, *CA4*, *CA7*, *ETHE1*, *MT1E*, *MT1F*, *MT1G*, *MT1H*, *MT1M*, *SCN7A*, *SCNN1A*, *SCNN1B*, *SCNN1G*, *SLC4A4*, *SQRDL* ve *TST* olduğu saptanmıştır.

Bu moleküler imzayı oluşturan 28 genin çoğunun kolorektal kanserde önemli rol oynadığını gösteren başka çalışmalar da mevcuttur. Matriks metaloproteinazler (MMP1 ve MMP3 genler), neredeyse bütün hücre-dışı matris bileşenlerinin kırma kapasitesine sahip olan proteazlardır (68). Bu genlerin kolorektal kanser dahil olmak üzere, pek çok kanserde tümör istilası ve metastazda görev aldığı belirtilmiştir (69). 2005 yılında Kwong ve arkadaşları tarafından yapılan çalışmada, *COL1A1*, *COL5A2* ve *COL12A1* gibi kollajen ailesi genlerinin yanı sıra, metation ailesinde yer alan *MT1E*, *MT1F*, *MT1G* ve *MT1H*

genlerinin kolorektal kansere özgül moleküler imzalar oldukları gösterilmiştir (70). Karbonik anhidrazların (*CA1*, *CA2*, *CA4*, *CA12*) kolorektal kanserde görev aldığı belirten çalışmalar da oldukça fazladır (71), (72), (73). Bununla birlikte *SCN7A* (74), *SLC4A4* (75) ve *TST* (76) genlerinin KRK örneklerine ait olduğu belirlenmiştir.

6.2. GRADELER- KONTROL ÇAKIŞMALARI

Kanser derecelendirmesi karsinogenez sürecini izlemek için kullanılmaktadır. Kolorektal kanserde meydana gelen farklı genetik değişiklikler, hücrelerin prekanseröz lezyondan malign duruma dönüşümünde rol almaktadır (77). Sonuç bölümünde ifade edildiği gibi, bu tezde keşif ve validasyon setlerinin arasındaki çakışan grade genlere bakıldığında, normaller ile kıyasladığında hastalarımızdaki grade 2'ye ait olan 890 genin % 66'sı (yani, 583 geni) valide edilirken grade 3' nün 460 geninden 311 gen (% 68) doğrulanabilmiştir.

Hücre döngüsüne ek olarak, p53 sinyal yolağı, Wnt yolağı ve kalsiyum sinyal yolağı yalnız grade 2 deki gen listelerine ait oldukları için bu yolaklar muhtemelen grade 2 ve 3 ayırımında kullanılabilir. Ayrıca yolakların çoğunun oldukça kapsayıcı yani onlarca gene bağlı olduğu için sadece bu çalışmada esas alınan kriterler veya parametreler ile grade arasındaki yolakların izole edilmesini zorlaştırmaktadır. Ayrıca Walther ve ark (2009) tarafından yapılan çalışmada Wnt sinyal yolağı adenomada tespit edilirken, p53 sinyal yolağının ise erken kolorektal kanser ortamında ortaya çıktığı belirlenmiştir (78). Aynı şekilde hücre döngüsü yolağının adenoma-kolorektal kanser geçişinde görev alması belirlenmiştir (79).

Meta-analiz kullanmak suretiyle zenginleşme skoru gibi tüm kriterleri karşılayan keşif setimizdeki grade 3'e özgül genler içinde *CLCA1*, *CLCA4*, *NR3C1*, *NR3C2*, *NR5A2* ve *PPARD* genleri yer almaktadır. Ayrıca, yalnız grade 2' ye bağlı olup tüm kriterleri karşılayan genler tespit edilmemiştir.

Kalsiyum bağlı klorür kanalı ailesinin üyeleri olan *CLCA1* ve *CLCA4* genleri çoğalma inhibitörleridir. Klorür kanal aksesuarı 1 (*CLCA1*) geninin ifade değerleri, histolojik derecelendirmesi ile ilişkili ve özellikle grade 3 ifade seviyesinde keskin bir düşüş olduğu gösterilmiştir (80). Bustin ve arkadaşları (2001) tarafından yapılan bir çalışmada, *CLCA4* gen ifadesinin kolorektal kanserde azalış gösterildiği rapor edilmiştir (81), ancak *bu genin grade 3 ile bağlantısını belirleyen ilk çalışma bu tez çalışmasıdır.*

Nükleer reseptör alt ailesi 3'e üye olan *NR3C1* geni, glikokortikoid reseptörünü kodlayan bir transkripsiyon faktörüdür. Lind ve ark. (2005) tarafından yapılan çalışmada *NR3C1* geninin, kolorektal kanser gelişiminde epigenetik olarak ifadesi azalan genlerden biri olarak tespit edilmiştir (82). Bulgularımızın aksine, yapılan önceki çalışmalarda, *NR3C2* ifadesinin azalması erken KRK ilerlemesinde potansiyel bir etken olduğu belirtilmektedir (79), (83). Nükleer reseptör alt ailesi 5'e üye olan *NR5A2* geninin varlığı diğer dokularda tespit edilmesine rağmen, bu genin bağırsak hücre yenilenmesinde rol oynamasının yanı sıra, KRK de önemli rol oynamaktadır (84). Ancak mevcut literatürde *NR5A2* geninin, kanser ilerlemesinde meydana gelen karmaşık tabloyu oluşturduğunun belirtilmesinin yanında, bu genin bazı çalışmalarda evre 3'e bağlı olduğu iddia edilmiştir.

Grade 3 imzasında spesifik olarak doğrulanmış diğer bir gen ise ifadesi azalan *PPARD* genidir. Peroksizom proliferator-aktif reseptör delta (*PPARD*) geninin, kolorektal kanser diferansiyasyon ve progresyonunda etkin olduğu inanılan bir nükleer hormon reseptördür ancak, farklı bulgular tümörgenezdeki özel rolünün açıklanmasını zorlaştırmaktadır (85). *Ayrıca PPARD geninin KRK'deki invazif fazı ile ilişkili olduğu ilk olarak bu tez çalışmasında işaret edilmiştir.*

Genellikle tümörgenezin invazif basamakları tümör baskılayıcı genlerin mutasyonlarıyla karakterize olduğu düşünülmektedir (54). Bu tez çalışması sayesinde *CLCA1*, *NR3C1* ve *PPARD* gibi kanserin invazif aşamasında rol oynayabilen genler bağımsız verisetleri kullanılarak ortaya çıkarılmıştır.

6.3. EVRELER- KONTROL ÇAKŞMALARI

Bu tez kapsamında tümör oluşması ve ilerlemesinde önemli rol oynayan genlerin belirlenmesi için, farklı evrelere ait tümör ve kontrol örnekleri arasındaki kıyaslamalar yapılmıştır. Sonuç bölümünde ifade edildiği gibi bu tezde keşif ve validasyon setlerinin arasında çakışan evre genlere bakıldığında, normal örnekler ile evre 2 hastalarımızın karşılaştırması sonucunda hastalara ait olan 598 genin % 61'i (yani, 364 gen) valide edilirken evre 3 hastalarına ait olan 784 genin % 68'i (yani 405 gen) doğrulanabilmiştir.

Bu çalışmada meta-analiz kullanmak suretiyle tüm kriterleri karşılayan keşif setimizdeki evre 2 hastalarına özgü genler içinde *COL4A1*, *COL27A1*, *PDE3A*, *PDE4D* ve *PDE9A* genleri yer almaktadır. Hücre döngüsü yolağının yalnız çakışan evre 2 deki gen listesine ait olduğu belirlenmiştir. Evre 3'e özgü genler içinde *ATP2A3*, *ATP2B1*, *MUC2* ve *MUC5B*

genleri yer almaktadır. Wnt sinyal ve Hippo sinyal yollarının yalnız çakışan evre 2 deki gen listesine ait olduğu belirlenmiştir

Hücre döngüsü yolağı kolon kanseri gelişimine yol açması için iki başka yolağın (p13K-Akt ve MAPK) tarafından aktive edildiği düşünülmektedir (86). Hippo sinyal yolağı ile ilişkili olan *TAZ* ve *YAP* genlerin KRK' de prognostik rolüne sahip oldukları belirlenmiştir. Başka bir deyimle hippo yolağı kolorektal kanserdeki ileri evreler ile ilişkilidir (87). Bununla birlikte, 2013 yılında Konsavage ve Yochum tarafından yapılan bir çalışmada, hippo ve wnt sinyal yollarının birbirine etkilendiği gösterilmiştir (88). Ayrıca Wnt sinyal yolağı, KRK' in hem erken aşamaları he invasif aşamalarında gözlenmektedir (89).

İntestinal mukus zarının önemli bir bileşeni olup (90), fonksiyonel p53 kaybına etki ettiği belirtilen münin 2 (*MUC2*) geninin ifade seviyesi kolorektal karsinomada azalış göstermektedir. Prognostik değere sahip olan *MUC2* geninin KRK' in ileri evrelerinde sınırlı olduğunu belirlenmiştir (91). *ATP2A3* ifadesinin azalması, Flores-Peredo ve ark (2016) tarafından yapılan bir çalışmada, kolorektal kanser ile bağlantılı olduğu belirlenirken (92), *yapılan bu tez çalışmasında ilk defa KRK'in evre 3 profili ile bağlantılı olduğunu gösterilmiştir.*

Siklik nükleotid fosfodiesterazlar (PDE'ler), tümör ilerlemesinde görev almaktadırlar ancak çeşitli kanser dokularında bu genlerin ifadesinde azalış gösterilmektedir (93). 2011 yılında Kim ve arkadaşları tarafından yapılan bir çalışmada, *COL4A1* geni evre 4 ile bağlantılı olduğu gösterilmiştir (94) ve keşif setinde evre 4 örneklerin olmadığı için bu bulgunun araştırılması mümkün olmamıştır.

Son olarak, bu tez çalışması kapsamında, KRK'in hem gradeler hem de evrelere ait bulgular elde edilmiştir. Derecelendirme (grading), tümör hücrelerinin diferansiyasyon derecesi ve atipisine dayanılarak, mikroskopik incelemeyle tümörün biyolojik davranışlarını öngörmeye dayanmaktadır. Gradelere göre değerlendirme yapıldığında, iyi, orta ve az differansiyasyon olarak farklılıklar gözlenirken, tümör değerlendirilmesinde, özellikle büyük tümörlerde makroskopik örneklemedeki yetersizliklere bağlı olarak derecelendirmede yanlışlıklar olabilmektedir. Bu nedenle, evrelendirmenin (staging), kanserin boyutuna ve metastaza ait ayrıntıları göstererek tümörün progresyonu veya hastanın prognozu ile ilgili bilgi edinilmesi açısından en güvenilir yöntem olduğu

görülmektedir. Dolayısıyla, ileride yapılması planlanan çalışmalarda, sınıflandırmaların öncelikle tümör evrelerine bağlı olarak yapılması gerektiğini göstermektedir.



Bu tez çalışmasında;

- Keşif setimizdeki var olan kolorektal kanser ile ilişkili gen bulgularının (özellikle gen imzaları) valide edilmesi için literatür araştırmasından elde edilen benzer mikrodizin çalışmaları ile karşılaştırılmıştır
- ANOVA temelli naive bir meta-analiz metodu kullanılmıştır.
- Keşif ve validasyon setleri arasındaki çakışan genlerin moleküler fonksiyonel kümeleme analizleri sonucu keşif setinde bulunan önemli kolorektal kanser yolları da belirlenmiştir.
- Böylece tümör örnekleri ve normallerinin kıyaslanması sonucu ortaya çıkan KKK' e özgü gen imzalarının yanı sıra, bu kanser türünün moleküler alt tiplerini birbirinden ayırabilen evreler ve gradelere özgü aday gen imzaları belirlenmiştir.

Bu çalışma ışığında elde edilen verilere göre ileride yapılması düşünülen çalışmalar;

- Tümör ve normallerin karşılaştırılmasında, uygulanan meta-analiz sonucu olarak elde edilen 28 gen imzasını eş-zamanlı PZR yöntemi ile ileri araştırmaların yapılması.
- Aynı zamanda evreler ve gradeler arasındaki karşılaştırmalarda KKK' da önemli rol oynadığı belirlenen *CLCA1*, *CLCA4*, *NR3C1*, *PPARD*, *ATP2A3*, *MUC2* ve *MUC5B* gibi biyobelirteç adaylarının bağımsız çalışmalarda incelenmesi amaçlanmaktadır.
- Tez çalışmanın devamı olarak kolorektal kanserde metabolik profillemeye bir çalışmanın yapılması düşünülmektedir.

7. KAYNAKLAR

1. Siegel R, Naishadham D, Jemal A. Cancer statistics, 2013. *CA Cancer J Clin.* 2013;63(1):11-30.
2. Fearon ER, Vogelstein B. A genetic model for colorectal tumorigenesis. *Cell.* 1990;61(5):759-767.
3. Srivastava S. Cancer biomarker discovery and development in gastrointestinal cancers: early detection research network-a collaborative approach. *Gastrointest Cancer Res.* 2007;1(4 Suppl 2):S60-63.
4. Wang Y, Jatkoe T, Zhang Y, et al. Gene expression profiles and molecular markers to predict recurrence of Dukes' B colon cancer. *J Clin Oncol.* 2004;22(9):1564-1571.
5. Kim J, Takeuchi H, Lam ST, et al. Chemokine receptor CXCR4 expression in colorectal cancer patients increases the risk for recurrence and for poor survival. *J Clin Oncol.* 2005;23(12):2744-2753.
6. Brown PO, Botstein D. Exploring the new world of the genome with DNA microarrays. *Nat Genet.* 1999;21(1 Suppl):33-37.
7. Petersen D, Chandramouli GV, Geoghegan J, et al. Three microarray platforms: an analysis of their concordance in profiling gene expression. *BMC Genomics.* 2005;6:63.
8. Larry V. Hedges IO. *Statistical Methods for Meta-Analysis.* 1985.
9. Lai Y, Eckenrode SE, She JX. A statistical framework for integrating two microarray data sets in differential expression analysis. *BMC Bioinformatics.* 2009;10 Suppl 1:S23.
10. Wren JD. A global meta-analysis of microarray expression data to predict unknown gene functions and estimate the literature-data divide. *Bioinformatics.* 2009;25(13):1694-1701.
11. Min Jou W, Haegeman G, Ysebaert M, Fiers W. Nucleotide sequence of the gene coding for the bacteriophage MS2 coat protein. *Nature.* 1972;237(5350):82-88.
12. Sanger F, Air GM, Barrell BG, et al. Nucleotide sequence of bacteriophage phi X174 DNA. *Nature.* 1977;265(5596):687-695.

13. Subramanian A, Tamayo P, Mootha VK, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A*. 2005;102(43):15545-15550.
14. Brazma A, Hingamp P, Quackenbush J, et al. Minimum information about a microarray experiment (MIAME)-toward standards for microarray data. *Nat Genet*. 2001;29(4):365-371.
15. Gene Expression Omnibus. <http://www.ncbi.nlm.nih.gov/geo/>.
16. ArrayExpress. <https://www.ebi.ac.uk/arrayexpress/>.
17. Garcia-Hernandez M, Berardini TZ, Chen G, et al. TAIR: a resource for integrated Arabidopsis data. *Funct Integr Genomics*. 2002;2(6):239-253.
18. Betel D, Wilson M, Gabow A, Marks DS, Sander C. The microRNA.org resource: targets and expression. *Nucleic Acids Res*. 2008;36(Database issue):D149-153.
19. Kreil DP, Russell RR, Russell S. Microarray oligonucleotide probes. *Methods Enzymol*. 2006;410:73-98.
20. Lipshutz RJ, Fodor SP, Gingeras TR, Lockhart DJ. High density synthetic oligonucleotide arrays. *Nat Genet*. 1999;21(1 Suppl):20-24.
21. Venter JC, Adams MD, Myers EW, et al. The sequence of the human genome. *Science*. 2001;291(5507):1304-1351.
22. Belder N. Sporadik kolorektal kanser vakalarında genom ebadında kopya sayısı değişimlerinin ve transkriptom profilinin belirlenmesi ile kanserin gelişmesinde ve ilerlemesinde etken yeni genlerin tanımlanması. 2013.
23. Schuchhardt J, Beule D, Malik A, et al. Normalization strategies for cDNA microarrays. *Nucleic Acids Res*. 2000;28(10):E47.
24. Ryden P, Andersson H, Landfors M, et al. Evaluation of microarray data normalization procedures using spike-in experiments. *BMC Bioinformatics*. 2006;7:300.
25. Yang YH, Dudoit S, Luu P, et al. Normalization for cDNA microarray data: a robust composite method addressing single and multiple slide systematic variation. *Nucleic Acids Res*. 2002;30(4):e15.

26. Beissbarth T, Fellenberg K, Brors B, et al. Processing and quality control of DNA array hybridization data. *Bioinformatics*. 2000;16(11):1014-1022.
27. Bolstad BM, Irizarry RA, Astrand M, Speed TP. A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. *Bioinformatics*. 2003;19(2):185-193.
28. Jolliffe IT. *Principal Component Analysis*. 1986.
29. Alter O, Brown PO, Botstein D. Singular value decomposition for genome-wide expression data processing and modeling. *Proc Natl Acad Sci U S A*. 2000;97(18):10101-10106.
30. Stekel DJ, Git Y, Falciani F. The comparison of gene expression from multiple cDNA libraries. *Genome Res*. 2000;10(12):2055-2061.
31. Dhammika Amaratunga JC. *Exploration and Analysis of DNA Microarray and Protein Array Data*. 2004.
32. Thomas PD, Mi H, Lewis S. Ontology annotation: mapping genomic regions to biological function. *Curr Opin Chem Biol*. 2007;11(1):4-11.
33. Huang da W, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc*. 2009;4(1):44-57.
34. Ramasamy A, Mondry A, Holmes CC, Altman DG. Key issues in conducting a meta-analysis of gene expression microarray datasets. *PLoS Med*. 2008;5(9):e184.
35. Sorlie T, Tibshirani R, Parker J, et al. Repeated observation of breast tumor subtypes in independent gene expression data sets. *Proc Natl Acad Sci U S A*. 2003;100(14):8418-8423.
36. Wirapati P, Sotiriou C, Kunkel S, et al. Meta-analysis of gene expression profiles in breast cancer: toward a unified understanding of breast cancer subtyping and prognosis signatures. *Breast Cancer Res*. 2008;10(4):R65.
37. Chen R, Khatri P, Mazur PK, et al. A meta-analysis of lung cancer gene expression identifies PTK7 as a survival gene in lung adenocarcinoma. *Cancer Res*. 2014;74(10):2892-2902.

38. Taminau J, Lazar C, Meganck S, Nowé A. Comparison of Merging and Meta-Analysis as Alternative Approaches for Integrative Gene Expression Analysis. *ISRN Bioinformatics*. 2014;2014:1-7.
39. Chan SK, Griffith OL, Tai IT, Jones SJ. Meta-analysis of colorectal cancer gene expression profiling studies identifies consistently reported candidate biomarkers. *Cancer Epidemiol Biomarkers Prev*. 2008;17(3):543-552.
40. D.J A. *Burger's Medicinal Chemistry and Drug Discovery* Vol 42003.
41. Barrett JC. Mechanisms of multistep carcinogenesis and carcinogen risk assessment. *Environ Health Perspect*. 1993;100:9-20.
42. Hoadley KA, Yau C, Wolf DM, et al. Multiplatform analysis of 12 cancer types reveals molecular classification within and across tissues of origin. *Cell*. 2014;158(4):929-944.
43. Cheah PY. Recent advances in colorectal cancer genetics and diagnostics. *Crit Rev Oncol Hematol*. 2009;69(1):45-55.
44. Parkin DM, Bray F, Ferlay J, Pisani P. Global cancer statistics, 2002. *CA Cancer J Clin*. 2005;55(2):74-108.
45. 2013 Yılı Türkiye Kanser İstatistikleri.
46. Potter JD. Colorectal cancer: molecules and populations. *J Natl Cancer Inst*. 1999;91(11):916-932.
47. Cruz-Bustillo Clarens D. Molecular genetics of colorectal cancer. *Rev Esp Enferm Dig*. 2004;96(1):48-59.
48. Kinzler KW, Vogelstein B. Cancer-susceptibility genes. Gatekeepers and caretakers. *Nature*. 1997;386(6627):761, 763.
49. Miyashita T, Krajewski S, Krajewska M, et al. Tumor suppressor p53 is a regulator of bcl-2 and bax gene expression in vitro and in vivo. *Oncogene*. 1994;9(6):1799-1805.
50. Dobrucalı A. Kolorektal kanserde sınıflama ve etyopatogenez. Kolon, rektum ve anorektal bölge hastalıkları. *Tasarım yayıncılık*. 2003:395-412.
51. Li ZS, Li Q. [The latest 2010 WHO classification of tumors of digestive system]. *Zhonghua Bing Li Xue Za Zhi*. 2011;40(5):351-354.

52. Frederick L. Greene M.D. DLPMD, Irvin D. Fleming M.D., April G. Fritz C.T.R., R.H.I.T., Charles M. Balch M.D., Daniel G. Haller M.D., Monica Morrow M.D. *AJCC Cancer Staging manual*. 6 ed2002.
53. Nannini M, Pantaleo MA, Maleddu A, Astolfi A, Formica S, Biasco G. Gene expression profiling in colorectal cancer using microarray technologies: results and perspectives. *Cancer Treat Rev*. 2009;35(3):201-209.
54. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. *Cell*. 2011;144(5):646-674.
55. Arpino G, Generali D, Sapino A, et al. Gene expression profiling in breast cancer: a clinical perspective. *Breast*. 2013;22(2):109-120.
56. Takami K, Yana I, Kurahashi H, Nishisho I. Multistep carcinogenesis in colorectal cancers. *Southeast Asian J Trop Med Public Health*. 1995;26 Suppl 1:190-196.
57. Jin Z, Jiang W, Wang L. Biomarkers for gastric cancer: Progression in early diagnosis and prognosis (Review). *Oncol Lett*. 2015;9(4):1502-1508.
58. Cardoso J, Boer J, Morreau H, Fodde R. Expression and genomic profiling of colorectal cancer. *Biochim Biophys Acta*. 2007;1775(1):103-137.
59. Normand SL. Meta-analysis: formulating, evaluating, combining, and reporting. *Stat Med*. 1999;18(3):321-359.
60. Dallas PB, Gottardo NG, Firth MJ, et al. Gene expression levels assessed by oligonucleotide microarray analysis and quantitative real-time RT-PCR -- how well do they correlate? *BMC Genomics*. 2005;6:59.
61. Vakiani E, Janakiraman M, Shen R, et al. Comparative genomic analysis of primary versus metastatic colorectal carcinomas. *J Clin Oncol*. 2012;30(24):2956-2962.
62. Hong Y, Ho KS, Eu KW, Cheah PY. A susceptibility gene set for early onset colorectal cancer that integrates diverse signaling pathways: implication for tumorigenesis. *Clin Cancer Res*. 2007;13(4):1107-1114.
63. Holmes E, Wilson ID, Nicholson JK. Metabolic phenotyping in health and disease. *Cell*. 2008;134(5):714-717.
64. Nam S, Park T. Pathway-based evaluation in early onset colorectal cancer suggests focal adhesion and immunosuppression along with epithelial-mesenchymal transition. *PLoS One*. 2012;7(4):e31685.

65. Coulson-Thomas VJ, Coulson-Thomas YM, Gesteira TF, et al. Colorectal cancer desmoplastic reaction up-regulates collagen synthesis and restricts cancer cell invasion. *Cell Tissue Res.* 2011;346(2):223-236.
66. Coulson-Thomas VJ, Gesteira TF, Coulson-Thomas YM, et al. Fibroblast and prostate tumor cell cross-talk: fibroblast differentiation, TGF-beta, and extracellular matrix down-regulation. *Exp Cell Res.* 2010;316(19):3207-3226.
67. Zhao L, Vogt PK. Class I PI3K in oncogenic cellular transformation. *Oncogene.* 2008;27(41):5486-5496.
68. Visse R, Nagase H. Matrix metalloproteinases and tissue inhibitors of metalloproteinases: structure, function, and biochemistry. *Circ Res.* 2003;92(8):827-839.
69. Hewitt RE, Leach IH, Powe DG, Clark IM, Cawston TE, Turner DR. Distribution of collagenase and tissue inhibitor of metalloproteinases (TIMP) in colorectal tumours. *Int J Cancer.* 1991;49(5):666-672.
70. Kwong KY, Bloom GC, Yang I, et al. Synchronous global assessment of gene and protein expression in colorectal cancer progression. *Genomics.* 2005;86(2):142-158.
71. Kivela AJ, Saarnio J, Karttunen TJ, et al. Differential expression of cytoplasmic carbonic anhydrases, CA I and II, and membrane-associated isozymes, CA IX and XII, in normal mucosa of large intestine and in colorectal tumors. *Dig Dis Sci.* 2001;46(10):2179-2186.
72. Zhang J, Tsoi H, Li X, et al. Carbonic anhydrase IV inhibits colon cancer development by inhibiting the Wnt signalling pathway through targeting the WTAP-WT1-TBL1 axis. *Gut.* 2016;65(9):1482-1493.
73. Mori M, Staniunas RJ, Barnard GF, Jessup JM, Steele GD, Jr., Chen LB. The significance of carbonic anhydrase expression in human colorectal cancer. *Gastroenterology.* 1993;105(3):820-826.
74. Ostasiewicz B, Ostasiewicz P, Dus-Szachniewicz K, Ostasiewicz K, Ziolkowski P. Quantitative analysis of gene expression in fixed colorectal carcinoma samples as a method for biomarker validation. *Mol Med Rep.* 2016;13(6):5084-5092.
75. Yang Y, Wu PP, Wu J, et al. Expression of anion exchanger 2 in human gastric cancer. *Exp Oncol.* 2008;30(1):81-87.

76. Ramasamy S, Singh S, Taniere P, Langman MJ, Eggo MC. Sulfide-detoxifying enzymes in the human colon are decreased in cancer and upregulated in differentiation. *Am J Physiol Gastrointest Liver Physiol*. 2006;291(2):G288-296.
77. Cattaneo E, Laczko E, Buffoli F, et al. Preinvasive colorectal lesion transcriptomes correlate with endoscopic morphology (polypoid vs. nonpolypoid). *EMBO Mol Med*. 2011;3(6):334-347.
78. Walther A, Johnstone E, Swanton C, Midgley R, Tomlinson I, Kerr D. Genetic prognostic and predictive markers in colorectal cancer. *Nat Rev Cancer*. 2009;9(7):489-499.
79. Pesson M, Volant A, Uguen A, et al. A gene expression and pre-mRNA splicing signature that marks the adenoma-adenocarcinoma progression in colorectal cancer. *PLoS One*. 2014;9(2):e87761.
80. Yang B, Cao L, Liu J, et al. Low expression of chloride channel accessory 1 predicts a poor prognosis in colorectal cancer. *Cancer*. 2015;121(10):1570-1580.
81. Bustin SA, Li SR, Dorudi S. Expression of the Ca²⁺-activated chloride channel genes CLCA1 and CLCA2 is downregulated in human colorectal cancer. *DNA Cell Biol*. 2001;20(6):331-338.
82. Lind GE, Kleivi K, Meling GI, et al. ADAMTS1, CRABP1, and NR3C1 identified as epigenetically deregulated genes in colorectal tumorigenesis. *Cell Oncol*. 2006;28(5-6):259-272.
83. Budinska E, Popovici V, Tejpar S, et al. Gene expression patterns unveil a new level of molecular heterogeneity in colorectal cancer. *J Pathol*. 2013;231(1):63-76.
84. Schoonjans K, Dubuquoy L, Mebis J, et al. Liver receptor homolog 1 contributes to intestinal tumor formation through effects on cell cycle and inflammation. *Proc Natl Acad Sci U S A*. 2005;102(6):2058-2062.
85. Tachibana K, Yamasaki D, Ishimoto K, Doi T. The Role of PPARs in Cancer. *PPAR Res*. 2008;2008:102737.
86. Chen J, Huang XF. The signal pathways in azoxymethane-induced colon cancer and preventive implications. *Cancer Biol Ther*. 2009;8(14):1313-1317.
87. Yuen HF, McCrudden CM, Huang YH, et al. TAZ expression as a prognostic indicator in colorectal cancer. *PLoS One*. 2013;8(1):e54211.

88. Konsavage WM, Jr., Yochum GS. Intersection of Hippo/YAP and Wnt/beta-catenin signaling pathways. *Acta Biochim Biophys Sin (Shanghai)*. 2013;45(2):71-79.
89. Basu S, Haase G, Ben-Ze'ev A. Wnt signaling in cancer stem cells and colon cancer metastasis. *F1000Res*. 2016;5.
90. Kufe DW. Mucins in cancer: function, prognosis and therapy. *Nat Rev Cancer*. 2009;9(12):874-885.
91. Lugli A, Zlobec I, Baker K, et al. Prognostic significance of mucins in colorectal cancer with different DNA mismatch-repair status. *J Clin Pathol*. 2007;60(5):534-539.
92. Flores-Peredo L, Rodriguez G, Zarain-Herzberg A. Induction of cell differentiation activates transcription of the Sarco/Endoplasmic Reticulum calcium-ATPase 3 gene (ATP2A3) in gastric and colon cancer cells. *Mol Carcinog*. 2016.
93. Savai R, Pullamsetti SS, Banat GA, et al. Targeting cancer with phosphodiesterase inhibitors. *Expert Opin Investig Drugs*. 2010;19(1):117-131.
94. Kim HJ, Kang UB, Lee H, et al. Profiling of differentially expressed proteins in stage IV colorectal cancers with good and poor outcomes. *J Proteomics*. 2012;75(10):2983-2997.

8. Ekler

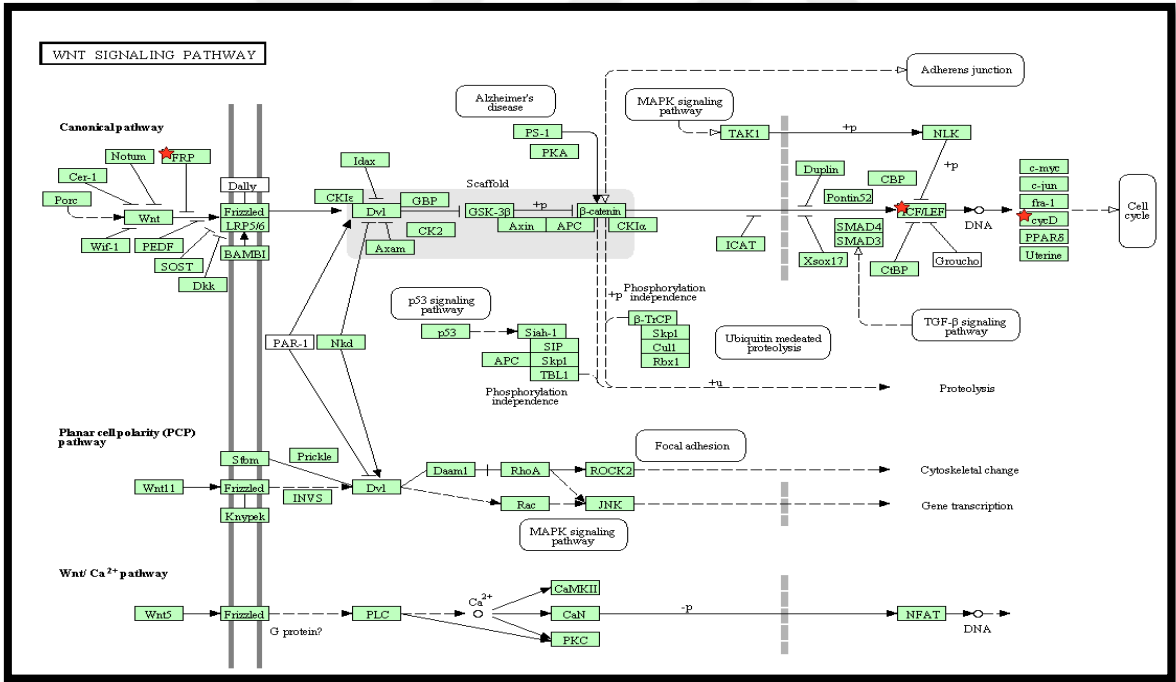
Ek 1. Meta-analizin sonucunda elde edilen 28 gen imzasının ifadesi azalan genlerinin kat deęişim deęerleri (n=20)

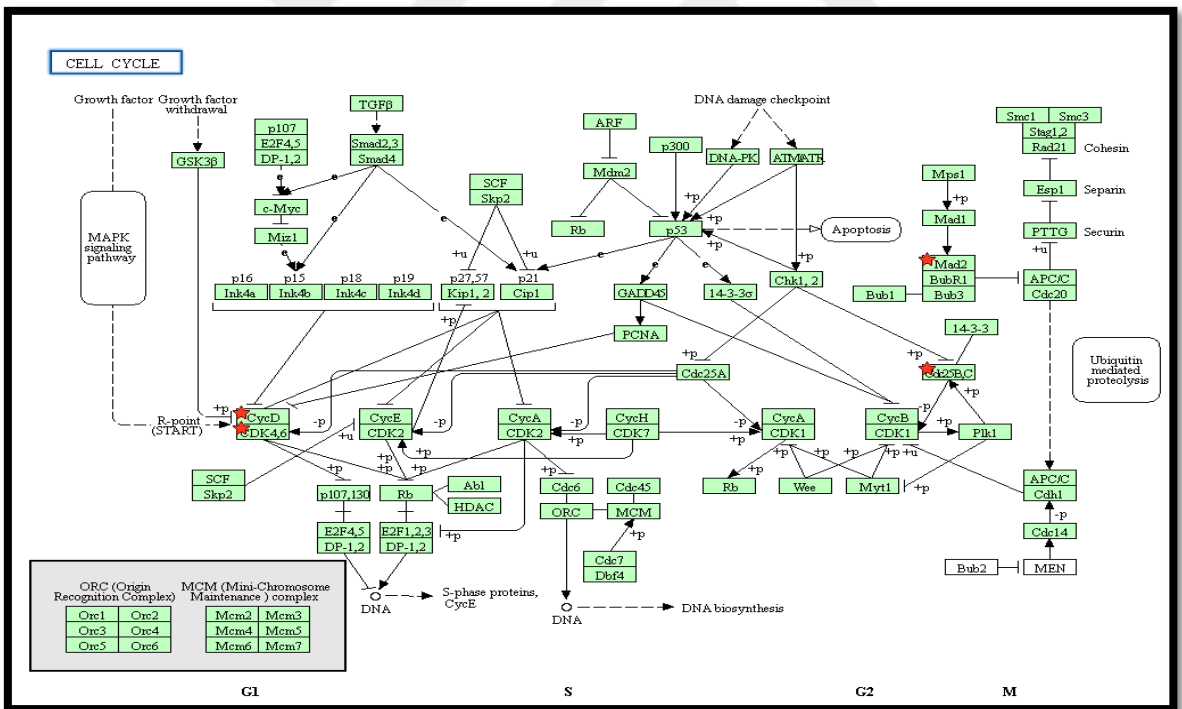
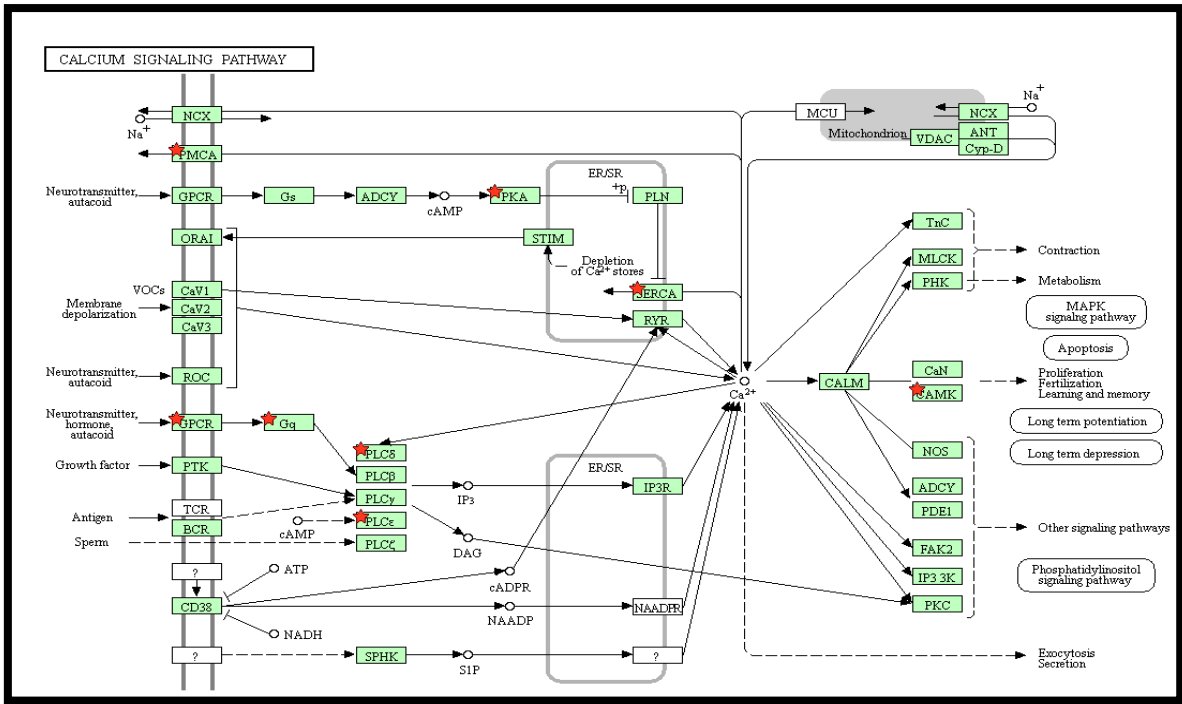
GENLER	YOLAK SAYISI	KAT DEęİŐİMİ
CA2	1	27,52
CA1	1	22
SCNN1B	2	12,55
SCL4A4	3	9,14
MT1M	1	8,31
CA4	1	6,46
CA7	1	5,7
CA12	1	5,09
MT1H	1	4,94
SCNN1G	2	4,86
ATP2A3	2	4,24
MT1E	1	3,77
MT1G	1	3,48
MT1F	1	3,45
ETHE1	1	3,21
SQRDL	1	2,87
SCN7A	1	2,8
ATP2B1	1	2,62
TST	4	2,5
SCNN1A	2	2,34

Ek 2. Meta-analizin sonucunda elde edilen 28 gen imzasının ifadesi artan genlerinin kat deęişim dięerleri (n=8)

GENLER	YOLAK SAYISI	KAT DEęİŐİŐİMİ
MMP3	3	5,58
MMP1	3	5,3
COL11A1	2	4,61
COL12A1	1	4,48
COL1A2	2	3,95
COL1A1	2	3,94
COL5A2	2	3,11
COL27A1	6	2,19

Ek 3: Grade ve Evre normaller ile karŐılaŐtırılmasında ön plana ıkan yolaklar





ÖZGEÇMİŞ

Adı Soyadı: Hakiimu KAWALYA

Doğum Yeri: MENGO (KAMPALA)

Doğum Tarihi: 26.01.1991

Medeni Hali: Bekar

Yabancı Dili: İngilizce ve Türkçe

Eğitim Durumu

Lise: Kibuli düz Lisesi (2008)

Lisans: Makerere Üniversitesi, Biyomedikal laboratuvar teknolojisi Bölümü (2013)

Tezden çıkan Yayınlar

Genomic meta-analysis and discovery of new biomarkers in colorectal cancer (2016). ML-03.01.1-002. H. KawalyaI, N. Belder, A. Kuzu, B. Savas, A. Ensari, H. Özdag.

SYMPOSIA	
📅 4.9.2016 ⌚ 09:00 - 11:00h	Hall C ♣
S.03.01.1. Systems biology Chair: Uri Alon, Israel	
Evolutionary tradeoffs and the geometry of gene expression space Alon, Uri	09:00h - 09:30h
Respect the noise: exquisite interaction of cellular noise and dynamics lead to novel biological function Khammash, Mustafa Hani	09:30h - 10:00h
Dissecting the complexity of Cancer Signalling Bluthgen, Nils	10:00h - 10:30h
An integrative approach to analyze dynamic transcriptional response of yeast cells to DNA damage Karabekmez, Muhammed Erkan	10:30h - 10:45h
Genomic meta-analysis and discovery of new biomarkers in colorectal cancer Kawalya, Hakiimu	10:45h - 11:00h



ML-03.01.1-002

Genomic meta-analysis and discovery of new biomarkers in colorectal cancer

H. Kawalya, N. Belder, A. Kuzu, B. Savas, A. Ensari,
H. Özdag

Ankara University Biotechnology Institute, Ankara, Turkey

Colorectal cancer (CRC) is one of the most prevalent tumors worldwide with the third highest mortality rate in developed countries. Early diagnosis is crucial to the treatment and prevention of CRC, yet the absence of clear symptoms at its onset makes early diagnosis almost impossible. To date, the most plausible diagnosis tool for screening this type of malignancy is the use of biomarkers, however there is need to identify new biomarkers due to inadequacies of the current ones. Gene expression profiling studies are crucial in the identification of probable CRC biomarkers hence a meta-analysis was performed comparing 98 samples of our discovery set against 1174 samples of a validation set from 10 studies of the GEO and Array express databases (Affymetrix platform) so as to validate our findings and derive more robust conclusions. Corresponding subgroups of both sets were analyzed separately after sample normalization using RMA methodology embedded in the Partek[®] Genomics Suite[™] 6.6 software and differentially expressed genes were determined at the $P \leq 0.001$ cut off with false discovery rates and enrichment score ≥ 1.3 was applied during clustering.

The results indicated a 60% validation of the ANOVA tumor vs normal gene-list of the discovery set. Meanwhile, DAVID bioinformatics tool analyses related significant enriched clusters to pathways such as Pathways in cancer, Focal adhesion and Aldosterone regulated sodium reabsorption. A 9-gene signature was highlighted as possessing diagnostic potential in CRC patients. Grade gene-list overlaps validated 66% of grade II vs normal and 68% of grade III vs normal of our cohort. Traces of pre-invasive genes such as CCND1, MYC, and CDK1 were identified in grade II while grade III had a complex composition of genes including PPARD, CDKN2A and GPX1.

Therefore this study may be useful in providing more insight into scantily studied genes believed to have a role in CRC grading.