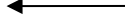


Adınızı soyadınızı giriniz

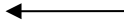


Tez kabul edildikten sonra yapılan **sabit ciltte sırt yazısı** bu şablona göre yazılacak. Yazılar tek satır olacak

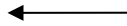
Cilt sırtı yazıların yönü yukarıdan aşağıya (sol yandaki gibi) olacak .



Tez, Yüksek Lisans'sa, YÜKSEK LİSANS TEZİ;  
Doktora ise DOKTORA TEZİ ifadesi kalacak



Tez Sınavının yapılacağı yılı yazınız



T.C.  
İSTANBUL ÜNİVERSİTESİ  
SAĞLIK BİLİMLERİ ENSTİTÜSÜ

YÜKSEK LİSANS TEZİ

YARIK ELAYAK YAPISAL BOZUKLUĞUNUN GENOMİK  
TEKNİKLER VE BİYOENFORMATİK YAKLAŞIMLARLA  
İNCELENMESİ

SADIK BARIŞ SALMAN

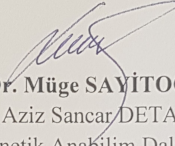
DANIŞMAN  
DOÇ. DR. SİBEL AYLİN UĞUR İŞERİ

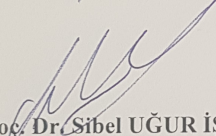
GENETİK ANABİLİM DALI  
GENETİK PROGRAMI

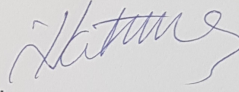
İSTANBUL-2019

**TEZ ONAYI****YÜKSEK LİSANS TEZİ ONAYI**

İstanbul Üniversitesi Sağlık Bilimleri Enstitüsü Genetik Anabilim Dalı, Genetik Programında Yüksek Lisans öğrencisi Sadık Barış SALMAN tarafından Doç. Dr. Sibel Aylın UĞUR İŞERİ'nin danışmanlığında hazırlanan “**Yarık Elayak Yapısal Bozukluğunun Genomik Teknikler ve Biyoenformatik Yaklaşımlarla İncelenmesi**” başlıklı tez aşağıdaki jüri üyeleri tarafından 30.07.2019 tarihinde yapılan Tez Savunma Sınavında başarılı bulunmuş ve Yüksek Lisans Tezi olarak kabul edilmiştir.

  
**Prof. Dr. Müge SAYITOĞLU**  
İ.Ü. Aziz Sancar DETAE  
Genetik Anabilim Dalı

  
**Doç. Dr. Sibel UĞUR İŞERİ**  
İ.Ü. Aziz Sancar DETAE  
Genetik Anabilim Dalı

  
**Doç. Dr. Özden HATIRNAZ NG**  
Acıbadem Mehmet Ali Aydınlar Üni. Tıp Fak.  
Tıbbi Biyoloji Anabilim Dalı

**BEYAN**

iii

**BEYAN**

Bu tez çalışmasının kendi çalışmam olduğunu, tezin planlanmasından yazımına kadar bütün safhalarda etik dışı davranışımın olmadığını, bu tezdeki bütün bilgileri akademik ve etik kurallar içinde elde ettiğimi, bu tez çalışmasıyla elde edilmeyen bütün bilgi ve yorumlara kaynak gösterdiğimi ve bu kaynakları da kaynaklar listesine aldığımı, yine bu tezin çalışılması ve yazımı sırasında patent ve telif haklarını ihlal edici bir davranışımın olmadığını beyan ederim.

Sadık Barış SALMAN

## İTHAF

Aileme ithaf ediyorum.

## TEŞEKKÜR

Tez çalışmasının bütün süreçlerinde desteğini esirgemeyen danışmanım Dr. Sibel Uğur İşeri'ye,

Tez çalışmasının gerçekleşmesinde her türlü desteği veren ve Dr. Aslı Tolun'a,

Yüksek lisansda farklı proje çalışmalarında sınırsız destek veren Dr. Uğur Özbek'e,

Bilgi ve deneyimlerini, tüm yoğunluklarına rağmen paylaşan, Dr. Nerses Bebek'e ve Dr. Betül Baykan'a,

Bana laboratuvarı ve bilimsel çalışmaları öğrenmemde her daim destek veren, Uzm. Bio. Yeşim Kesim'e, Dr. Emrah Yücesan'a ve Dr. Özkan Özdemir'e,

Çalışma arkadaşlarım, Bio. Garen Hayranyan'a, Uzm. Dr. Hande Akçakaya'ya, Uzm. Bio. Sevcan Mercan'a, ve Uzm. Mol. Bio. Cemre Örnek'e,

Her konuda çok değerli geri bildirim veren Dr. Yücel Erbilgin'e, Dr. Müge Sayitoğlu'na, Dr. Özden Hatırnaz Ng'ye

Laboratuvardaki her anı eğlenceli kılan Uzm. Mol. Bio. Khusan Khodzhaev'e ve Uzm. Bio. Didem Altındirek'e,

Lisans döneminden bu yana arkadaşlığı için Mol. Bio. Eralp Emre Yılmaz'a,

Tez sürecinde çalıştığım TÜBİTAK projesi (214S222), Aktif Gen'e ve GenEra Diagnostik'e

Hayatım boyunca arkamda olan annem Selma Salman'a ve babam Veysel Salman'a

Tezin son zamanlarında her zaman yanımda olan Büşra Ay'a çok teşekkür ederim.

Bu çalışma, İstanbul Üniversitesi Bilimsel Araştırma Projeleri Birimi (Proje No: TYL-2018-30315), TÜBİTAK 1001 projesi (114Z829) desteklenmiştir.

## İÇİNDEKİLER

TEZ ONAYI.....	İİ
BEYAN.....	İİİ
İTHAF.....	İV
TEŞEKKÜR.....	V
İÇİNDEKİLER.....	VI
TABLolar LİSTESİ.....	X
ŞEKİLLER LİSTESİ.....	Xİİİ
SEMBOLLER / KISALTMALAR LİSTESİ.....	XVİ
ÖZET.....	XVİİ
ABSTRACT.....	XVİİİ
1. GİRİŞ VE AMAÇ.....	1
2. GENEL BİLGİLER.....	2
2.1. Yeni Nesil Dizileme.....	2
2.1.1. YND Verisinin İşlenmesi.....	4
2.1.2. YND Verisinin Bilgisayar Ortamında Temsili.....	4
2.1.2.1. FASTQ Dosya Formatı.....	5
2.1.2.2. SAM ve BAM Dosya Formatları.....	6
2.1.2.3. BED Dosya Formatı.....	6
2.1.2.4. VCF Dosya Formatı.....	7
2.1.3. YND Analizlerinde Akış Hatları.....	7
2.2. Yarı El/Ayak Bozukluğu.....	8
2.2.1. Uzun Gelişimi.....	8
2.2.2. Yarı El/Ayak Bozukluğu Tanımı ve Sıklığı.....	11
2.2.3. SHFM Genetiği.....	12
2.3. Genom Boyu Genotipleme.....	12
2.3.1. SNP dizilimi.....	13
2.3.1.1. BAF ve LRR.....	13

2.3.1.2. CNV Tespiti.....	13
2.3.2. Homozigot Haritalama.....	14
2.4. Bağlantı Analizi.....	14
2.5. Ekzom Dizileme.....	17
2.6. Sanger Dizileme ile Gen Taraması.....	19
2.7. Gerçek Zamanlı Analizler.....	21
2.7.1. Yüksek Çözünürlüklü Erime Analizi.....	21
2.7.2. qPZR ile CNV Doğrulaması.....	22
3. GEREÇ VE YÖNTEM.....	24
3.1. Gereç.....	24
3.1.1. Çalışma Kapsamında İncelenen Aileler.....	24
3.1.2. Aile MA134.....	24
3.1.3. Aile MA135.....	27
3.1.4. Aile MA136.....	29
3.1.5. Aile MA154.....	31
3.1.6. Çalışmada Kullanılan Laboratuvar Cihazları.....	31
3.1.7. Çalışmada Kullanılan Çözeltiler.....	32
3.1.8. Çalışma Kapsamında Hizmet Alınan Firmalar.....	33
3.1.9. Çalışmada Kullanılan Bilgisayarlar.....	33
3.1.10. Çalışmada Kullanılan Elektronik Veri tabanları.....	34
3.1.11. Çalışmada Kullanılan Yazılımlar.....	34
3.2. Yöntem.....	36
3.2.1. Ailelerin Analizlerinde İzlenen Yöntem.....	36
3.2.1.1. Aile MA134 İnceleme Yöntemi.....	37
3.2.1.2. Aile MA135 İnceleme Yöntemi.....	39
3.2.1.3. Aile MA136 İnceleme Yöntemi.....	40
3.2.1.4. Aile MA154 İnceleme Yöntemi.....	41
3.2.2. Akış Hattı Yazılımının Programlanması.....	42
3.2.2.1. Akış Hattı Yazılımının Yapısı.....	42



3.2.2.2. Akış Hattı Yapılandırma Dosya Yapısı.....	43
3.2.2.3. Akış Hattı Yazılımının Çalıştırılması.....	46
3.2.3. Vcf Sonrası Analizler Yazılımının Programlanması.....	48
3.2.3.1. Vcf Sonrası Analizler Yazılımının Yapısı.....	49
3.2.3.2. Vcf Sonrası Analizler Yazılımının Çalıştırılması.....	51
3.2.4. Geliştirilen Yazılımların Test Edilmesi.....	54
3.2.5. Genom Boyu Genotipleme Analizi.....	54
3.2.5.1. SNP Dizilimi Verisinin Analizi.....	55
3.2.5.2. CNV Bölgelerinin Tespiti.....	57
3.2.5.3. Homozigot Haritalama.....	59
3.2.6. Bağlantı Analizi.....	60
3.2.7. Ekzom Dizileme.....	63
3.2.8. Aday Genlerin Sanger Dizileme ile Taranması.....	69
3.2.9. Varyantların İsimlendirilmesi.....	71
3.2.10. HRMA ve Sanger Dizileme ile Ailesel Ayrışım.....	71
3.2.11. Gerçek Zamanlı PZR ile CNV'lerin doğrulanması.....	71
4. BULGULAR.....	73
4.1. Aile MA134.....	73
4.1.1. Genom Boyu SNP Dizilimi Verisinin Analizi.....	73
4.1.2. Aday Genlerin Sanger Dizileme ile Taranması.....	74
4.1.3. HRMA ve Sanger Dizileme ile Ailesel Ayrışım.....	75
4.2. Aile MA135.....	80
4.2.1. Genom Boyu SNP Genotipleme için SNP Dizilimi Verisinin Analizi.....	80
4.2.2. Bağlantı Analizi.....	85
4.2.3. Ekzom Dizileme.....	85
4.3. Aile MA136.....	86
4.3.1. Genom Boyu SNP Genotipleme için SNP Dizilimi Verisinin Analizi.....	86
4.3.2. Aday Genlerin Sanger Dizileme ile Taranması.....	87
4.3.3. Ekzom Dizileme.....	88

4.4. Aile MA154.....	91
4.4.1. Genom Boyu SNP Genotipleme için SNP Dizilimi Verisinin Analizi.....	91
4.4.2. Gerçek Zamanlı PZR ile CNV'lerin doğrulanması.....	94
4.5. Akış Hattı Deneme Çalışmaları.....	95
5. TARTIŞMA.....	97
5.1. Aile MA134 ve MA136: <i>WNT10B</i> .....	97
5.2. Aile MA135.....	105
5.3. Aile MA154: 10q24.....	107
5.4. Akış Hattı ve VCF Sonrası Analizler Yazılımı.....	110
5.5. Sonuç ve Öneriler.....	111
KAYNAKLAR.....	113
ETİK KURUL KARARI.....	123
İNTİHAL RAPORU İLK SAYFASI.....	126
ÖZGEÇMİŞ.....	127

## TABLOLAR LİSTESİ

Tablo 2-1: Yeni nesil platformlarında kullanılan metotlar ve azami çıktıları.....	3
Tablo 2-2: FASTQ formatında bir okuma bloğunun yazım kuralları.....	5
Tablo 2-3: SHFM OMIM fenotipler ve gen/lokusları.....	12
Tablo 3-1: Çalışmada kullanılan cihazlar ve üretici firmalar.....	32
Tablo 3-2: Çalışmada kullanılan çözelti ve kimyasallar.....	32
Tablo 3-3: Çalışma kapsamında hizmet alınan firma/kurumlar.....	33
Tablo 3-4: Çalışmada kullanılan bilgisayarlar.....	33
Tablo 3-5: Çalışmada kullanılan veritabanları.....	34
Tablo 3-6: Yazılımların geliştirilmesinde kullanılan python paketleri.....	34
Tablo 3-7: SNP dizilimi analizlerinde kullanılan yazılımlar.....	35
Tablo 3-8: ED analizinde kullanılan yazılımlar.....	35
Tablo 3-9: Çalışma yazımında kullanılan yazılımlar.....	36
Tablo 3-10: Projeye özgü argümanlar bölümü.....	44
Tablo 3-11: İşleme özgü argümanlar bölümü.....	45
Tablo 3-12: Girdi ve çıktılarının yerini tutan özel değişkenler.....	46
Tablo 3-13: Anotasyon komutuyla beraber kullanılacak argümanlar ve parametreleri.....	51
Tablo 3-14: Filtreleme komutuyla beraber kullanılacak argümanlar ve parametreleri.....	52
Tablo 3-15: Karşılaştırma bayraklarının işlevleri.....	53
Tablo 3-16: Aile içi analiz komutuyla beraber kullanılacak argümanlar ve parametreleri.....	54
Tablo 3-17: 062817 kodlu ilk SNP dizilimi grubunda çalışılan bireyler.....	56
Tablo 3-18: 100417 kodlu ikinci SNP dizilimi grubunda çalışılan bireyler.....	56
Tablo 3-19: 190415 kodlu üçüncü SNP dizilimi grubunda çalışılan bireyler.....	56
Tablo 3-20: Dizilemede çalışmalarında kullanılan cihaz ve kitler.....	66
Tablo 3-21: MA134 ailesinde Sanger dizileme ile dizilenen bölgeler.....	69
Tablo 3-22: MA136 ailesinde Sanger dizileme ile dizilenen bölgeler.....	70
Tablo 3-23: Sanger dizileme öncesi yapılan PZR için iniş(touchdown) protokolü.....	70
Tablo 3-24: HRMA ile taranan bölge primerleri.....	71
Tablo 3-25: Gerçek zamanlı CNV doğrulaması için kullanılan primerler.....	72

Tablo 4-1: MA134 ailesinde havuzlanarak SNP dizilimi yapılan örneklerin Plink analizi sonucu tespit edilen homozigot bölgeler.....	74
Tablo 4-2: HRMA erime eğrilerine göre bireylerin grupları ve fenotipleri.....	77
Tablo 4-3: HRMA erime eğrilerine göre homozigot bireylerin grupları ve fenotipleri..	80
Tablo 4-4: MA135 ailesinde havuzlanarak SNP dizilimi yapılan örneklerin Plink analizi sonucu tespit edilen homozigot bölgeler.....	81
Tablo 4-5: MA135 ailesinde etkilenmiş beş bireyde ortak tespit edilen homozigot bölgeler.....	81
Tablo 4-6: MA135 ailesinde etkilenmemiş iki bireyde ortak tespit edilen homozigot bölgeler.....	82
Tablo 4-7: İkinci SNP çalışmasında MA135 ailesinde 4 etkilenmiş bireyde tespit edilen CNV.....	82
Tablo 4-8: Üçüncü SNP çalışmasında MA135 ailesinde 1 etkilenmiş birey ve 2 etkilenmemiş bireylerde tespit edilen CNV.....	83
Tablo 4-9: A3 bireyinin, ED sonucu elde edilen 72.516 varyantın adım adım filtrelenmesi.....	86
Tablo 4-10: MA136 ailesinde etkilenmiş iki bireyde ortak tespit edilen homozigot bölgeler.....	87
Tablo 4-11: MA136 ailesinde Sanger dizileme ile <i>WNT10B</i> geninin taranması.....	88
Tablo 4-12: X3 bireyinin, ED sonucu elde edilen 51.975 varyantın adım adım filtrelenmesi.....	88
Tablo 4-13: MA154 ailesinde etkilenmiş iki kardeşle ortak tespit edilen homozigot bölgeler.....	91
Tablo 4-14: Üçüncü SNP çalışmasında MA135 ailesinde 1 etkilenmiş birey ve 2 etkilenmemiş bireylerde tespit edilen CNV.....	92
Tablo 4-15: Birinci deneme çalışmasında, ED sonucu elde edilen 84.806 varyantın adım adım filtrelenmesi.....	95
Tablo 4-16: İkinci çalışmada, birleşik heterozigotluk açısından incelenen 30.797 varyantın adım adım filtrelenmesi.....	96
Tablo 5-1: Literatürde bugüne kadar <i>WNT10B</i> geninde tespit edilen patojenik varyasyonlar.....	100
Tablo 5-2: X3 bireyi ED verisinin filtrelenmesi sonucu elde edilen homozigot varyantlar.....	106

Tablo 5-3: X3 bireyi ED verisinin filtrelenmesi sonucu elde edilen novel heterozigot varyantlar.....	107
Tablo 5-4: 10q24 bölgesinde protein kodlayan genlerin sembolleri.....	109
Tablo 5-5: Komut satır arayüzlü diğer akış hattı yazılımları.....	111



## ŞEKİLLER LİSTESİ

Şekil 2-1: Üç uzuv ekseninde, erken uzuv örüntülenmesini yöneten moleküler etkileşimler.....	9
Şekil 2-2: Omurgalı uzvunun iskeletinin üç temel yapısı ve gelişimi. <sup>127</sup> .....	10
Şekil 2-3: Parmakların oluşumu ve segmentasyonu modeli. <sup>25</sup> .....	11
Şekil 2-4: Karşılıklı parça değişimi ile rekombinant gamet oluşumu <sup>128</sup> .....	15
Şekil 2-5: Sturtevant'ın <i>D. melanogaster</i> gen haritası <sup>128</sup> .....	16
Şekil 2-6: Farklı yakalama metotlarının tasviri. <sup>129</sup> .....	18
Şekil 2-7: dATP ve ddATP şematik gösterimi.....	19
Şekil 2-8: Sanger dizileme basamakları.....	20
Şekil 2-9: Florişi erime eğrisi ile genotiplerin ayrımı.....	21
Şekil 2-10: HRMA ile homozigot wt, homozigot mt ve heterozigot ayrımı.....	22
Şekil 3-1: Şekil 3-1: MA134 aile ağacı.....	26
Şekil 3-2: MA135 aile ağacı.....	28
Şekil 3-3: Şekil 3-3: MA136 aile ağacı.....	30
Şekil 3-4: MA154 aile ağacı.....	31
Şekil 3-5: Ailelerin incelemesinde izlenen genel yöntem akış şeması.....	37
Şekil 3-6: MA134 ailesinin incelenmesinde izlenen yöntem akış şeması.....	38
Şekil 3-7: MA135 ailesinin incelenmesinde izlenen yöntem akış şeması.....	39
Şekil 3-8: MA136 ailesinin incelenmesinde izlenen yöntem akış şeması.....	40
Şekil 3-9: MA154 ailesinin incelenmesinde izlenen yöntem akış şeması.....	41
Şekil 3-10: Akış hattı yazılımının dosya sistemi üzerindeki yapısını gösteren tree komutu çıktısı.....	42
Şekil 3-11: Kalite kontrol ve genoma hizalama akış hattı yapılandırma dosyası.....	47
Şekil 3-12: Akış hattının prova çalıştırılması sonucu.....	48
Şekil 3-13: Vcf sonrası analizler yazılımının dosya sistemi üzerindeki yapısını gösteren tree komutu çıktısı.....	49
Şekil 3-14: GenomeStudio yazılımına SNP çalışmalarının tanıtılması.....	55
Şekil 3-15: GenomeStudio yazılımında LRR ve BAF değerlerinin görseli.....	57
Şekil 3-16: CNVpartition yazılımının varsayılan seçenekleri.....	58
Şekil 3-17: CNV analizlerinin CNV Region Report yazılımı ile grafikleştirilmesi.....	58
Şekil 3-18: Homozigot bölgelerin GenomeStudio yazılımı üzerinden gösterimi.....	59

Şekil 3-19: Farklı SNP dizilimi çalışmalarının çıktılarının birleştirilmesi için kullanılan betik.....	60
Şekil 3-20: MA135 ailesinde her bireyin verilen SNP belirteçindeki genotiplerini içeren dosya.....	61
Şekil 3-21: Aile bireylerini ve bireylerin birbirleri arası ilişkilerini gösteren aile ağacı dosyası.....	61
Şekil 3-22: SNP belirteçlerinin cM cinsinden genomik pozisyonlarının bulunduğu harita dosyası.....	62
Şekil 3-23: Bağlantı analizi sonucu elde edilen LOD skor grafiği.....	63
Şekil 3-24: Ham veri analizi sonrası elde edilen varyantların adım adım filtrelenmesi.....	68
Şekil 4-1: <i>WNT10B</i> geninde Sanger dizileme ile tespit edilen patojenik varyasyonun kromatogram görseli.....	75
Şekil 4-2: İlk HRMA aşamasında sinyallerin erime noktalarına göre ayrımı.....	76
Şekil 4-3: HRMA ile sınıflandırılan genotiplerin Sanger dizileme ile ayrımı.....	78
Şekil 4-4: İkinci HRMA aşamasında sinyallerin erime noktalarına göre ayrımı.....	79
Şekil 4-5: MA135 ailesinde etkilenmiş 4 bireyde tespit edilen CNV'nin GenomeStudio üzerinde gösterimi. CNV bölgesi mavi ile işaretlenmiştir.....	84
Şekil 4-6: Ekzom dizileme sonucu <i>WNT10B</i> geninde tespit edilen NM_003394.4:c.343C>T:p.Arg115Ter varyasyonunun IGV ile gösterimi.....	89
Şekil 4-7: MA136 ailesinde tespit edilen NM_003394.4:c.343C>T:p.Arg115Ter varyasyonunun Sanger dizileme ile aile içerisindeki ayrımının gösterimi.....	90
Şekil 4-8: MA154 ailesinde etkilenmiş 2 bireyde chr10:102.924.567-103.404.601 bölgesinde tespit edilen artış CNV'sinin GenomeStudio üzerinde gösterimi. CNV bölgesi eflatun ile işaretlenmiştir.....	93
Şekil 4-9: MA154 ailesinde CNV artışının gerçek zamanlı PZR ile doğrulanması.....	94
Şekil 5-1: MA134 ve MA136 ailelerinde <i>WNT10B</i> geninde tespit edilen varyasyonların protein dizisi üzerindeki etkileri.....	99
Şekil 5-2: Tez kapsamında tespit edilen NM_003394.4:c.994C>G varyasyonunun GnomAD veri tabanında popülasyonlar arası alel sıklığının gösterimi.....	101
Şekil 5-3: <i>WNT10B</i> geninde tespit edilen varyasyonlarının denk geldiği bölgelerin ortolog ve homolog proteinler üzerinde korunmuşluğunun gösterimi.....	102
Şekil 5-4: <i>WNT10B</i> amino asit modifikasyonlarının Uniprot veri tabanında gösterimi.....	104

Şekil 5-5: Decipher veri tabanında bildirilen patojenik CNV'ler ile Y3 ve Y5 bireylerinde tespit edilen CNV'lerin UCSC Genome Browser'da gösterimi.....	108
Şekil 5-6: Fare embriyosunun 11. gününde gelişen uzvunda Dac protein anlatımı.....	110





**SEMBOLLER / KISALTMALAR LİSTESİ**

- A: Adenin  
AER: Apical ectodermal ridge (Apikal ektodermal sırt)  
ASCII: American Standard Code for Information Interchange (Bilgi Değişimi İçin Amerikan Standart Kodlama Sistemi)  
ACMG: American College of Medical Genetics (Amerikan Medikal Genetik Koleji)  
BAF: B alel frekansı  
bç: Baz çifti  
C: Sitozin  
CLI: Command Line Interface (Komut satır arayüzü)  
cM: Santimorgan  
CNV: Kopya sayısı değişimi  
DNA: Deoksiribonükleik asit  
G: Guanin  
IGV: Integrative Genomics Viewer (bütünleştirici genomik gösterici)  
kb: Kilobaz (1000 baz çifti)  
LOD: Logarithm of odds ratio (şansın logaritması)  
LOH: Heterozigotluğun kaybı  
LRR: LogR kesiri  
MAF: Minör alel frekansı  
Mb: Megabaz (1 milyon baz çifti)  
mt: Mutant  
OD: Otozomal dominant  
OMIM: Online Mendelian Inheritance in Man (İnsanda çevrimiçi mendelyen kalıtım)  
OR: Otozomal resesif  
PZR: Polimeraz zincir reaksiyonu  
RAM: Random-access memory (Rastgele-erişim hafıza)  
SHFM: Split hand/foot malformation (yarık el/ayak malformasyonu)  
SNP: Single Nucleotide Polymorphism (Tek nükleotid polimorfizmi)  
ED: Ekzom dizileme  
UTR: Kodlanmayan bölge  
wt: Wild Type (Yabani tip)  
YND: Yeni nesil dizileme  
ZPA: Zone of Polarizing Activity (Polarize aktivite alanı)

## ÖZET

Salman, S.B. (2019). Yarık El/Ayak Yapısal Bozukluğunun Genomik Teknikler ve Biyoenformatik Yaklaşımlarla İncelenmesi. İstanbul Üniversitesi Sağlık Bilimleri Enstitüsü, Genetik ABD. Yüksek Lisans Tezi. İstanbul.

Yarık el/ayak bozukluğu, gelişim sırasında el ve ayaklardaki orta parmakların eksikliği ile meydana gelen nadir genetik bir bozukluktur. Akraba evliliği oranının yüksek olduğu bölgelerde Mendelyen otozomal resesif (OR) kalıtım modeline göre kalıtılan hastalıkların görülme sıklığının artmasına sebep olmaktadır. Yeni nesil dizileme ve dizilim teknolojileri ile OR kalıtım gösteren birçok genetik hastalığın altında yatan gen ve bölge belirlenebilmiştir. Tez kapsamında OR kalıtım modeline uygun 4 aileye mensup toplam 41 birey incelenmiştir. Ailelerde SNP dizilimi, homozigot haritalama, bağlantı analizi ve ekzom dizileme yöntemlerinin farklı kombinasyonları kullanılarak ilişkili aday varyantların tespiti amaçlanmıştır. MA134 ailesinde daha önce hastalık ile ilişkilendirilmiş *WNT10B* geninde NM\_003394.4::c.994C>G:p.Arg332Gly patojenik varyantı belirlenmiştir. MA136 ailesinde yine *WNT10B* geninde NM\_003394.4::c.343C>T:p.Arg115Ter patojenik varyantı belirlenmiştir. MA154 ailesinde kromozom 10q24 bölgesinde kopya sayısı varyasyonu açısından tek kopya bir patojenik artış tespit edilmiştir. MA135 ailesinde yapılan çalışmalar sonucu patojenik bir varyant ya da bölge tespit edilememiştir. Çalışma kapsamında yukarıdaki 4 ailenin analizinin yanı sıra yeni nesil dizileme veri analizlerinin basamak basamak yapılandırılarak akış hattı oluşturan ve ham veri analizleri sonrası elde edilen VCF formatının nitelendirilmesini, filtrelenmesini ve aile içi analizlerin yapılmasını sağlayan iki yazılım geliştirilmiştir. Bağlantılar: <https://pypi.org/project/seq-pigeon/> ve <https://pypi.org/project/seq-dove/>

Anahtar Kelimeler : Yarık El/Ayak Yapısal Malformasyonu, Gelişim Genetiği, Biyoenformatik, Yeni Nesil Dizileme, SNP dizilimi, WNT10B

Bu çalışma, İstanbul Üniversitesi Bilimsel Araştırma Projeleri Birimi Proje No: 30315 ve TÜBİTAK 1001 projesi (114Z829) tarafından desteklenmiştir.

## ABSTRACT

Salman, S.B. (2019). Study of Split Hand/Foot Malformation with Genomic Techniques and Bioinformatic Approaches. İstanbul University, Institute of Health Science, Department of Genetics . Master's Thesis. İstanbul.

Split hand/foot malformation is a rare genetic disorder characterized by missing middle digits in hand and foot. Genetic diseases that inherited according to Mendelian autosomal recessive (AR) model have a higher prevalence in regions with higher consanguineous marriages. Many of the genetic disorders with AR inheritance has been identified with next generation sequencing and array technologies. In this study total of 41 individuals from 4 families with consanguineous marriages were studied using different combinations of SNP array, homozygous mapping, linkage analysis, and exome sequencing methods to identify pathogenic candidate variants. Analyses performed on MA134 family revealed NM\_003394.4:c.994C>G:p.Arg332Gly pathogenic variant in SHFM related *WNT10B* gene. Analyses performed on MA136 family revealed NM\_003394.4:c.343C>T:p.Arg115Ter pathogenic also on *WNT10B* gene. Analyses performed on MA154 family showed CNV gain on SHFM related 10q24 region. No pathogenic variants or genomic regions have been discovered by analyses performed on MA135. Alongside analyses of 4 families, two software has been developed. First one is to automatizing pipelines used in next generation sequencing and second one for annotating VCF format after raw data analyses as well as filtering variants and in family analyses. Links: <https://pypi.org/project/seq-pigeon/> and <https://pypi.org/project/seq-dove/>

Key Words: SHFM, Developmental Genetics, Bioinformatics, NGS, SNP array

The present work was supported by the Research Fund of İstanbul University Project No. 30315 and TUBITAK 1001 Project No. 114Z829.

## 1. GİRİŞ VE AMAÇ

Yarık el/ayak bozukluğu (SHFM, Split Hand/Foot Malformation) eşiği bütün falanksların eksikliği ve metatarsalın kısmi eksikliğidir. Orta parmakların eksikliği el ve ayakta sıklıkla yarık görünümü yaratır. SHFM prevalansı her 100.000 doğumda 1 ila 9 arasındadır ve konjenital uzuv yetersizliklerinin %1-5'ini oluşturur.

Gelişimsel bir hastalık olan SHFM'nin altta yatan nedenlerinin araştırılması embriyonun gelişimini sırasında meydana gelen genetik olayların anlaşılmasına olanak sağlar.

Yeni nesil dizileme ve dizilim teknolojileri ile genetik hastalıkların etiyolojileri anlaşılmaya başlamıştır. Ancak yeni nesil teknolojilerin getirdiği verinin analizi ve yorumlanması yüksek hacminden dolayı zorlu olmuştur. Bu zorlukları aşmak için tez çalışmasında veri işleme adımlarının otomasyonu ve standartizasyonu için yazılım paketleri geliştirilmiştir.

Tez çalışmasında da SHFM hastalığından etkilenmiş bireylerin bulunduğu, akraba evlilikleri yüksek olan dört ailede toplam 41 birey, yeni nesil genomik ve moleküler genetik yöntemleri ile incelenmiştir. Bu yöntemlerin sonrasında yapılan analizler ile SHFM'den sorumlu olan olası gen ya da gen bölgesinin tespit edilmesi amaçlanmıştır.

## 2. GENEL BİLGİLER

### 2.1. Yeni Nesil Dizileme

Yeni nesil dizileme (YND); genom dizileme, genom yeniden-dizileme, transkriptom profillemesi (RNA-seq), DNA-protein etkileşimleri (ChIP dizileme), ve epigenom tasvifi kapsayan yüksek hacimli bir dizileme metodudur<sup>1</sup>. Yeniden dizileme, tek bir bireyin genomu türün diğer bireylerinde var olan genom varyasyonunu yansıtmayacağı için gereklidir.

Canlılarda özelliklerin nesilden nesile aktarımının aydınlatılması<sup>2</sup> ve DNA molekülünün yapısının anlaşılması<sup>3</sup> üzerine, nesilden nesile aktarılan bu genetik bilginin DNA üzerinde nasıl kodlandığı bir sonraki keşif konusu oldu. Kalıtılan özelliklerin tezahürünü; ebeveynlerden çocuklara aktarılan, DNA molekülünü oluşturan adenin (A), timin (T), sitozin (C), ve guanin (G) nükleotitlerinin dizilimleri belirler. Başta zincir sonlandırma metoduna dayanan Sanger dizileme<sup>4</sup> olmak üzere, bu nükleotit dizilimini gösterecek farklı enzimatik<sup>5,6</sup> ve kimyasal<sup>7</sup> yöntemler 1970 ortaları ve sonlarında geliştirilmeye başlandı. Bu dizileme metodlarının en etkili ve yaygın olarak kullanılanı olan Sanger dizileme projeleri, tek genleri ya da 5000 civarı çok küçük genomların dizilenmesini hedefledi. Sanger dizileme ile daha büyük genomların dizilenmesi, zaman alan ve masraflı bir işlemdi ve İnsan Genom Projesinin öncü olduğu büyük genomik projelerin gereksinimleriyle 2005 yılında ilk çok yüksek sayıda paralelli DNA dizileme metodu geliştirildi<sup>8</sup>.

Düşük maliyetli dizilemeye artan talep, dizileme işlemini paralelleştirerek aynı anda milyonlarca dizi üretebilen yüksek hacimli dizileme teknolojilerin geliştirilmesine ön ayak olmuştur. Yüksek hacimli dizileme teknolojileri, DNA dizileme maliyetini, standart zincir sonlandırma metodlarında mümkün olabileceğinden daha da azaltmayı hedefler. Çok yüksek hacimli dizileme sistemlerinde 500.000 dizileme işlemi aynı anda yapılabilir<sup>9</sup>. Bu gibi teknolojiler bütün insan genomunun bir gün gibi kısa bir sürede dizilenebilmesine imkan sağlar.

Dizileme için Illumina, Roche, Applied Biosystems ve yenileri de gelen farklı firmalar tarafından geliştirilen ve farklı teknolojilere dayanan platformlar mevcuttur<sup>10</sup> (Tablo 2-1).

**Tablo 2-1: Yeni nesil platformlarında kullanılan metotlar ve azami çıktıları.**

Enstrüman	Amplifikasyon Metodu	Dizileme Metodu	Azami Çıktı
<b>Roche GS-FLX</b>	Emülsiyon PZR	Pirodizileme	1,5M okuma/koşum(1 gün) 500 baz/okuma
<b>Life Technologies Ion Torrent</b>	Emülsiyon PZR	PostLight dizileme	100k okuma/koşum(2 saat) 100 baz/okuma
<b>Life Technologies SOLiD 5500XL</b>	Emülsiyon PZR	Ligasyon ile dizileme	4G okuma/koşum(7 gün) 75 baz/okuma
<b>Illumina HiSeq 2000</b>	Köprü amplifikasyonu	Sentez ile dizileme	6G okuma/koşum(8 gün) 100 baz/okuma
<b>Pacific Biosciences</b>	Tek molekül	Gerçek zamanlı dizileme	100k okuma/koşum(2 saat) > 1000 baz/okuma

Bu platformlardan en yaygını Illumina tarafından geliştirilen ve sentez ile dizileme yöntemine dayanan dizileyicilerdir. Sentez ile dizileme metodunda, adaptörler ile sabitlenmiş dizilemeye hazırlanan DNA kütüphanesi, denatüre edilerek, tek iplik olarak akış-hücrelerine aşılır. Sonrasında köprü amplifikasyonu ile klonal DNA parçaları içeren kümeler oluşturulur. Dizilemeden önce kütüphane düzleştirme enzimi<sup>11</sup> yardımı ile tek ipliğe ayrılır. Tek ipliğe ayrılmış kütüphane üzerine eklenen; kesilebilir, birbirinden farklı florışıl boya ve çıkarılabilir blok bulunan dört çeşit nükleotit (ddATP, ddGTP, ddCTP, ddTTP) ile basamak basamak sentez gerçekleşirken, nükleotitlerden gelen florışı sinyalleri cihaz tarafından yakalanır<sup>12</sup>. Aynı bölgeye düşen DNA parçalarının birden fazla dizilenmesi ile tek bir bölge yüksek derinlikte okunabilir<sup>13</sup>.

Bu yüksek hacimli dizileme yöntemleri, düşük maliyetli olmaları ve yüksek hızda çalışabilmeleri nedeniyle birçok araştırma alanında, klinik testlerde ve hatta direkt halka sunulabilen bir ürün olarak uygulama bulmuştur. YND teknolojisi ile beraber genetiğin çalışıldığı alanlar ve çalışılan örnek sayısı da yüksek hacimlere

ulaşmıştır. Dizilemedeki bu üssel artış ile ortaya çıkan devasa genomik veri beraberinde bu verinin işlenmesinin standardizasyonu ve otomasyonu gereksinimlerini getirmiştir.

### 2.1.1. YND Verisinin İşlenmesi

Biyoenformatik, YND'de verinin üretilmesi, analizi, saklanması ve yorumunda önemli bir yer tutmaktadır. Bu veri işleme adımları; dizilerin üretimi, dizilerin işlenmesi ve sonuçların yorumlanması olarak üç gruba ayrılabilir.

**Dizi üretimi;** dizileme cihazından gelen sinyallerin çözümlenmesi, kalite skorlarının verilmesi ve sinyallerin dizi işleminde kullanılabilecek FASTQ formatına çevrimini içerir. Bu adımda kullanılan yazılımlar dizileme cihazını üreten şirketler tarafından geliştirilmektedir.

**Dizi işlenmesi,** üretim basamaklarından gelen FASTQ formatındaki okumaların; referans genoma hizalanması, varyant çağırma öncesi ön-işleme, varyant çağırma ve yanlış pozitif varyantların filtrelenmesini içerir. Bu adımlarda her adım için geliştirilen farklı yazılımlar mevcuttur. Bu yazılımlar farklı kombinasyonlar ile bir araya getirilerek farklı analiz akış hatları elde edilebilir.

**Sonuçların yorumlanması,** diziden elde edilen bilgiyi biyolojik veya klinik bir çerçeve içerisinde anlamlandırdığı için en önemli adımdır. Bu grup; sonuçlara ek açıklamaların (anotasyon) getirilmesi, sonuçların doğrulanması, yorumlanması ve raporlanması adımlarından oluşur<sup>14</sup>.

### 2.1.2. YND Verisinin Bilgisayar Ortamında Temsili

YND verisinin, bilgisayar ortamında farklı temsilleri bulunabilir. Ancak bunlar arasında en yaygın kullanılanları, dizi ve varyant çağırma verisinin saklandığı düz metin dosyaları ve hizalama verisinin saklandığı BAM formatındaki dosyalardır.

### 2.1.2.1. FASTQ Dosya Formatı

FASTQ, dizilerin ve phred kalite değerlerinin bir arada, öz ve yoğun bir biçimde depolandığı dosya formatıdır. FASTQ formatı, düz metin dosyası olup her okuma bloklar halinde belli kurallar ile yazılır (Tablo 2-2).

**Tablo 2-2: FASTQ formatında bir okuma bloğunun yazım kuralları**

Sıra	Satır Anlamı	Düzenli İfadeler ile Satır İçeriği
1	Tanımlayıcı Satır	@[A-Za-z0-9_.-:~]+
2	Dizi Satırı	[A-Za-z\n\.\~]+
3	İkinci Tanımlayıcı Satır	+[A-Za-z0-9_.-:~]+
4	Kalite Değerleri Satırı	[!~\n]+

Üçüncü satır ‘+’ işaretinden sonra tanımlayıcı bir ifade içerebilir. Bu ifade ilk satırdaki ‘@’ işaretinden sonra yazılan ifade ile aynı olmalıdır. İlk satır bu ifadeyi kesinlikle içermelidir. Dizi ve kalite satırlarının uzunlukları birbirine eşit olmalıdır. Her kalite değeri kendisi ile aynı sırada gelen nükleotidin phred kalite skorunu ifade eder.

### Phred Kalite Skorları

Phred kalite skoru, aynı sıradaki bazın okunmasında hata olma olasılığını gösteren değerdir. Verilen hata olasılığı  $P$  (1-2) için phred kalite skoru  $Q$  (1-1) aşağıdaki formül ile hesaplanabilir<sup>15</sup>.

$$Q = -10 \log_{10}(P) \quad (1-1)$$

Tersi,

$$P = 10^{(-Q/10)} \quad (1-2)$$



### 2.1.2.2. SAM ve BAM Dosya Formatları

#### SAM Dosya Formatı

SAM, “Sequence Alignment/Map (Dizi Hizalama/Harita)” temsil eder. SAM formatı, ihtiyari olarak bir başlık ve hizalama bölümleri bulunduran ‘\t’ karakteri ile ayrılmış düz metin dosyasıdır. Eğer başlık bulunuyorsa hizalamalardan önce gelmelidir. Başlık satırları ‘@’ işareti ile başlarken hizalama satırları başlamaz.

Her hizalama satırı, haritalama pozisyonu gibi on bir zorunlu alan ve esneklik için farklı sayıda ihtiyari satırlar ya da hizalayıcıya özgü bilgiler içerebilir. Bu özellikler SAM/BAM dosyalarının 1.6’ncı versiyonları için geçerlidir. Her SAM/BAM dosyası versiyon numarasını ‘@HD VN’ etiketi ile belirtebilir.

#### BAM Dosya Formatı

BAM, SAM dosyasının BGZF formatında sıkıştırılmış halidir. BAM formatı, normalde sabit disk üzerinde büyük yer kaplayan SAM formatının saklanması, taşınması ve işlenmesini kolaylaştırmaktır<sup>16</sup>.

### 2.1.2.3. BED Dosya Formatı

BED, “Browser Extensible Data (Tarayıcı Genişletilebilir Veri)” temsil eder. BED formatı, ihtiyari olarak başlık içerebilir. Başlık ‘track’ ya da ‘browser’ kelimeleri ile başlamalıdır. BED satırları üç zorunlu ve ek olarak dokuz ihtiyari, ‘\t’ ya da beyaz boşluk karakteri ile ayrılmış alan içerir. Satırlardaki alan sayısı bütün dosya boyunca tutarlı olmalıdır.

Gerekli olan ilk üç alan, kromozom ve gösterilen özelliğin kromozom üzerindeki başlangıç ve bitiş noktalarıdır<sup>17</sup>.

#### 2.1.2.4. VCF Dosya Formatı

VCF formatı; meta bilgi satırları, bir başlık satırı ve her biri genomik bir pozisyon hakkında bilgi içeren satırlardan oluşan, genelde sıkıştırılmış olarak depolanan düz metin dosyasıdır. Ayrıca her pozisyon için her örneğin genotip bilgisini barındırabilir.

#### Meta Bilgi Satırları

Dosya meta satırları ‘##’ ile başlamalı ve anahtar=değer çifti olmalıdır. VCF dosyasının gövdesinde kullanılan ‘INFO’, ‘FILTER’ ve ‘FORMAT’ alanlarını tanımlayan meta satırların olması önemlidir. Kullanılan VCF formatının versiyonunu belirten ‘fileformat’ bilgisi her zaman ilk satırda belirtilmelidir.

#### Başlık Satır Formatı

Başlık satırı sekiz zorunlu kolonu belirtir. Bunlar sırasıyla, ‘#CHROM’, ‘POS’, ‘ID’, ‘REF’, ‘ALT’, ‘QUAL’, ‘FILTER’ ve ‘INFO’ kolonlarıdır. Eğer genotip bilgisi bulunuyorsa bu kolonlardan sonra bir ‘FORMAT’ başlığı ve istenilen sayıda örnek kolon ismi eklenebilir. Başlık satırındaki kolon isimleri ‘\t’ karakteri ile ayrılmıştır.

#### Veri Satırları

Her veri girdisi için sekiz zorunlu alan bulunur. Bütün alanlar ‘\t’ karakteri ile ayrılmıştır. Bütün durumlarda eksik değerler ‘.’ ile gösterilir. Zorunlu alanlar sırasıyla; kromozom ismi, genomik pozisyon, tanımlayıcı isim, referans baz, alternatif baz, kalite değeri, filtreleme durumu, ek bilgi alanlarıdır<sup>18</sup>.

#### 2.1.3. YND Analizlerinde Akış Hatları

Analiz sırasında biyolojik veri, farklı işlemlerden geçerek, 2.1.2 bölümünde bahsedilen formatlar başta olmak üzere birtakım değişimlere uğrar. Bu değişim işlemlerini gerçekleştiren Unix-uyumlu farklı yazılımlar mevcuttur. Farklı YND

teknolojilerinin farklı alanlarda yaygınlaşması ile çok yönlü ve güçlü akış hattı yazılımlarına gereksinim de artmıştır. Akış hatları, YND analizindeki, zaman alan ve yüklü parametre gerektiren, genoma hizalama ve anotasyon gibi adımların otomatizasyonu ve standardizasyonunu sağlar. Basit bir ekzom akış hattı ham dizi dosyalarından varyant çağırma dosyasına 12 adımda ulaşabilir ancak çoğu durumda varyant sonrası analizler ve raporlama gerektirebilmektedir<sup>19</sup>.

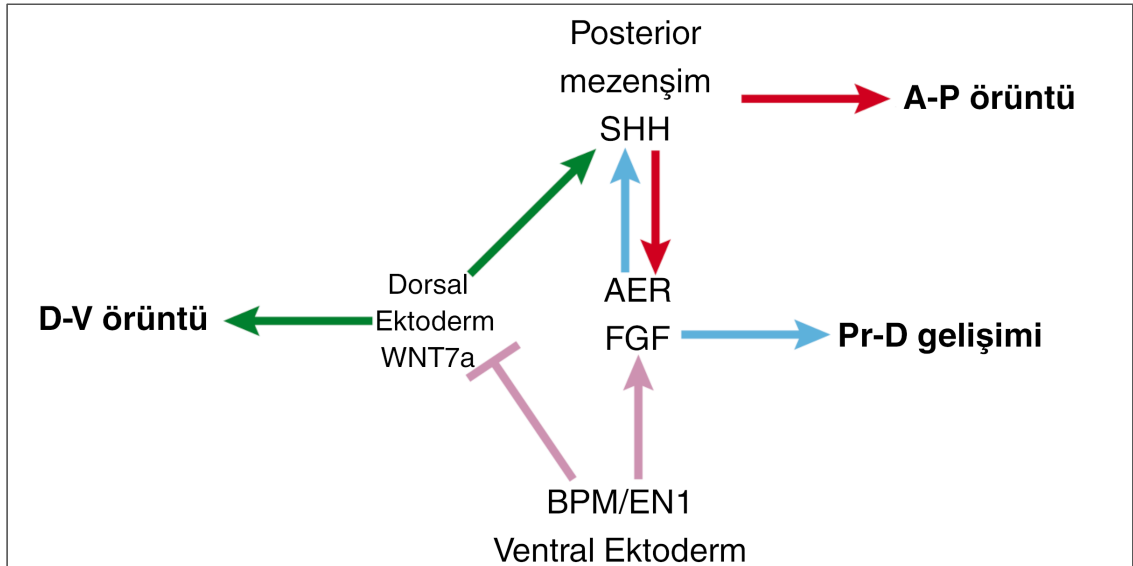
## 2.2. Yarık El/Ayak Bozukluğu

### 2.2.1. Uzun Gelişimi

Uzun gelişimi uzaysal ve zamansal olarak kontrol edilen, hücrel farklılaşma, örüntü oluşumu, ve doku morfogenezi olaylarının arasındaki etkileşime dayanır. Mutasyonlar ve teratojenler uzun morfolojisi üzerinde, hayatta kalma ve üreme üzerinde etkisi olmadan, ağır malformasyonlar yaratabileceği için insanda gözükten en yaygın konjenital anomalilerden biri uzun bozukluklarıdır. Bu durum sonucunda geniş yelpazeli bir uzun varyasyonu insan popülasyonunda gözlenebilir.

Omurgalı uzun gelişimi, uzun tomurcuğu oluşumu, erken uzun örüntülenmesi ve geç uzun morfogenezi olmak üzere üç fazdan meydana gelir<sup>20</sup>. Uzun tomurcuğu oluşumu fazında, uzun bir kat ektoderm hücreleri ile kaplı, morfolojik olarak homojen mezenşim hücrelerinden oluşan küçük bir tomurcuk olarak embriyonun yanından, lateral plaka mezoderminden çıkar<sup>21</sup>. İnsanda, uzun tomurcuğu kollar için 26. ve 28. günlerde ve bacaklar için ise iki gün sonrasında meydana gelir<sup>22</sup>. Erken uzun örüntülenmesi fazında, uzun özgün modellenmesi, fibroblast büyüme faktörleri (FGFs, Fibroblast Growth Factors), kemik morfogenetik proteinleri (BMPs, Bone Morphogenetic Proteins), WNT ve hedgehog sinyal molekülleri ve homeobox proteinleri gibi çeşitli sinyal moleküllerinin ve transkripsiyon faktörlerinin uzaysal ve zamansal anlatımı ile koordine edilir(Şekil 2-1)<sup>23</sup>. Altta yatan mezenşimal hücrelerden kaynaklanan indüktif sinyal ile uzun tomurcuğunun en üst noktasındaki ektodermal hücreler farklılaşarak apikal ektodermal sırtı (AER, Apical Ectodermal Ridge) meydana getirir. AER, alt-yan mezenşim hücrelerinin -gelişim alanı (PZ, progress Zone)- farklılaşmadan çoğalmalarına devam etmeleri için FGF anlatımı ve salgısı yapar. AER

ve PZ arasındaki etkileşim, uzvun proksimo-distal (Pr-D, omuz-parmak) polaritesini belirler. AER'den salgılanan FGF aynı zamanda, antero-posteriör (A-P, baş parmak-serçe parmak) polaritesini kontrol eden posterior mezenşimdeki polarizasyon aktivite alanında (ZPA, Zone of Polarizing Activity) sonic hedgehog (SHH) anlatımının devamlılığını sağlar<sup>24</sup>. Dorso-ventral (D-V, elin tersi- avuç) polarizasyonu, ventral ektodermden BMPs ve Engrailed1 ve dorsal ektodermden WNT7a, AER olmayan ektoderm sinyalleşmesi ile sağlanır. Geç uzuv morfogenezini sırasında uzuv, iskelet, kaslar, tendonlar, damarlar gibi işlevsel organlara olgunlaşır.

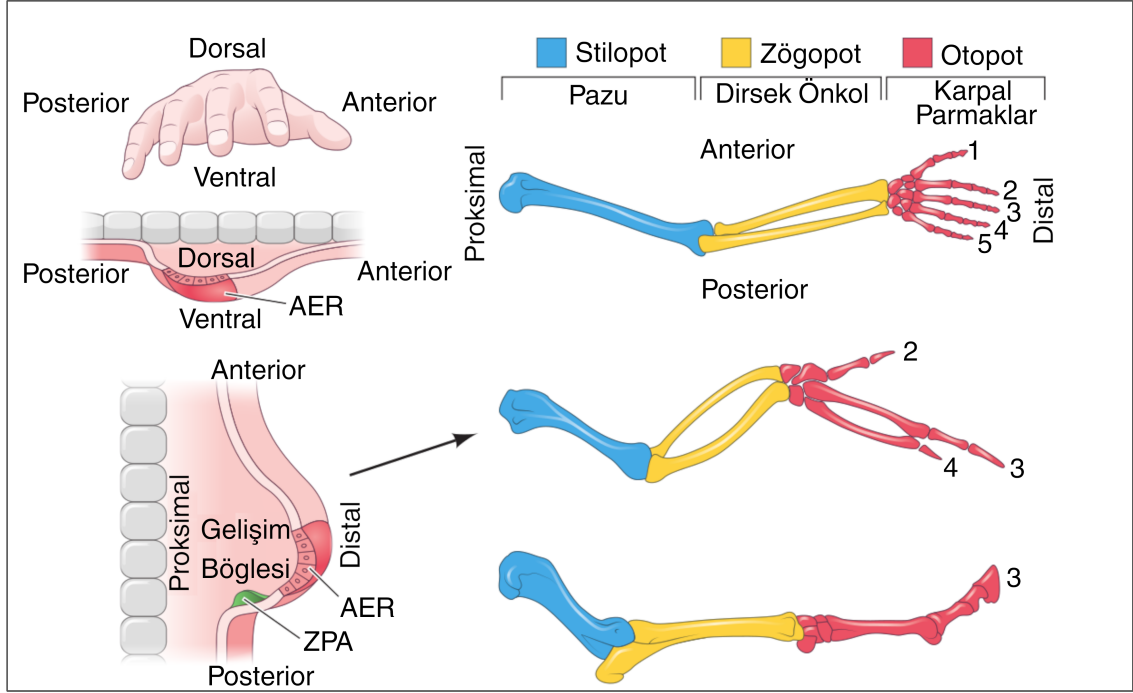


**Şekil 2-1: Üç uzuv ekseninde, erken uzuv örüntülenmesini yöneten moleküler etkileşimler.**

Pr-D eksenini AER'den salgılanan FGF kontrolünde, A-P eksenini posterior mezenşimden salgılanan SHH kontrolünde, D-V eksenini ventral ektodermden salgılanan BMPs ve EN1 ve dorsal ektodermden salgılanan WNT7a kontrolündedir.<sup>21</sup>

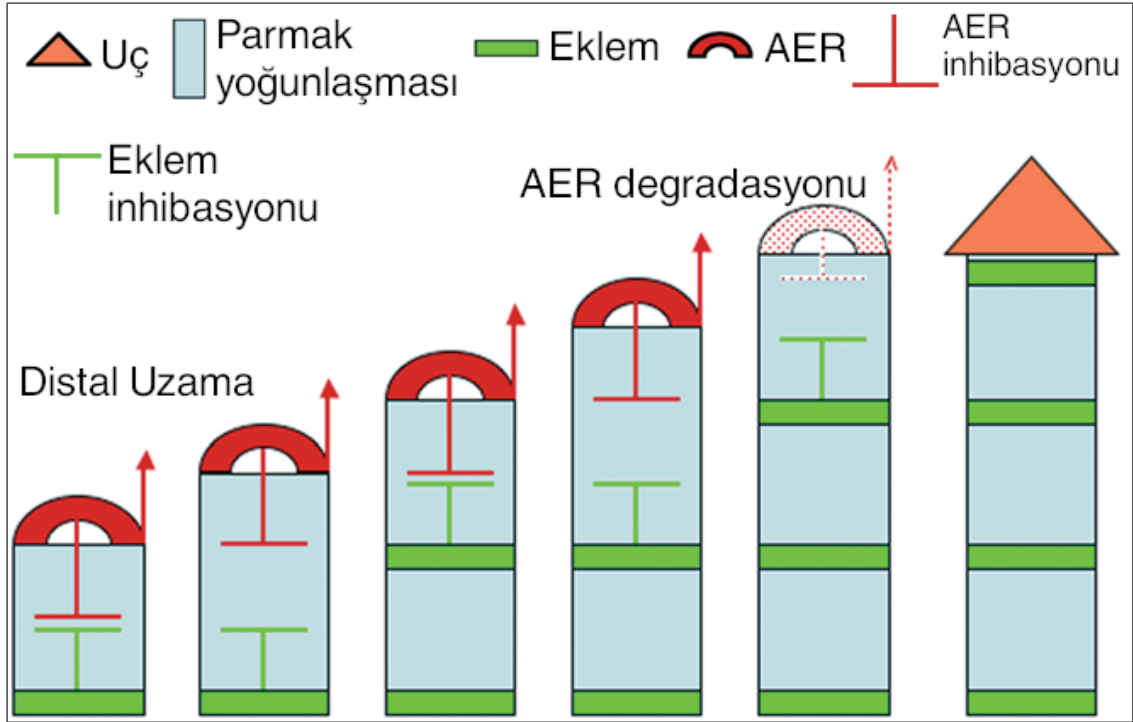
Omurgalı uzvunun iskeleti, stilopot (pazu ya da uyluk), zögopot (ön kol ya da baldır) ve otopot (el ya da ayak) olmak üzere üç temel yapıdan oluşur (Şekil 2-2). Bu bölümlenme Pr-D ekseninde koordine olur. PZ'de bulunan mezenşimal hücreler, vücudun yakınından uzağına doğru uzuv oluşumunu başlatır. Bu durumda ilk örüntülenen stilopot, sonrasında zögopot ve en son olarak otopottur. Parmaklar ise

örüntüsü belirlenen en son yapılardır. Uzun gelişiminin ucunda olan parmaklar A-P ekseninde oluşmaktadır ve zamanla parmaklar arası boşluklar ile ayrılırlar.



Şekil 2-2: Omurgalı uzvunun iskeletinin üç temel yapısı ve gelişimi.<sup>127</sup>

AER başta olan segmentasyon modelinde, AER'nin hem aktive edici hem de baskılayıcı rolü bulunur (Şekil 2-3). FGF sinyali ile parmakların distal doğru büyümesini sağlarken aynı zamanda distal eklem oluşumunu baskılar. Aynı zamanda AER'nin varlığı parmak ucu oluşumunu engeller. Eklem oluşumu ancak hücreler bu ve daha önceden oluşmuş eklemün inhibasyonundan çıktığı vakit gerçekleşir. Her iki inhibasyondan da kaçıldığında muhtemelen Wnt14 tarafından indüklenen eklem belirteçlerinin anlatımı olur. AER'den gelen FGF sinyalinin kesilmesi üzerine son eklem distal bölgede oluşur ve uçta son falanksın oluşumunu sağlayacak moleküler program indüklenir<sup>25</sup>.



Şekil 2-3: Parmakların oluşumu ve segmentasyonu modeli.<sup>25</sup>

### 2.2.2. Yarı El/Ayak Bozukluğu Tanımı ve Sıklığı

Yarı el/ayak, el ya da ayaklarda 1. ve 5. parmaklar hariç, parmaklardaki uzunlamasına eksikliklerdir.

Yarı el/ayak bozukluğu (SHFM, Split Hand/Foot Malformation) eşiği bütün falanksların eksikliği ve metatarsalın kısmi eksikliğidir. Orta parmakların eksikliği el ve ayakta sıklıkla yarı görünümü yaratır. Ektrodaktili (Yunanca ektroma “anormallik” ve daktylos “parmak”), yarı el/yarı ayak için sıkça kullanılan bir terim olmasına rağmen falanks hipoplazisi ve ampütasyon için daha uygundur ve yarı el/yarı ayağı tanımlamasında kullanılması tercih edilmez<sup>26</sup>.

SHFM prevalansı her 100.000 doğumda 1 ila 9 arasındadır ve konjenital uzuv yetersizliklerinin %1-5’ini oluşturur<sup>27-29</sup>.

### 2.2.3. SHFM Genetiği

Klinik olarak heterojen olan SHFM, sendromik ya da izole (non-sendromik) gözükabilir ve aile içerisinde bile fenotipik olarak heterojenite gösterebilir. SHFM fenotipik olduğu gibi genetik açıdan da yüksek heterojenite göstermektedir (Tablo 2-3). Tek gen mutasyonları ya da kromozomal yeniden düzenlemeleri sonucu ortaya çıkan 75'den fazla sendromik ya da sendrom ile ilişki SHFM tanımlanmıştır<sup>30</sup>.

**Tablo 2-3: SHFM OMIM fenotipler ve gen/lokusları**

Lokasyon	Fenotip	Kalıtım	Gen	Lokus
2q31	SHFM5		<i>SHFM5</i>	-
3q28	SHFM4	OD	<i>TP63</i>	-
7q21.2-q21.3	SHFM1	OD	-	SHFM1
7q21.3	SHFM1D, SHFM1, işitme kaybı ile	OR	<i>DLX5</i>	-
10q24	SHFM3, gen duplikasyon sendromu	OD	-	SHFM3
12q13.12	SHFM6	OR	<i>WNT10B</i>	-
Xq26	SHFM2	XL	-	SHFM2

### 2.3. Genom Boyu Genotipleme

Genom boyu genotipleme (GBG), genom üzerindeki polimorfik lokusların allelik bakımdan sınıflandırılması ile genomun tamamına genel bir bakış sağlar. Günümüzde genom boyu genotipleme mikrodizilim ya da YND ile yapılabilir. GBG, birden fazla genom arasındaki küçük varyasyonların karşılaştırılması sağlayan bir yöntemdir. Klinikteki en önemli uygulaması hastalık yatkınlığı ve kişiye özel ilaç direncinin belirlenmesidir. Aynı şekilde GBG, büyük kohortlarda, boy, zeka gibi farklı fenotiplerin genotip ile ilişkilendirilmesi amaçlı kullanılır. Nadir görülen fenotiplerinde belli bir genotip ile ilişkilendirilmesi amacıyla GBG verisi, genetik bağlantı analizlerinde kullanılabilir.

### 2.3.1. SNP dizilimi

İnsan genomunun varyasyonlarının detaylı bir şekilde haritalanması ve teknolojik SNP'lerin yüksek güvenilirlikte çağrılabilmesi yüksek çözünürlükte GBG çalışmalarını mümkün kılmıştır<sup>31</sup>. Çalışmada, Illumina'nın boncuk çipli SNP dizileme platformu kullanılmıştır. Illumina SNP çipleri, her biri tekrarlı, rastgele düzenlenmiş boncuklar bulunan dikdörtgen şeritlerden oluşur. Genotiple için bu boncuklar, SNP bölgesine homoloji gösteren 50mer proplar içerir ve iki allel (A, B) kırmızı ve yeşil boya kullanılarak ayırt edilir<sup>32</sup>.

Çip içerisinde bulunan bir SNP'in, allel spesifik propların renk ve ışınma şiddetlerinin ölçümlerinden elde edilen sinyaller normalize edildikten sonra, heterozigotluk kaybı (LOH; Loss of Heterozygosity) ve kopya sayısı değişiklikleri (CNV; Copy Number Variation) incelemesinde kullanılacak olan B allel frekansı (BAF) ve LogR kesiri (LRR, Log R Ratio) metrikleri hesaplanır.

#### 2.3.1.1. BAF ve LRR

Total hidridizasyon yoğunluğunu (R), referans seviyeye ( $R_0$ ) bağlı ifade eden  $\log_2 R/R_0$  oranı (LRR) ve A ve B bağlı sinyal değerlerini ifade eden B allel frekansı (BAF) SNP dizilimi verisinin analizinde kullanılan standart metriklerdir<sup>33</sup>. BAF,  $N_B$ , B alleli ve  $N_A$  A alleli olmak üzere,  $N_B/(N_A+N_B)$  değerlerinin bir yaklaşımıdır ve sırasıyla AA, AB ve BB'nin karşılığı olan 0, 0,5 ve 1 olarak standart üç gruba ayrılır<sup>34</sup>.

#### 2.3.1.2. CNV Tespiti

CNV analizleri GenomeStudio yazılımına ek olarak geliştirilen CNVpartition yazılımı tarafından gerçekleştirilir. Bu yazılım CNV'lerin genelinin 0 ve 4 kopya arasında olduğunu varsayarak, homozigot delesyon, heterozigot delesyon, dizigot (normal durum), trizigot (bir ek kopya), ve tetrazigot (iki ek kopya) olmak üzere beş seçenek bulundurur. CNVpartition, LRR ve BAF değerlerini, basit iki değişkenli Gauss dağılımı olarak mümkün olan on dört kopya genotipi için modeller. Her çalışılan her



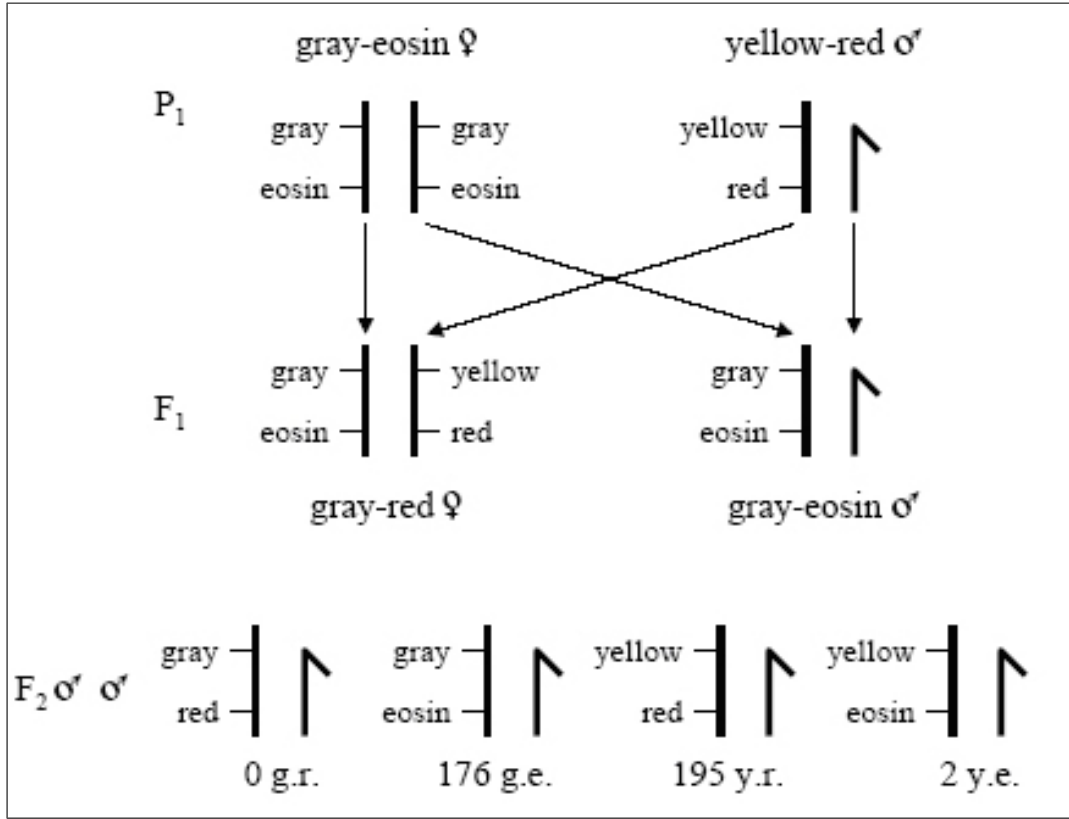
lokus için, o lokusun LRR ve BAF değerleri modellenen on dört genotipin değerleri ile karşılaştırılarak, ön kopya sayısı hesaplanır<sup>35</sup>.

### 2.3.2. Homozigot Haritalama

Homozigot haritalama, bir çiftin, çocuklarına geçirdikleri, aynı kökenden gelen özdeş kromozom parçalarının incelenebileceği bir yöntemdir. Çocuk, her iki bireyden özdeş parçasını aldığı kromozom bölgesi için otozigottur. Otozigotluk akraba evliliği olan aileler içerisinde birden fazla bireyin genetik verisi mevcut ise tespit edilebilir<sup>36</sup>. Otozigotluğun incelenmesi için bağlantılı belirteçler setler halinde ele alınmalıdır. Beraber kalıtılan bu belirteç setinin tek bir homolog kromozom üzerindeki genotip haritası haplotip olarak isimlendirilir<sup>37</sup>. Homozigot haritalama, çoğunlukla akraba evliliklerinde, otozigot hastalık bölgenin bulunmasını, komşu bölgelerin de homozigot olacağı nedeniyle, ebeveynlerden gelen bölgelerin haplotiplerinin karşılaştırması ile gerçekleştirir<sup>38</sup>. Bu bölgelerin Mendelyen resesif kalıtım modeline göre, ailedeki sağlıklı bireylerin heterozigot ya da wt homozigot, hasta bireylerin ise mt homozigot olması beklenir.

### 2.4. Bağlantı Analizi

Genetik bağlantı, aynı kromozom üzerinde bulunan iki lokusun, bir sonraki kuşağa beraber aktarılmasıdır. Mendel'in bezelye deneyleri sırasında aralarında bağlantı bulunmayan özelliklere denk gelmesi ile "bağımsız dağılım" ilkesine varmıştır<sup>39</sup>. Ancak Morgan'ın *D. melanogaster* deneyleri ile birlikte fiziksel olarak birbirine yakınlığı arttıkça lokusların karşılıklı parça değişimi sırasında farklı kromatitlere dağılmadığı ve beraber kalıtılma olasılığının arttığı görülmüştür<sup>40</sup> (Şekil 2-4). Birbirine fiziksel olarak yakın lokus arasında rekombinasyon olasılığı düşük olacağından, beraber kalıtılma olasılığı yüksek olan iki lokus *bağlantılı* olarak isimlendirilir. Farklı kromozomlar üzerindeki lokuslar tam olarak *bağlantısızdır*.

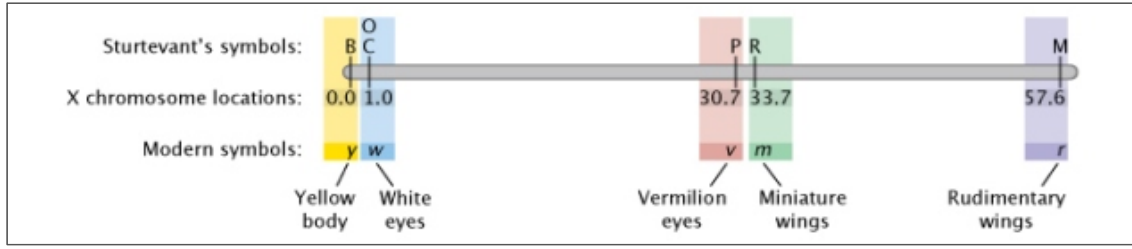


**Şekil 2-4: Karşılıklı parça değişimi ile rekombinant gamet oluşumu**<sup>128</sup>.

Sturtevant'ın gri-eozin dişi ve sarı-kırmızı erkek ebeynlerden elde edilen bireyleri çaprazlaması. g.e. gray-eosin, g.r. gray-red, y.r. yellow-red, y.e. yellow-esoin.

Nature Education'da belirtildiği üzere şekil, üzerinde herhangi bir değişiklik yapmadan kullanılmıştır ve şeklin tezde kullanılması uygundur. (<https://www.nature.com/scitable/topicpage/thomas-hunt-morgan-genetic-recombination-and-gene-496>)

Morgan'ın öğrencisi Sturtevant, *D. melanogaster* çalışmalarında, bağlantı lokusların doğrusal uzaklıklarının düzenlenmesi ile bağlantı haritasını çıkardı<sup>41</sup>(Şekil 2-5).Bağlantı haritasının yapımında iki özellik arasında 0,01 ya da %1 rekombinasyon frekansını gösteren “harita birimi” daha sonradan santimorgan (cM, centimorgan) olarak isimlendirildi.



**Şekil 2-5: Sturtevant'ın *D. melanogaster* gen haritası<sup>128</sup>.**

Sturtevant'ın gen haritasında, 6 özellik doğrusal kromozom üzerinde B fenotipine olan bağlı uzaklıklarına göre yerleştirilmişlerdir. Bu fenotipler, sarı vücut(B), beyaz gözler(C,O), kırmızı gözler (P), küçük kanatlar(R) ve gelişmemiş kanatlardır(M)

Nature Education'da belirtildiği üzere şekil, üzerinde herhangi bir değişiklik yapmadan kullanılmıştır ve şeklin tezde kullanılması uygundur. (<https://www.nature.com/scitable/topicpage/thomas-hunt-morgan-genetic-recombination-and-gene-496>)

İki lokus arasındaki rekombinasyon frekansı, aynı zamanda 0 ile 1 arasında olasılıksal değer alan rekombinasyon kesiri " $\theta$ " olarak da isimlendirilir.  $\theta$ , 0 ile 1 arasında değer alabilmesine karşın kromozom üzerinde Poisson sürecine bağlı birden fazla karşılıklı parça değişimi olayı meydana geldiğinden dolayı gerçekte hiç bir zaman, aynı kromozomlar için rekombinasyon kesiri olan  $\frac{1}{2}$  oranını geçmez<sup>42</sup>.

Bağlantı analizinin amacı, bir aile içerisinde ilgilenilen bir ya da birden fazla fenotipin ortaya çıkması ile ilişkili genetik bir mekanizmanın varlığını ya da yokluğunu aydınlatmak ve eğer böyle bir mekanizma var ise bireyleri genotiplerine göre sınıflandırmaktır<sup>43</sup>. Bağlantının varlığı aile içerisinde gözlenen iki özelliğin ilişkili olması temelindeki istatistiksel analizlere dayanır. Bağlantının varlığı sorgulanırken iki hipotez sınanır; bağlantının olmaması ( $\theta = \frac{1}{2}$ ; sıfır hipotezi) ve belli bir rekombinasyon frekansı ile bağlantı olması ( $\theta < \frac{1}{2}$ ; alternatif hipotez)<sup>44</sup>. İki özellik arasındaki bağlantının varlığı ya da yokluğunu belirleyen kriter, iki özelliğin dağılım modelinin, bağlantı olan durumda ( $\theta = 0$ ) gözlenme ihtimalinin, bağlantı olmayan durumda ( $\theta = \frac{1}{2}$ ) gözlenme olasılığına oranıdır<sup>45</sup> (2-1). Bu olasılık oranı "L" ile gösterilir ve raporlanırken lod skoru diye isimlendirilen 10 tabanlı logaritması verilir.

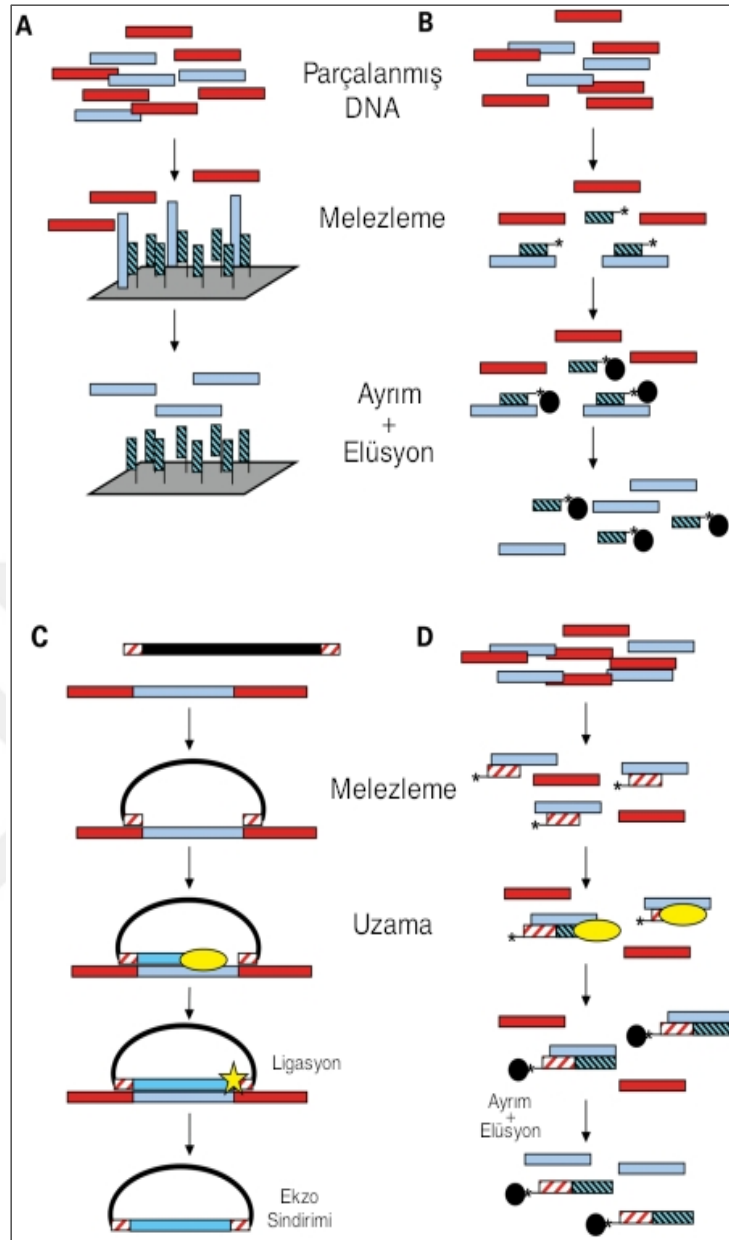
$$LOD = Z = \log_{10} \frac{\text{bağlantılı olasılığı}}{\text{bağlantısız olasılığı}} = \log_{10} \frac{(1-\theta)^{NR} \times \theta^R}{0.5^{(NR+R)}} \quad (2-1)$$

R: Rekombinant NR: Rekombinant olmayan

Bağlantının varlığının kabul edilebilmesi için lod skor en az 3 (olasılık oranı 1000:1) olmalıdır. LOD skoru analizi iki özellik yanı sıra çoklu genetik belirteç kullanılarak gerçekleştirilebilir. Çok noktalı analizde ise fenotip ile ilişkili olabilecek bölgelerin daha kesin bağlantı haritası çıkartılabilir<sup>46</sup>. Bağlantı analizleri, çok miktarda hesap gerektiren analizler olduğu için bu analizleri gerçekleştirecek Allegro<sup>47</sup>, Genehunter<sup>48</sup> gibi bilgisayar programları ve bu programları kullanımını kolaylaştıran EasyLinkage<sup>49</sup> gibi grafik kullanıcı arayüzleri mevcuttur.

## 2.5. Ekzom Dizileme

Ekzom dizileme (ED), Bölüm 2.1’de anlatılan YND yönteminin özelleşmiş ve hedefli bir tipidir. ED, insan genomunun, proteine dönen yaklaşık 30 milyon baz ya da %1’lik<sup>50</sup> kısmını oluşturan yaklaşık 180.000 ekzonu hedefler<sup>51</sup> ve dizileme öncesi bu kodlayan kısmın diğer genomik bölgelerden ayrılması için genom dizilemeden farklı adımlar içerir. Hedeflenen ekzonların yakalanmasında, Katı-faz melezlemesi<sup>52</sup>, sıvı-faz melezlemesi<sup>53</sup>, moleküler inversiyon propları<sup>54</sup> (MIP) ve primer uzama ile yakalama<sup>55</sup> (PEC), gibi metotlar Şekil 2-6’te geliştirilmiştir. Hedeflenen bölgeler daha sonra bir YND platformunda dizilenir.



**Şekil 2-6: Farklı yakalama metodlarının tasviri.**<sup>129</sup>

Mavi bölgeler yakalanmak istenilen, kırmızı ise istenmeyen genomik bölgeleri göstermektedir. A) Katı-faz melezmesi. B) Sıvı-faz melezmesi. C) MIP. D) PEC.

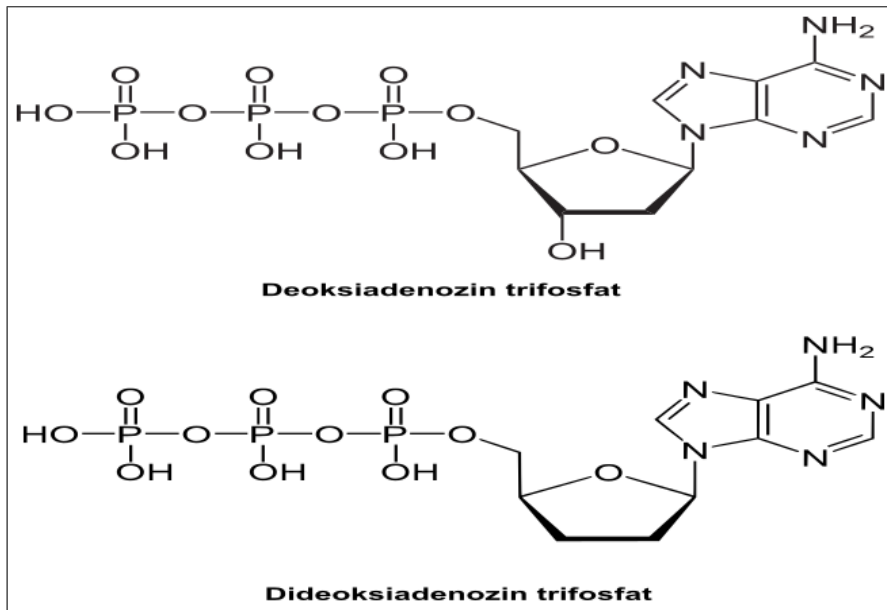
ED, genomun yalnızca %1'ine odaklanmasına rağmen Mendelyen kalıtım gösteren hastalıkların genetik etiyojilerinin aydınlatılmasında büyük rol oynamıştır. Bunun en büyük sebebi genetik kökenli hastalıkların %85'inin, genomdaki protein kodlayan bölgelerde meydana gelen mutasyonlar sonucu ortaya çıkmasındandır<sup>56</sup>. Tek

gene bağılı hastalıklar tek tek nadir olsalar bile, toplu olarak düşünüldüklerinde 5856<sup>57</sup> gibi yüksek bir sayıya tekabül ederler. Bugün itibari ile 10.000'i aşkın gen üzerinde 256070 bildirilmiş varyasyon bulunmaktadır<sup>58</sup>. Gelişmekte olan antisense terapi gibi farklı farmokolojik yöntemlerle klinikte uygulama bulan varyasyon sayısı artmaktadır.

İlk uygulanmasından<sup>51</sup> bu yana, çalışmalarda kısıtlı hasta olmasına rağmen ED ile birçok hastalığın nedeni olarak literatüre bildirilen varyasyon sayısı üssel olarak artmıştır<sup>59-61</sup>. Otozomal resesif hastalıkların altında yatan genetik sorunlar, aralarında akrabalık bulunmayan bireylerin<sup>62-64</sup> ya da aynı aileden bireylerin<sup>65-67</sup> bulunduğu çalışmalarda bazı durumlarda da homozigot haritalama kullanılarak<sup>68-70</sup> tanımlanmıştır.

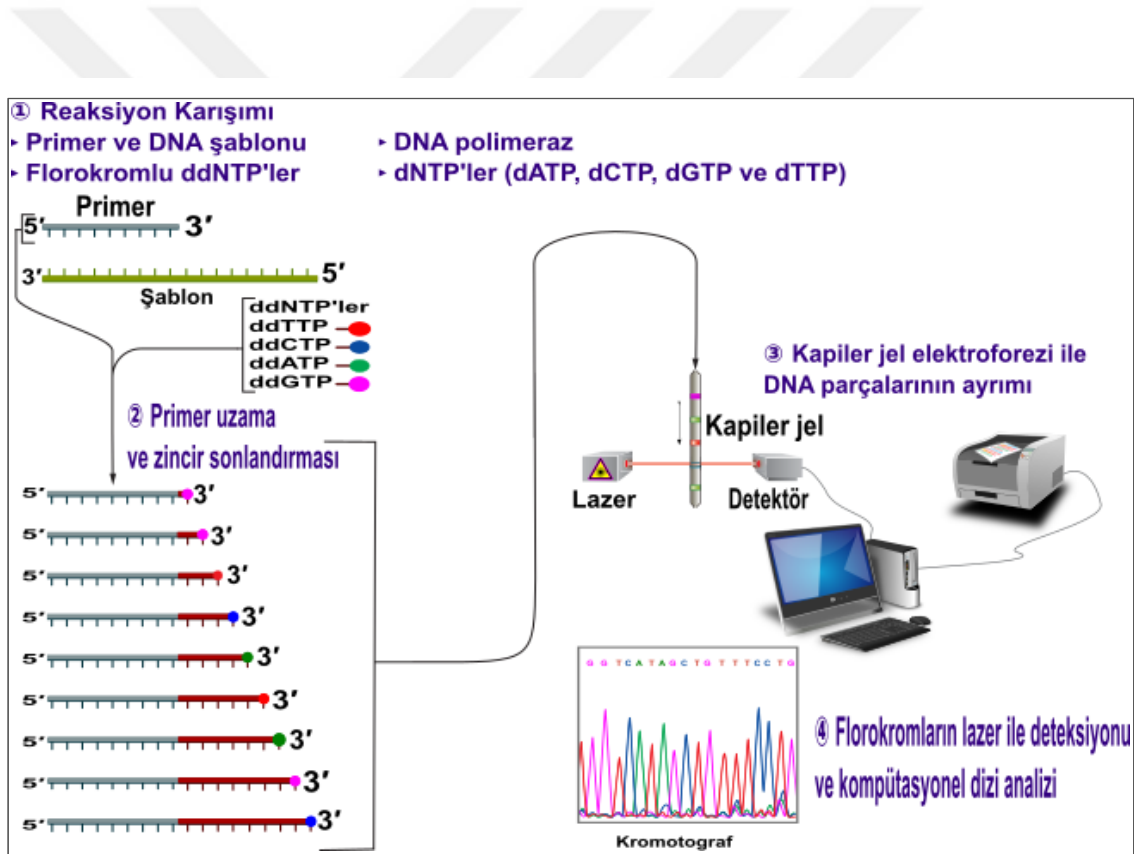
## 2.6. Sanger Dizileme ile Gen Taraması

Sanger dizileme, zincir sonlandırma metoduna dayanan 1977 yılında Sanger tarafından icat edilen<sup>4</sup> ve Applied Biosystems tarafından ticarileştirilen DNA dizilme metodudur. Yöntem PZR ile senteze dayalıdır ancak PZR'den farklı olarak reaksiyon sırasında normal nükleotitler ile beraber florokrom ile işaretlenmiş dideoksi nükleotitler(ddNTP) de kullanılır (Şekil 2-7).



Şekil 2-7: dATP ve ddATP şematik gösterimi.

Dideoksi nükleotiler, polimerazın sentez sırasında fosfodiester bağının kurulmasında kullanılan 3' karbona bağlı OH grubundan yoksundurlar. Bu OH grubunun eksiliği nedeniyle polimeraz bir sonraki nükleotidi ekleyemez ve polimerizasyon son bulur. Her dört bazın ddNTP'si karışım içerisinde bulunduğu için polimerizasyon rastgele farklı noktalarda sonlanarak farklı uzunluklarda DNA parçaları ürün olarak oluşur. Bu DNA ürünleri daha sonra kapiler jel elektroforezi ile ayrılarak florokromların renklerine göre nükleotitler ayırt edilerek dizileme gerçekleştirilir (Şekil 2-8).



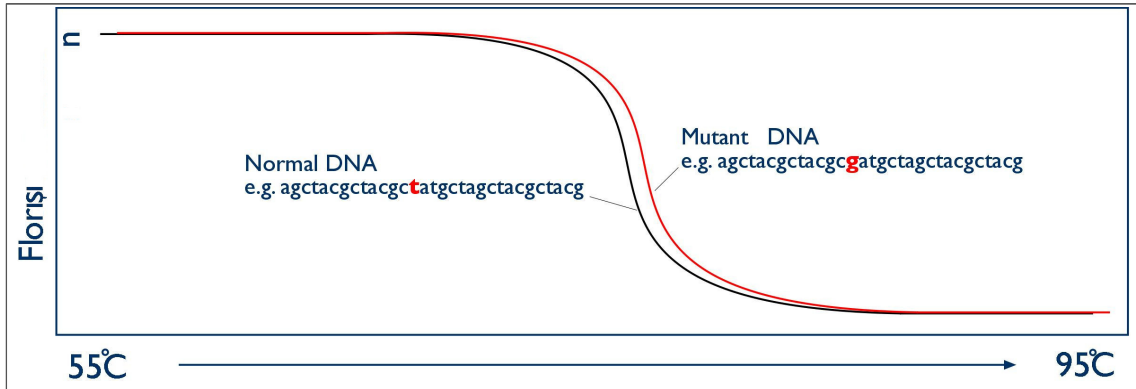
Şekil 2-8: Sanger dizileme basamakları.

Araştırma ve klinik çalışmalarda hala altın standart<sup>71</sup> olarak kullanılan bu yöntem çok uzun olmayan genlerin veya genom üzerindeki küçük bölgelerin dizilenmesi için idealdir. Hedeflenen genlerin ekzonlarına özgü tasarlanan primerler ile RNA ve proteine geçen ya da kırılmaya etki edebilecek varyasyonlar taranabilir.

## 2.7. Gerçek Zamanlı Analizler

### 2.7.1. Yüksek Çözünürlüklü Erime Analizi

Yüksek çözünürlüklü erime analizi (HRMA, High Resolution Melting Analysis), iki iplikli DNA molekülündeki, mutasyonların, polimorfizmlerin ve epigenetik değişimlerin belirlenmesinde kullanılan bir yöntemdir<sup>72</sup>. PZR ile birden fazla örnekte aynı anda genotiplemeyi, dizileme ya da SNP diziline karşın daha hızlı ve uygun maliyetli şekilde gerçekleştirebilir. Genotip ayırımı, florışı DNA boyası ve etiketlenmemiş oligonükleotit propların bulunduğu ortamda gerçekleşen asimetrik PZR ile yapılır<sup>73</sup>. Yöntemin temelinde, DNA'nın ısı ile denatürasyonu yani iki iplikli DNA'nın ısıtılarak bazlar arası hidrojen bağlarının kopması vardır. İncelenen dizilerde bulunan varyasyonlara bağlı olarak değişen florışıma erime eğrilerinin incelenmesi ile farklı genotiplerin ayırımı yapılır (Şekil 2-9).

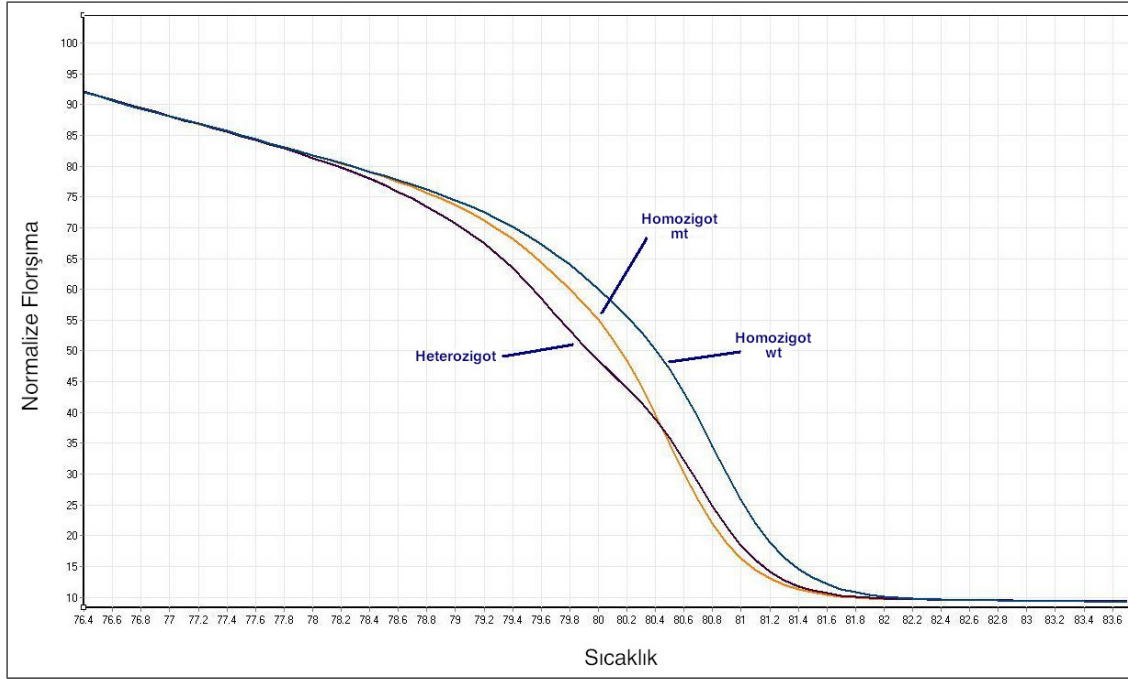


Şekil 2-9: Florışı erime eğrisi ile genotiplerin ayırımı.

Farklı homozigot allel varyantları ayrılırken florışı erime eğrisinin sıcaklığa bağlı kaydığı gözükürken heterozigot durumdaki örneklerin florışı erime eğrilerinin şekillerinin değiştiği gözlenir (Şekil 2-10). Bu yüzden heterozigot genotipin homozigot genotiplerden ayırımı daha verimli şekilde yapılabilirken homozigot wt ve mt genotiplerin kendi içleri içinde ayırımı bazı durumlarda mümkün olmayabilir. Bu



durumlarda çalışılan homozigot DNA'lara, belli bir oranda homozigot wt, sağlıklı kontrol DNA'sı katıştırılarak (spike-in), mt homozigot bulunan tüplerde kısmi heterozigotluk yaratılması ile, homozigot heterozigot ayrımı yapar gibi çalışılabilir.



**Şekil 2-10: HRMA ile homozigot wt, homozigot mt ve heterozigot ayrımı.**

### 2.7.2. qPZR ile CNV Doğrulaması

Kantitatif gerçek zamanlı PZR, gen anlatımını analizlerine benzer bir şekilde kopya sayısı varyasyonlarının kantifikasyonu içinde kullanılabilir. Bu yöntem ilk olarak Charcot-Marie-Tooth hastalığındaki kaybın tespitinden<sup>74</sup> sonra, kistik fibroz<sup>75</sup>, ve nöropati<sup>76</sup> gibi çeşitli birçok Mendelyen hastalığın ve farklı kanser tiplerindeki<sup>77-79</sup> CNV'lerin tespitinde kullanılmıştır. Bu yöntem gerçek zamanlı PZR<sup>80,81</sup> ve PZR sırasında serbest kalan 5' raportör boya ve 3' soğurucu boya bulunduran proplardan yararlanır. PZR'nin ilerleyişi gerçek zamanlı olarak florışma miktarının ölçümü ile ve florışmanın seçilen eşik döngü sayısına ( $C_T$ ) ulaşması ile takip edilir. Farklılıkların kontrolünün sağlanması için çalışma sırasında, genomik DNA girdileri eşit olmalı ve

aynı zamanda kopya sayısı bilinen referans genomik bir bölge ve sağlıklı bir kontrol ile çalışılmalıdır.

Elde edilen verinin analizi karşılaştırmalı  $C_T$  yöntemi ile gerçekleştirilebilir. Bu yöntemde kat değişimi yani CNV değeri  $2^{-\Delta\Delta C_T}$  formülü ile hesaplanır<sup>82</sup>. Bu formülün açılımı formül 2-2'de gösterilmiştir.

$$2^{-\Delta\Delta C_T} = \left[ \frac{(\text{İlgilenilen Genin } C_T \text{ değeri} - \text{İç Kontrol } C_T \text{ Değeri}) \text{ Örnek A}}{(\text{İlgilenilen Genin } C_T \text{ değeri} - \text{İç Kontrol } C_T \text{ Değeri}) \text{ Örnek B}} \right] \quad (2-2)$$

Örnek B referans olarak kullanılan sağlıklı kontrol ve örnek A CNV için test edilen örneklerdir. Aynı şekilde ilgilenilen gen CNV için test edilen gen ve iç kontrol ise CNV olmadığı bilinen referans genidir.

### 3. GEREÇ VE YÖNTEM

#### 4. Gereç

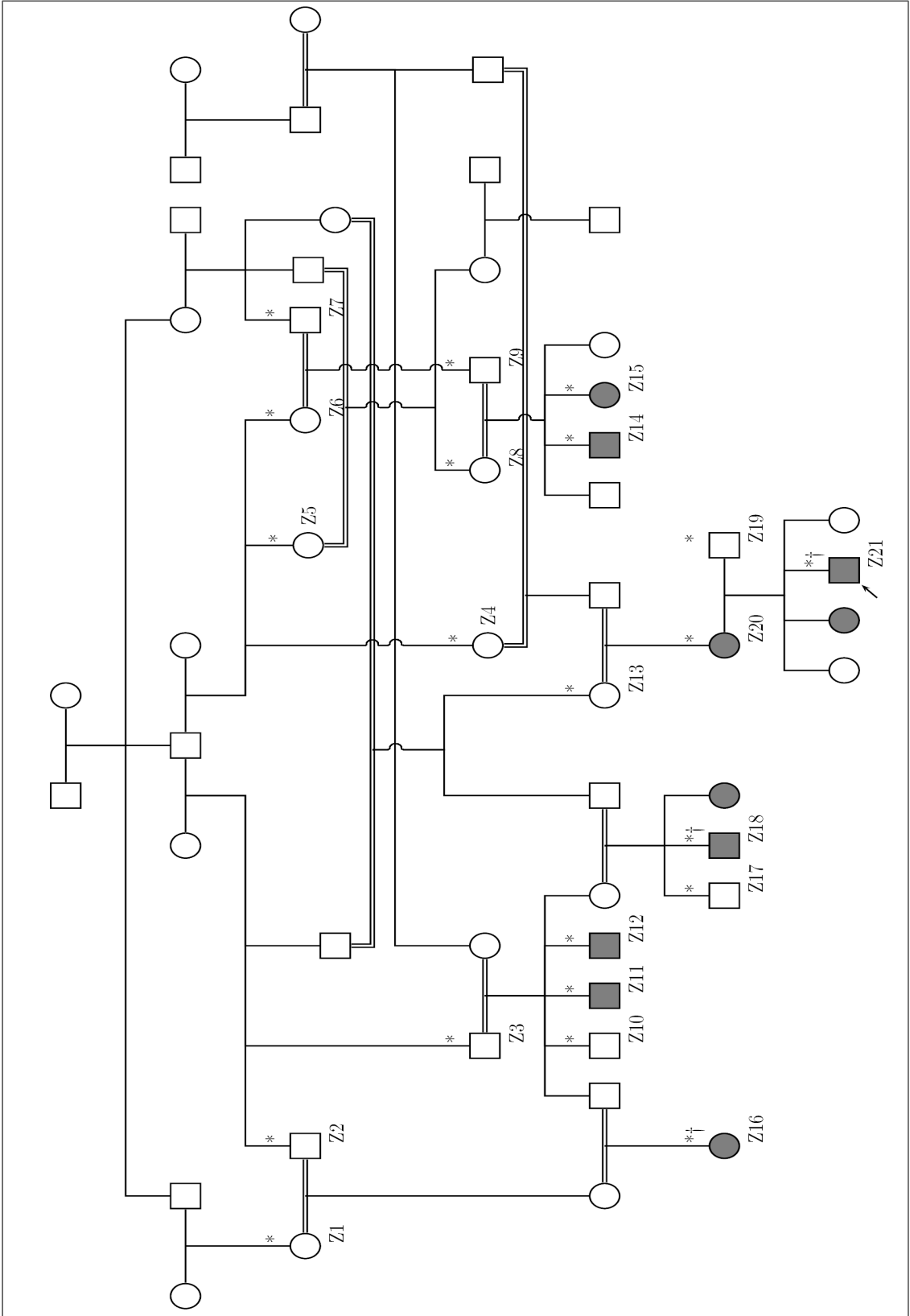
##### 4.1.1. Çalışma Kapsamında İncelenen Aileler

Çalışma, İstanbul Üniversitesi İstanbul Tıp Fakültesi Klinik Araştırmalar Etik Kurulu 16.03.2018 tarihli 101525 sayılı izni, İstanbul Teknik Üniversitesi İnsan Araştırmaları Etik Kurulu 28.8.2014 tarihli MBG.22/2014 sayılı izni, ve Quaid-i-Azam Üniversitesi Hayvansal Bilimleri Bölümü Etik Kurulu 27.8.2014 tarihli HG-148 sayılı izni kapsamında gerçekleştirilmiştir.

Çalışma kapsamında SHFM fenotipi görülen, imzalı gönüllü onam formu alınmış, dört aileye mensup 41 birey incelenmiştir. Aile ve birey isimlendirmeleri sıfırdan yapılmayıp gelecekte referansının kolay olması ve daha önceki isimlendirmeler tutarlılığı açısından projedeki isimleri ile aynı bırakılmıştır. Pakistan'dan gelen numaralandırılmış DNA örnekleri ileri çalışmalarda kullanılmak üzere -20°C buzdolaplarında tutulmuştur. Çalışmadaki bireylerin klinik değerlendirmesi Pakistan Quaid-i-Azam Üniversitesi Hayvan Bilimleri Bölümü tarafından gerçekleştirilmiştir. Hasta bireylere ek olarak, ailedeki sağlıklı kardeşler ve ebeveynler klinik olarak değerlendirilmiştir. Çalışmadaki bütün aileler için aile ağaçları çizilmiş ve etkilenmiş ve normal bireylerin ayrımı net olarak yapılmıştır. Bireylerin ayrıntılı klinik incelemeleri dahilinde X-ray görüntüleri bulunmaktadır. Aile ağaçlarında DNA örneği bulunan bireyler ‘\*’, DNA örnekleri karıştırılarak SNP dizilimi çalışılan bireyler ‘†’, tek başına SNP dizilimi çalışılan bireyler ‘‡’, ve ekzom dizileme yapılan bireyler ‘§’ sembolleri ile gösterilmiştir.

##### 4.1.2. Aile MA134

Ailede kırk dokuz birey ve altı jenerasyon incelenmiştir (Şekil 3-1). Aile içindeki toplam yirmi bir evliliğin on biri akraba evliliğidir. Bu altı jenerasyonun üçünde hastalık ayırım göstermiştir. Beş erkek, beş kadın olmak üzere toplam on bireyde SHFM görülmektedir.



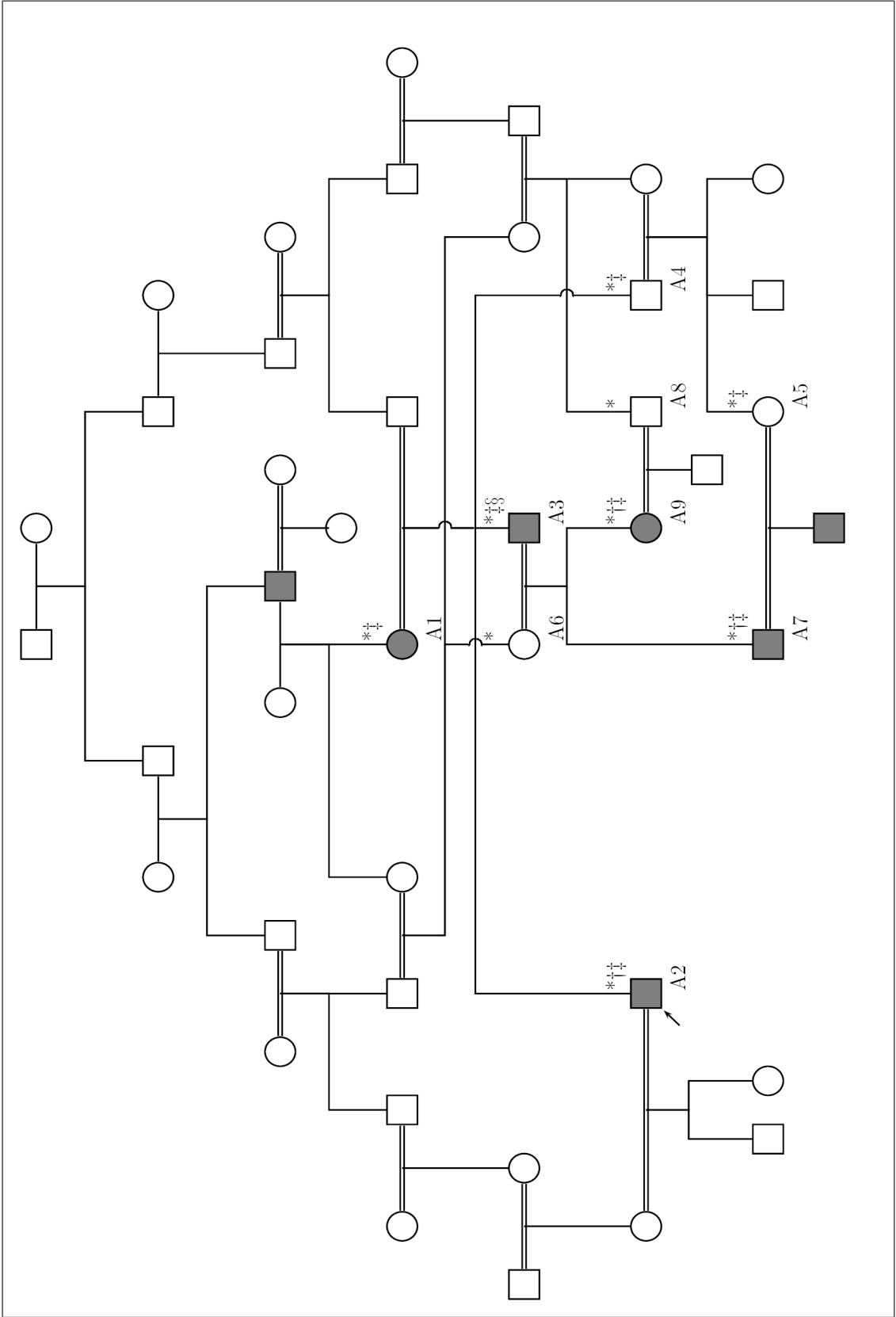
**Şekil 3-1: Şekil 3-1: MA134 aile ağacı.**

DNA örneği olan bireyler \* işareti ile belirtilmiştir. SNP dizilimi çalışılan bireyler † ile belirtilmiştir. DNA örnekleri karıştırılarak SNP dizilimi çalışılan bireyler † ile gösterilmiştir.

#### 4.1.3. Aile MA135

Ailede kırk iki birey ve sekiz jenerasyon incelenmiştir (Şekil 3-2). Aile içindeki toplam on sekiz evliliğin on dördü akraba evliliğidir. Bu sekiz jenerasyonun beşinde hastalık gözükmiştir. Beş erkek, üç kadın olmak üzere toplam sekiz bireyde SHFM görülmektedir.





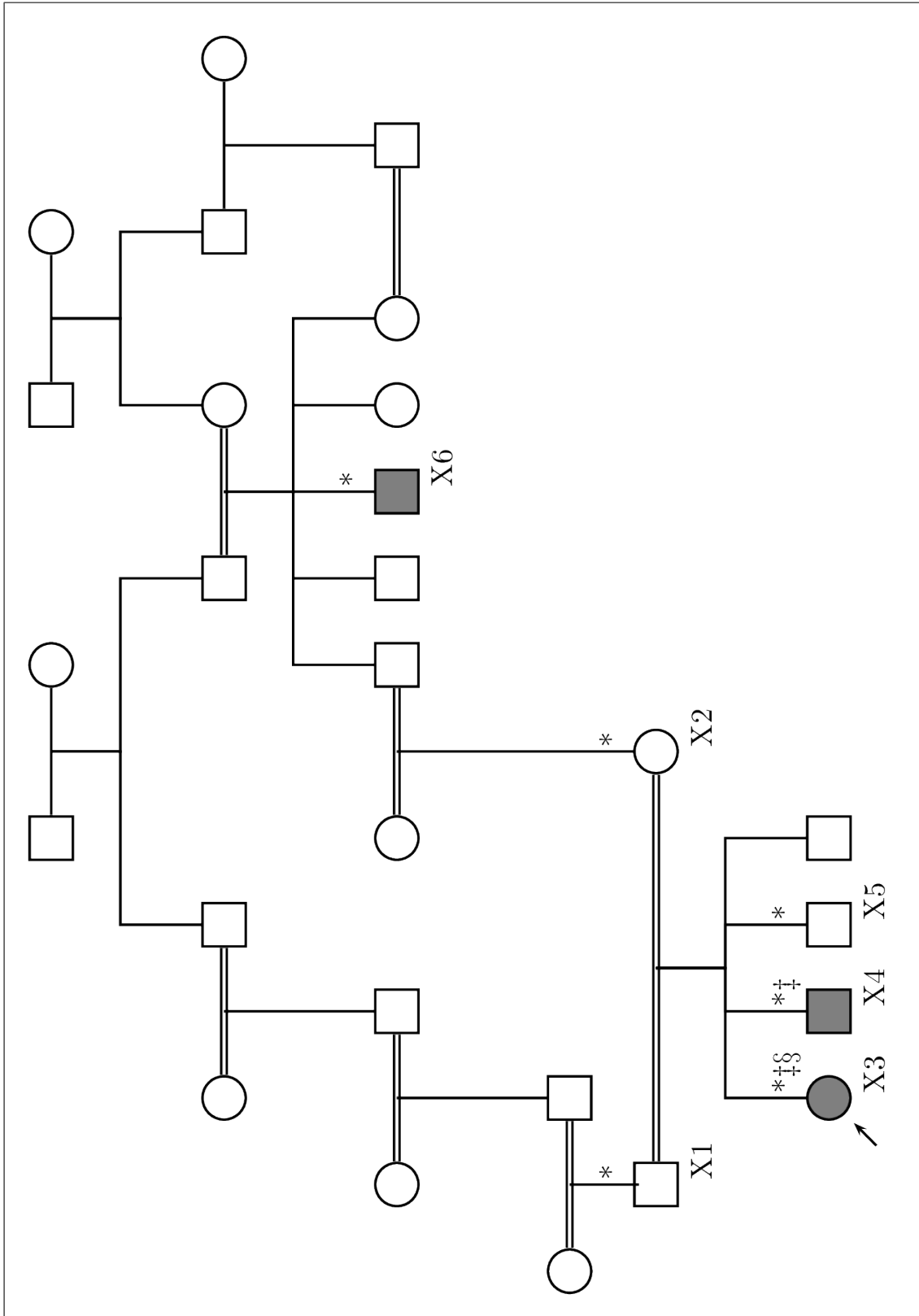
**Şekil 3-2: MA135 aile ağacı.**

DNA örneği olan bireyler \* işareti ile belirtilmiştir. SNP dizilimi ile gen bireyler † ile belirtilmiştir. DNA örneği karıştırılarak SNP dizilimi çalışılan bireyler † ile gösterilmiştir. Ekzom dizileme yapılan bireyler § ile gösterilmiştir.

#### 4.1.4. Aile MA136

Ailede yirmi yedi birey ve altı jenerasyon incelenmiştir (Şekil 3-3). Aile içindeki toplam on evliliğin yedisi akraba evliliğidir. Bu altı jenerasyonun ikisinde hastalık görülmektedir. İki erkek ve bir kadın olmak üzere toplam üç bireyde SHFM görülmektedir.





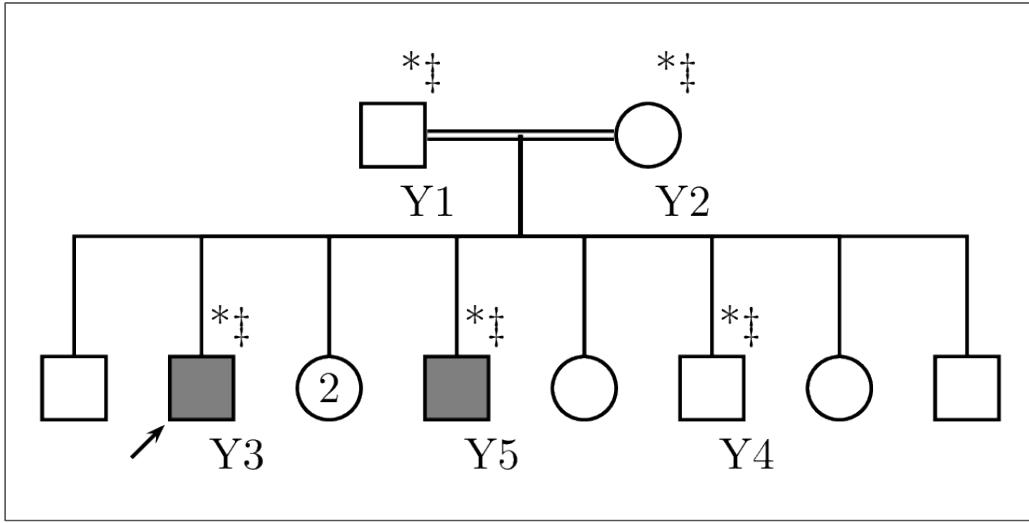
**Şekil 3-3: Şekil 3-3: MA136 aile ağacı.**

DNA örneği olan bireyler \* işareti ile belirtilmiştir. SNP dizilimi çalışılan bireyler ‡ ile belirtilmiştir. Ekzom dizileme yapılan bireyler § ile gösterilmiştir.



#### 4.1.5. Aile MA154

Aile, akraba evliliği yapmış çift ve onların dört kadın beş erkek çocuklarından oluşmaktadır (Şekil 3-4). SHFM iki erkek kardeşte görülmektedir.



**Şekil 3-4: MA154 aile ağacı.**

DNA örneği olan bireyler \* işareti ile belirtilmiştir. SNP dizilimi çalışılan bireyler † ile belirtilmiştir.

#### 4.1.6. Çalışmada Kullanılan Laboratuvar Cihazları

Çalışma kapsamında kullanılan cihazlar Tablo 3-1’de gösterilmiştir.

**Tablo 3-1: Çalışmada kullanılan cihazlar ve üretici firmalar.**

<b>Cihaz Adı</b>	<b>Firma</b>
Hassas terazi	Shimadzu
Otomatik pipetler	Gilson, Eppendorf
Buzdolabı +4°C	Sanyo
Derin dondurucu (-20°C, -80°C)	Uğur, Heraeus Sepatech
Soğutmalı santrifüj	Eppendorf
Masa üstü mini santrifüj	Hettich, Eppendorf
Elektroforez aleti	ThermoFisher Scientific
Güç kaynağı	Stratagene
PZR cihazı	BioRad
Vorteks	Kermanlar, Yellowline
Flow kabin	ThermoFisher Scientific
Spektrofotometre	Nanodrop
Hibridizasyon fırını	Heraeus
Jel görüntüleme sistemi	UVP Imaging System

#### 4.1.7. Çalışmada Kullanılan Çözeltiler

Çalışma kapsamında kullanılan çözeltiler Tablo 3-2’de gösterilmiştir.

**Tablo 3-2: Çalışmada kullanılan çözelti ve kimyasallar.**

Çözelti adı	Firma/Ülke
Etidyum Bromür	Sigma, Almanya
Agaroz	Invitrogen, İsviçre
DMSO	Sigma, Almanya
Betain	Promega, ABD
MgCl <sub>2</sub>	Roche, Almanya
dNTP seti (dATP, dTTP, dCTP, dGTP)	Roche, Almanya
Taq Polimeraz	Roche, Almanya
DreamTaq Green PZR Master Mix	ThermoFisher Scientific, ABD
LightCycler 480 High Resolution Melting Master	Roche, Almanya

#### 4.1.8. Çalışma Kapsamında Hizmet Alınan Firmalar

Çalışma kapsamında hizmet alınan firma ve kurumlar Tablo 3-3’te gösterilmiştir.

**Tablo 3-3: Çalışma kapsamında hizmet alınan firma/kurumlar.**

Alınan Hizmet	Firma/Kurum Adı
SNP dizilimi (Illumina Infinium OmniExpress 700k)	Macrogen
SNP dizilimi (Illumina Infinium OmniExpress 700k)	Yale Center for Genome Analysis
Ekzom Dizileme (Illumina, IDT xGen exome kit)	Yale Center for Genome Analysis

#### 4.1.9. Çalışmada Kullanılan Bilgisayarlar

Çalışma kapsamında kullanılan bilgisayar ve amaçları Tablo 3-4'te gösterilmiştir.

**Tablo 3-4: Çalışmada kullanılan bilgisayarlar.**

Kullanım Amacı	İşlemci	RAM	İşletim Sistemi
ED analizleri	Intel Core i7-5930K CPU @ 3.50GHz	64 Gb	Ubuntu 18.04
SNP dizilimi analizleri	Intel Core i7-5930K CPU @ 3.50GHz	64 Gb	Microsoft Windows 10
Bağlantı analizleri	Intel Core 2 Quad CPU Q6660 @ 2.40GHz	2 GB	Microsoft Windows XP
Programların Yazım	Intel Core i5-2520M CPU @ 2.50GHz	8 GB	Debian Testing

#### 4.1.10. Çalışmada Kullanılan Elektronik Veri tabanları

Çalışmada analiz amaçlı kullanılan elektronik veri tabanları Tablo 3-5'te verilmiştir.

**Tablo 3-5: Çalışmada kullanılan veritabanları.**

Veritabanı	Bağlantı	Kullanım Amacı
UCSC	<a href="https://genome.ucsc.edu/">https://genome.ucsc.edu/</a>	Referans genom
RefSeq	<a href="https://ncbi.nlm.nih.gov/refseq/">https://ncbi.nlm.nih.gov/refseq/</a>	Referans transkriptom
Ensembl	<a href="https://ensembl.org/vep">https://ensembl.org/vep</a>	Gen/frekans anotasyonu
dbSNP	<a href="https://ncbi.nlm.nih.gov/snp/">https://ncbi.nlm.nih.gov/snp/</a>	SNP anotasyonu
Decipher	<a href="https://decipher.sanger.ac.uk">https://decipher.sanger.ac.uk</a>	CNV veri tabanı
gnomAD	<a href="https://gnomad.broadinstitute.org/">https://gnomad.broadinstitute.org/</a>	SNP frekans anotasyonu
OMIM	<a href="https://ncbi.nlm.nih.gov/snp/">https://ncbi.nlm.nih.gov/snp/</a>	Hastalık anotasyonu
Clinvar	<a href="https://ncbi.nlm.nih.gov/clinvar/">https://ncbi.nlm.nih.gov/clinvar/</a>	Varyant önem anotasyonu
HPO	<a href="https://hpo.jax.org/">https://hpo.jax.org/</a>	Fenotip anotasyonu
Uniprot	<a href="https://www.uniprot.org/">https://www.uniprot.org/</a>	Protein veri tabanı
Robetta	<a href="http://rosetta.bakerlab.org/">http://rosetta.bakerlab.org/</a>	Protein Modellemesi

#### 4.1.11. Çalışmada Kullanılan Yazılımlar

Çalışmada yazılım geliştirilmesinde kullanılan python paketleri Tablo 3-6'da verilmiştir.

**Tablo 3-6: Yazılımların geliştirilmesinde kullanılan python paketleri.**

Paket ismi	Kullanıldığı yazılım	İşlevi
responses v0.10.5	Dove	Ensembl VEP API bağlantısı
pandas v0.23.4	Dove	Vcf ve anotasyon dosyalarının işlenmesi

Çalışmada SNP diziliminin analizinde kullanılan yazılımlar Tablo 3-7'de verilmiştir.

**Tablo 3-7: SNP dizilimi analizlerinde kullanılan yazılımlar.**

Yazılım İsmi	İşlevi
GenomeStudio v2.0.4	SNP dizilimi verisinin analizi
cnvPartition CNV Analysis v3.2.0	GenomeStudio içinde CNV analizi
CNV Region Report v2.1.2	GenomeStudio CNV analizlerinin raporlanması
PLINK Input Report v2.1.4	GenomeStudio'dan SNP dizilimi verisinin Plink'e uygun formatta çıktı alınması
Plink v2.0	GBG verisinden homozigot bölgelerin belirlenmesi
EasyLinkage v5.082	Çeşitli bağlantı analiz programları için grafik arayüz
Allegro v2.0f	Bağlantı analizleri ve haplotiplerin oluşturulması
Pedcheck v1.00	Bağlantı analizi öncesi Mendelyen uyumsuzlukların tespiti
HaploPainter v1.043	Haplotiplerin aile ağacı üzerinde gösterimi

YND verisinin analizinde kullanılan yazılımlar Tablo 3-8’de verilmiştir.

**Tablo 3-8: ED analizinde kullanılan yazılımlar.**

Yazılım İsmi	İşlevi
FastQC v0.11.8	Fastq dosyaları kalite kontrolü
Bwa v0.7.1	Referans genoma hizalama
Picard v2.19.2	BAM dosyalarının işlenmesi
GATK v3.6	Kalite skorları kalibrasyonu/Varyantları çağırma/Okuma derinliği kapsamı hesaplanması
VEP release 95.0	Varyantların anotasyonu
Picus v 0.04	Varyantların ACMG yönergesine göre sınıflandırılması
IGV v2.5.3	BAM dosyasında hizalanan okumaların görselleştirilmesi
Mutalyzer v2.0.30	Varyantların isimlendirilmesinin kontrolü
Primer-BLAST	Primer tasarımı
CLC Main Workbench v5.02	Sanger dizileme sonucu kromatogramların görselleştirilmesi
Clustalo v1.2.4-2	Çoklu protein dizi hizalaması

Çalışmanın yazımında kullanılan yazılımlar Tablo 3-9’da verilmiştir.

**Tablo 3-9: Çalışma yazımında kullanılan yazılımlar.**

Yazılım İsmi	İşlevi
Latex pst-pdgr v0.4	Aile ağaçların çizimi
Libre Office v6.1.5.2	Tez yazımı
GIMP v2.10	Tez görsellerinin hazırlanması

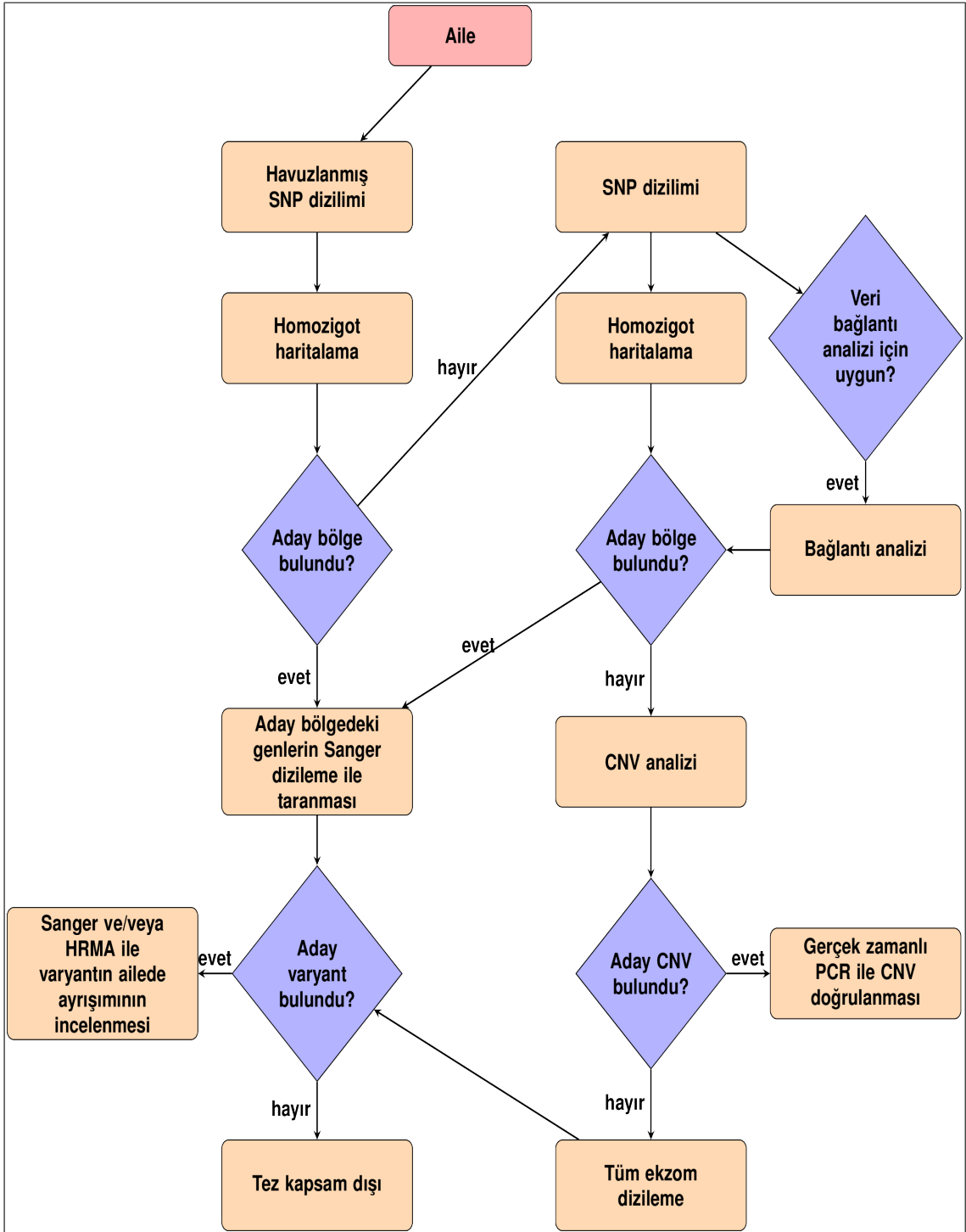
## 4.2. Yöntem

### 4.2.1. Ailelerin Analizlerinde İzlenen Yöntem

Aile ağacı incelendiğinde ailenin farklı kollarında, birbirlerine yakın derecelerde akraba olan sağlıklı ebeveynlerin, her iki cinsiyetten de hasta çocuklarının olduğu görülmektedir. Ebeveynlerin sağlıklı olması, hastalığın cinsiyet ayrımı yapmıyor olması, cinsiyet kromozomları bağlı kalıtıma karşı delillerdir. Hastalığın ailenin birden fazla kolunda akraba evlilikleri sonucu ortaya çıkması otozomal dominant bir modele karşı delil oluşturur. Öte yandan bu gözlemler otozomal resesif bir kalıtım modelini desteklemektedir<sup>83</sup>.

Otozomal resesif kalıtım modeli ve hastalığın ailenin farklı kollarında ortaya çıkmış olması bu farklı kollardaki bireylerin örneklerinin havuzlanarak çalışılabilmesini ideal bir yöntem yapar. Homozigot haritalamada, hastalık ile ilişki bölge bireyler arasında değişmeden paylaşılacağından homozigot olarak gözükecek, hastalık ile ilişkili olmayan genom bölgeleri ise ailenin farklı kollarında meydana gelen rekombinasyon olayları sonucu homozigotluğunu kaybetmiş olacaktır.

Ailelerin analizlerinde uygulanan genel yöntem Şekil 3-5’de gösterilmiştir.

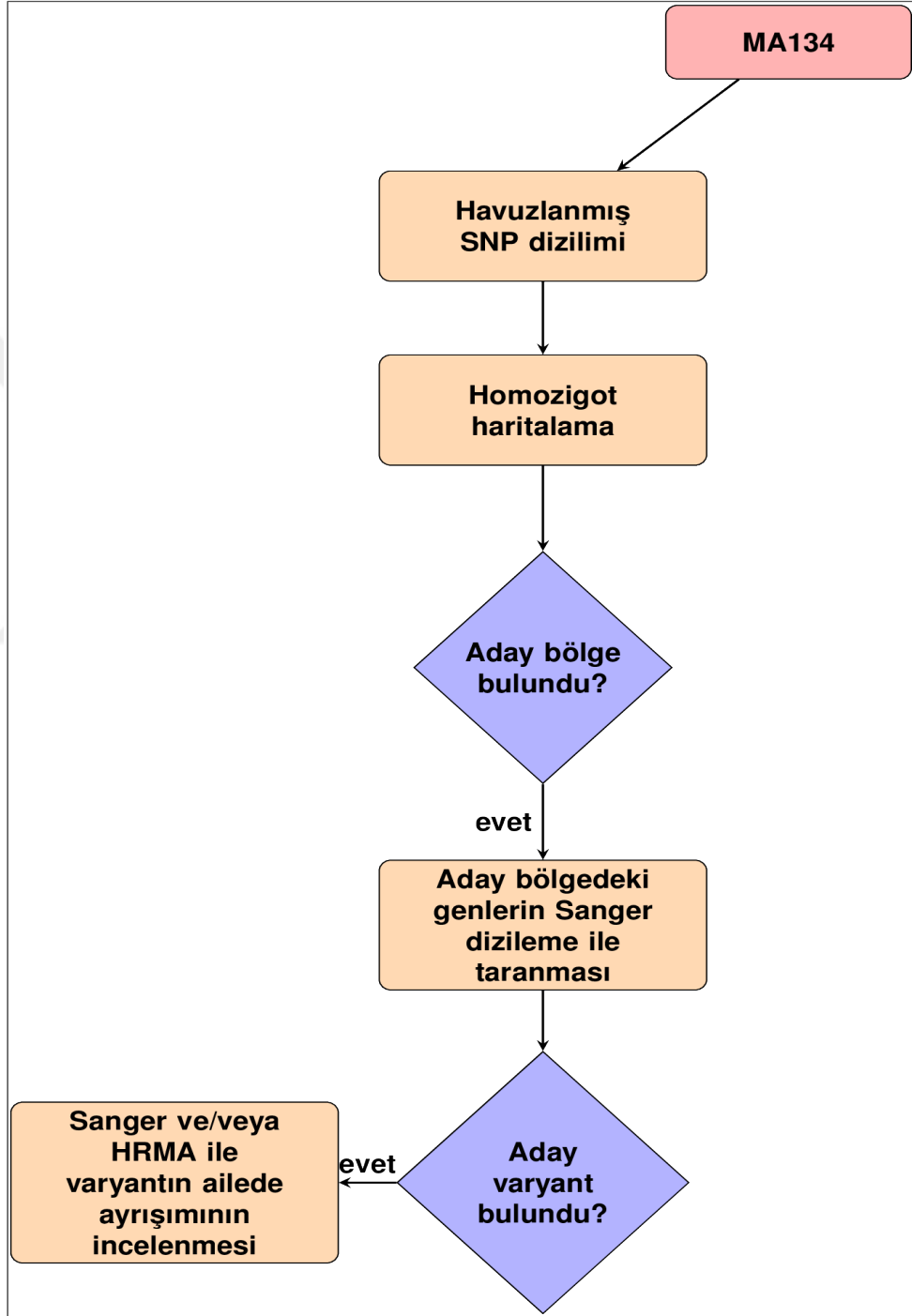


Şekil 3-5: Ailelerin incelemesinde izlenen genel yöntem akış şeması.



#### 4.2.1.1. Aile MA134 İnceleme Yöntemi

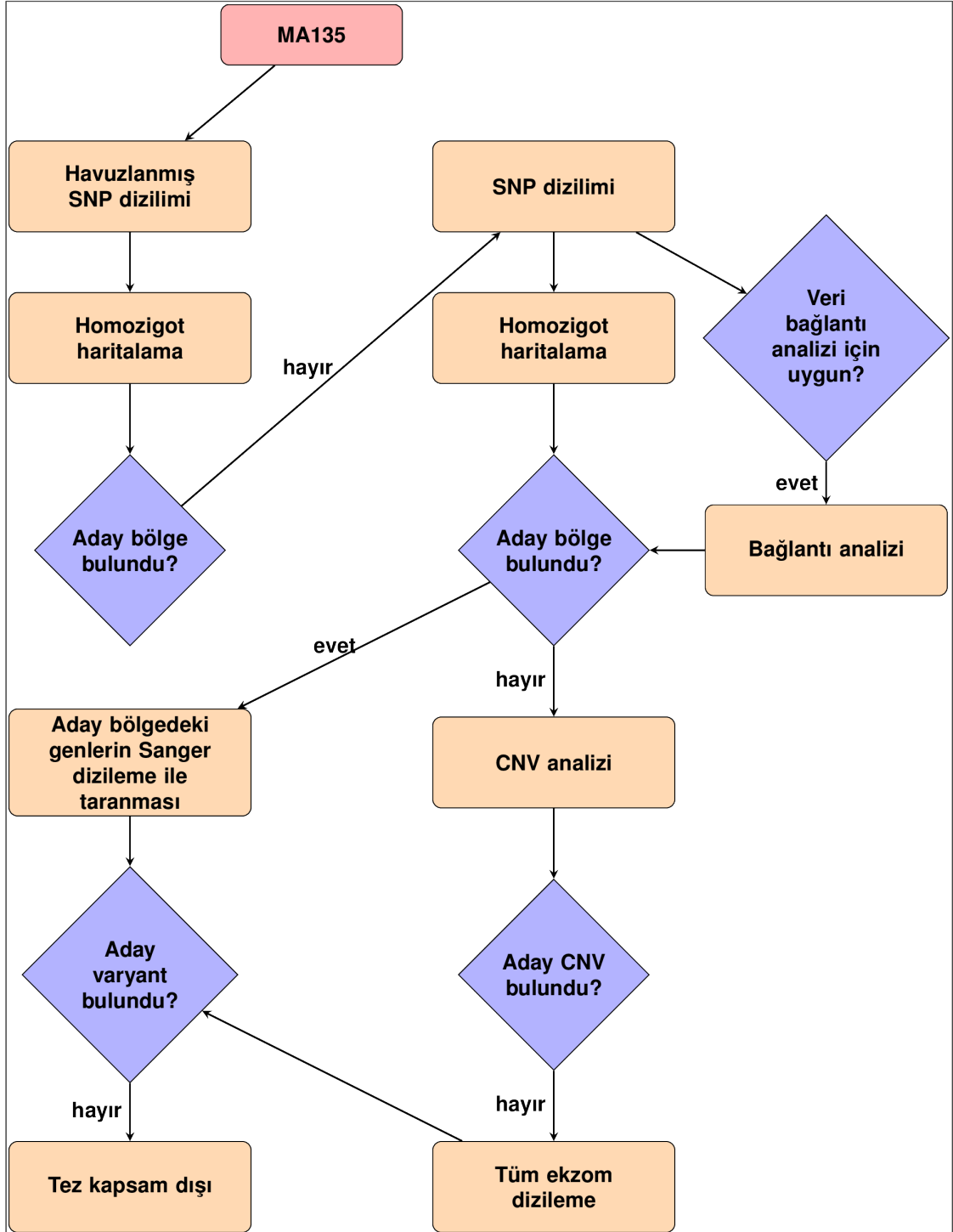
MA134 ailesinin incelenmesinde izlenen yöntem akışı Şekil 3-6' da verilmiştir.



Şekil 3-6: MA134 ailesinin incelenmesinde izlenen yöntem akış şeması.

#### 4.2.1.2. Aile MA135 İnceleme Yöntemi

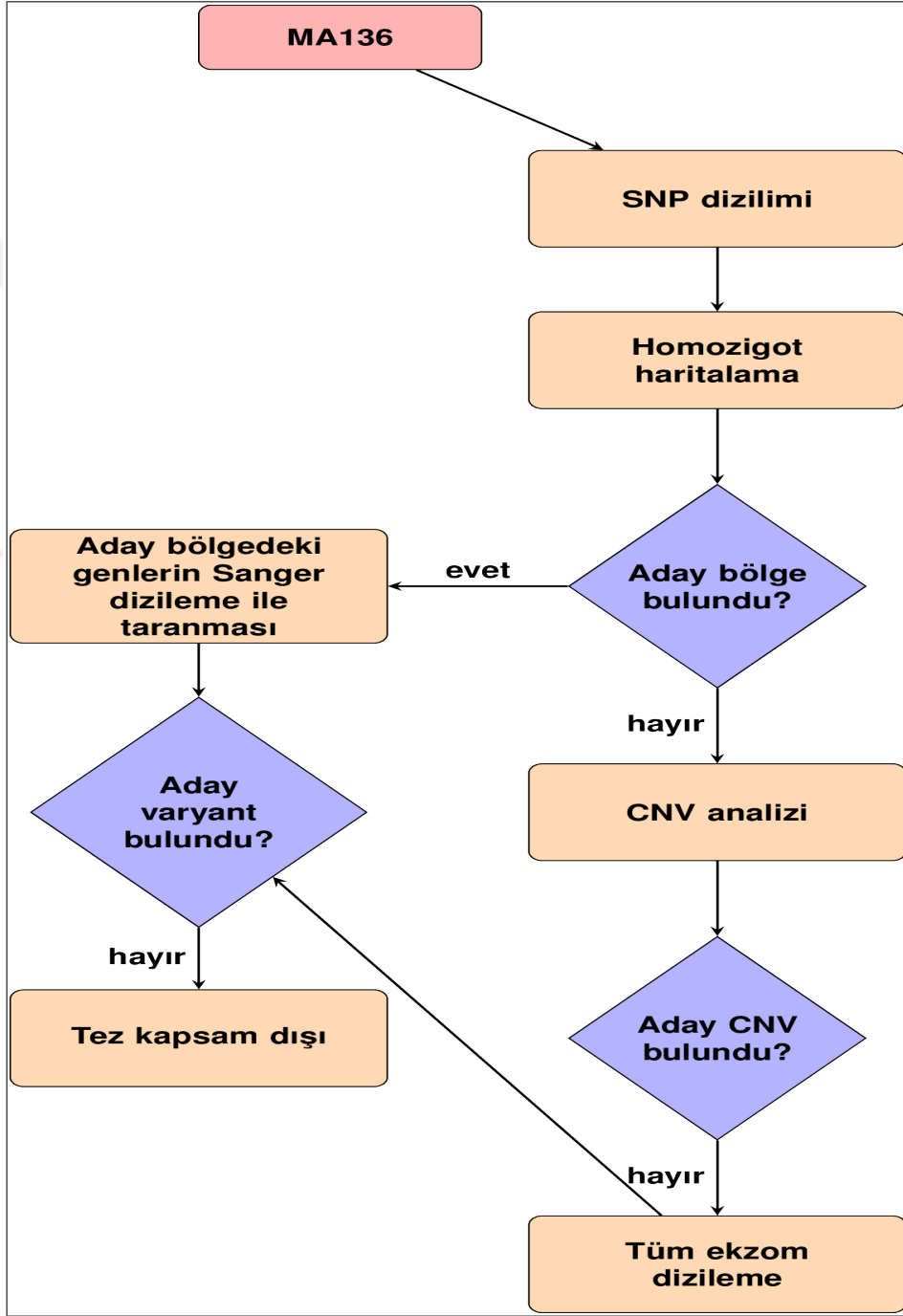
MA135 ailesinin incelenmesinde izlenen yöntem akışı Şekil 3-7' de verilmiştir.



Şekil 3-7: MA135 ailesinin incelenmesinde izlenen yöntem akış şeması.

#### 4.2.1.3. Aile MA136 İnceleme Yöntemi

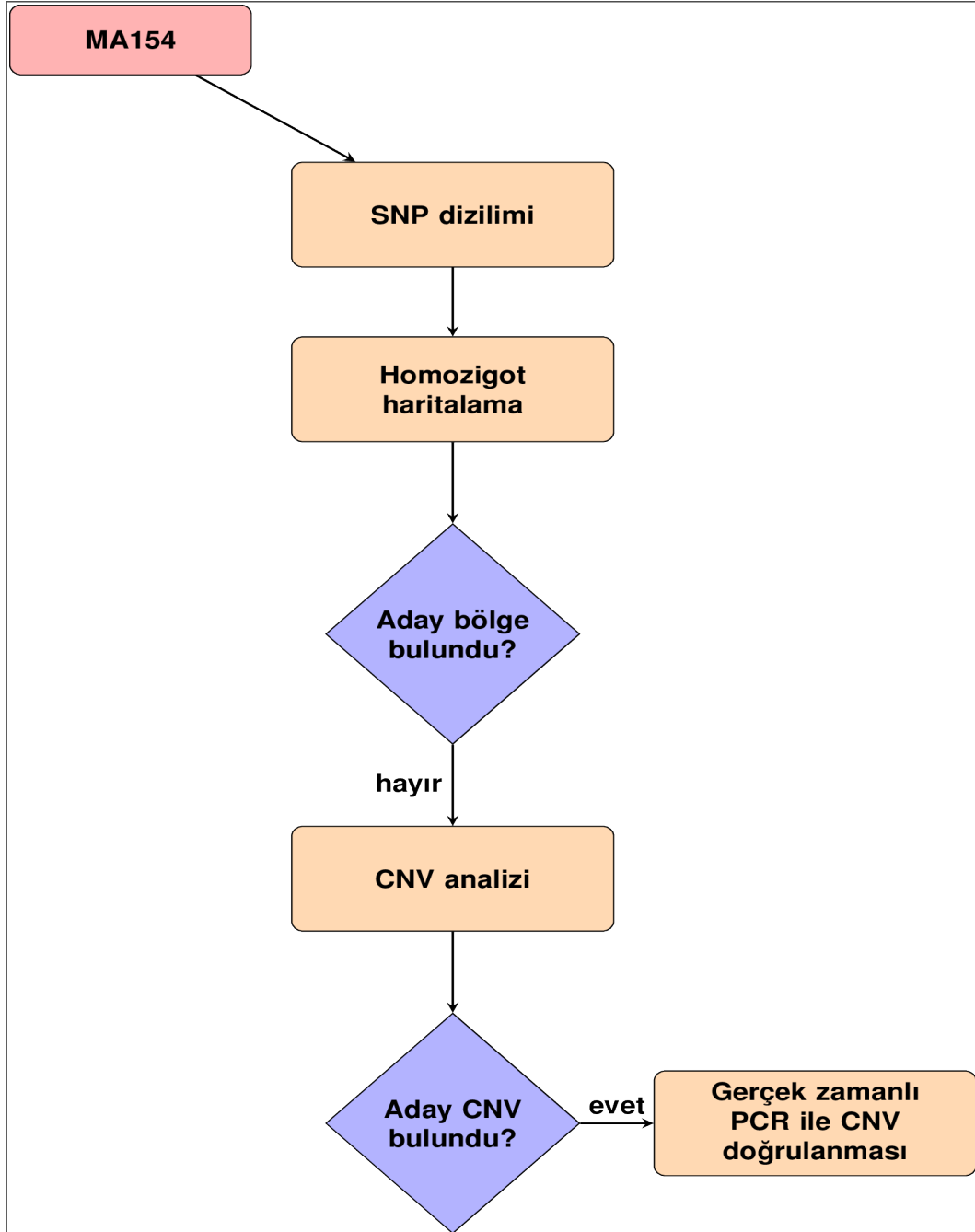
MA135 ailesinin incelenmesinde izlenen yöntem akışı Şekil 3-8' de verilmiştir.



Şekil 3-8: MA136 ailesinin incelenmesinde izlenen yöntem akış şeması.

#### 4.2.1.4. Aile MA154 İnceleme Yöntemi

MA135 ailesinin incelenmesinde izlenen yöntem akışı Şekil 3-9' da verilmiştir.



Şekil 3-9: MA154 ailesinin incelenmesinde izlenen yöntem akış şeması.

## 4.2.2. Akış Hattı Yazılımının Programlanması

Akış hattı yazılımı, hazırlanan akış hattı yapılandırma dosyasını okuyarak, bu dosya içerisinde belirtilen tanımlamalara uygun şekilde işlemleri gerçekleştirmek şeklinde planlanmıştır. İşlemlerin argümanları ve parametreleri ve çalışma sıraları yapılandırma dosyası üzerinden verilir.

### 4.2.2.1. Akış Hattı Yazılımının Yapısı

Yazılım, pigeon.py, pipeline.py, parseconfig.py, ve runpipe.py olmak üzere dört temel dosyaya ayrılmıştır (Şekil 3-10).

```
$ tree pigeon
pigeon
├── core
│   ├── __init__.py
│   └── pipeline.py
├── examples
│   ├── example.conf
│   └── __init__.py
├── __init__.py
├── pigeon.py
├── utils
│   ├── copyexamplefiles.py
│   ├── __init__.py
│   ├── parseconfig.py
│   └── runpipe.py
```

Şekil 3-10: Akış hattı yazılımının dosya sistemi üzerindeki yapısını gösteren tree komutu çıktısı.

### **pigeon.py İşlevi**

Bu dosya, kullanıcının komut satırı ile programa erişimini sağlar. Kullanıcının verdiği argümanlar burada ayrıştırılarak uygun işlemler çağrılır.

### **parseconfig.py İşlevi**

Bu dosya, yazılıma sağlanan yapılandırma dosyası içerisinde geçen argümanları ayrıştırarak kullanıma hazır hale getirir.

### **pipeline.py İşlevi**

Bu dosya, yazılımın temeli olup sağlanan argümanları bir araya getirerek akış hattında kullanılacak komutları oluşturur.

### **runpipe.py İşlevi**

Bu dosya, akış hattı çalıştırıldığında pipeline.py tarafından oluşturulan komutları kabuğa(shell) geçirerek çalıştırılmasını sağlar.

#### **4.2.2.2. Akış Hattı Yapılandırma Dosya Yapısı**

Yapılandırma dosyalarının ayrıştırılması python configparser modülü ile gerçekleştirilir ve dosya yapısı bu modülün dökümantasyonuna uygun olarak geliştirilmiştir. Bu formata ek olarak okunabilirliği arttırmak üzere argümanlar içerisinde yeni satır eklenebilir.

Yapılandırma dosyası, işlenecek dosyaların, sonuçların yazılacağı klasörün, çalışacak yazılımların sırasının, her yazılımın kendine özgü argümanlarının ve analizlerde kullanılacak ek verilerin tanımlanabileceği düz metin dosyasıdır. Bu dosya kendi içerisinde, PROJECT, GENERAL, INCLUDE, PIPELINE isimli dört tanımlı bölüm ve kullanıcının her yazılım için ekleyebileceği bölümlerden meydana gelir. Her bölüm ismi, [PROJECT] gibi köşeli parantezler içerisinde belirtilir. Kullanıcı yorumları satıra '#' karakteri ile başlayarak eklenebilir.

## PROJECT Bölümü

Bu bölüm içerisinde çalışmalar arası değişebilecek argümanlar tanımlanır (Tablo 3-10). İsteğe bağlı `project_name` değişkeninin olmaması durumunda çalıştırma tarih ve zamanı proje ismi olarak `output_dir` değişkeninin olmaması durumunda ise komutun çalıştırıldığı klasör çıktıların yazılacağı klasör olarak kullanılır.

**Tablo 3-10: Projeye özgü argümanlar bölümü.**

Argüman	Gereklilik	İşlevi
<code>project_name</code>	İsteğe bağlı	Çalışmaya özgü isim
<code>output_dir</code>	İsteğe bağlı	Çalışma sırasında çıktı dosyalarının yazılacağı klasör
<code>input_files</code>	Zorunlu	Akış hattında analizin yapılacağı girdi dosyaları
<code>input_names</code>	İsteğe bağlı	Girdi dosyalarının isimleri

## GENERAL Bölümü

Bu bölüm içerisinde çalışmalar arası değişmeyecek argümanlar tanımlanır. Burada referans genom, popülasyon varyasyon verisi ya da kullanılan yazılımların gerektirdiği herhangi bir dosya ya da linke bağlantı verilebilir. Verilen dosyaların, dosya sistemi üzerindeki tam yolları verilmelidir. Bu kısımda tanımlanan değişkenler daha sonra işlem bölümlerinde, `{GENERAL:değişken}` formatı ile kullanılabilir.

## INCLUDE Bölümü

Bu bölüm içerisinde diğer akış hattı yapılandırma dosyaları eklenebilir. Böylece birden fazla akış hattında kullanılan ya da tekrarlı kısımlar modüllere çevrilerek kullanılabilir. Bu kısımda tanımlanan değişkenler daha sonra `[INCLUDE:değişken]` formatında bölüm olarak kullanılabilir.

## PIPELINE Bölümü

Bu bölüm içerisinde tanımlanan her işlemin çalıştırma sırası pipeline argümanı içerisine yazılır. Bununla beraber işlemlerin kabuk üzerinde çalıştıracak yazılım isimleri ve konteynerler de bu kısımda isteğe bağlı olarak tanımlanabilir. Bu kısımda tanımlanan yazılım isim değişkenleri işlem kısmının tool argümanı olarak, konteyner değişkenleri de işlem kısımlarında container argümanı olarak kullanılabilir.

## İşlem Bölümü

İşlem bölümünde çalıştırılacak her işlemin girdisini, çıktısını, işlem argüman ve parametrelerini işleme özel olarak verilir (Tablo 3-11). Bu bölümlerin isimlendirmesi kullanıcı tarafından yapılır.

**Tablo 3-11: İşleme özgü argümanlar bölümü.**

Argüman	Gereklilik	İşlevi
<b>tool</b>	Zorunlu	Çalışacak yazılımın kabuğun tanıdığı ismi
<b>input_from</b>	Zorunlu	Çalışacak işlemin girdisini nereden alacağı
<b>args</b>	Zorunlu	İşlemin argüman ve parametreleri
<b>dump_dir</b>	İsteğe bağlı	İşlem çıktısının yazılacağı klasör
<b>ext</b>	İsteğe bağlı	İşlem çıktı dosyalarının uzantısı
<b>suffix</b>	İsteğe bağlı	İşlem çıktı dosya isimlerine ek
<b>input_multi</b>	İsteğe bağlı	İşlemin girdileri çoklu alması
<b>container</b>	İsteğe bağlı	İşlemin çalışacağı konteyner
<b>sub_tool</b>	İsteğe bağlı	İşlemin alt komutu
<b>input_dir</b>	İsteğe bağlı	İşleme bağlı girdi klasörü
<b>secondary_input</b>	İsteğe bağlı	İşleme ikinci çeşit girdi belirteci
<b>secondary_in_dir</b>	İsteğe bağlı	İşleme bağlı ikincil girdinin klasörü
<b>input_flag_repeat</b>	İsteğe bağlı	İşlemin girdi argümanını her girdi için tekrarlama
<b>output_dir</b>	İsteğe bağlı	İşleme bağlı çıktı klasörü
<b>remove</b>	İsteğe bağlı	İşlemden sonra çıktı dosyalarını kaldır
<b>secondary_output</b>	İsteğe bağlı	İkincil çıktı çeşidi belirteci
<b>secondary_dump_dir</b>	İsteğe bağlı	İkincil çıktıların yazılacağı klasör
<b>secondary_ext</b>	İsteğe bağlı	İkincil çıktı dosya uzantısı
<b>secondary_suffix</b>	İsteğe bağlı	İkincil çıktı dosya isimlerine ek



İşlemlerin kullanacağı girdi ve çıktı dosyalarının isimlendirmesi otomatik olarak gerçekleştirilir ve işlemlere tanıtılması her işlem bölümündeki args isimli argümanın içinde yapılır (Tablo 3-12).

**Tablo 3-12: Girdi ve çıktıların yerini tutan özel değişkenler.**

Özel Değişken	Gereklilik	İşlevi
<b>input_placeholder</b>	Zorunlu	İşlemin girdi dosyasının yerini tutar
<b>output_placeholder</b>	İsteğe bağlı	İşlemin çıktı dosyasının yerini tutar.
<b>secondary_in_placeholder</b>	İsteğe bağlı	İşlemin ikincil girdi dosyasının yerini tutar.
<b>secondary_out_placeholder</b>	İsteğe bağlı	İşlemin ikincil çıktı dosyasının yerini tutar.

#### 4.2.2.3. Akış Hattı Yazılımının Çalıştırılması

Yazılımın çalıştırılması için ilk olarak yapılandırma dosyasının hazırlanması gereklidir. Örnek olarak Bölüm 4.2.2.2'ye uygun olarak Fastq dosyalarını fastqc yazılımı ile kalite kontrolü ve bwa yazılımı ile genoma hizalama yapacak bir yapılandırma dosyası Şekil 3-11'de verilmiştir.

```
example.conf
[PROJECT]
project_name = example
output_dir = /home/bar/example/output
input_files = /home/bar/example/input/A_1.fastq
              /home/bar/example/input/A_2.fastq
              /home/bar/example/input/B_1.fastq
              /home/bar/example/input/B_2.fastq
input_names = A B

[GENERAL]
reference_genome = /home/bar/references/hg19.fa

[PIPELINE]
pipeline = QUALITYCHECK ALIGNMENT
bwa = bwa
fastqc = fastqc

[QUALITYCHECK]
tool = ${PIPELINE:fastqc}
input_from = input_files
args = input_placeholder
      --outdir=${PROJECT:output_dir}/fastqc
      --extract
dump_dir = fastqc

[ALIGNMENT]
tool = ${PIPELINE:bwa}
sub_tool = mem
input_from = input_files
args = -t 8
      ${GENERAL:reference_genome}
      input_placeholder > output_placeholder
named = True
suffix = _bwa
ext = sam
input_multi = paired
```

Şekil 3-11: Kalite kontrol ve genoma hizalama akış hattı yapılandırma dosyası.

Yapılandırma dosyası hazırlandıktan sonra komut satırında ‘pigeon’ komutuyla yazılım çalıştırılabilir. Yazılıma ‘-c’ argümanı ile hazırlanan yapılandırma dosyası tanıtılabilir ve analize başlamadan önce ‘-d’ argümanı ile kabuğa gönderilecek komutlar prova edilebilir (Şekil 3-12). Herhangi bir problem yoksa ‘-d’ argümanı kaldırılarak analize başlanabilir.

```
$ pigeon -c example.conf -d
fastqc /home/bar/example/input/A_1.fastq --outdir=/home/bar/example/output/fastqc --extract
fastqc /home/bar/example/input/A_2.fastq --outdir=/home/bar/example/output/fastqc --extract
fastqc /home/bar/example/input/B_1.fastq --outdir=/home/bar/example/output/fastqc --extract
fastqc /home/bar/example/input/B_2.fastq --outdir=/home/bar/example/output/fastqc --extract
bwa mem -t 8 /home/bar/references/hg19.fa /home/bar/example/input/A_1.fastq /home/bar/example/in
put/A_2.fastq > /home/bar/example/output/A_1_bwa.sam
bwa mem -t 8 /home/bar/references/hg19.fa /home/bar/example/input/B_1.fastq /home/bar/example/in
put/B_2.fastq > /home/bar/example/output/B_1_bwa.sam
```

**Şekil 3-12: Akış hattının prova çalıştırılması sonucu.**

Akış hattı yazılımının alabileceği bütün argümanlar ‘pigeon -h’ komutu ile komut satırı üzerinde görülebilir.

#### 4.2.3. Vcf Sonrası Analizler Yazılımının Programlanması

Vcf sonrası analizler yazılımında varyant çağırma adımından sonra elde edilen varyantların, anotasyonu, esnek bir şekilde filtrelenebilmesi ve ailesel çalışmalarda aile içi karşılaştırmaların yapılabilmesi hedeflenmiştir.

Varyant çağırma sonrasında elde edilen dosya varyantın genomik pozisyonu, değişimi ve allellerin okuma derinliklerini içerir. Sadece bu bilgilerden yola çıkılarak bir varyantın anlamlandırılması mümkün değildir. Bu yüzden elde edilen varyantların, farklı genetik işaretçilere bağlı pozisyonu, popülasyonlardaki görülme sıklığı, RNA ve protein aşamalarında neden olabileceği bozukluklar gibi bilgilerin anotasyon aşaması ile

eklenmesi gerekmektedir. Böylece varyantın fenotip üzerinde yaratabileceği etkiler, genomik, transkriptomik ve proteomik bağlamda popülasyon bilgisi ile birlikte irdelenebilir.

Varyantların anotasyonu sonrasında elde edilen çok sayıdaki varyantın oluşturduğu büyük verinin içerisinde ilgilenecek fenotip ile ilişkili olabilecek varyantları ayırtmak gereklidir. Bu ayırtma varyantların belli kriterler üzerinden elenmesi ile sağlanır.

Bir diğer varyant eleme yolu ise aile içerisindeki kalımdan yola çıkarak varyantların elenmesidir. Nadir varyantların ilişkilendirilmesinde çalışmalar aile odaklıdır ve probant ile beraber ebeveynler ve/veya etkilenmiş/normal kardeşler dizilenebilir. Bu durumda ilişkili olabilecek varyantların elde edilmesi için Mendelyen kalıtım modeline uygun şekilde probant ve akrabalarının varyasyon verisi karşılaştırılabilir.

#### 4.2.3.1. Vcf Sonrası Analizler Yazılımının Yapısı

Yazılım, `dove.py`, `onlinevariantannotation.py`, `variantfilter.py`, `infamilyanalysis.py`, `vcf.py`, ve `ensemblapi.py` olmak üzere altı temel dosyaya ayrılmıştır (Şekil 3-13). Yazılımda VCF formatı gibi veri tabloları ile etkileşimde `pandas`<sup>84</sup> paketi kullanılmıştır.

```
$ tree dove
dove
├── core
│   ├── infamilyanalysis.py
│   ├── __init__.py
│   ├── onlinevariantannotation.py
│   └── variantfilter.py
├── dove.py
├── __init__.py
└── utils
    ├── bed.py
    ├── ensemblapi.py
    ├── __init__.py
    ├── updateclinvar.py
    ├── updatehpo.py
    ├── updateomim.py
    └── vcf.py
```

Şekil 3-13: Vcf sonrası analizler yazılımının dosya sistemi üzerindeki yapısını gösteren `tree` komutu çıktısı.

#### `dove.py`

Bu dosya, kullanıcının komut satırı ile programa erişimini sağlar. Kullanıcının verdiği argümanlar burada ayrıştırılarak uygun işlemler çağrılır.

**onlinevariantannotation.py**

Bu dosya, Vcf dosyasının VEP, OMIM, ClinVar ve HPO bilgileri ile anotasyonunu yapar. Ensembl sunucusuna gidecek varyantlar burada formatlanır ve sunucudan gelen veri burada ayrıştırılır.

**variantfilter.py**

Bu dosya, verilen Vcf ya da anotasyon dosyasını alarak kullanıcının belirlediği kriterlere uygun olarak varyantlar elenmesini gerçekleştirir.

**infamilyanalysis.py**

Bu dosya, probant varyantlarının, ebeveynlerinden kalıtımını ve kardeşlerdeki ayrılımını göz önüne alarak kullanıcının sağladığı Mendelyen kalıtım modeline uygun elemesini gerçekleştirir.

**vcf.py**

Bu dosya, Vcf formatındaki dosyaların, okunması, ayrıştırılması ve analizlerde kullanılabilir hale getirilmesini sağlar.

**ensemblapi.py**

Bu dosya, anotasyonu yapılacak varyantları Ensembl sunucusuna gönderir ve sunucunu gönderdiği yanıtı alır. Ensembl'in API ile etkileşimde responses paketi kullanılmıştır.

**4.2.3.2. Vcf Sonrası Analizler Yazılımının Çalıştırılması**

Yazılımın çalıştırılmasında temel olarak üç farklı alt komut kullanılır. Anotasyon 'dove OVA', varyant filtreleme 'dove VF', ve aile içi analizler 'dove IFA' ile gerçekleştirilebilir.

## Anotasyon

Anotasyon komutuyla beraber yazılımına verilebilecek argümanlar Tablo 3-13'te gösterilmiştir. Anotasyona yeni kolon eklemesinde kullanılan '-c' argümanı '+' sembolü ile kullanılırsa '+' sembolünden sonra yazılan her kolon varsayılan kolonların ardına eklenir. Eğer kolon isimleri '+' sembolü kullanılmadan '-c' argümanına verilirse bütün kolonlar baştan tanımlanmış olur. Anotasyonu mümkün olan bütün kolon isimleri '-h' komutu ile görülebilir.

**Tablo 3-13: Anotasyon komutuyla beraber kullanılacak argümanlar ve parametreleri.**

Argüman	Gereklilik	Parametre	İşlevi
-h			Yardımlı sayfasını yazar
-i	Zorunlu	Vcf girdi dosyası	Vcf dosyasını yazılıma tanıtır
-o	Zorunlu	Anotasyon çıktı dosyası	Çıktı dosyasının isimlendirmesi
-j	İsteğe bağlı	Vep komut satırı yazılımının json çıktı dosyası	Çevrimdışı anotasyon için kullanılır
-q	Zorunlu	Sayma sayısı	Ensembl sunucusuna her seferde gönderilecek varyant sayısı
-k	İsteğe bağlı	Sayma sayısı	Her seferinde Vcf'ten okunacak varyant sayısı
-c	İsteğe bağlı	Anotasyona eklenecek kolon isimleri	İstenilen kolonların anotasyona eklenmesini sağlar
-m	İsteğe bağlı		OMIM anotasyon kolonu ekler
-p	İsteğe bağlı		HPO anotasyon kolonu ekler
-l	İsteğe bağlı		ClinVar anotasyon kolonu ekler
-r	İsteğe bağlı		Anotasyon işlemi yarıda kalırsa kaldığı yerden devamı sağlar

## Filtreleme

Anotasyon komutuyla beraber yazılımına verilebilecek argümanlar Tablo 3-14'te gösterilmiştir.

**Tablo 3-14: Filtreleme komutuyla beraber kullanılabilir argümanlar ve parametreleri.**

Argüman	Gereklik	Parametre	İşlevi
-h			Yardımlı sayfasını yazar
-i	Zorunlu	Anotasyon ya da Vcf girdi dosyası	Anotasyon ya da Vcf dosyasını yazılıma tanıtır
-o	Zorunlu	Anotasyon ya da Vcf çıktı dosyası	Çıktı dosyasının isimlendirmesi
-b	İsteğe bağlı	Bed girdi dosyası	Bed dosyasını yazılıma tanıtır
-c	İsteğe bağlı	Kolon ismi ve filtreleme parametreleri	Filtrelenecek kolonları ve o kolonların nasıl filtreleneceğini belirtir
-d	İsteğe bağlı	Kolon isimleri	Tekrarlı satırların atılmasında hangi kolonlara bakılacağı

Filtrelenecek kolonlar yazılıma tanıtılırken '-c kolon\_ismi karşılaştırma\_bayrağı filtreleme\_parametresi' formatında verilir. Kolon ismi Vcf veya anotasyon dosyası içerisindeki filtrelenmek istenilen kolon ismidir. Karşılaştırma bayrağı filtrelenecek kolon içerisindeki değerler ile filtreleme parametresinin nasıl karşılaştırılacağını belirleyen önceden tanımlanmış değerlerden biri olmalıdır (Tablo 3-15). Filtreleme parametresi eğer 'le', 'lt', 'ge', veya 'gt' bayrakları ile veriliyorsa tek bir numerik değer olmalıdır. Diğer bayraklar ile birlikte birden fazla filtreleme parametresi boşluk ile ayrılarak yazılabilir.



**Tablo 3-15: Karşılaştırma bayraklarının işlevleri.**

Karşılaştırma bayrağı	Anlamı	İşlevi
eq	eşittir	Filtreleme parametresi ile eşleşmeyen satırları eler
ne	eşit değildir	Filtreleme parametresi ile eşleşen satırları eler
le	küçük ya da eşittir	Filtreleme parametresinden büyük olan satırları eler
lt	küçüktür	Filtreleme parametresinden büyük ya da eşit olan satırları eler
ge	büyük ya da eşittir	Filtreleme parametresinden küçük olan satırları eler
gt	büyüktür	Filtreleme parametresinden küçük ya da eşit olan satırları eler
in	İçinde bulunur	Filtreleme parametresini içinde bulundurmeyen satırları eler
ex	İçinde bulunmaz	Filtreleme parametresini içinde bulunduran satırları eler

### Aile İçi Analiz

Anotasyon komutuyla beraber yazılımına verilebilecek argümanlar Tablo 3-16'da gösterilmiştir.

Aile içi analizler akrabaların anotasyon dosyaları verilirken akrabaların probanda yakınlığının belirlenebilmesi için '-r kuşak derece cinsiyet fenotip anotasyon\_dosyası' formatında verilir. Kuşak bilgisi probant 0. kuşak kabul edilerek, probanda bağıl olarak yazılır. Probandtan önceki kuşaklar '-' ile ifade edilir ve her önceki kuşak için bir çıkartılır. Probandtan sonraki kuşaklar ise '+' ile ifade edilir ve her sonraki kuşak için bir eklenir. Derece verilen akrabanın probanda yakınlık derecesini ifade eder. Cinsiyet, önceden tanımlı 'male', 'female' ya da 'undefined' değerleri olabilir. Fenotip, önceden tanımlı 'normal', 'affected' değerleri olabilir.

**Tablo 3-16: Aile içi analiz komutuyla beraber kullanılacak argümanlar ve parametreleri.**

Argüman	Gereklilik	Parametre	İşlevi
-h			Yardımlı sayfasını yazar
-i	Zorunlu	Anotasyon girdi dosyası	Probandın anotasyon dosyasını yazılıma tanıtır
-o	Zorunlu	Anotasyon çıktı dosyası	Çıktı dosyasının isimlendirmesi
-r	Zorunlu	Akraba anotasyon dosyası	Akrabaların anotasyon dosyasını yazılıma tanıtır
-a	Zorunlu	Analiz tipi	Anne baba çocuk ya da kardeşler arası analiz olduğunu belirtir
-p	İsteğe bağlı	Kalıtım Şekli	Seçilen Mendelyen kalıtım şekline göre analizi gerçekleştirir

#### 4.2.4. Geliştirilen Yazılımların Test Edilmesi

Akış hattı ve Vcf sonrası analizler yazılımlarının test edilmesi için, daha önceden grup içerisinde çalışılmış ve yayınlanmış iki ED verisi seçildi. Bu test çalışmalarının ilkinde intellektüel yeti kaybı ve epilepsi fenotipi gözükten bir bireye ait ekzom dizileme verisi<sup>85</sup> ve ikincisinde ise çene distonisi fenotipi gözükten bir ailede anne, baba ve probanda ait ekzom dizileme verileri<sup>86</sup> kullanılmıştır.

#### 4.2.5. Genom Boyu Genotipleme Analizi

Yeterli konsantrasyona sahip, kalite kontrol aşamaların geçen örneklerin, 700.000'den fazla SNP belirteci içeren Illumina InfiniumOmniExpress-24 v1-2 ve v1-3 ile genotipleme hizmeti alınmıştır. SNP çipi okunduktan sonra elde edilen .idat dosya formatının okunması ve ileri ki homozigot haritalama ve bağlantı analizleri aşamasında kullanılacak formatta verisin elde edilmesi GenomeStudio yazılımında gerçekleştirilmiştir. Aynı zamanda SNP dizilimi verisinde, GenomeStudio yazılımı kullanılarak, CNV ve LOH bölgeleri hem kantitatif hem de görsel olarak incelenmiştir.

#### 4.2.5.1. SNP Dizilimi Verisinin Analizi

SNP dizilimi çalışmaları farklı zamanlarda üç ayrı grup olarak yapılmıştır. GenomeStudio yazılımı ile analizlerin gerçekleştirilebilmesi için “idat” formatında olan ham veriler, örnek isimleri, çip adı ve örneklerin çipteki pozisyonlarını bulandıran bir örnek dosyası ile yazılıma tanıtıldı (Şekil 3-14). Aynı zamanda çipte kullanılan SNP işaretçilerinin pozisyon, referans dizi niteliklerini içeren anotasyon dosyası (manifest) ve SNP işaretçilerini kümeleyen (cluster) dosya tanıtıldı.

```

A [Header]
INVESTIGATOR_NAME, BeadStudio User
PROJECT_NAME, Asli Tolun_062817
EXPERIMENT_NAME, Asli Tolun_062817
DATE, 7/5/2017 8:53
[Manifests]
A, InfiniumOmniExpress-24v1-2_A1
[Data]
Sample_ID, SentrixBarcode_A, SentrixPosition_A
T445, 201686270015, R04C02
T446, 201686270015, R05C02
T447, 201686270015, R06C02
T448, 201686270015, R07C02

B [Header]
INVESTIGATOR_NAME, BeadStudio User
PROJECT_NAME, AsliTolunOminEx24_100417
EXPERIMENT_NAME, AsliTolunOminEx24_100417
DATE, 10/5/2017 14:53
[Manifests]
A, InfiniumOmniExpress-24v1-2_A1
[Data]
Sample_ID, SentrixBarcode_A, SentrixPosition_A
T455, 201682390063, R02C01
T456, 201682390063, R03C01
T457, 201682390063, R04C01
T458, 201682390063, R05C01
T459, 201682390063, R06C01
T460, 201682390063, R07C01
T461, 201682390063, R08C01
T462, 201682390063, R09C01

C [Header]
INVESTIGATOR_NAME, SUI
PROJECT_NAME, 190415
EXPERIMENT_NAME, SUI_190415
DATE, 15/04/2019 8:53:00 AM
[Manifests]
A, InfiniumOmniExpress-24v1-3_A1
[Data]
Sample_ID, SentrixBarcode_A, SentrixPosition_A
A1, 203440940039, R04C02
A4, 203440940039, R05C02
X3, 203440940039, R06C02
A5, 203440940039, R07C02

```

**Şekil 3-14: GenomeStudio yazılımına SNP çalışmalarının tanıtılması.**

A) İlk SNP dizilimi grubu örnek dosyası. B) İkinci SNP dizilimi grubu örnek dosyası. C) Üçüncü SNP dizilimi grubu örnek dosyası.

Analizin sonraki aşamaları için kullanılan BAF ve LRR değerleri GenomeStudio yazılımı tarafından idat dosyalarında bulunan her SNP'in sinyal yoğunluk değeri kullanılarak hesaplanır.

İlk seferde MA134 ve MA135 ailesinden üçer birey havuzlanarak ve MA154 ailesinden iki hasta kardeş çalışıldı (Tablo 3-17).

**Tablo 3-17: 062817 kodlu ilk SNP dizilimi grubunda çalışılan bireyler.**

Çalışılan Bireyler	Aile	Örnek Numarası	Çip Üzerindeki Pozisyonu
Z16, Z18, Z21 (Havuzlanarak)	MA134	T445	R04C02
A2, A7, A9 (Havuzlanarak)	MA135	T446	R05C02
Y3	MA154	T447	R06C02
Y5	MA154	T448	R07C02

İkinci seferde MA135 ailesinde hasta dört birey, MA154 ailesinden anne, baba, sağlıklı bir kardeş, ve MA136 ailesinden hasta bir birey çalışıldı (Tablo 3-18).

**Tablo 3-18: 100417 kodlu ikinci SNP dizilimi grubunda çalışılan bireyler.**

Çalışılan Bireyler	Aile	Örnek Numarası	Çip Üzerindeki Pozisyonu
A2	MA135	T455	R02C01
A3	MA135	T456	R03C01
A7	MA135	T457	R04C01
A9	MA135	T458	R05C01
Y1	MA154	T459	R06C01
Y2	MA154	T460	R07C01
Y4	MA154	T461	R08C01
X4	MA136	T462	R09C01

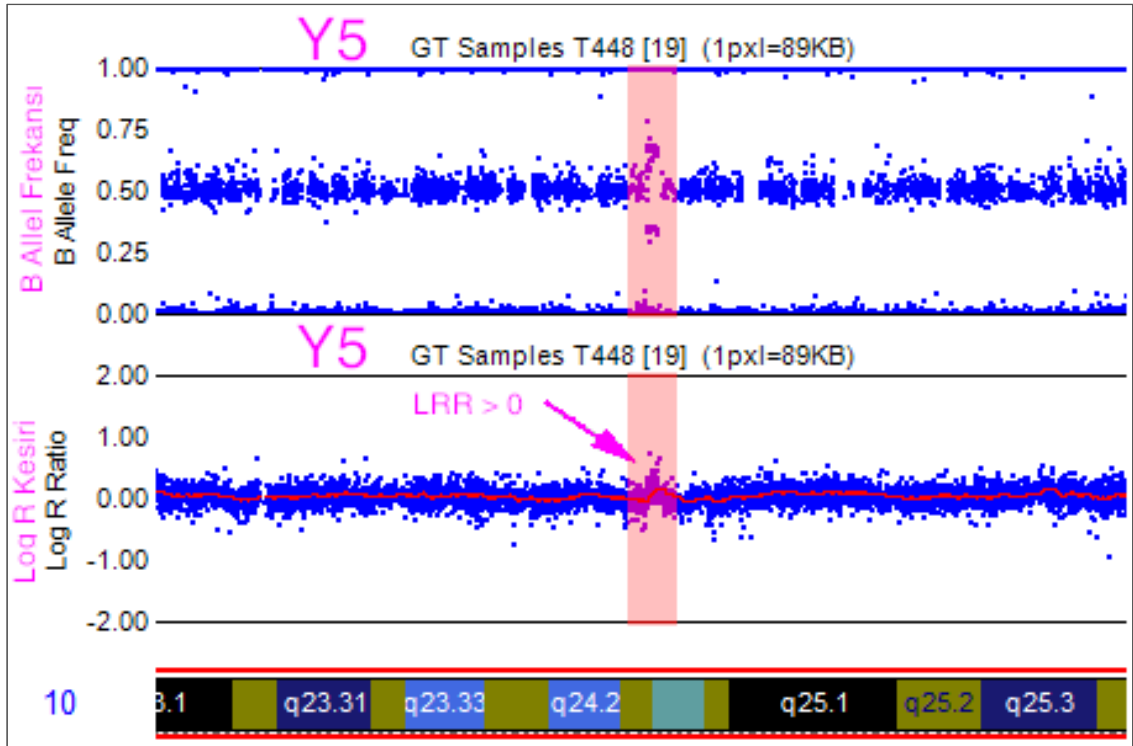
Üçüncü seferde, MA135 ailesinden iki sağlıklı, bir hasta birey, ve MA136 ailesinden bir hasta birey daha çalışıldı (Tablo 3-19).

**Tablo 3-19: 190415 kodlu üçüncü SNP dizilimi grubunda çalışılan bireyler.**

Çalışılan Bireyler	Aile	Örnek Numarası	Çip Üzerindeki Pozisyonu
A1	MA135	A1	R04C02
A4	MA135	A4	R05C02
A5	MA135	A5	R07C02
X3	MA136	X3	R06C02

#### 4.2.5.2. CNV Bölgelerinin Tespiti

Bireyler GenomeStudio'ya tanıtıldıktan sonra LRR ve BAF değerleri ile analizler gerçekleştirildi. CNV'lerin tespiti için LRR değerleri kullanıldı. LRR değeri iki kopyada 0 iken CNV bölgelerinde 0'dan sapar. Bir bölgede, LRR değerinin 0'dan büyük olması o bölgede bir kazanç (Şekil 3-15), 0'dan küçük olması ise bir kayıp olduğunu gösterir.



Şekil 3-15: GenomeStudio yazılımında LRR ve BAF değerlerinin görseli.

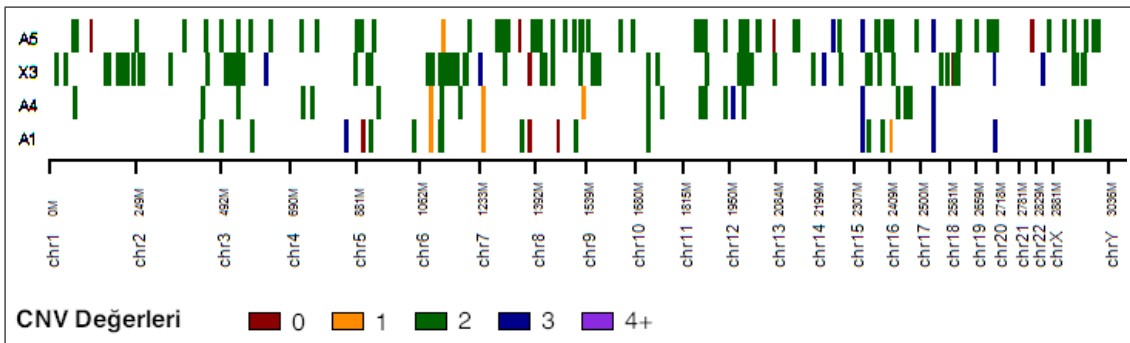
Pembe ok ile gösterilen ve kırmızı kare ile belirtilen bölgede LRR değeri 0'dan büyüktür ve bu bölgede bir kazanç olduğu gözlenmektedir.

CNV bölgelerinin görsel analizinin yanı sıra genom boyu kantitatif bir analiz, GenomeStudio yazılımının eklentisi olan cnvPartition v3.2.0 yazılımı ile gerçekleştirildi. Analiz yapılırken yazılımın varsayılan seçenekleri kullanılmıştır(Şekil 3-16).

Options	
Confidence Threshold	35
Detect Extended Homozygosity	True
Exclude Intensity Only	False
GC Wave Adjust	False
Include Sex Chromosomes	True
Minimum Homozygous Region Size	1000000
Minimum Probe Count	3

Şekil 3-16: CNVpartition yazılımının varsayılan seçenekleri.

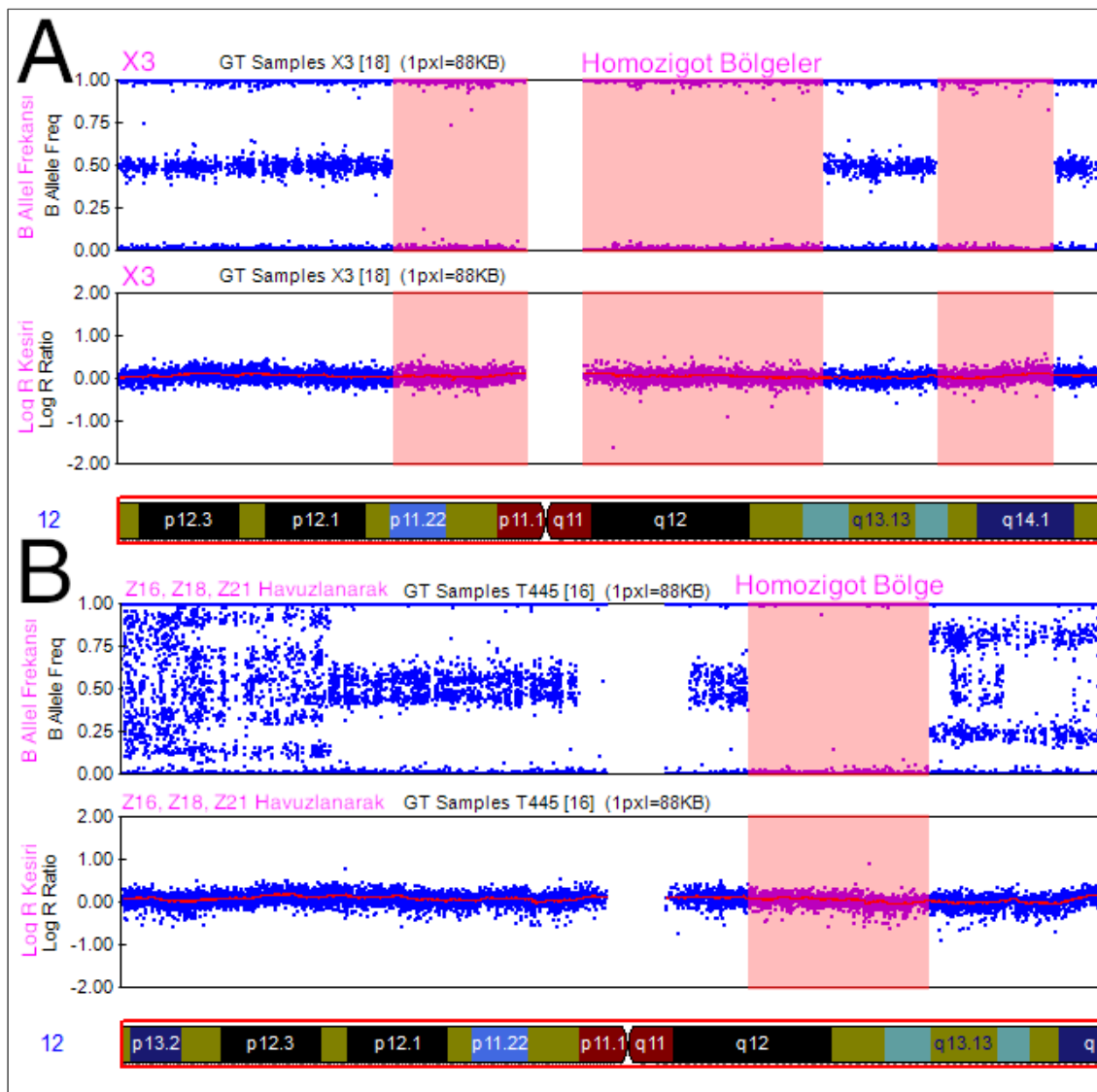
CNVpartition analizinin sonuçları yine başka bir GenomeStudio eklentisi olan CNV Region Report v2.1.2 yazılımı ile görselleştirildi (Şekil 3-17).



Şekil 3-17: CNV analizlerinin CNV Region Report yazılımı ile grafikleştirilmesi.

#### 4.2.5.3. Homozigot Haritalama

Homozigot bölgelerin incelenmesi için ilk olarak GenomeStudio yazılımın eklentisi olan PLINK Input Report ile veri GenomeStudio'dan metin formatında dışarıya alınmıştır. Daha sonrasında Plink<sup>87</sup> yazılımının '--homozyg group' komutu kullanılarak ortak homozigot bölgeler çekilmiştir. Daha sonra yine GenomeStudio görseli üzerinde bu bölgeler pembe ile görselleştirilmiştir (Şekil 3-18).



**Şekil 3-18: Homozigot bölgelerin GenomeStudio yazılımı üzerinden gösterimi.**

A) Tek bir bireyin SNP verisindeki homozigot bölgelerin gösterimi. B) Havuzlanarak çalışılmış örneklerin SNP verisindeki homozigot bölgenin gösterimi.

#### 4.2.6. Bağlantı Analizi

Bağlantı analizi için homozigot bölgelerin plink ile incelenmesi için kullanılan PLINK Input Report eklentisinin metin formatındaki çıktısı kullanılmıştır. MA135 ailesi için yapılan bağlantı analizlerinde ilk olarak kullanılan çiplerin farkından dolayı ortaya çıkan uyumsuz olabilecek veriler bir araya getirilmiştir (Şekil 3-19).

```
import pandas as pd

IkinciSNP = pd.read_csv('MA-135_all_f_linkage.snp', sep='\t')
UcuncuSNP = pd.read_csv('MA135_A1_A4_A5.snp', sep='\t')

IkinciSNP = IkinciSNP[['SNP_ID', 'A2', 'A3', 'A7', 'A9']]
UcuncuSNP = UcuncuSNP[['SNP_ID', 'A1', 'A4', 'A5']]

MA135_7birey = pd.merge(IkinciSNP, UcuncuSNP, on='SNP_ID')
MA135_7birey.to_csv('MA135_7inds.snp', sep='\t', index=False)
```

**Şekil 3-19: Farklı SNP dizilimi çalışmalarının çıktılarının birleştirilmesi için kullanılan betik.**

Sonuçta MA135 ailesinde SNP dizilimi yapılan bütün bireylerin genotiplerini içeren bir dosya elde edilmiştir (Şekil 3-20).



SNP ID	A2	A7	A9	A3	A1	A4	A5
rs10000911	BB	AA	BB	AB	AB	AB	AB
rs10000988	AB	AB	BB	AB	BB	AB	AB
rs10002181	AB	AA	AA	AA	AB	AA	AB
rs10005140	BB	BB	BB	BB	BB	BB	BB
rs10005242	AB	AB	AA	AB	AB	AA	AB
rs10010434	BB	BB	BB	BB	BB	BB	BB
rs10011256	AA	AA	AA	AA	AA	AA	AA
rs1001131	AA	AA	AA	AA	AA	AA	AB
rs1001291	AB	AB	AB	AA	AB	AB	AB
rs10013558	AA	AA	AA	AA	AA	AA	AA
rs10013782	AA	AA	AB	AA	AA	AA	AB
rs10013797	AA	BB	AB	BB	AB	AB	BB
rs10017777	AA	AA	AA	AA	AA	AA	AA
rs10018142	BB	BB	BB	BB	AB	BB	AB
rs10020809	AB	AA	AB	AA	AB	AB	AA
rs10026870	BB	BB	BB	BB	BB	BB	BB
rs10027926	BB	BB	BB	BB	BB	AB	AB
rs10036552	BB	BB	BB	BB	BB	BB	BB
rs10039022	BB	BB	BB	BB	BB	BB	BB
rs10039039	BB	AB	AB	BB	BB	BB	BB

Şekil 3-20: MA135 ailesinde her bireyin verilen SNP belirteçindeki genotiplerini içeren dosya.

Bu genotipler ile beraber aile üyesi olan bireylerin birbiri arasındaki ilişkilerinin bulunduğu “ped” uzantılı aile ağacı dosyası oluşturulmuştur (Şekil 3-21).

MA - 135	101	0	0	2	1
MA - 135	102	0	0	1	1
MA - 135	201	0	0	1	1
MA - 135	202	102	101	2	1
MA - 135	A1	102	101	2	2
MA - 135	204	0	0	1	1
MA - 135	205	201	202	2	1
MA - 135	A2	204	A1	1	2
MA - 135	A3	204	A1	1	2
MA - 135	A4	204	A1	1	1
MA - 135	A7	A3	205	1	2
MA - 135	A9	A3	205	2	2
MA - 135	301	0	0	2	1
MA - 135	A5	A4	301	2	1
MA - 135	401	A7	A5	1	2

Şekil 3-21: Aile bireylerini ve bireylerin birbiri arasındaki ilişkilerini gösteren aile ağacı dosyası.

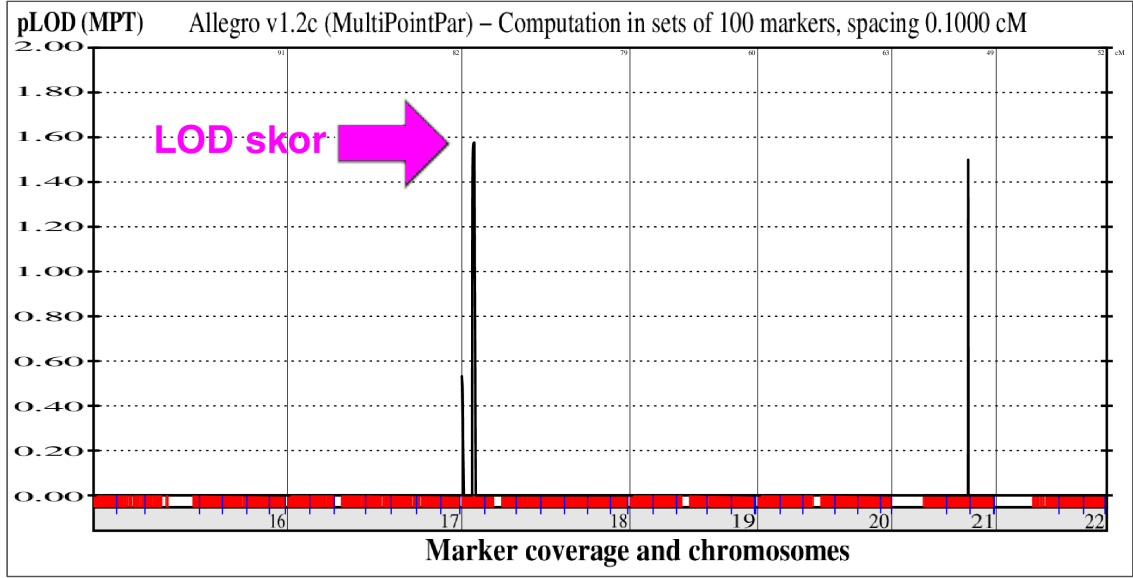
Bağlantı analizinde kullanılacak SNP işaretçilerinin yerlerini belirten “map” uzantılı harita dosyası GenomeStudio yazılımına verilen manifest dosyasından oluşturuldu (Şekil 3-22).

```
# Map name: Illumina_700k_hg19Mb
# Projects: SNP
# Results folder abbreviation: Ill700KHg19
#_dbSNP_ID CHR Sex-Averaged_cM
-----
rs4477212 1 0.082154
rs3094315 1 0.752566
rs3131972 1 0.752721
rs12562034 1 0.768448
rs12124819 1 0.776546
rs11240777 1 0.798959
rs6681049 1 0.800007
rs4970383 1 0.838555
rs4475691 1 0.846808
rs7537756 1 0.854250
rs13302982 1 0.861808
rs1110052 1 0.873558
rs2272756 1 0.882033
rs17160698 1 0.887162
rs3748597 1 0.888659
```

**Şekil 3-22: SNP belirteçlerinin cM cinsinden genomik pozisyonlarının bulunduğu harita dosyası.**

Dosyalarının hazırlanmasından sonra analizler EasyLinkage grafik arayüzü aracılığıyla Allegro yazılımı kullanılarak gerçekleştirildi. Bağlantı analizine başlanmadan önce Mendelyen uyumsuzluk gösteren SNP belirteçleri PedCheck<sup>88</sup> yazılımı ile elendi.

Bağlantı analizi; otozomal dominant, otozomal resesif, X'e bağlı dominant ve X'e bağlı resesif Mendelyen kalıtım modelleri için, 0,1cM belirteç aralığı ile enformatif olmayan belirteçler çıkartılarak gerçekleştirildi. Analizler sonucu LOD skoru grafikleri easyLinkage arayüzü ile oluşturuldu (Şekil 3-23).



Şekil 3-23: Bağlantı analizi sonucu elde edilen LOD skor grafiği.

#### 4.2.7. Ekzom Dizileme

SHFM çalışması dahilinde çalışılan iki birey ve akış hattı yazılımların sınanması için dört birey ekzom dizileme ile çalışılmıştır. SHFM çalışmasındaki iki birey için YCGA'dan (Yale Center for Genome Analysis), akış analizinin denemesindeki bir birey için MacroGen firmasından diğer üç birey için OGT (Oxford Genome Technologies) firmasından hizmet alınmıştır.

FASTQ formatındaki ham verilerden analiz edilecek formata getirmesine kadar ki bütün işlemlerin otomatizasyonu tez kapsamında geliştirilen akış hattı yazılımı pigeon ile yapılmıştır. Varyantların anotasyonu, filtrelenmesi ve aile içi analizleri yine tez kapsamında geliştirilen VCF sonrası analizler yazılımı dove ile gerçekleştirilmiştir.

Akış hattında okumaların kalite kontrolü için fastqc<sup>89</sup>, hizalanması için bwa<sup>90</sup>, SAM/BAM formatlarının manipülasyonu için picard<sup>91</sup>, phred skorlarının rekaliibrasyonu ve varyant çağırma için GATK<sup>92</sup> yazılımları kullanılmıştır. Bütün analizlerde ortak olarak kullanılan akış hattı yapılandırması aşağıdaki gibidir.

```
[GENERAL]
reference_genome = ucsc.hg19.fasta
knownSNP       = hg19BroadInstitute/dbsnp_138.hg19.vcf
knownINDEL     = Mills_and_1000G_gold_standard.indels.hg19.vcf
refSeq         = geneTrack.refSeq

[PIPELINE]
pipeline = fastqc

                                mem
                                SortSam
                                RemoveSam
                                MarkDuplicates
                                RemoveSortedBam
                                BuildBamIndex
                                BaseRecalibrator
                                PrintReads
                                HaplotypeCaller
                                OVAHaplotypeCaller
                                geneCov
fastqc = /media/bar/2tb/NGSpipeline/tools/FastQC/fastqc
bwa    = /media/bar/2tb/NGSpipeline/tools/bwa/bwa
picard = /media/bar/2tb/NGSpipeline/tools/picard/picard.jar
GATK   = /media/bar/2tb/NGSpipeline/tools/GATK/GenomeAnalysisTK.jar -T
dove   = dove
rm     = rm
java   = java

[fastqc]
tool = ${PIPELINE:fastqc}
input_from = input_files
args = input_placeholder
      --outdir=${PROJECT:output_dir}/${PROJECT:project_name}/fastqc
      --extract
dump_dir = fastqc

[mem]
tool = ${PIPELINE:bwa}
sub_tool = mem
input_from = input_files
args = -t 16
      -R "@RG\tID:foo\tSM:foo\tPL:illumina\tLB:lib1\tPU:foo"
      ${GENERAL:reference_genome}
input_placeholder > output_placeholder

suffix = _bwa
ext = sam
input_multi = paired

[SortSam]
tool = ${PIPELINE:picard}
sub_tool = SortSam
input_from = mem
args = I=input_placeholder
      O=output_placeholder
      SORT_ORDER=coordinate

suffix = _picardSort
ext = bam
```

```

container = ${PIPELINE:java}

[RemoveSam]
tool      = ${PIPELINE:rm}
input_from = mem
args      = input_placeholder

[MarkDuplicates]
tool      = ${PIPELINE:picard}
sub_tool  = MarkDuplicates
input_from = SortSam
args      = ASSUME_SORTED=True
           REMOVE_DUPLICATES=True
           I=input_placeholder
           O=output_placeholder
           M=markdup_metrics.txt
suffix    = _picardMarkDup
ext       = bam
container = ${PIPELINE:java}

[RemoveSortedBam]
tool      = ${PIPELINE:rm}
input_from = SortSam
args      = input_placeholder

[BuildBamIndex]
tool      = ${PIPELINE:picard}
sub_tool  = BuildBamIndex
input_from = MarkDuplicates
args      = I=input_placeholder
           O=output_placeholder
ext       = bai
container = ${PIPELINE:java}

[BaseRecalibrator]
tool      = ${PIPELINE:GATK}
sub_tool  = BaseRecalibrator
input_from = MarkDuplicates
args      = -R ${GENERAL:reference_genome}
           -I input_placeholder
           -L ${PROJECT:bed_file}
           -knownSites ${GENERAL:knownSNP}
           -knownSites ${GENERAL:knownINDEL}
           -o output_placeholder
paired_output = True
dump_dir     = GATKvarcall
suffix       = _recal_data
ext          = table
container    = ${PIPELINE:java}

[PrintReads]
tool      = ${PIPELINE:GATK}
sub_tool  = PrintReads
input_from = BaseRecalibrator
input_multi = paired
args      = -R ${GENERAL:reference_genome}
           -I input_placeholder
           -L ${PROJECT:bed_file}
           -BQSR secondary_in_placeholder
           -o output_placeholder
dump_dir  = GATKvarcall/bam
suffix    = _recal_reads
ext       = bam
container = ${PIPELINE:java}
secondary_in_placeholder = True

[HaplotypeCaller]
tool      = ${PIPELINE:GATK}

```

```

sub_tool = HaplotypeCaller
input_from = PrintReads
args = -R ${GENERAL:reference_genome}
      -I input_placeholder
      -L ${PROJECT:bed_file}
      --genotyping_mode DISCOVERY
      -stand_emit_conf 1
      -stand_call_conf 20
      -o output_placeholder
dump_dir = GATKvarcall/vcf
suffix = _haplotypcaller
ext = vcf
container = ${PIPELINE:java}
secondary_out_placeholder = True
secondary_ext = bam
secondary_suffix = _bamout
secondary_dump_dir = bamout

```

```

[OVAHaplotypeCaller]
tool = ${PIPELINE:dove}
sub_tool = OVA
input_from = HaplotypeCaller
args = -i input_placeholder
      -o output_placeholder
      -q 199
      --omim --hpo
dump_dir = GATKvarcall/annotations
suffix = _ensembl_vep
ext = csv

```

```

[geneCov]
tool = ${PIPELINE:GATK}
sub_tool = DepthOfCoverage
input_from = PrintReads
args = -R ${GENERAL:reference_genome}
      -I input_placeholder
      -L ${PROJECT:bed_file}
      -o output_placeholder
      --calculateCoverageOverGenes ${GENERAL:refSeq}
      -ct 1
dump_dir = GATKvarcall/DepthOfCoverage
suffix = _gene_summary
ext = table
container = ${PIPELINE:java}

```

Farklı çalışmalarda kullanılan cihazlar ve yakalama kitleri Tablo 3-20’de verilmiştir.

**Tablo 3-20: Dizilemede çalışmalarında kullanılan cihaz ve kitle**

Çalışma/Örnek	Kurum/Firma	Cihaz	Yakalama kiti
A4	YCGA	Illumina HiSeq 2500 Rapid	IDT xGen exome kit
X3	YCGA	Illumina HiSeq 2500 Rapid	IDT xGen exome kit
Test 1	Macrogen	Illumina HiSeq 2000	SureSelect Human All Exon V6
Test 2	OGT	Illumina HiSeq 2000	SureSelect Human All Exon V5

Veri analizi sonucu elde edilen varyantlar adım adım şekildeki pigeon yapılandırma dosyası kullanılarak filtrelenmiştir (Şekil 3-24). Filtreme öncesi varyantlar picus<sup>93</sup> yazılımı ile ACMG yönergelerine göre belli kanıtlar kullanılarak sınıflandırılmıştır. Elde edilen varyantlar, BAM formatından okumalar IGV<sup>94</sup> ile tekrardan görsel kontrolü için görselleştirilmiştir.



```

[PIPELINE]
pipeline = protein_coding maf impact vus hom het

[protein_coding]
tool = dove
sub_tool = VF
input_from = input_files
args = -i input_placeholder
      -o output_placeholder
      -c biotype eq protein_coding
suffix = _protein_coding

[maf]
tool = dove
sub_tool = VF
input_from = protein_coding
args = -i input_placeholder
      -o output_placeholder
      -c gnomad lt 0.05
suffix = _maf

[impact]
tool = dove
sub_tool = VF
input_from = maf
args = -i input_placeholder
      -o output_placeholder
      -c impact eq HIGH
suffix = _impact

[vus]
tool = dove
sub_tool = VF
input_from = impact
args = -i input_placeholder
      -o output_placeholder
      -c significance in "'uncertain_significance': 1"
suffix = _vus

[hom]
tool = dove
sub_tool = VF
input_from = vus
args = -i input_placeholder
      -o output_placeholder
      -c GT eq 1/1
suffix = _hom

[het]
tool = dove
sub_tool = VF
input_from = vus
args = -i input_placeholder
      -o output_placeholder
      -c GT eq 0/1
suffix = _het

```

Şekil 3-24: Ham veri analizi sonrası elde edilen varyantların adım adım filtrelenmesi.



#### 4.2.8. Aday Genlerin Sanger Dizileme ile Taranması

Varyasyonları taşıyan genomik bölgeleri çoğaltacak primerler NCBI’ın Primer-BLAST<sup>95</sup> yazılımı kullanılarak tasarlandı. Tasarlanan primerlerin in-siliko ortamda test edilmesi UCSC’nin in-silico PCR<sup>96</sup> aracı ile gerçekleştirilmiştir. Sanger dizileme sonrası kromatogramların görselleştirilmesi CLC yazılımı ile gerçekleştirilmiştir.

Homozigot haritalama sonucu MA134 ailesinde homozigot bölgede bulunan *WNT10B* geni Tablo 3-21’de verilen primerler ile taranmıştır.

**Tablo 3-21: MA134 ailesinde Sanger dizileme ile dizilenen bölgeler.**

HGNC gen sembolü	RefSeq eşi	İncelenen varyasyon/bölge	Primer F/R	Uzunluk (bp)	Tm (°C)
<i>WNT10B</i>	NM_003394.4	Ekzon2	GTGTCTGATTGGGCAAGGTT/ CTCATTGCTTAGAGCCCTGG	491	57.5
<i>WNT10B</i>	NM_003394.4	Ekzon3	GGAGAGTTGGAGGGGTCTG/ GAAACCATCCCTTCCCGC	405	55.8
<i>WNT10B</i>	NM_003394.4	Ekzon4	TGCCTGTCAACCTTACCTCC/ TAACCAGGCCTCAAAGCTG	470	53.5
<i>WNT10B</i>	NM_003394.4	Ekzon5	TGTGCCTCTGTGTTCTGTCC/ GAAATCAGAGCAAAGGGCTG	598	56

Homozigot haritalama sonucu MA136 ailesinde homozigot bölgede bulunan *WNT10B* geni Tablo 3-22’de verilen primerler ile taranmıştır.

**Tablo 3-22: MA136 ailesinde Sanger dizileme ile dizilenen bölgeler.**

HGNC gen sembolü	RefSeq eşi	İncelenen varyasyon/bölge	Primer F/R	Uzunluk (bp)	Tm (°C)
<i>WNT10B</i>	NM_003394.4	Ekzon1	CCTGTCTCTTCCTCGTGTGC/ CTTCTTCAGTGGGGCCAAGC	566	62,5
<i>WNT10B</i>	NM_003394.4	Ekzon2	GTGTCTGATTGGGCAAGGTT/ CTCATTGCTTAGAGCCCTGG	491	57,5
<i>WNT10B</i>	NM_003394.4	Ekzon3	GGAGAGTTGGAGGGGTCTG/ GAAACCATCCCTTCCCGC	405	55,8
<i>WNT10B</i>	NM_003394.4	Ekzon4	TGCCTGTCAACCTTACCTCC/ TAACCAGGCCTCAAAAGCTG	470	53,5
<i>WNT10B</i>	NM_003394.4	Ekzon5	TGTGCCTCTGTGTTCTGTCC/ GAAATCAGAGCAAAGGGCTG	598	56

Zincir tepkimesinde kullanılan çözelti miktarları üretici firmanın yönergelerine uygun kullanılmış ve bütün primerler en uygun koşulları için DMSO ve betain ile optimize edilmiştir. PZR sırasında kullanılan protokol Tablo 3-23’de verilmiştir.

**Tablo 3-23: Sanger dizileme öncesi yapılan PZR için iniş(touchdown) protokolü.**

İniş sıcaklıkları her primerin erime sıcaklığına uygun olarak değiştirilmiştir.

Adım	Sıcaklık	Süre	Tekrar
<b>Erime</b>	95C	5 dk	1X
	94C	45 sn	
<b>İniş (Touchdown)</b>	61-->53 (0,5C)	45 sn	16X
	72C	1 dk	
	94C	45 sn	
<b>PZR</b>	52C	45 sn	20X
	72C	1 dk	
<b>Son Uzama</b>	72C	15 dk	1X

#### 4.2.9. Varyantların İsimlendirilmesi

Sanger dizileme ya da ED sonucu bulunan bütün varyantlar GRCh37 referans genomuna göre HGNC (HUGO Gene Nomenclature Committee) nomenklatürüne uygun olarak isimlendirilmiştir. Tüm isimlendirmelerin uygunluğu Mutalyzer<sup>97</sup> tarafından kontrol edilmiştir.

#### 4.2.10. HRMA ve Sanger Dizileme ile Ailesel Ayırışım

Varyasyonları taşıyan genomik bölgeleri çoğaltacak primerler NCBI'ın Primer-BLAST yazılımı kullanılarak tasarlandı (Tablo 3-24). Tasarlanan primerlerin in-silico ortamda test edilmesi UCSC'nin in-silico PCR aracı ile gerçekleştirilmiştir. Sanger dizileme sonrası kromatogramların görselleştirilmesi CLC yazılımı ile gerçekleştirilmiştir.

**Tablo 3-24: HRMA ile taranan bölge primerleri**

HGNC gen sembolü	RefSeq eşi	İncelenen varyasyon/bölge	Primer F/R	Uzunluk (bç)	Tm (°C)
<i>WNT10B</i>	NM_003394.4	Ekzon5	GCCTCTCAGGAGAGCTGGT/ ACAGCACAGGCTGCCACA	140	57

#### 4.2.11. Gerçek Zamanlı PZR ile CNV'lerin doğrulanması

Varyasyonları taşıyan genomik bölgeleri çoğaltacak primerler NCBI'nin Primer-BLAST yazılımı kullanılarak tasarlandı (Tablo 3-25). Tasarlanan primerlerin in-siliko ortamda test edilmesi UCSC'nin in-silico PCR aracı ile gerçekleştirilmiştir. Reaksiyonun gerçek zamanlı takibi ve grafiklerin oluşturulması LightCycler 480 yazılımı tarafından gerçekleştirilmiştir.

**Tablo 3-25: Gerçek zamanlı CNV doğrulaması için kullanılan primerler.**

HGNC gen sembolü	İncelenen bölge	Primer F/R	Uzunluk (bp)	Tm (°C)
KIF1C NM_006612.6	Ekzon1	CACTGTAGGCGAGGAGGAAG/ GCTTCACCTCCTGCAGAATC	131	60
-	Chr10:102,9-103,3Mb	GTGGCCAGAGTGGAAGAATTAC/ ATCCACTACCCAGCAAACCC	201	60.5

## 5. BULGULAR

Çalışma kapsamında SHFM hastalığından etkilenmiş dört aile çalışılmıştır. Çalışmada yer alan ailelerdeki fenotipler aynı olmasına rağmen yöntemdeki farklılaşmalardan dolayı her ailenin bulguları kendi başlıkları içerisinde sunulmuştur.

### 5.1. Aile MA134

#### 5.1.1. Genom Boyu SNP Dizilimi Verisinin Analizi

Ailenin ayrı kollarından üç ayrı bireyin DNA'ları karıştırılarak, Illumina platformunda, 700k SNP dizilimi ile genotipleme gerçekleştirildi. Ayrı kollardan etkilenmiş bireylerin DNA'sı karıştırılması, fenotip ile ilişkili homozigot bölgelere odaklanmayı sağlamıştır. Aynı aile içerisinde özellikle homozigot IBD bölgelerini azaltmayı sağlayacağı için bu aileler için uygun bir stratejidir.

GenomeStudio yazılımı ile ham veriler işlenerek, BAF ve LRR değerleri kullanılarak tüm kromozomlar homozigotluk açısından görsel ve kantitatif olarak incelendi. Üç bireyin DNA'sı havuzlanarak çalışıldığı için CNV analizi mümkün değildir. GenomeStudio yazılımının, Plink Input Report eklentisi ile genotiplerin düz metin çıktısı kullanılarak plink yazılımı ile homozigot bölgeler gösterildi (Tablo 5-1).

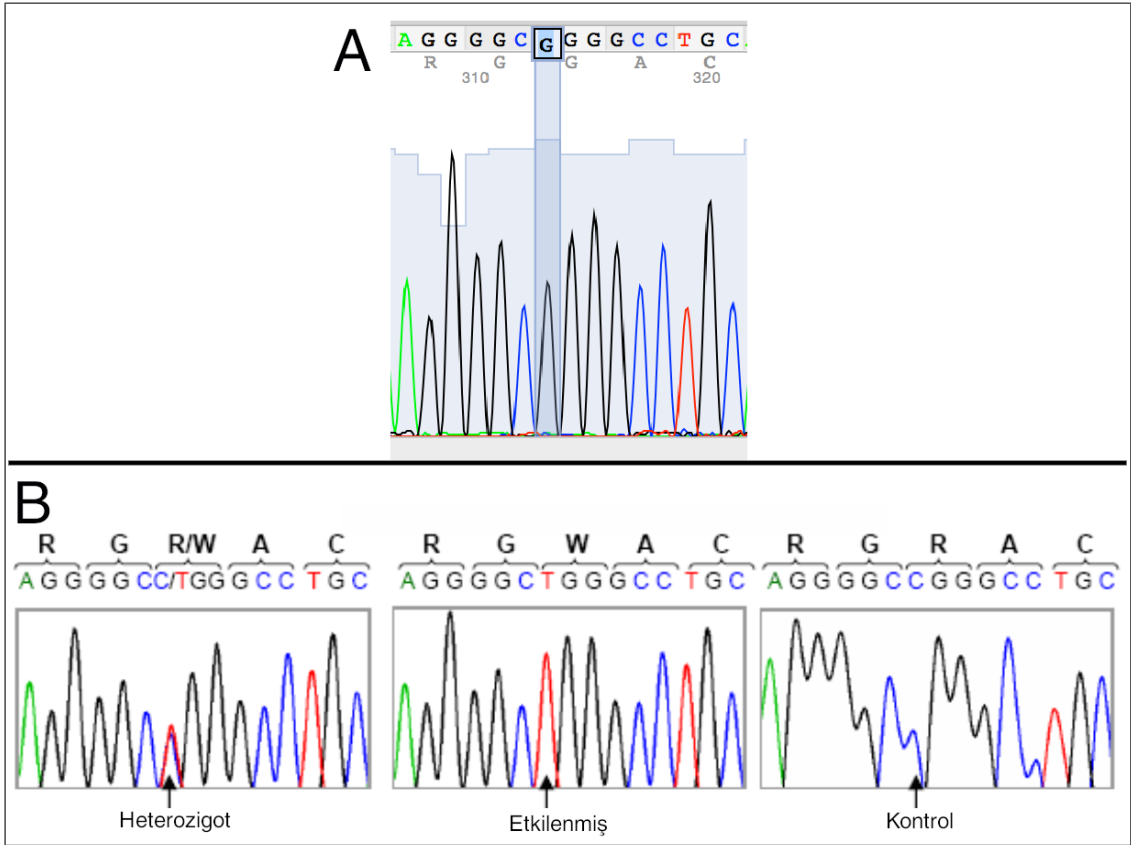
**Tablo 5-1: MA134 ailesinde havuzlanarak SNP dizilimi yapılan örneklerin Plink analizi sonucu tespit edilen homozigot bölgeler.**

CHR	SNP1	SNP2	POS1	POS2	KB	NSNP	DENSITY
3	rs11719995	rs9324	51.593.827	52.825.585	1.231.759	208	5922
7	rs846433	rs2158324	119.502.677	120.519.615	1.016.939	104	9778
10	rs10883551	rs10786792	102.633.059	106.166.570	3.533.512	790	4473
12	rs1500072	rs9705474	33.285.193	34.826.574	1.541.382	196	7864
12	rs4002730	rs3759139	37.876.400	39.087.609	1.211.210	100	12112
12	rs10785280	rs10735832	42.105.409	51.596.763	9.491.355	2062	4603
14	rs9630389	rs1315731	66.862.929	67.938.472	1.075.544	106	10147
14	rs2896063	rs2540879	88.360.957	91.708.195	3.347.239	892	3753
23	rs34565429	rs5915179	48.931.481	50.056.814	1.125.334	100	11253
23	rs5978179	rs4826274	53.676.588	55.248.090	1.571.503	136	11555
23	rs9723973	rs5965201	61.719.290	65.657.652	3.938.363	128	30768
23	rs6616255	rs5987498	100.807.909	102.191.083	1.383.175	123	11245
23	rs12390885	rs669237	153.356.015	154.916.845	1.560.831	208	7504

Tespit edilen bu on üç bölgeden kromozom 12 üzerinde 42.105.409 (rs10785280) – 51.596.763(rs10735832) bölgeleri arasında bulunan, daha önce literatürde birçok SHFM vakası ile ilişkilendirilmiş *WNT10B*<sup>98-104</sup> geninde bulunabilecek varyasyonların araştırılması için genin Sanger dizileme ile taranması uygun görüldü.

### 5.1.2. Aday Genlerin Sanger Dizileme ile Taranması

Sanger dizileme sonucunda daha önce yüksek lisans tez danışmanım Uğur İşeri S. Tolun A. tarafından tespit edilen<sup>103</sup> NM\_003394.4:c.994C>T:p.Arg332Trp ile aynı noktada NM\_003394.4::c.994C>G:p.Arg332Gly varyasyonu tespit edilmiştir (Şekil 5-1).

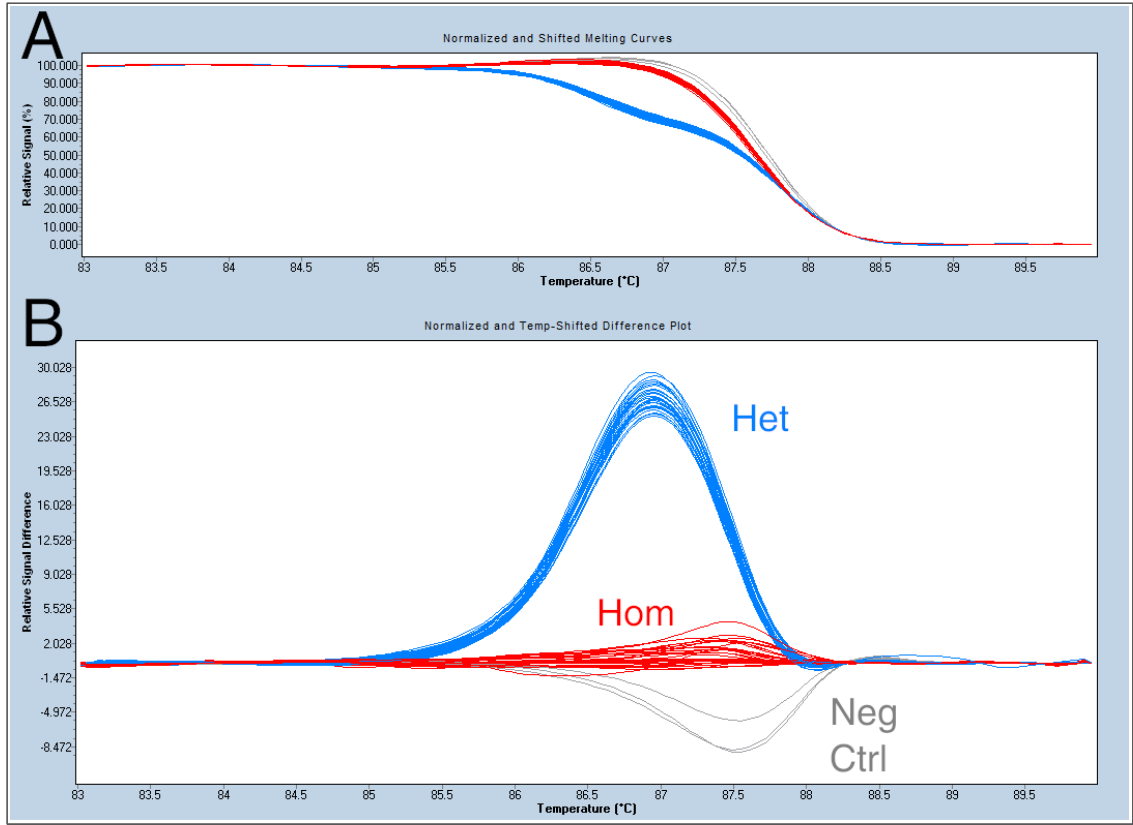


**Şekil 5-1: *WNT10B* geninde Sanger dizileme ile tespit edilen patojenik varyasyonun kromatogram görseli**

A) MA134 ailesinde tespit edilen NM\_003394.4::c.994C>G varyasyonu. B) Ugur İseri S., Tolun A. tarafından tespit edilen NM\_003394.4:c.994C>T varyasyonu.

### 5.1.3. HRMA ve Sanger Dizileme ile Ailesel Ayrışım

Sanger dizileme ile tarama sonucu bulunan mutasyonun aile içindeki ayrışımı, iki HRMA ve bir Sanger dizileme çalışmaları ile tespit edildi. İlk HRMA çalışmasında örnekler, normalize ve sıcaklık-kaydırılmış sinyallerine göre gruplandılar (Şekil 5-2).



**Şekil 5-2: İlk HRMA aşamasında sinyallerin erime noktalarına göre ayrımı.**

A) Normalize ve sıcaklık-kaydırılmış sinyallerin erime eğrilerinin gösterimi. B) Normalize ve sıcaklık-kaydırılmış sinyal değerlerinin seçilen bireye farklarının bağlı gösterimi.

Bu gruplama sonucu mutasyon lokusunda homozigot (wt/wt ya da mt/mt) ve heterozigot (Taşıyıcı) olmaları açısından iki sınıfa ayrıldılar (Tablo 5-2). Bu gruplandırma sonucu bireylerin fenotipleri ve grupları açısından üç durum gözlenmektedir. Hasta bireyler sadece birinci grupta gözlenirken, sağlıklı bireyler hem bir hem de ikinci grup içerisinde gözlenmektedir. HRMA, homozigot ve heterozigot genotip ayrımını daha iyi yaptığı göz önüne alındığında, birinci gruba düşen sağlıklı bireylerin genotipinin wt/wt, ikinci grubu oluşturan sağlıklı bireylerin genotipinin wt/mt, ve birinci gruba düşen hasta bireylerin genotipinin mt/mt olabileceği söylenebilir.

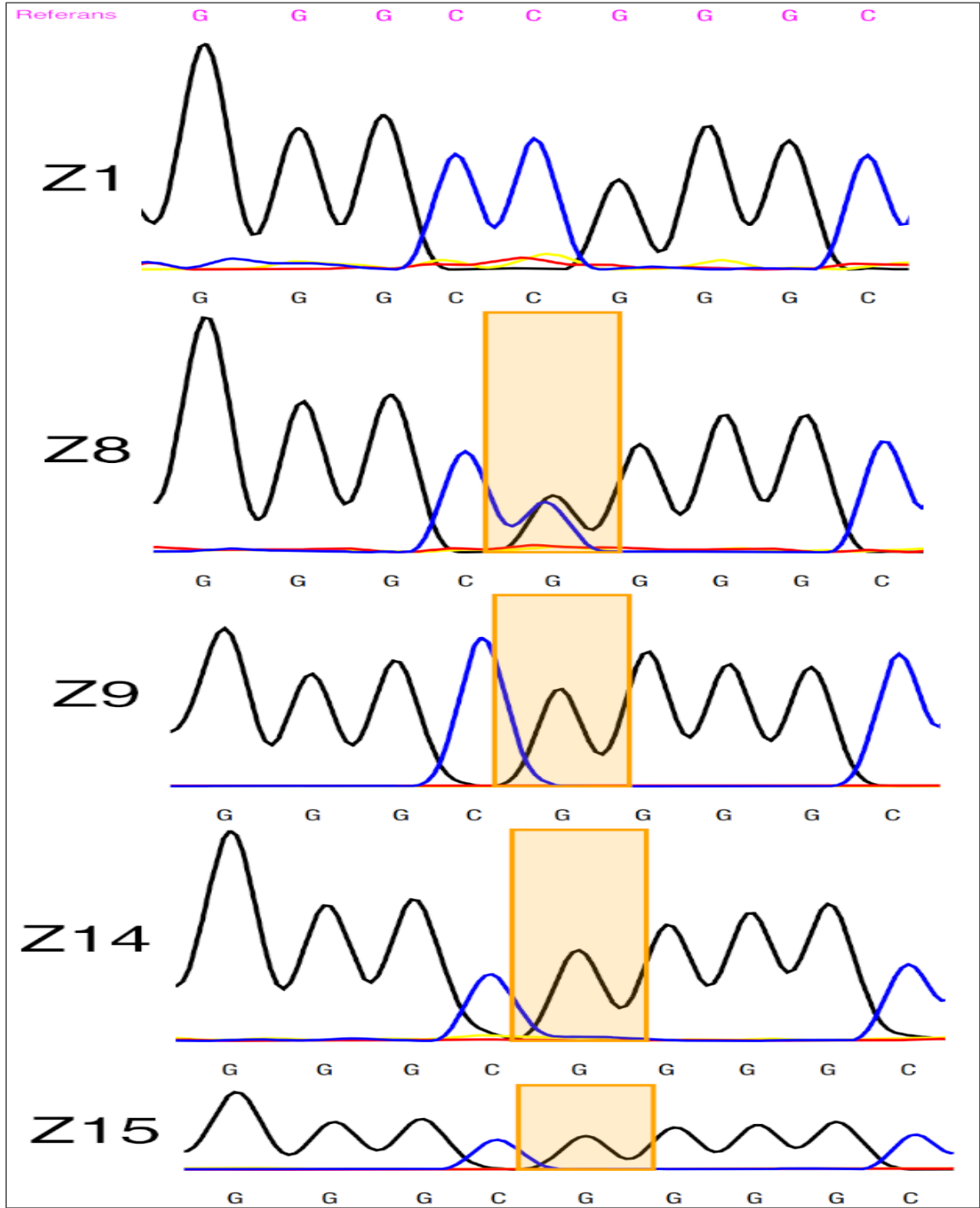


**Tablo 5-2: HRMA erime eğrilerine göre bireylerin grupları ve fenotipleri.**

<b>Birey</b>	<b>Grup</b>	<b>Fenotip</b>	<b>Birey</b>	<b>Grup</b>	<b>Fenotip</b>
<b>Z1</b>	1	Sağlıklı	<b>Z12</b>	1	Hasta
<b>Z2</b>	2	Sağlıklı	<b>Z13</b>	2	Sağlıklı
<b>Z3</b>	2	Sağlıklı	<b>Z14</b>	1	Hasta
<b>Z4</b>	1	Sağlıklı	<b>Z15</b>	1	Hasta
<b>Z5</b>	2	Sağlıklı	<b>Z16</b>	1	Hasta
<b>Z6</b>	2	Sağlıklı	<b>Z17</b>	2	Sağlıklı
<b>Z7</b>	2	Sağlıklı	<b>Z18</b>	1	Hasta
<b>Z8</b>	2	Sağlıklı	<b>Z19</b>	2	Sağlıklı
<b>Z9</b>	2	Sağlıklı	<b>Z20</b>	1	Hasta
<b>Z10</b>	2	Sağlıklı	<b>Z21</b>	1	Hasta
<b>Z11</b>	1	Hasta			

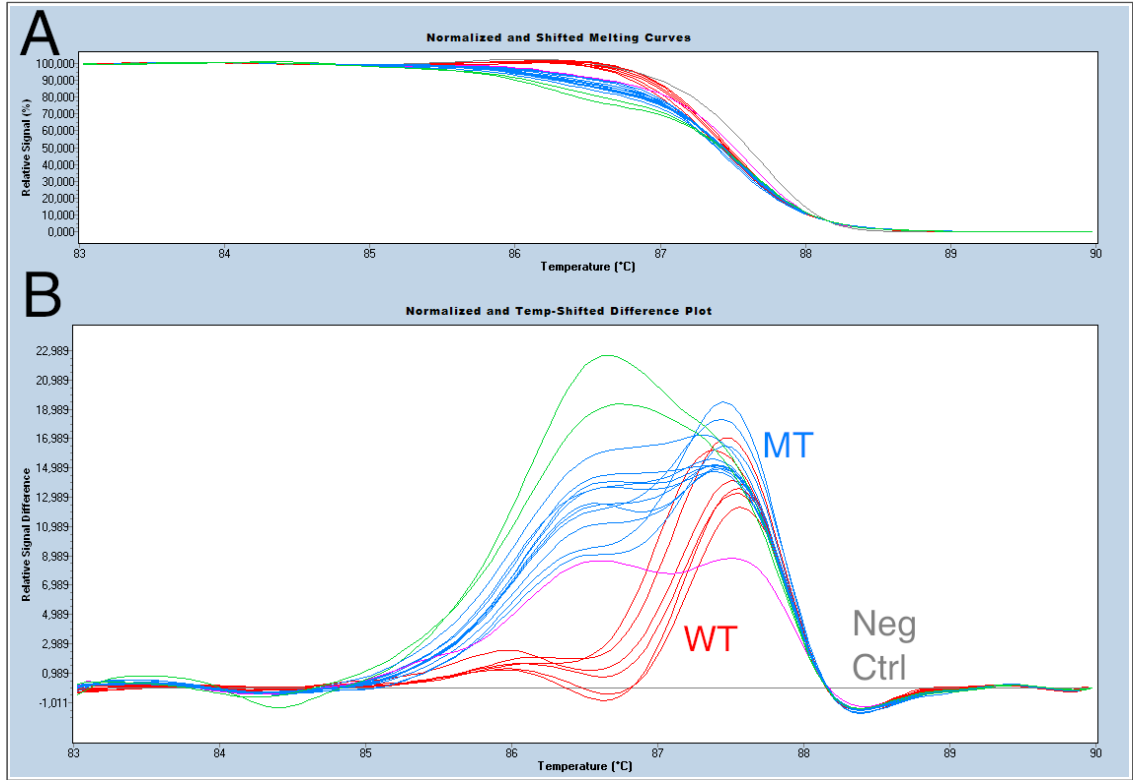
Sınıflandırma sonucu ayrılan gruplardan seçilen bireylerin genotipleri Sanger dizilime ile kontrol edildi (Şekil 5-3). Her üç durumunda örneklenmesi amacıyla Sanger dizilemeye, birinci gruba düşen sağlıklı birey Z1, ikinci gruba düşen sağlıklı birey Z8 ve birinci gruba düşen hasta birey Z15 seçildi.

Klinisyen tarafından gönderilen aile ağacında Z9 bireyinin sağlıklı gözükmesi ancak HRMA sonucunda birinci grupta olup çocuklarının da hasta olması resesif kalıtım modeli ile uyuşmadığından, Z9 ve Z15'in hasta kardeşi olan Z14'te Sanger dizileme ile dizilendi. Klinisyenden Z9'un hasta olduğu doğrulanması sonrası Z9'un genotipinin de mt/mt olduğu Sanger dizileme ile doğrulandı.



**Şekil 5-3: HRMA ile sınıflandırılan genotiplerin Sanger dizileme ile ayrımı.**

Birinci gruba düşen homozigot genotipli (wt/wt ve mt/mt) sağlıklı ve hasta bireylerin ayrımı ikinci bir HRMA çalışması ile gerçekleştirildi ve örnekler, normalize ve sıcaklık-kaydırılmış sinyallerine göre gruplandılar (Şekil 5-4).



**Şekil 5-4: İkinci HRMA aşamasında sinyallerin erime noktalarına göre ayrımı.**

A) Normalize ve sıcaklık-kaydırılmış sinyallerin erime eğrilerinin gösterimi. B) Normalize ve sıcaklık-kaydırılmış sinyal değerlerinin seçilen bireye farklarının bağlı gösterimi.

Bu çalışmada ayrımın sağlanması her örneğe genotipinin wt/wt olduğu bilinen sağlıklı kontrol DNA'sı toplam hacimin yüzde onu olacak kadar eklendi. Bu durum, wt/wt genotipli bireylerde bir fark yaratmazken, genotipi mt/mt olan bireylere wt DNA eklendiği için kısmi heterozigotluk yaratmaktadır. Böylece iki homozigot grup yerine bir homozigot bir kısmi heterozigot grup elde edilir ve bu grup arasında oluşan erime noktası farkı HRMA ile gözlenebilir (Tablo 5-3).

**Tablo 5-3: HRMA erime eğrilerine göre homozigot bireylerin grupları ve fenotipleri.**

Birey	Grup	Fenotip
Z1	2	Sağlıklı
Z4	2	Sağlıklı
Z11	1	Hasta
Z12	1	Hasta
Z14	1	Hasta
Z15	1	Hasta
Z16	1	Hasta
Z18	1	Hasta
Z20	1	Hasta
Z21	1	Hasta

## 5.2. Aile MA135

### 5.2.1. Genom Boyu SNP Genotipleme için SNP Dizilimi Verisinin Analizi

Ailenin ayrı kollarından üç ayrı bireyin DNA'ları karıştırılarak, Illumina platformunda, 700k SNP dizilimi ile genotipleme gerçekleştirildi. Etkilenmiş bireylerin DNA'sı karıştırılması, fenotip ile ilişkili homozigot bölgelere odaklanmayı sağlamıştır. Aynı aile içerisinde özellikle homozigot IBD bölgelerini azaltmayı sağlayacağı için bu aileler için uygun bir stratejidir.

GenomeStudio yazılımı ile ham veriler işlenerek, BAF ve LRR değerleri kullanılarak tüm kromozomlar homozigotluk açısından görsel ve kantitatif olarak incelendi. Üç bireyin DNA'sı karıştırıldığı için CNV analizi yapılmadı. GenomeStudio yazılımının, Plink Input Report eklentisi ile genotiplerin düz metin çıktısı kullanılarak plink yazılımı ile homozigot bölgeler gösterildi (Tablo 5-4).

**Tablo 5-4: MA135 ailesinde havuzlanarak SNP dizilimi yapılan örneklerin Plink analizi sonucu tespit edilen homozigot bölgeler.**

CHR	SNP1	SNP2	POS1	POS2	KB	NSNP	DENSITY
3	rs35623035	rs11919418	48.630.252	50.233.215	1.602.964	264	6072
4	rs10517277	rs6849320	33.466.803	34.765.595	1.298.793	116	11196
5	rs17692694	rs3805685	130.330.467	131.528.153	1.197.687	122	9817
6	rs6927204	rs9481581	114.859.899	115.922.400	1.062.502	129	8236
11	rs7947391	rs1638560	66.186.882	67.220.175	1.033.294	197	5245
16	rs1428106	rs14920	66.536.257	67.691.515	1.155.259	235	4916
23	rs36047913	rs12835276	20.615.923	21.752.008	1.136.086	101	11248
23	rs4131729	rs12690009	51.697.194	53.871.623	2.174.430	122	17823
23	rs6612250	rs4384157	55.035.557	58.483.247	3.447.691	152	22682
23	rs1352015	rs907150	65.843.843	68.143.998	2.300.156	142	16198
23	rs6616255	rs7876320	100.807.909	102.025.971	1.218.063	112	10876
23	rs5987171	rs669237	153.123.411	154.916.845	1.793.435	255	7033

Tespit edilen homozigot bölgelerde SHFM ile ilişkili olabilecek bir gen ya da bölgeye rastlanmadı. Daha sonraki SNP çalışmalarında MA135 ailesinden bireyler, DNA örnekleri karıştırılmadan çalışıldı ve aynı şekilde plink yazılımı ile homozigot haritalama yapıldı(Tablo 5-5).

**Tablo 5-5: MA135 ailesinde etkilenmiş beş bireyde ortak tespit edilen homozigot bölgeler.**

CON: paylaşılan en küçük ortak bölge. UNION: paylaşılan en büyük ortak bölge.

IID	CHR	SNP1	SNP2	BP1	BP2	KB	NSNP
A2	3	rs9882962	rs11919418	48.573.311	50.233.215	1.659.905	276
A3	3	rs2255532	rs1061474	48.624.611	50.144.951	1.520.341	244
A7	3	rs7631574	rs2006199	48.663.046	50.241.621	1.578.576	252
A9	3	rs7631574	rs1061474	48.663.046	50.144.951	1.481.906	231
A1	3	rs13060933	rs2005557	48.606.701	49.701.298	1.094.598	181
CON	3	rs7631574	rs2005557	48.663.046	49.701.298	1.038.253	160
UNION	3	rs9882962	rs2006199	48.573.311	50.241.621	1.668.311	278

Ancak aynı homozigot bölge aynı zamanda etkilenmemiş bireyler olan A4 ve A5'de de tespit edildi(Tablo 5-6).

**Tablo 5-6: MA135 ailesinde etkilenmemiş iki bireyde ortak tespit edilen homozigot bölgeler.**

<b>IID</b>	<b>CHR</b>	<b>SNP1</b>	<b>SNP2</b>	<b>POS1</b>	<b>POS2</b>	<b>KB</b>	<b>NSNP</b>	<b>DENSITY</b>
<b>A4</b>	3	rs9882962	rs11919418	48.573.311	50.233.215	1.659.905	276	6014
<b>A5</b>	3	rs9882962	rs2006199	48.573.311	50.241.621	1.668.311	278	6001

Bireylerin DNA'ları karıştırılmadan yapılan SNP dizilimi incelemeleri sürecinde CNV analizleri GenomeStudio yazılımının eklentisi olan CNVpartition ile gerçekleştirildi. İkinci SNP dizilimi çalışmasında çalışılan MA135 ailesine ait dört bireyde chr15:24.525.505-24.668.007 bölgesinde artış tespit edildi (Tablo 5-7). Tespit edilen CNV aynı zamanda GenomeStudio üzerinde de gösterildi (Şekil 5-5).

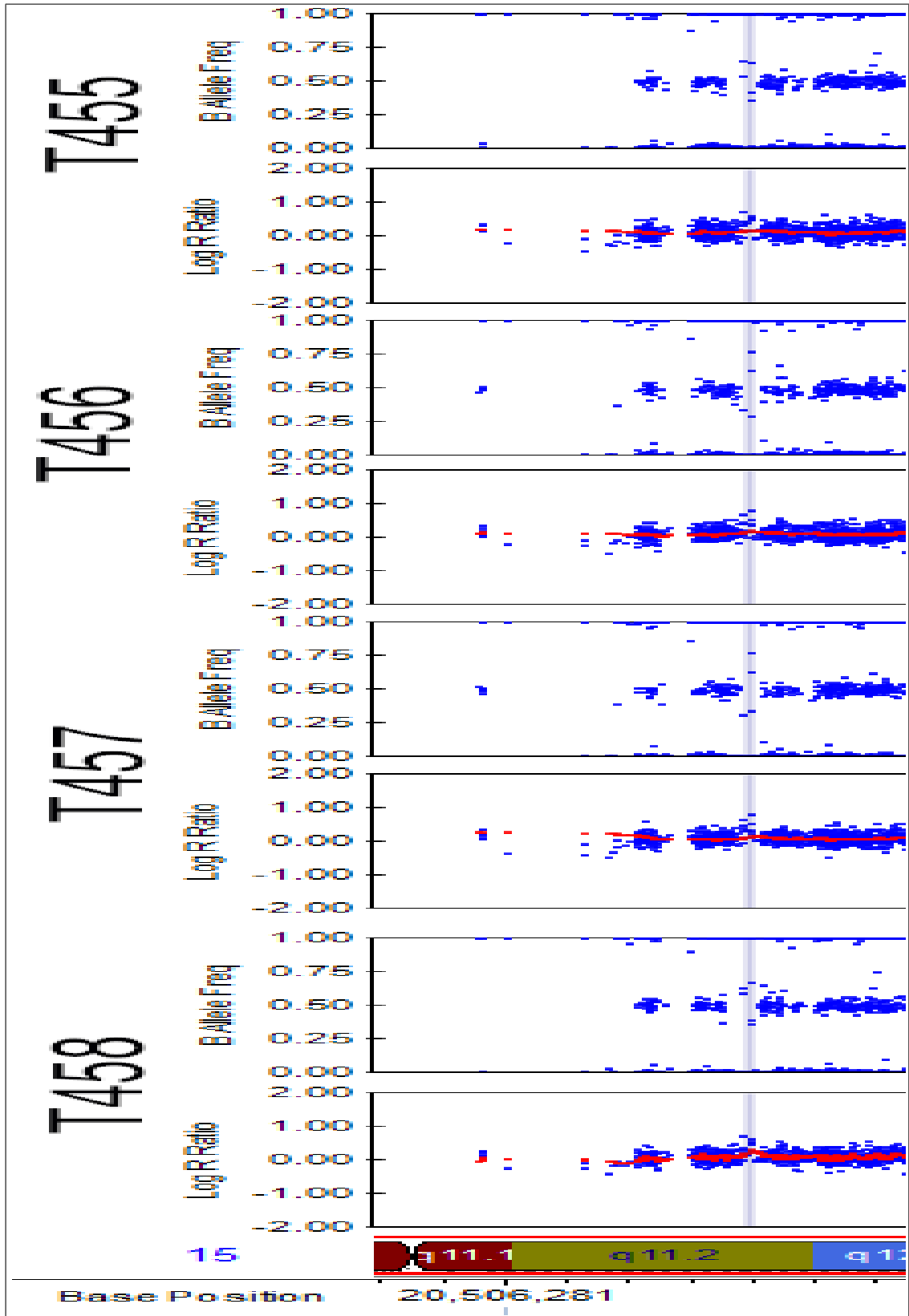
**Tablo 5-7: İkinci SNP çalışmasında MA135 ailesinde 4 etkilenmiş bireyde tespit edilen CNV.**

<b>Birey kodu</b>	<b>Kromozom</b>	<b>Başlangıç</b>	<b>Bitiş</b>	<b>CNV Değeri</b>	<b>CNV Güvenilirliği</b>
<b>T455</b>	15	24525505	24668007	3	40.38526
<b>T456</b>	15	24526247	24668007	3	96.00147
<b>T457</b>	15	24526247	24668007	3	90.26974
<b>T458</b>	15	24525505	24668007	3	52.00919

Bir sonraki SNP çalışmasında MA135 ailesinden bir etkilenmiş ve iki etkilenmemiş birey daha çalışıldı. Bu bireylerin kromozom 15 üzerinde tespit edilen CNV açısından analizleri sonucunda etkilenmiş ve etkilenmemiş her bireyde bu CNV tespit edildi (Tablo 5-8).

**Tablo 5-8: Üçüncü SNP çalışmasında MA135 ailesinde 1 etkilenmiş birey ve 2 etkilenmemiş bireylerde tespit edilen CNV.**

<b>Birey kodu</b>	<b>Kromozom</b>	<b>Başlangıç</b>	<b>Bitiş</b>	<b>CNV Değeri</b>	<b>CNV Güvenilirliği</b>
<b>A1</b>	15	24.521.050(bp)	24.668.007(bp)	3	90.0502
<b>A4</b>	15	24.521.050(bp)	24.668.007(bp)	3	61.72486
<b>A5</b>	15	24.525.505(bp)	24.668.007(bp)	3	57.57182



Şekil 5-5: MA135 ailesinde etkilenmiş 4 bireyde tespit edilen CNV'nin GenomeStudio üzerinde gösterimi. CNV bölgesi mavi ile işaretlenmiştir.

### 5.2.2. Bağlantı Analizi

MA135 ailesinde 6 bireyin genotip verisi kullanılarak bağlantı analizleri gerçekleştirildi. Mendelyen OR kalıtım modeli için Allegro yazılımı ile tam penetrans, sadece enformatif belirteçler kullanılarak ve belirteçler arası mesafe 0,1cM olacak şekilde analizler gerçekleştirildi. Mendelyen OD kalıtım modeli için yine Allegro yazılımı ile 0.9 penetrans, sadece enformatif belirteçler kullanılarak ve belirteçler arası mesafe 0.1cM olacak şekilde gerçekleştirildi. Analizler tüm otozomal ve cinsiyet kromozomları için gerçekleştirildi. Analizler sonucunda anlamlı LOD skoru olan bölge tespit edilmedi.

### 5.2.3. Ekzom Dizileme

MA135 ailesinden A3 bireyinin ekzom dizilemesinin ıslak laboratuvar kısmı için YCGA kurumundan hizmet alındı. Ham verinin analizleri yöntem kısmında anlatıldığı gibi tez kapsamında geliştirilen yazılımlar ile kendi laboratuvarımızda gerçekleştirildi. Analizler, Ensembl'in kanonik olarak kabul ettiği transkriptler üzerinden yapılmıştır. Elde edilen 72.516 varyantın %87,5'i 20X, %34,6'sı 50X okuma derinliğinin üzerindedir.

Elde edilen nitelendirilmiş varyantların analizi tündengelemlim yaklaşımı ile incelenmiş ve çeşitli filtreleme yaklaşımları kullanılarak hastalık ile ilişkili olabilecek olası aday varyantlara ulaşılmıştır. Varyantların isimlendirilmesi Ensembl VEP anotasyonundan gelen HGVS isimleri ile yapılmıştır. Filtrelemeler sonucunda elde edilen varyant sayıları Tablo 5-9'da gösterilmiştir.

**Tablo 5-9: A3 bireyinin, ED sonucu elde edilen 72.516 varyantın adım adım filtrelenmesi.**

Filtreleme adımı	Kalan varyant sayısı
Protein kodlayan transkript üzerindeki varyantlar	67.028
GnomAD üzerinde maf değeri 0.05'den küçük varyantlar	18.163
Traskript üzerinde yüksek etkili varyantlar	242
ACMG sınıflandırmasına göre benin olmayan varyantlar	191
Heterozigot varyantlar	174
Homozigot varyantlar	15



### 5.3. Aile MA136

#### 5.3.1. Genom Boyu SNP Genotipleme için SNP Dizilimi Verisinin Analizi

Ailenin iki etkilenmiş bireyleri ile Illumina platformunda, 700k SNP dizilimi ile genotipleme gerçekleştirildi. GenomeStudio yazılımı ile ham veriler işlenerek, BAF ve LRR değerleri kullanılarak tüm kromozomlar homozigotluk açısından görsel ve kantitatif olarak incelendi. GenomeStudio yazılımının, Plink Input Report eklentisi ile genotiplerin düz metin çıktısı kullanılarak plink yazılımı ile homozigot bölgeler gösterildi (Tablo 5-10).

**Tablo 5-10: MA136 ailesinde etkilenmiş iki bireyde ortak tespit edilen homozigot bölgeler.**

X3 ve X4 bireylerinde paylaşılan en büyük ortak bölgeler verilmiştir.

CHR	SNP1	SNP2	BP1	BP2	KB	NSNP
1	rs1752020	rs1040623	15.572.133	18.646.231	3.074.099	848
1	rs10493064	rs795047	35.369.014	36.637.708	1.268.695	179
2	rs28672994	rs4907254	95.395.757	97.549.668	2.153.912	228
2	rs13035933	rs4521021	203.469.960	204.816.575	1.346.616	204
3	rs6796787	rs17334102	11.200.110	64.919.103	53.718.994	13788
3	rs6796787	rs17334102	11.200.110	64.919.103	53.718.994	13788
3	rs6796787	rs17334102	11.200.110	64.919.103	53.718.994	13788
4	rs6843954	rs3903261	185.907.334	190.963.766	5.056.433	1404
5	rs1276219	rs7293455	30.930.971	46.399.093	15.468.123	3383
5	rs1276219	rs7293455	30.930.971	46.399.093	15.468.123	3383
6	rs6907076	rs2477234	20.200.420	33.894.694	13.694.275	9230
7	rs6460721	rs6951030	71.854.104	73.133.241	1.279.138	137
8	rs2410326	rs968001	15.820.384	19.047.187	3.226.804	1543
11	rs11828371	rs4881847	47.943.058	51.591.253	3.648.196	256
12	rs2194821	rs9705474	27.976.946	34.826.574	6.849.629	1764
12	rs4002730	rs584878	37.876.400	50.324.567	12.448.168	2661
12	rs7970885	rs10877760	56.031.273	62.282.739	6.251.467	1282
15	rs8033248	rs7171671	31.672.580	37.485.094	5.812.515	1711
18	rs9966660	rs4800833	18.561.020	25.494.414	6.933.395	1878
18	rs9966660	rs4800833	18.561.020	25.494.414	6.933.395	1878

Tespit edilen bu yirmi bölgeden kromozom 12 üzerinde 37.876.400 (rs4002730) – 50.324.567 (rs584878) bölgeleri arasında bulunan, daha önce literatürde birçok SHFM vakası ile ilişkilendirilmiş *WNT10B*<sup>5,12,17,40,41,90,91</sup> geninde bulunabilecek varyasyonların araştırılması için genin Sanger dizileme ile taranması uygun görüldü.

### 5.3.2. Aday Genlerin Sanger Dizileme ile Taranması

MA136 ailesinden seçilen X3 bireyinin DNA ile *WNT10B* geninin 2., 3., 4., 5. ve kodlamayan 1. ekzonları Sanger dizileme ile tarandı (Tablo 5-11). Tarama sonucunda hastalık ile ilişkili olabilecek herhangi bir varyasyon tespit edilmedi. Bunun üzerine ekzom dizileme yapılması uygun görüldü.

**Tablo 5-11: MA136 ailesinde Sanger dizileme ile *WNT10B* geninin taranması.**

Ekzon	Başlangıç-Bitiş (NM_003394.4)	Taranan Bölge Başlangıç-Bitiş
Ekzon 1 (UTR)	1-281 (hepsi 5' UTR)	-114 - +14
Ekzon 2 (UTR)	282-395 (40 5' UTR bitişi)	-203 - +63
Ekzon 3	396-658	30 - +89
Ekzon 4	659-1032	11 - +51
Ekzon 5 (UTR)	1033-2246 (1492 3' UTR başlangıcı)	-64 - 523

### 5.3.3. Ekzom Dizileme

MA136 ailesinden X3 bireyinin ekzom dizilemesinin ıslak laboratuvar kısmı için YCGA kurumundan hizmet alındı. Ham verinin analizleri yöntem kısmında anlatıldığı gibi tez kapsamında geliştirilen yazılımlar ile kendi laboratuvarımızda gerçekleştirildi. Analizler, Ensembl'in kanonik olarak kabul ettiği transkriptler üzerinden yapılmıştır. Elde edilen 51.975 varyantın %96.7'si 20X okuma derinliğinin üzerinde, %67.1'i 50X okuma derinliğinin üzerindedir.

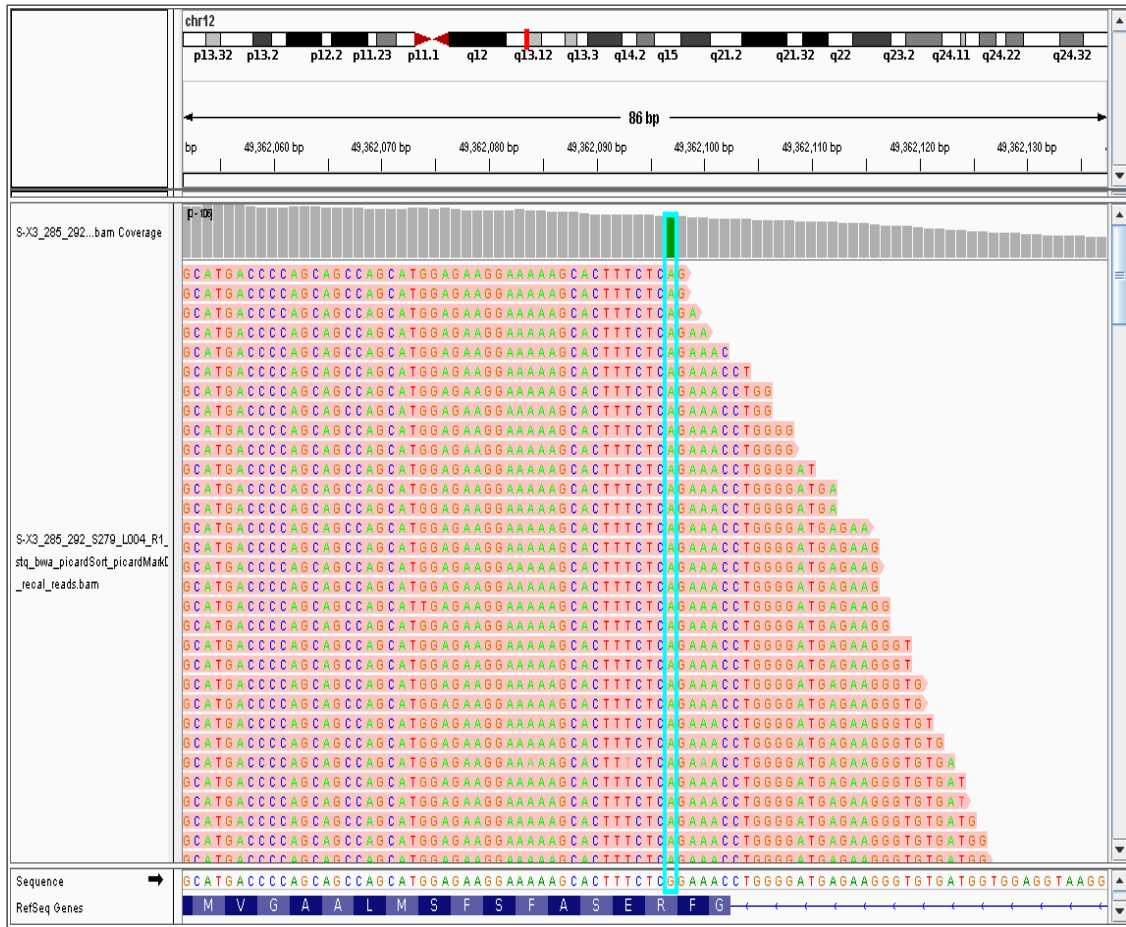
Elde edilen nitelendirilmiş varyantların analizi tümdengelim yaklaşımı ile incelenmiş ve çeşitli filtreleme yaklaşımları kullanılarak hastalık ile ilişkili olabilecek olası aday varyantlara ulaşılmıştır. Varyantların isimlendirilmesi Ensembl VEP

anotasyonundan gelen HGVS isimleri ile yapılmıştır. Filtrelemeler sonucunda elde edilen varyantlar sayıları Tablo 5-12’de gösterilmiştir.

**Tablo 5-12: X3 bireyinin, ED sonucu elde edilen 51.975 varyantın adım adım filtrelenmesi.**

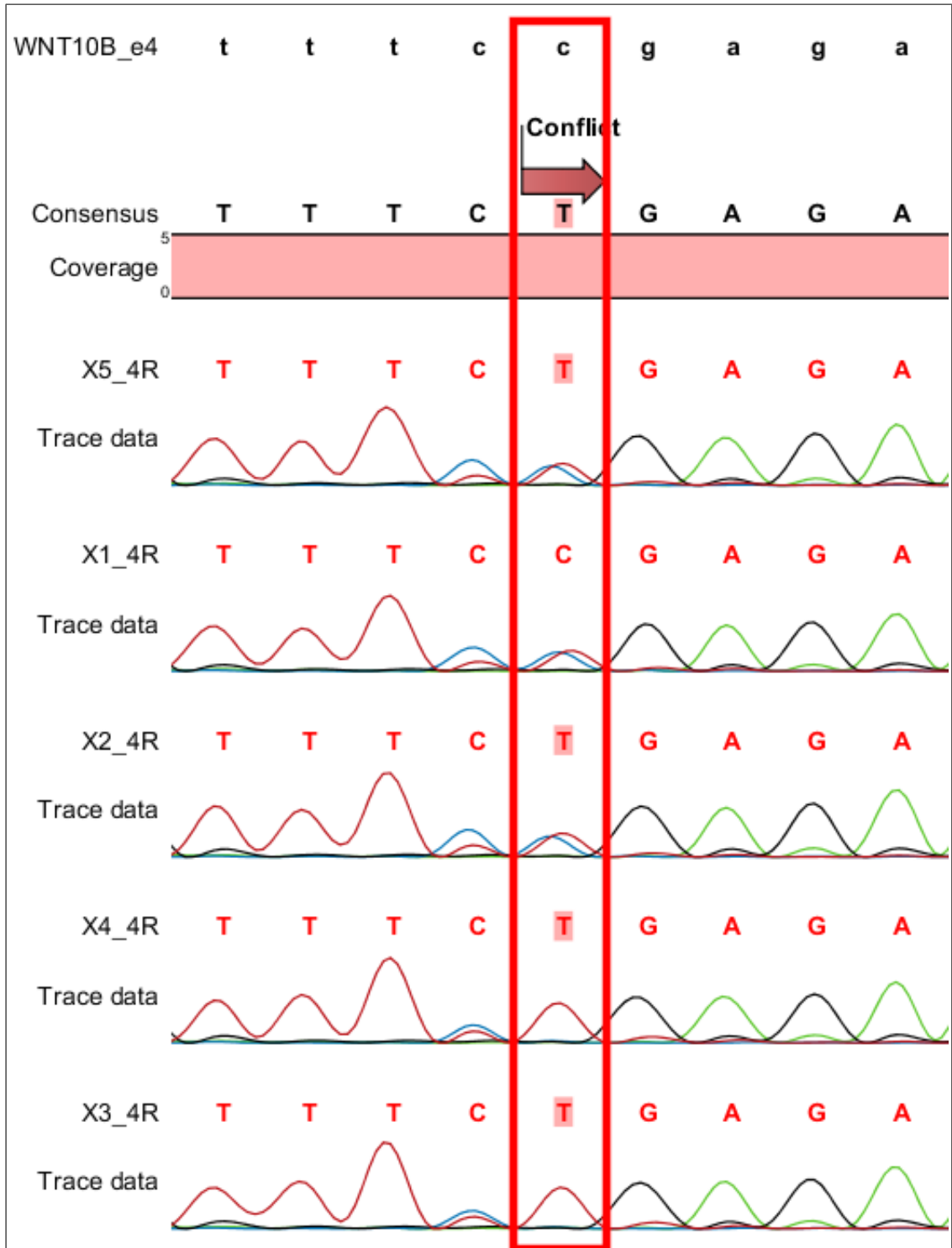
<b>Filtreleme adımı</b>	<b>Kalan varyant sayısı</b>
<b>Protein kodlayan transkript üzerindeki varyantlar</b>	47.749
<b>GnomAD üzerinde maf değeri 0.05’den küçük varyantlar</b>	15.050
<b>Traskript üzerinde yüksek etkili varyantlar</b>	234
<b>ACMG sınıflandırmasına göre benin olmayan varyantlar</b>	187
<b>Heterozigot varyantlar</b>	161
<b>Homozigot varyantlar</b>	24

Filtreleme ile beraber aynı zamanda *WNT10B* geni öncelikli olarak kontrol edilmiştir ve hem de filtrelerle de elde edilen, hastalık ile ilişkili olabilecek NM\_003394.4:c.343C>T:p.Arg115Ter değişimi tespit edilmiştir (Şekil 5-6).



**Şekil 5-6:** Ekzom dizileme sonucu *WNT10B* geninde tespit edilen NM\_003394.4:c.343C>T:p.Arg115Ter varyasyonunun IGV ile gösterimi.

Tespit edilen varyasyonun aile içindeki ayrımı Sanger dizileme ile doğrulanmıştır. X6 bireyinin DNA'sının miktarı ve saflığı uygun olmadığı için çalışılamamıştır.



Şekil 5-7: MA136 ailesinde tespit edilen NM\_003394.4:c.343C>T;p.Arg115Ter varyasyonunun Sanger dizileme ile aile içerisindeki ayrımının gösterimi.

## 5.4. Aile MA154

### 5.4.1. Genom Boyu SNP Genotipleme için SNP Dizilimi Verisinin Analizi

Ailedeki iki etkilenmiş ve bir sağlıklı kardeşler ve ebeveynler ile Illumina platformunda, 700k SNP dizilimi ile genotipleme gerçekleştirildi. GenomeStudio yazılımı ile ham veriler işlenerek, BAF ve LRR değerleri kullanılarak tüm kromozomlar homozigotluk açısından görsel ve kantitatif olarak incelendi. GenomeStudio yazılımının, Plink Input Report eklentisi ile genotiplerin düz metin çıktısı kullanılarak plink yazılımı ile homozigot bölgeler gösterildi (Tablo 5-13).

**Tablo 5-13: MA154 ailesinde etkilenmiş iki kardeşte ortak tespit edilen homozigot bölgeler.**

CON: paylaşılan en küçük ortak bölge. UNION: paylaşılan en büyük ortak bölge.

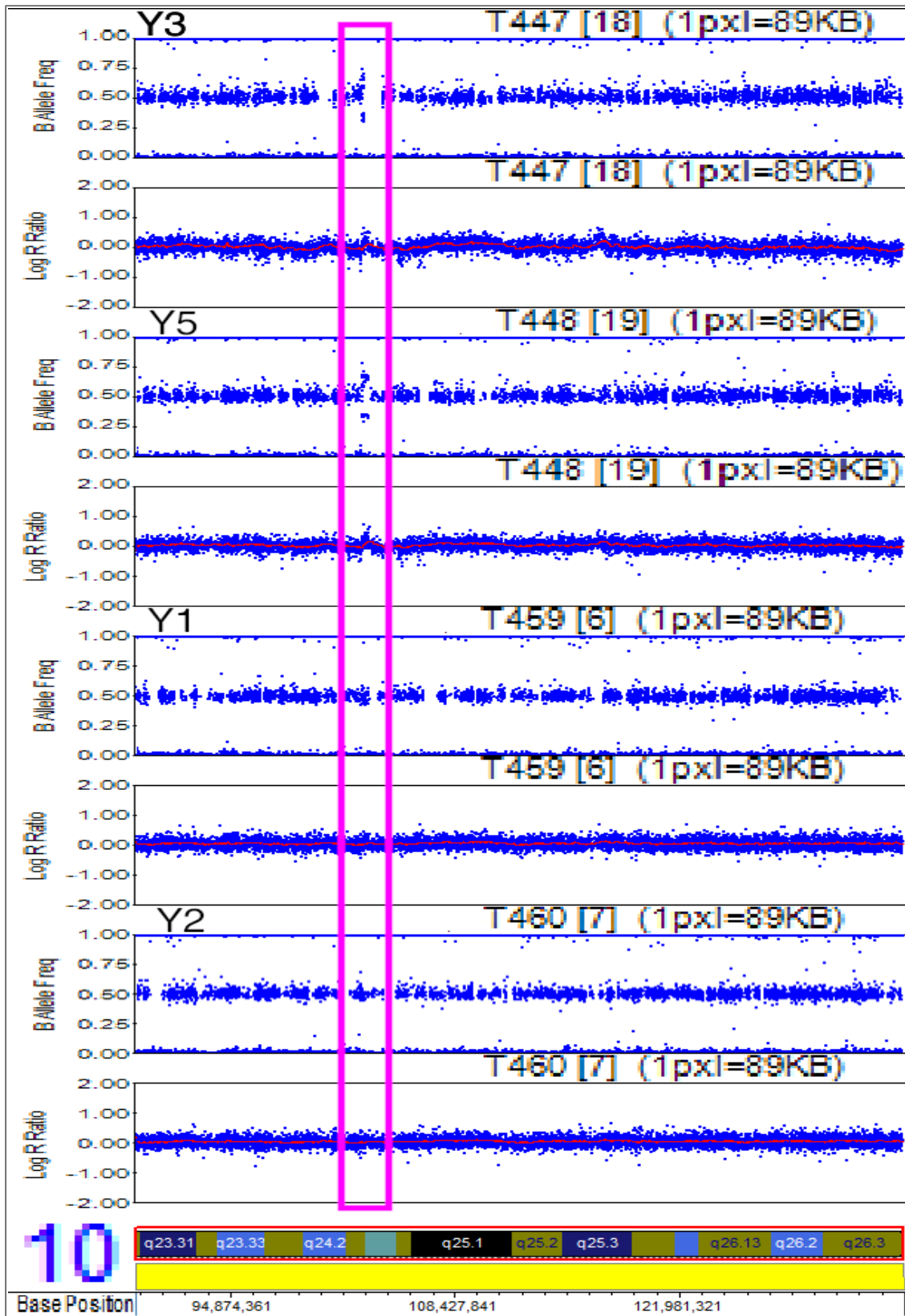
ID	PHE	CHR	SNP1	SNP2	BP1	BP2	KB	NSNP
T447	2	5	rs17713413	rs6868554	129.651.668	131.370.827	1.719.159	171
T448	2	5	rs17713413	rs6868554	129.651.668	131.370.827	1.719.159	171
CON	2:0	5	rs17713413	rs6868554	129.651.668	131.370.827	1.719.159	171
UNION	2:0	5	rs17713413	rs6868554	129.651.668	131.370.827	1.719.159	171
T447	2	5	rs165185	rs10071369	139.144.766	140.249.284	1.104.519	186
T448	2	5	rs165185	rs10071369	139.144.766	140.249.284	1.104.519	186
CON	2:0	5	rs165185	rs10071369	139.144.766	140.249.284	1.104.519	186
UNION	2:0	5	rs165185	rs10071369	139.144.766	140.249.284	1.104.519	186
T447	2	10	rs816651	rs1869210	601.114	4.808.620	4.207.507	1841
T448	2	10	rs816651	rs1869210	601.114	4.808.620	4.207.507	1841
CON	2:0	10	rs816651	rs1869210	601.114	4.808.620	4.207.507	1841
UNION	2:0	10	rs816651	rs1869210	601.114	4.808.620	4.207.507	1841
T447	2	17	rs1050223	rs2176134	14.111.915	22.242.355	8.130.441	1649
T448	2	17	rs2127994	rs2176134	14.298.594	22.242.355	7.943.762	1564
CON	2:0	17	rs2127994	rs2176134	14.298.594	22.242.355	7.943.762	1564
UNION	2:0	17	rs1050223	rs2176134	14.111.915	22.242.355	8.130.441	1649
T447	2	17	rs7211389	rs1003916	25.311.244	32.458.659	7.147.416	1510
T448	2	17	rs7211389	rs1003916	25.311.244	32.458.659	7.147.416	1510
CON	2:0	17	rs7211389	rs1003916	25.311.244	32.458.659	7.147.416	1510
UNION	2:0	17	rs7211389	rs1003916	25.311.244	32.458.659	7.147.416	1510

MA154 ailesinde hasta iki kardeşin ortak homozigot bölgeler incelendiğinde hastalık ile ilişkili olabilecek bir gen ya da bölge tespit edilmedi.

SNP dizilimi incelemeleri sürecinde CNV analizleri GenomeStudio yazılımının eklentisi olan CNVpartition ile gerçekleştirildi. İlk ve İkinci SNP dizilimi çalışmasında çalışılan MA154 ailesine etkilenmiş iki kardeşte chr10:102.924.567-103.404.601 bölgesinde artış tespit edildi (Tablo 5-14). Tespit edilen CNV aynı zamanda GenomeStudio üzerinde de gösterildi (Şekil 5-8). Daha önce SHFM3 ile de ilişkilendirilmiş olan CNV bölgesinin ikinci bir yöntem olarak gerçek zamanlı kantitatif PZR ile doğrulanması uygun görüldü.

**Tablo 5-14: Üçüncü SNP çalışmasında MA135 ailesinde 1 etkilenmiş birey ve 2 etkilenmemiş bireylerde tespit edilen CNV.**

Birey kodu	Kromozom	Başlangıç	Bitiş	CNV Değeri	CNV Güvenilirliği
Y3	10	102.924.567	103.320.642	3	222.3983
Y5	10	102.924.567	103.404.601	3	439.7837

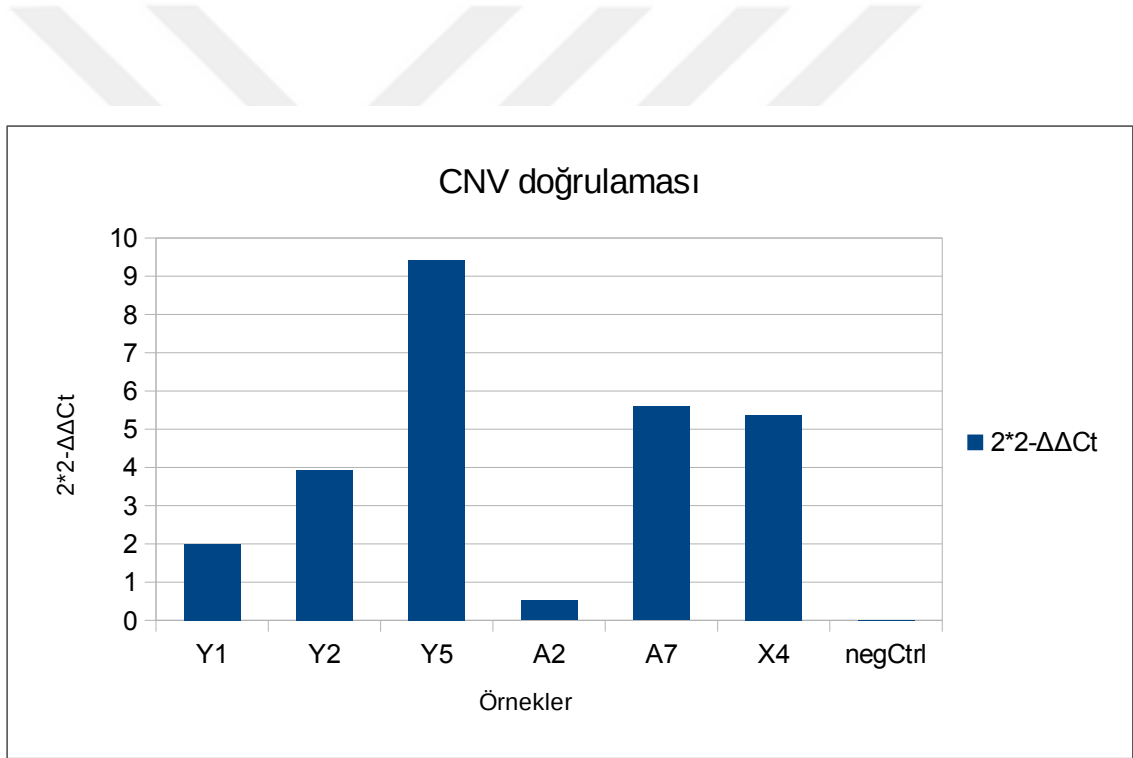


Şekil 5-8: MA154 ailesinde etkilenmiş 2 bireyde chr10:102.924.567-103.404.601 bölgesinde tespit edilen artış CNV'sinin GenomeStudio üzerinde gösterimi. CNV bölgesi eflatun ile işaretlenmiştir.



#### 5.4.2. Gerçek Zamanlı PZR ile CNV'lerin doğrulanması

Etkilenmiş kardeşlerden Y5 ve ebeveynler arasındaki CNV'lerin karşılaştırmalı delta  $C_T$  değerleri kullanılarak ölçülmesi için dizayn edilen deneyde sağlıklı kontrol olarak SNP dizilimi çalışmalarında yer alan diğer bireylerden seçilen DNA örnekleri kullanıldı. Diğer etkilenmiş kardeş Y3'ün DNA mevcut olmadığından çalışmaya katılmadı. Çalışmada referans gen olarak laboratuvarında diğer çalışmalarda da daha önceden kullanılmış olan *KIF1C* kullanıldı. Ancak çalışma sonrası anlamlı bir sonuç tespit edilmedi (Şekil 5-9).



**Şekil 5-9: MA154 ailesinde CNV artışının gerçek zamanlı PZR ile doğrulanması.**

Birçok kez tekrarlanan CNV doğrulama deneylerinde bir hasta ve sağlıklı kardeşin DNA örneklerinin olmaması ve elde olan DNA örneklerinin de kalitesinden dolayı anlamlı bir sonuca varılamamıştır.

### 5.5. Akış Hattı Deneme Çalışmaları

Akış hattı yazılımlarının denenmesi için laboratuvarımızda daha önceden çalışılmış ve yayınlanmış bulguları olan ve bu sırada analizleri farklı yazılımlar ile gerçekleştirilen iki örnek çalışma seçilmiştir. İlk denemede çalışmasında, epilepsi ile beraber intellektüel yeti kaybı fenotipi gösteren bireyin ekzom dizileme verisi çalışılmıştır<sup>85</sup>. Bu bireyde SNP dizilimi ile gerçekleştirilen bağlantı analizleri ile desteklenen *DUSIL* geninde aday bir varyanta ulaşılmıştır. İkinci örnekte ise anne, baba ve çocuk olarak üçlü ekzom dizileme olarak çalışılan ve sonuç olarak *PANK2* geninde iki yanlış anlamlı varyasyon birleşik heterozigot olarak bulunan aile çalışılmıştır<sup>86</sup>.

Birinci çalışmada, elde edilen 84.806 varyantın %86'sı 20X, %52'si 50X derinlik ile kapsanmıştır. Akraba evliliği göz önünde bulundurularak Mendeliyen OR bir kalıtımda varyantın keşfi için kullanılacak bir filtreleme yöntemi izlenmiştir (Tablo 5-15).

**Tablo 5-15: Birinci deneme çalışmasında, ED sonucu elde edilen 84.806 varyantın adım adım filtrelenmesi.**

Filtreleme adımı	Kalan varyant sayısı
Protein kodlayan transkript üzerindeki varyantlar	62.449
GnomAD üzerinde maf değeri 0.05'den küçük varyantlar	20.119
Traskript üzerinde yüksek etkili varyantlar	109
ACMG sınıflandırmasına göre benin olmayan varyantlar	101
Heterozigot varyantlar	89
Homozigot varyantlar	8

İkinci çalışmada anne ve babanın ekzom verilerinin en verimli şekilde değerlendirilebilmesi üzere filtrelemeler önce aile içi analiz yapıldı. Elde edilen 30.796 varyantın, %90'ı 20X, %57'si 50X derinlikte kapsanmıştır. Probandın varyantları birleşik homozigotluk açısından değerlendirildi ve filtrelemeler bu değerlendirme sonrasında gerçekleştirildi (Tablo 5-16).

**Tablo 5-16: İkinci çalışmada, birleşik heterozigotluk açısından incelenen 30.797 varyantın adım adım filtrelenmesi.**

<b>Filtreleme adımı</b>	<b>Kalan varyant sayısı</b>
<b>Protein kodlayan transkript üzerindeki varyantlar</b>	23.197
<b>GnomAD üzerinde maf değeri 0.05'den küçük varyantlar</b>	5.948
<b>Traskript üzerinde yüksek etkili varyantlar</b>	993
<b>ACMG sınıflandırmasına göre benin olmayan varyantlar</b>	891
<b>Heterozigot varyantlar</b>	890

## 6. TARTIŞMA

YND'nin araştırma ve klinik uygulamalarda yüksek hacme ulaşmış olması beraberinde yüksek miktarda veri getirmiş ve bu verilerin analizini kolaylaştıracak ve standartlaştıracak yöntemler gerektirmiştir. Bu tezin biyoenformatik kısmında bu analizlere yardımcı olacak iki yazılım geliştirilmiştir.

Tez çalışması kapsamında YND ile birlikte, SNP dizilimi, bağlantı analizi gibi çeşitli genomik teknolojiler ve analiz yöntemleri Sanger dizileme, gerçek zamanlı PZR gibi temel moleküler genetik yöntemlerle kullanılarak SHFM gözlenen dört ailede bu gelişimsel bozukluğun genetik etiyojisi araştırılmıştır. Hastalık, isimleri aynı olmasına rağmen, aileler arası ve hatta aile içinde bile fenotipik heterojenitesi yüksektir. Tez kapsamındaki ailelerin içerisindeki ve Pakistan genelindeki yüksek akraba evliliği oranı göze alındığında hastalığın Mendelyen OR kalıtım gösteren genetik bölgelerden köken alması beklenmektedir. Tez çalışmasında gerçekleştirilen analizler bu duruma öncelik verilerek yapılmıştır. Ancak SHFM, fenotipik heterojenitenin yanı sıra yüksek genetik heterojenite (Tablo 2-3) de göstermektedir.

Çalışma kapsamında, dört aileden etkilenmiş ve etkilenmemiş toplam 41 birey incelenmiştir. Bu dört ailelerden üçünde (MA134, MA136, MA154) fenotip ile ilişkili bir varyant bulunmuş MA135 ailesinde SHFM fenotipine neden olabilecek bir varyant bu tez çalışması kapsamında tespit edilememiştir. MA134 ve MA136 ailelerinde her ikisinde de bulgu *WNT10B* üzerinde olduğu için ikisinin beraber tartışılması uygun görülmüştür. MA135, MA154 ve yazılımlar kendi başlarına ele alınarak açıklanacaktır.

### 6.1. Aile MA134 ve MA136: *WNT10B*

Klinik tanısı Bölüm 4.1.1'de verilen aileler için gerçekleştirilen tüm incelemeler sonrasında MA134 ailesinde, NM\_003394.4::c.994C>G:p.Arg332Gly (rs121918349) varyantı, MA136 ailede ise NM\_003394.4::c.343C>T:p.Arg115Ter (rs775699954) varyantı, fenotipi açıklayan aday varyantlar olarak tespit edilmiştir. *WNT10B* geni ters

zincirden okunduğu için atasal alel genomik gösterimlerde G, transkript gösterimlerde C'dir.

MA134'te tespit edilen rs121918349 varyantı proteinin sonlarına doğru yanlış anlamlı amino asit değişimine yol açarken, MA136'da tespit edilen rs7171671 varyantı proteinin daha başlarında dur kodonuna neden olmaktadır (Şekil 6-1).



Şekil 6-1: MA134 ve MA136 ailelerinde *WNT10B* geninde tespit edilen varyasyonların protein dizisi üzerindeki etkileri.

A) MA134 NM\_003394.4::c.994C>G varyasyonu. B) MA136 NM\_003394.4::c.343C>T varyasyonu.

MA134 ailesinde homozigot olarak tespit edilen rs121918349 çok alelli bir bölgedir. Atasal alleli G iken diğer üç nükleotide (A, T, C) heterozigot olarak değişimi de farklı popülasyonlarda gözlenmiştir. Tez kapsamında tespit edilen G alelli Kuzey-Batı Avrupa uyruklu popülasyonda heterozigot olarak bir bireyde gözlenmişken (Şekil 6-2) T alleli Latin uyruklu iki bireyde heterozigot olarak A alelli ise Doğu Asya popülasyonunda tek bir kişide heterozigot olarak tespit edilmiştir. MA136'da tespit edilen rs7171671 ise iki alellidir ve yine heterozigot olarak 1 Afrika, ve 8 Kuzey-Batı Avrupa uyruklu bireylerde gözlenmiştir. Dünya genelinde heterozigot olarak rastgele gözükken bu aleller Pakistan gibi akraba evliliği oranı yüksek popülasyonlarda homozigot formda genellikle LoF mekanizmaları ile hastalıklara yol açmaktadır<sup>105</sup>. *WNT10B* geninin 2008 yılında hastalık ile ilk ilişkilendirildiğinden bu zamana kadar da genellikle akraba evliliği olan birkaç diğer ailede de farklı varyasyonlar tanımlanmıştır (Tablo 6-1).

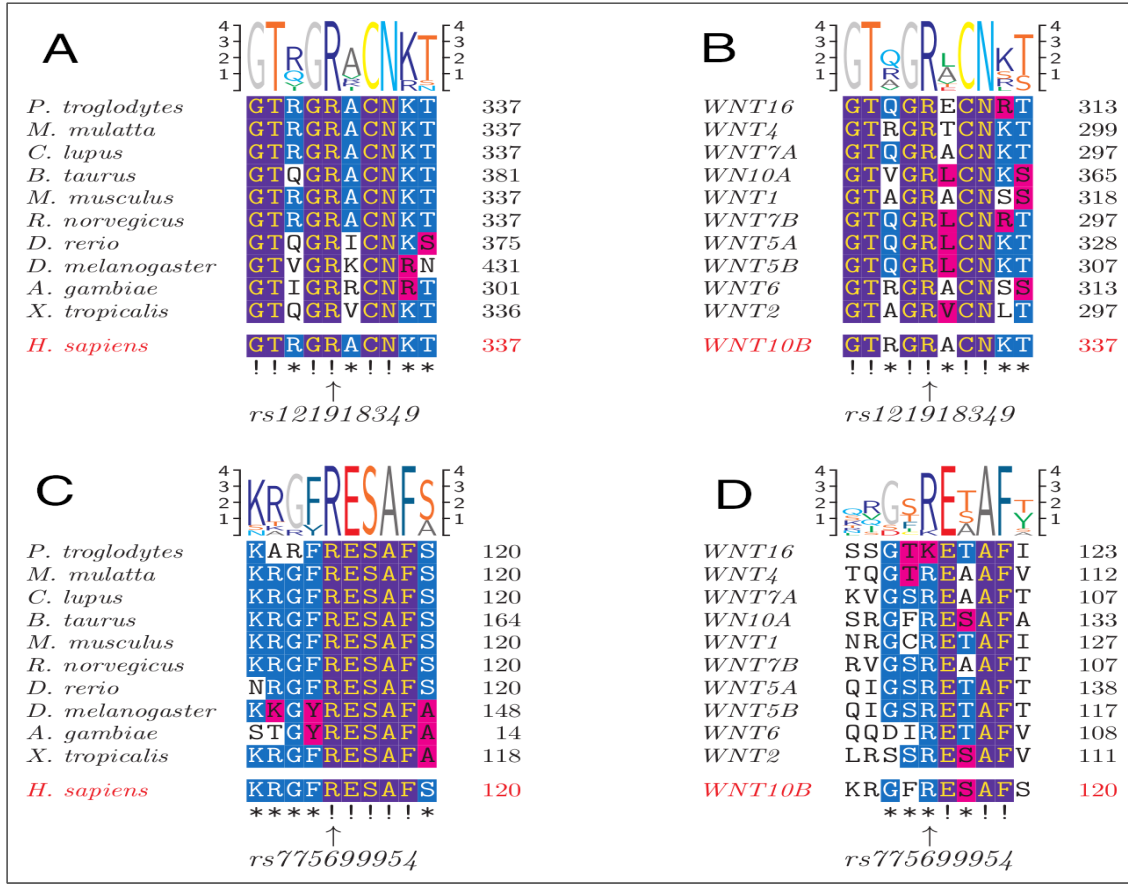
**Tablo 6-1: Literatürde bugüne kadar *WNT10B* geninde tespit edilen patojenik varyasyonlar**

Değişim (NM_003394.4)	Hastalığa yol açan genotip
c.994C>T;p.Arg332Trp <sup>103</sup>	Homozigot
c.1165_1168delAAGT;p.Lys388delin36 <sup>98</sup>	Homozigot
c.300_306dupAGGGCGG;p.Leu103Argfs*53 <sup>98</sup>	Homozigot
c.968C>G;p.Thr329Arg <sup>102</sup>	Homozigot
c.458_461dupAGCA;p.Asp155Alafs*47 <sup>99</sup>	Homozigot
c.695_697delACA;p.Asn232del <sup>101</sup>	Homozigot
c.388-1G>C <sup>106</sup>	Homozigot
c.661C>T;986C>A;p.Arg221Trp;Thr329Lys <sup>107</sup>	Birleşik Heterozigot?(Tek ebeveyn)
c.460C>G;p.Gln154* <sup>104</sup>	Homozigot
c.817delG;p.Ala273Argfs*3 <sup>100</sup>	Homozigot
c.746G>T;817dupG;p.Cys249Phe;Ala273Glyfs*14 <sup>100</sup>	Birleşik Heterozigot
c.949T>A;p.Phe317Ile <sup>100</sup>	Homozigot

Population		Allele Count	Allele Number	Number of Homozygotes	Allele Frequency
	Overall	1	113348	0	0.000008822
	North-western European	1	42078	0	0.00002377
	Bulgarian	0	2658	0	0.000
	Estonian	0	238	0	0.000
▼	European (non-Finnish)	0	30814	0	0.000
	Other non-Finnish European	0	30814	0	0.000
	Southern European	0	11486	0	0.000
	Swedish	0	26074	0	0.000
	Female	1	49866	0	0.00002005
	Male	0	63482	0	0.000
▶	African	0	16130	0	0.000
▶	Latino	0	34584	0	0.000
▶	Ashkenazi Jewish	0	10056	0	0.000
▶	East Asian	0	18378	0	0.000
▶	European (Finnish)	0	21602	0	0.000
▶	Other	0	6130	0	0.000
▶	South Asian	0	30614	0	0.000
	<b>Total</b>	<b>1</b>	<b>250842</b>	<b>0</b>	<b>0.000003987</b>

Şekil 6-2: Tez kapsamında tespit edilen NM\_003394.4:c.994C>G varyasyonunun GnomAD veri tabanında popülasyonlar arası alel sıklığının gösterimi.

Tespit edilen varyasyonların *WNT10B* proteinin korunmuşluğu üzerinde olabilecek etkilerinin görselleştirilmesi için clustalo yazılımı ile *WNT10B* geninin hem homolog hem ortolog protein dizileri hizalandı. Her iki varyasyon da hem türler arasında hem de insandaki diğer *WNTB* genleri arasında yüksek korunmuşluk göstermektedir (Şekil 6-3).



**Şekil 6-3: *WNT10B* geninde tespit edilen varyasyonlarının denk geldiği bölgelerin ortolog ve homolog proteinler üzerinde korunmuşluğunun gösterimi.**









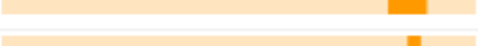

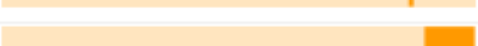
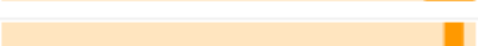

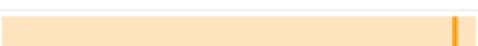

A-B) rs121918349 varyasyonunun denk geldiği bölgenin homolog ve ortolog proteinler üzerinde gösterimi. C-D) rs121918349 varyasyonunun denk geldiği bölgenin homolog ve ortolog proteinler üzerinde gösterimi.

İlk *D. melanogaster*'da keşfedilen Wnt genlerinden, insanda ve farede 19, *D. melanogaster* 7, ve *C. elegans*'da 5 tane bulunur<sup>108</sup>. Wnt ailesindeki sinyal proteinleri birçok farklı gelişimsel işlevi düzenler ve yetişkin organizmalarda homeostaziyi sağlar. Wnt proteinlerinin, hedef hücrelerini etkilemeleri, diğer sinyal yollarından daha karmaşık ve yüksek derecede geri bildirim duyarlıdır. Wnt sinyal aktivitesi yüksek derecede pleotropi göstererek mitotik uyarma, farklılaşma, polderide değişimler ve hücre bağıllığında farklılaşmayı etkiler<sup>109</sup>.



Wnt proteinlerinin sinyalleri, uzun sistin zengin bölge (CRD; cysteine-rich domain) ismi verilen uzun amino asit uzantıları olan yedi membran geçen Frizzled proteinleri tarafından alınır<sup>110,111</sup>. Wnt yolak aktivasyonunun bir diğer özelliği de sitoplazmik  $\beta$ -katenin seviyesindeki artıştır. Wnt sinyalinin yokluğunda  $\beta$ -katenin serin/teorin kinazlar, Casein Kinase<sup>112,113</sup> ve GSK-3<sup>114</sup> tarafından fosforillenerek degrade olur. Wnt sinyali  $\beta$ -katenin fosforillenmesini inhibe ederek, degradasyonun önüne geçer. Artan  $\beta$ -katenin seviyeleri, nükleusda birikmesine ve LEF/TCF transkripsiyon faktörleri ile kompleks kurmasına yol açar<sup>115</sup>.

*WNT10B*'nin amino asit dizisi sistein açısından zengindir ve beklenin<sup>116</sup> iki katı sistein içerir. İki sistein aminoasidi, disülfid bağları kurarak proteinin üç boyutlu yapısını şekillendirir. Sinyal molekülünün reseptör tarafından tanınmasında üç boyutlu yapısında ki herhangi bir değişiklik, sinyal molekülünün tanınmamasına yol açabilir. *WNT10B* amino asit dizisindeki yüksek sistein oranı ve disülfid bağ kurulması (Şekil 6-4) amino asit dizisinin yüksek derece korunmuş olmasını (Şekil 6-3) ve amino asit dizisi üzerinde etki yaratacak genetik varyasyonlarında proteinin yapısının bozulmasına ve reseptörler tarafından tanınmamasına yol açabilir. MA136 ailesindeki dur kodunu yaratan varyasyon ise çok daha kısa bir protein ürününe ya da RNA aşamasında NMD'e yol açabilir.

Feature key	Position(s)	Graphical view
Modified residue <sup>i</sup>	46	
Disulfide bond <sup>i</sup>	83 ↔ 94	
Glycosylation <sup>i</sup>	93	
Disulfide bond <sup>i</sup>	136 ↔ 144	
Disulfide bond <sup>i</sup>	146 ↔ 199	
Disulfide bond <sup>i</sup>	247 ↔ 261	
Disulfide bond <sup>i</sup>	249 ↔ 256	
Lipidation <sup>i</sup>	253	
Disulfide bond <sup>i</sup>	318 ↔ 349	
Disulfide bond <sup>i</sup>	334 ↔ 344	
Glycosylation <sup>i</sup>	335	
Disulfide bond <sup>i</sup>	348 ↔ 388	
Disulfide bond <sup>i</sup>	364 ↔ 379	
Disulfide bond <sup>i</sup>	366 ↔ 376	
Disulfide bond <sup>i</sup>	371 ↔ 372	

Şekil 6-4: *WNT10B* amino asit modifikasyonlarının Uniprot veri tabanında gösterimi.

Özellikle farklı gruplar tarafından tespit edilen rs121918349 varyasyonunun bulunduğu bölge, hastalık için hotspot mutasyon bölgesi olabilir ve *WNT10B*'nin RYK/ROR2 reseptörleri ile etkileşimini bozduğu tahmin edilmektedir<sup>117</sup>. Aynı zamanda Frizzled reseptörü bağlanma bölgesi ile aynı bölgeye düşün varyasyonlar<sup>107</sup>  $\beta$ -katenin yolağında aksaklıklara yol açabilir.

Uzuv gelişimi sırasında *Wnt10b* anlatımı ilk fare ön uzuvlarında gösterildi Summerhurst ve ark. tarafından gösterildi<sup>118</sup>. Anlatımı, proksimal uzuv ektoderminde ve en yüksek olarak AER'de gözükmektedir. AER, uzuvun proksimo-distal büyümesi ve örüntüsü yanında uzvun antero-posterior örüntüsü ile ilişkili polarizasyon aktivite alanı (ZPA, Zone of Polarizing Activity) ile de sıkı etkileşim içerisindedir. AER'de anlatımı

yapılan Fgf4, pozitif geri besleme ile ZPA içerisindeki Shh anlatımını sağlar<sup>119,120</sup>. Bu yüzden AER'in yapısında ya da korunmasındaki değişiklikler, ZPA'da Shh anlatım gradiyentini bozarak, parmak sayısının ve şeklinin belirlenmesini etkiler.

Çalışma kapsamında incelenen MA134 ve MA136 ailelerinde ortak olarak *WNT10B* geninde proteinin işlevini bozan iki farklı varyasyon tespit edilmiştir. *WNT10B* geni daha önce literatürde hastalık ile ilişkili olduğu ve her iki varyasyonun da uzuv gelişimi sırasında AER'de anlatımı olan Wnt10b sinyal proteininde yaratacağı yapısal değişiklikler parmak oluşumundaki gradiyentlerde bozuklukluğa sebep olacağı tahmin edildiğinden dolayı bu varyasyonların hastalık ile ilişkili olduğu öne sürülebilir.



## 6.2. Aile MA135

Klinik tanısı Bölüm 4.1.1’de verilen MA135 ailesi için gerçekleştirilen tüm incelemeler sonrasında, daha önceden SHFM ile ilişkilendirilmiş ya da novel, hastalığa neden olabilecek herhangi bir gen ya da bölge tespit edilmemiştir.

Aile ağacındaki kalıtım şekli dikkate alındığında, OD kalıtım bir modeli de bu aile için düşünülmüştür. Ancak bağlantı analizleri sonucunda anlamlı bir bölge elde edilemediği için ekzom dizilemede elde edilen varyantların bir yerden sonra filtrelenmesi mümkün olmamıştır. ED sonucu tespit edilen varyantlar içerisinde protein üzerinde ya da düzenleyici etkiye sahip bir varyant SHFM ile ilişkili olabilir.

SHFM’nin genetik heterojinitesi düşünüldüğünde SNP dizilime ve ED sonucu tespit edilemeyecek, tekrar artışı gibi genetik ya da epigenetik değişimlerde altta yatan sebeplerden olabilir.

Aşağıdaki Tablo 6-2 ve Tablo 6-3’ de listelenen varyantların ve listelenmeyen kodlama yapmayan varyantların SHFM oluşumundan sorumlu yolaklar üzerinde düzenleyici etki ya da bu yolaklar içerisinde yer alan genlerin anlatımını etkileyen motifler üzerinde etki yaratabilir<sup>121</sup>.

**Tablo 6-2: X3 bireyi ED verisinin filtrelenmesi sonucu elde edilen homozigot varyantlar.**

20X’den düşük okunan varyantlar çıkarılmıştır.

Gen Sembolü	Hgvsc	Hgvsp	Id
-	XM_005264681.1:c.403_404dup	XP_005264738.1:p.Arg136ProfsTer75	rs11407328
<b>ALG1L2</b>	NM_001136152.1:c.469del	NP_001129624.1:p.Leu157TrpfsTer5	-
<b>ZNF595</b>	NM_182524.2:c.627C>A	NP_872330.1:p.Tyr209Ter	-
<b>PCDHB9</b>	NM_019119.3:c.1143dup	NP_061992.2:p.Phe382IlefsTer417	-
<b>PCDHB9</b>	NM_019119.4:c.1143dup	NP_061992.3:p.Ter382Ile	-
<b>UNC93A</b>	NM_018974.3:c.625+1G>C	-	CS134281
<b>FOXD4L5</b>	NM_001126334.1:c.1215del	NP_001119806.1:p.Trp406GlyfsTer21	COSM1462695
<b>OR11G2</b>	NM_001005503.1:c.687dup	NP_001005503.1:p.Gly230ArgfsTer74	rs55781225
<b>MADCAM1</b>	NM_130760.2:c.699_700insAAC ACCACCTCCCC	NP_570116.2:p.Asp234AsnfsTer77	rs758941486
<b>MADCAM1</b>	NM_130760.2:c.700_701insGAG CCTCCCA	NP_570116.2:p.Asp234GlyfsTer64	rs768810399
-	XM_003960028.1:c.1629C>A	XP_003960077.1:p.Cys543Ter	rs782112565

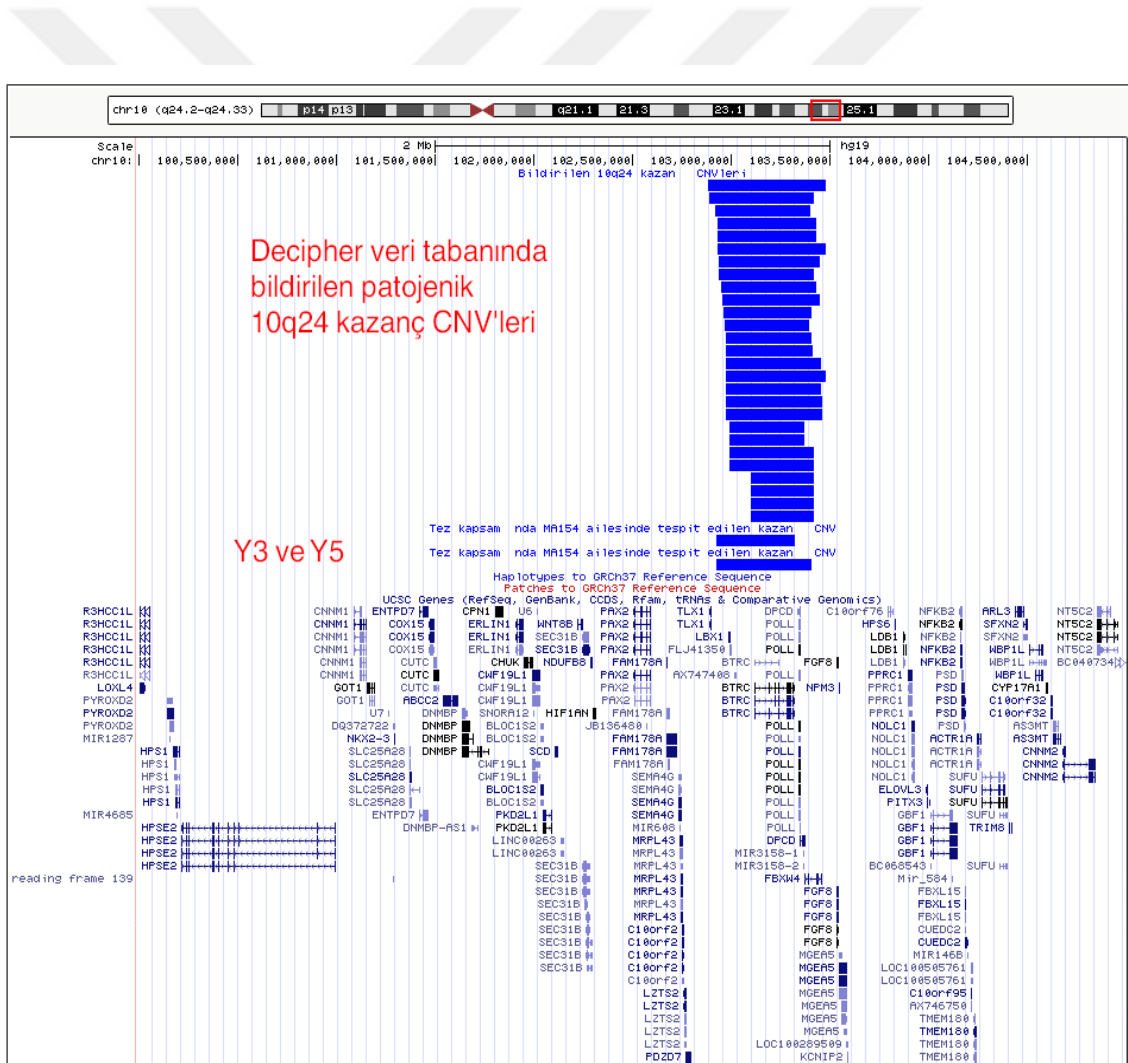
**Tablo 6-3: X3 bireyi ED verisinin filtrelenmesi sonucu elde edilen novel heterozigot varyantlar.**

20X'den düşük okunan varyantlar çıkarılmıştır.

Gen Sembolü	Hgvsc	Hgvsp
<b>PDE4DIP</b>	XM_005272981.1:c.2353C>T	XP_005273038.1:p.Arg785Ter
<b>FIGN</b>	XM_005246661.1:c.1075_1076insTTT TTTTA	XP_005246718.1:p.Tyr359PhefsTer45
<b>RBM48</b>	XM_005250636.1:c.1051C>T	XP_005250693.1:p.Arg351Ter
<b>MUC3A</b>	XM_001725354.6:c.3319C>T	XP_001725406.4:p.Gln1107Ter
<b>OR8U1</b>	NM_001005204.1:c.686_687del	NP_001005204.1:p.His229LeufsTer3
<b>BCL2L14</b>	XM_005253482.1:c.1017_1018del	XP_005253539.1:p.Phe339LeufsTer13
<b>SKA3</b>	NM_145061.5:c.*-2_*- 1insAAGAAAACT	-
<b>GRIN2A</b>	NM_001134407.2:c.3321_3322insCG GCG	NP_001127879.1:p.Thr1108ArgfsTer15
<b>LOC388282</b>	NM_001278081.1:c.140_141insAGGT GGGC	NP_001265010.1:p.Ter48_Ter50delinsGlyGlyLeu
<b>HPR</b>	XM_005255924.1:c.33_34del	XP_005255981.1:p.Tyr12HisfsTer12
<b>SREBF1</b>	NM_001005291.2:c.2304+1_2304+2in sGGGAGGGTGGGGGATTGGGGG TGGGGGGGTGGGGCTGGGGGGT GGCAGGGGTGTGGGGTGGCGGG GG	-
<b>MAP2K3</b>	XM_005256720.1:c.140+2T>A	
<b>MAP2K3</b>	XM_005256720.1:c.520C>T	XP_005256777.1:p.Gln174Ter
<b>TPTE</b>	NM_199261.2:c.1521-1G>T	
<b>ITSN1</b>	NM_003024.2:c.4180_4181insTCGCC GTATCATT	NP_003015.2:p.Glu1394ValfsTer9
<b>ITSN1</b>	NM_003024.2:c.4197_4204del	NP_003015.2:p.Asn1399LysfsTer30

### 6.3. Aile MA154: 10q24

Klinik tanısı Bölüm 4.1.1’de verilen MA154 ailesi için gerçekleştirilen tüm incelemeler sonrasında daha önceden SHFM3 ile ilişkilendirilmiş<sup>122-124</sup> 10q24(chr10:102.924.567-103.404.601) bölgesinde kazanç CNV’si tespit edilmiştir. CNV’lerin tutulduğu Decipher veri tabanında ise 10q24 bölgesinde patojenik olarak, 27 kazanç CNV’si bildirilmiştir (Şekil 6-5). CNV lokusunun küçültülmesi hedeflendiğinde, Decipher üzerinde başlangıç pozisyonu 103.098.408 olan 4 birey bulunmaktadır. Bitiş pozisyonu açısından Y3 bireyi daha önce bildirilen 103.366.970 pozisyonundan 46.328bç daha öndedir. Böylece bu çalışma sonucu hastalıkla ilişki lokus 103.098.408-103.320.642 pozisyonları arasında 222.234bç bir bölgeye indirgenir.



Şekil 6-5: Decipher veri tabanında bildirilen patojenik CNV'ler ile Y3 ve Y5 bireylerinde tespit edilen CNV'lerin UCSC Genome Browser'da gösterimi

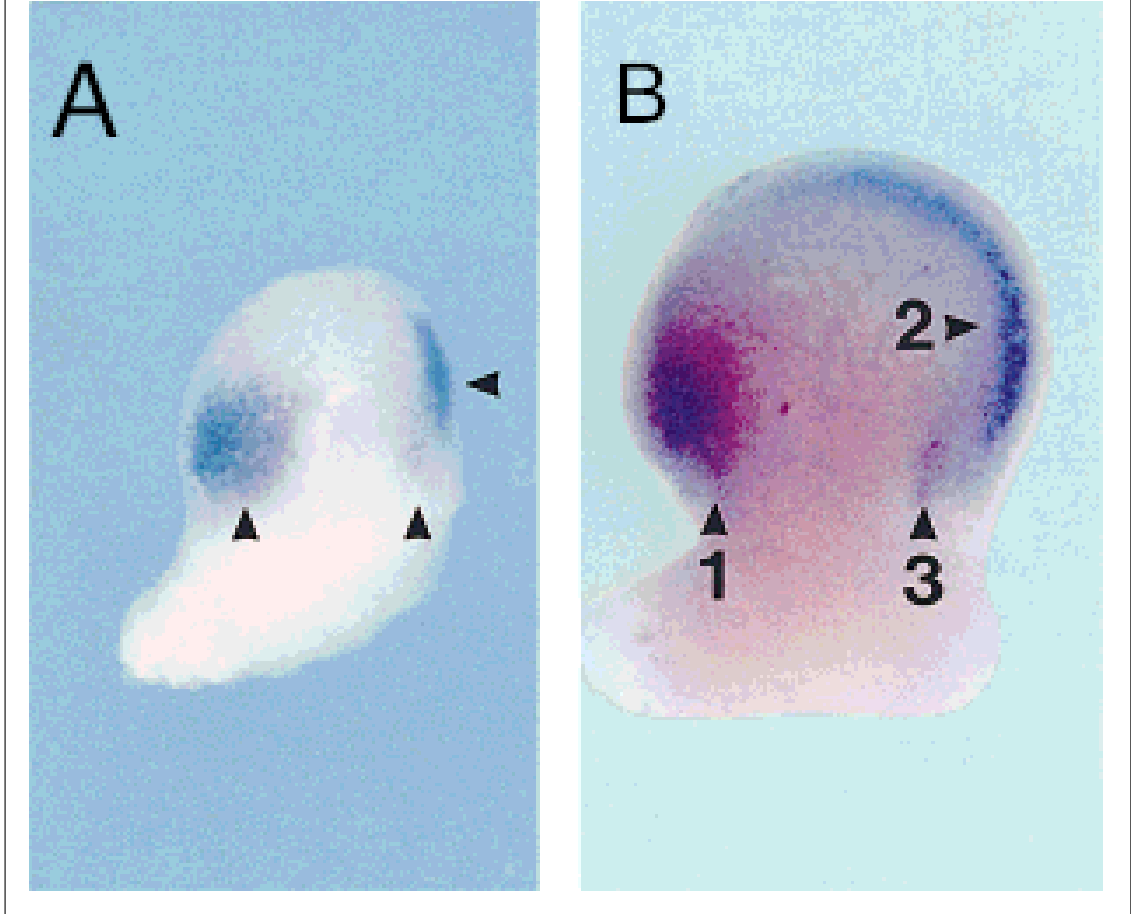
10q24 bölgesinde ncRNA, mirRNA genleride dahil olmak üzere 179-226 arası bulunmakta ve protein kodlayan ise 117 gen bilinmektedir (Tablo 6-4).

**Tablo 6-4: 10q24 bölgesinde protein kodlayan genlerin sembolleri.**

HGNC gen sembolleri					
<i>COX15</i>	<i>CCNJ</i>	<i>SH3PXD2A</i>	<i>CALHM2</i>	<i>ANKRD2</i>	<i>CALHM3</i>
<i>ABCC2</i>	<i>DNTT</i>	<i>STN1</i>	<i>ARL3</i>	<i>R3HCC1L</i>	<i>CALHMI</i>
<i>RRP12</i>	<i>HPS1</i>	<i>SLF2</i>	<i>ENTPD1</i>	<i>HIF1AN</i>	<i>PDZD7</i>
<i>MRPL43</i>	<i>DNMBP</i>	<i>ELOVL3</i>	<i>CYP17A1</i>	<i>NDUFB8</i>	<i>SLIT1</i>
<i>ALDH18A1</i>	<i>ERLIN1</i>	<i>NKX2-3</i>	<i>INA</i>	<i>BTRC</i>	<i>CC2D2B</i>
<i>PSD</i>	<i>PKD2L1</i>	<i>CUTC</i>	<i>TAF5</i>	<i>POLL</i>	<i>BLOC1S2</i>
<i>SLK</i>	<i>TLX1</i>	<i>PYROXD2</i>	<i>PPRC1</i>	<i>DPCD</i>	<i>LCOR</i>
<i>WNT8B</i>	<i>TWNK</i>	<i>CNNM1</i>	<i>CNNM2</i>	<i>HPS6</i>	<i>OPALIN</i>
<i>SEC31B</i>	<i>LZTS2</i>	<i>TCTN3</i>	<i>PDCD11</i>	<i>NOLC1</i>	<i>ENTPD7</i>
<i>PAX2</i>	<i>SFXN3</i>	<i>AVP11</i>	<i>MMS19</i>	<i>WBP1L</i>	<i>OGA</i>
<i>NT5C2</i>	<i>KAZALD1</i>	<i>ARMH3</i>	<i>PI4K2A</i>	<i>BORCS7</i>	<i>LDB1</i>
<i>TM9SF3</i>	<i>FBXW4</i>	<i>KCNIP2</i>	<i>MARVELD1</i>	<i>MORN4</i>	<i>C10orf62</i>
<i>NFKB2</i>	<i>FGF8</i>	<i>GOT1</i>	<i>ZFYVE27</i>	<i>TRIM8</i>	<i>CHUK</i>
<i>CWF19L1</i>	<i>NPM3</i>	<i>CPN1</i>	<i>GOLGA7B</i>	<i>ZDHHC16</i>	<i>ARHGAP19</i>
<i>SEMA4G</i>	<i>PITX3</i>	<i>C10orf95</i>	<i>SLC25A28</i>	<i>EXOSC1</i>	<i>AS3MT</i>
<i>BLNK</i>	<i>GBF1</i>	<i>SFRP5</i>	<i>PIK3AP1</i>	<i>PGAM1</i>	<i>RPEL1</i>
<i>TLL2</i>	<i>FBXL15</i>	<i>ACTR1A</i>	<i>PCGF6</i>	<i>HPSE2</i>	<i>HOGA1</i>
<i>SORBS1</i>	<i>CUEDC2</i>	<i>MFSD13A</i>	<i>SFXN2</i>	<i>ATP5MD</i>	
<i>CRTAC1</i>	<i>SUFU</i>	<i>LOXL4</i>	<i>FRAT1</i>	<i>ZNF518A</i>	
<i>SCD</i>	<i>NEURL1</i>	<i>LBX1</i>	<i>UBTD1</i>	<i>FRAT2</i>	

Bu genler arasından SHFM ile ilişkili olabilecek ve fenotipe sebebiyet verebilecek gen *FBXW4* (*DAC*) genidir. Bu gen, WD-40 protein-protein bağlanma domaini ile spesifik hedef proteinlerini ubiquitinasyon aracılığı ile degradasyona götüren, F-box/WD-40 gen ailesinin bir üyesidir. Farede, *FBXW4*'e yüksek benzerlik gösteren homologu *Dac*, gelişen uzvun AER'inin korunmasından rol aldığı ve fare geninde bozuklukların orta parmakların eksikliğine ve metakarpal/metakarsal kemiklerin az gelişmesine ya da eksikliğine sebep olmaktadır<sup>125</sup>.

Dac, uzuv gelişimi sırasında dinamik bir şekilde düzenlenir. İlk Dac mRNA'sı embriyonun 11. gününde gelişen uzvun üç farklı bölgesinde tespit edilebilir (Şekil 6-6)<sup>126</sup>.



**Şekil 6-6: Fare embriyosunun 11. gününde gelişen uzvunda Dac protein anlatımı.**

1)Anterior bölgeden ortaya içeriye doğru uzanan anlatımı. 2) Posterior şeritte anlatımı 3) ZPA'nın proksimalinde anlatımı.

#### 6.4. Akış Hattı ve VCF Sonrası Analizler Yazılımı

Bu iki yazılım genomik verinin işlenmesinde ve yorumlanmasında standartların takibini ve analizlerde ihtiyaç duyulan esnekliği sağlayarak veri analizi aşamalarını ulaşılabilir hale getirir. Her iki deneme çalışmasında da, kendi orijinal çalışmalarında ulaşılan sonuçlara ulaşılabilirdi. Aynı zamanda aile içi analizler ve farklı filtreleme gereksinimleri de VCF sonrası analizler yazılımında karşılandı.



Akış hattı yazılımının temel fonksiyonu, düz metin dosyası olarak verilen yapılandırmayı okuyarak çalıştırılacak komutların oluşturulmasıdır. Komut oluşturmasının yanı sıra bu komutları çağırarak analizleri gerçekleştirebilir. Herhangi başka bir fonksiyonu olmaması yazılımı kullanmayı kolaylaştırır. Değişik yazılımları karşılaştırırken tasarımları göz önüne alınmalıdır<sup>19</sup>. Tez kapsamında geliştirilen akış hattı yazılımı diğerleri ile karşılaştırıldığında, açık sentakslı, yapılandırma dosyası ile biçimlendirilen ve CLI ile etkileşilen bir yazılımdır (Tablo 6-5).

**Tablo 6-5: Komut satır arayüzlü diğer akış hattı yazılımları.**

Sentakslık	Biçim	Etkileşim	Örnek
İmalı	Düzenli	CLI	Snakeman, Nextflow
İmalı	Sınıf	CLI	Luigi
Açık	Düzenli	CLI	Ruffus, bpipe
Açık	Yapılandırma	CLI	Pegasus

VCF sonrası analizler yazılımının, VCF anotasyonu, filtrelenmesi ve aile içi analizler olmak üzere üç fonksiyonu bulunur. Bu fonksiyonlar kendi içlerinde, farklı analiz ihtiyaçlarını karşılamak üzere kullanıcıya esneklik tanıyacak farklı analizler gerçekleştirebilir.

## 6.5. Sonuç ve Öneriler

Bu çalışmanın kapsamında incelenen dört SHFM ailesinin üçünde altta yatan genetik varyasyonlar belirlendi. Üç ailenin ikisinde *WNT10B* geninde patojenik varyasyonlar , birinde ise 10q24 bölgesinde artış CNV'si tespit edildi. MA134 ailesinde NM\_003394.4:c.994C>G:p.Arg332Gly, MA136 ailesinde NM\_003394.4:c.343C>T:p.Arg115Ter varyasyonu tespit edildi. MA154 ailesinde 10. kromozomda 102.924.567–103.404.601 bölgeleri arasında 480.034 bç'lik kazanç CNV'si tespit edilmiştir.

Hastalık ile ilişkili herhangi bir gen ya da lokus tespit edilmeyen MA135 ailesi için ise genomdaki düzenleyici bölgelerin, tekrar artışlarının ya da transposon hareketlerinin tespiti için genom dizileme bir sonraki adım için uygun olabilir.

Ek olarak NGS ve diğer veri analizlerinde uygulaması olan iki yazılım geliştirildi. Bu yazılımlardan biri biyolojik verinin analizinde kullanılan akışhatlarının yapılmasını sağlamaktadır. Diğeri ise NGS verisinden elde edilen varyant bilgisinin nitelendirilmesini, filtrelenmesini ve aile içi analizlerinin gerçekleştirilmesini sağlamaktadır.

**KAYNAKLAR**

1. Behjati S, Tarpey PS. What is next generation sequencing? *Arch Dis Child Educ Pract Ed.* 2013;98(6):236-238. doi:10.1136/archdischild-2013-304340
2. Morgan TH. SEX LIMITED INHERITANCE IN DROSOPHILA. *Science.* 1910;32(812):120-122. doi:10.1126/science.32.812.120
3. Watson JD, Crick FH. Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. *Nature.* 1953;171(4356):737-738.
4. Sanger F, Nicklen S, Coulson AR. DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci U S A.* 1977;74(12):5463-5467.
5. Barnes WM. DNA sequencing by partial ribosubstitution. *J Mol Biol.* 1978;119(1):83-99.
6. Sanger F, Coulson AR. A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase. *J Mol Biol.* 1975;94(3):441-448.
7. Maxam AM, Gilbert W. A new method for sequencing DNA. *Proc Natl Acad Sci USA.* 1977;74(2):560-564. doi:10.1073/pnas.74.2.560
8. Margulies M, Egholm M, Altman WE, et al. Genome sequencing in microfabricated high-density picolitre reactors. *Nature.* 2005;437(7057):376-380. doi:10.1038/nature03959
9. Tucker T, Marra M, Friedman JM. Massively Parallel Sequencing: The Next Big Thing in Genetic Medicine. *The American Journal of Human Genetics.* 2009;85(2):142-154. doi:10.1016/j.ajhg.2009.06.022
10. Quail MA, Smith M, Coupland P, et al. A tale of three next generation sequencing platforms: Comparison of Ion Torrent, Pacific Biosciences and Illumina MiSeq sequencers. *BMC Genomics.* 2012;13:341. doi:10.1186/1471-2164-13-341
11. Mardis ER. The impact of next-generation sequencing technology on genetics. *Trends Genet.* 2008;24(3):133-141. doi:10.1016/j.tig.2007.12.007
12. Liu L, Li Y, Li S, et al. Comparison of Next-Generation Sequencing Systems. *J Biomed Biotechnol.* 2012;2012. doi:10.1155/2012/251364
13. Muzzey D, Evans EA, Lieber C. Understanding the Basics of NGS: From Mechanism to Variant Calling. *Curr Genet Med Rep.* 2015;3(4):158-165. doi:10.1007/s40142-015-0076-8
14. Oliver GR, Hart SN, Klee EW. Bioinformatics for clinical next generation sequencing. *Clin Chem.* 2015;61(1):124-135. doi:10.1373/clinchem.2014.224360

15. FASTQ Format. <http://maq.sourceforge.net/fastq.shtml>. Accessed May 5, 2019.
16. SAMv1.pdf. <http://samtools.github.io/hts-specs/SAMv1.pdf>. Accessed May 5, 2019.
17. Genome Browser FAQ. <https://genome.ucsc.edu/FAQ/FAQformat.html#format1>. Accessed May 5, 2019.
18. VCFv4.2.pdf. <https://samtools.github.io/hts-specs/VCFv4.2.pdf>. Accessed May 5, 2019.
19. Leipzig J. A review of bioinformatic pipeline frameworks. *Brief Bioinform.* 2017;18(3):530-536. doi:10.1093/bib/bbw020
20. Yang Y. Wnts and wing: Wnt signaling in vertebrate limb development and musculoskeletal morphogenesis. *Birth Defects Res C Embryo Today.* 2003;69(4):305-317. doi:10.1002/bdrc.10026
21. Niswander L. Pattern formation: Old models out on a limb. *Nat Rev Genet.* 2003;4(2):133-143. doi:10.1038/nrg1001
22. Manouvrier-Hanu S, Holder-Espinasse M, Lyonnet S. Genetics of limb anomalies in humans. *Trends Genet.* 1999;15(10):409-417.
23. Tickle C. Making digit patterns in the vertebrate limb. *Nature Reviews Molecular Cell Biology.* 2006;7(1):45. doi:10.1038/nrm1830
24. Ingham PW, Placzek M. Orchestrating ontogenesis: Variations on a theme by sonic hedgehog. *Nat Rev Genet.* 2006;7(11):841-850. doi:10.1038/nrg1969
25. Casanova JC, Sanz-Ezquerro JJ. Digit morphogenesis: Is the tip different? *Development, Growth & Differentiation.* 2007;49(6):479-491. doi:10.1111/j.1440-169X.2007.00951.x
26. Biesecker LG, Aase JM, Clericuzio C, Gurrieri F, Temple IK, Toriello H. Defining Morphology: Hands and Feet. *Am J Med Genet A.* 2009;149A(1):93-127. doi:10.1002/ajmg.a.32596
27. Bedard T, Lowry RB, Sibbald B, Kiefer GN, Metcalfe A. Congenital limb deficiencies in Alberta—a review of 33 years (1980-2012) from the Alberta Congenital Anomalies Surveillance System (ACASS). *Am J Med Genet A.* 2015;167A(11):2599-2609. doi:10.1002/ajmg.a.37240
28. Froster-Iskenius UG, Baird PA. Limb reduction defects in over one million consecutive livebirths. *Teratology.* 1989;39(2):127-135. doi:10.1002/tera.1420390205
29. Stoll C, Alembik Y, Dott B, Roth M-P. Associated malformations in patients with limb reduction deficiencies. *Eur J Med Genet.* 2010;53(5):286-290. doi:10.1016/j.ejmg.2010.07.012

30. Elliott AM, Evans JA, Chudley AE. Split hand foot malformation (SHFM). *Clin Genet.* 2005;68(6):501-505. doi:10.1111/j.1399-0004.2005.00530.x
31. McCarroll SA, Kuruvilla FG, Korn JM, et al. Integrated detection and population-genetic analysis of SNPs and copy number variation. *Nature Genetics.* 2008;40(10):1166-1174. doi:10.1038/ng.238
32. Steemers FJ, Chang W, Lee G, Barker DL, Shen R, Gunderson KL. Whole-genome genotyping with the single-base extension assay. *Nat Methods.* 2006;3(1):31-33. doi:10.1038/nmeth842
33. Peiffer DA, Le JM, Steemers FJ, et al. High-resolution genomic profiling of chromosomal aberrations using Infinium whole-genome genotyping. *Genome Res.* 2006;16(9):1136-1148. doi:10.1101/gr.5402306
34. Staaf J, Vallon-Christersson J, Lindgren D, et al. Normalization of Illumina Infinium whole-genome SNP data improves copy number estimates and allelic intensity ratios. *BMC Bioinformatics.* 2008;9:409. doi:10.1186/1471-2105-9-409
35. Marenne G, Rodríguez-Santiago B, Closas MG, et al. Assessment of Copy Number Variation Using the Illumina Infinium 1M SNP-Array: A Comparison of Methodological Approaches in the Spanish Bladder Cancer/EPICURO Study. *Hum Mutat.* 2011;32(2):240-248. doi:10.1002/humu.21398
36. Broman KW, Weber JL. Long homozygous chromosomal segments in reference families from the centre d'Etude du polymorphisme humain. *Am J Hum Genet.* 1999;65(6):1493-1500. doi:10.1086/302661
37. Clark AG. Inference of haplotypes from PCR-amplified samples of diploid populations. *Mol Biol Evol.* 1990;7(2):111-122. doi:10.1093/oxfordjournals.molbev.a040591
38. Lander ES, Botstein D. Homozygosity mapping: A way to map human recessive traits with the DNA of inbred children. *Science.* 1987;236(4808):1567-1570.
39. Blixt S. Why didn't Gregor Mendel find linkage? *Nature.* 1975;256(5514):206. doi:10.1038/256206a0
40. Morgan TH. RANDOM SEGREGATION VERSUS COUPLING IN MENDELIAN INHERITANCE. *Science.* 1911;34(873):384. doi:10.1126/science.34.873.384
41. Sturtevant AH. The linear arrangement of six sex-linked factors in *Drosophila*, as shown by their mode of association. *Journal of Experimental Zoology.* 1913;14(1):43-59. doi:10.1002/jez.1400140104
42. Ott J. Estimation of the recombination fraction in human pedigrees: Efficient computation of the likelihood for human linkage studies. *Am J Hum Genet.* 1974;26(5):588-597.

43. Elston RC, Stewart J. A general model for the genetic analysis of pedigree data. *Hum Hered.* 1971;21(6):523-542. doi:10.1159/000152448
44. Risch N. Genetic linkage: Interpreting lod scores. *Science.* 1992;255(5046):803-804.
45. Morton NE. Sequential tests for the detection of linkage. *Am J Hum Genet.* 1955;7(3):277-318.
46. Lander ES, Green P. Construction of multilocus genetic linkage maps in humans. *Proc Natl Acad Sci USA.* 1987;84(8):2363-2367. doi:10.1073/pnas.84.8.2363
47. Gudbjartsson DF, Jonasson K, Frigge ML, Kong A. Allegro, a new computer program for multipoint linkage analysis. *Nat Genet.* 2000;25(1):12-13. doi:10.1038/75514
48. Kruglyak L, Daly MJ, Reeve-Daly MP, Lander ES. Parametric and nonparametric linkage analysis: A unified multipoint approach. *Am J Hum Genet.* 1996;58(6):1347-1363.
49. Hoffmann K, Lindner TH. easyLINKAGE-Plus—Automated linkage analyses using large-scale SNP data. *Bioinformatics.* 2005;21(17):3565-3567. doi:10.1093/bioinformatics/bti571
50. Harrow J, Nagy A, Reymond A, et al. Identifying protein-coding genes in genomic sequences. *Genome Biol.* 2009;10(1):201. doi:10.1186/gb-2009-10-1-201
51. Ng SB, Turner EH, Robertson PD, et al. Targeted capture and massively parallel sequencing of 12 human exomes. *Nature.* 2009;461(7261):272-276. doi:10.1038/nature08250
52. Herman DS, Hovingh GK, Iartchouk O, et al. Filter-based hybridization capture of subgenomes enables resequencing and copy-number detection. *Nat Methods.* 2009;6(7):507-510. doi:10.1038/nmeth.1343
53. Gnirke A, Melnikov A, Maguire J, et al. Solution hybrid selection with ultra-long oligonucleotides for massively parallel targeted sequencing. *Nat Biotechnol.* 2009;27(2):182-189. doi:10.1038/nbt.1523
54. Porreca GJ, Zhang K, Li JB, et al. Multiplex amplification of large sets of human exons. *Nat Methods.* 2007;4(11):931-936. doi:10.1038/nmeth1110
55. Briggs AW, Good JM, Green RE, et al. Targeted retrieval and analysis of five Neandertal mtDNA genomes. *Science.* 2009;325(5938):318-321. doi:10.1126/science.1174462
56. Choi M, Scholl UI, Ji W, et al. Genetic diagnosis by whole exome capture and massively parallel DNA sequencing. *Proc Natl Acad Sci USA.* 2009;106(45):19096-19101. doi:10.1073/pnas.0910672106

57. RESERVED IU--AR. Orphanet. <http://www.orpha.net>. Accessed May 11, 2019.
58. HGMD®. <http://www.hgmd.cf.ac.uk/ac/stats.php>. Accessed May 11, 2019.
59. Chen JM, Férec C, Cooper DN. Revealing the human mutome. *Clin Genet*. 2010;78(4):310-320. doi:10.1111/j.1399-0004.2010.01474.x
60. Kuhlenbäumer G, Hullmann J, Appenzeller S. Novel genomic techniques open new avenues in the analysis of monogenic disorders. *Hum Mutat*. 2011;32(2):144-151. doi:10.1002/humu.21400
61. Patel K, Larson C, Hargreaves M, et al. Community screening outcomes for diabetes, hypertension, and cholesterol: Nashville REACH 2010 project. *J Ambul Care Manage*. 2010;33(2):155-162. doi:10.1097/JAC.0b013e3181dd4619
62. Gilissen C, Arts HH, Hoischen A, et al. Exome sequencing identifies WDR35 variants involved in Sensenbrenner syndrome. *Am J Hum Genet*. 2010;87(3):418-423. doi:10.1016/j.ajhg.2010.08.004
63. Hoischen A, van Bon BWM, Gilissen C, et al. De novo mutations of SETBP1 cause Schinzel-Giedion syndrome. *Nat Genet*. 2010;42(6):483-485. doi:10.1038/ng.581
64. Ng SB, Bigam AW, Buckingham KJ, et al. Exome sequencing identifies MLL2 mutations as a cause of Kabuki syndrome. *Nat Genet*. 2010;42(9):790-793. doi:10.1038/ng.646
65. Anastasio N, Ben-Omran T, Teebi A, et al. Mutations in SCARF2 are responsible for Van Den Ende-Gupta syndrome. *Am J Hum Genet*. 2010;87(4):553-559. doi:10.1016/j.ajhg.2010.09.005
66. Glazov EA, Zankl A, Donskoi M, et al. Whole-exome re-sequencing in a family quartet identifies POP1 mutations as the cause of a novel skeletal dysplasia. *PLoS Genet*. 2011;7(3):e1002027. doi:10.1371/journal.pgen.1002027
67. Ng SB, Buckingham KJ, Lee C, et al. Exome sequencing identifies the cause of a mendelian disorder. *Nat Genet*. 2010;42(1):30-35. doi:10.1038/ng.499
68. Bolze A, Byun M, McDonald D, et al. Whole-exome-sequencing-based discovery of human FADD deficiency. *Am J Hum Genet*. 2010;87(6):873-881. doi:10.1016/j.ajhg.2010.10.028
69. Çalışkan M, Chong JX, Uricchio L, et al. Exome sequencing reveals a novel mutation for autosomal recessive non-syndromic mental retardation in the TECR gene on chromosome 19p13. *Hum Mol Genet*. 2011;20(7):1285-1289. doi:10.1093/hmg/ddq569
70. Walsh T, Shahin H, Elkan-Miller T, et al. Whole exome sequencing and homozygosity mapping identify mutation in the cell polarity protein GPSM2 as the cause of nonsyndromic hearing loss DFNB82. *Am J Hum Genet*. 2010;87(1):90-94. doi:10.1016/j.ajhg.2010.05.010

71. Beck TF, Mullikin JC, Biesecker LG. Systematic Evaluation of Sanger Validation of NextGen Sequencing Variants. *Clin Chem*. 2016;62(4):647-654. doi:10.1373/clinchem.2015.249623
72. Gundry CN, Vandersteen JG, Reed GH, Pryor RJ, Chen J, Wittwer CT. Amplicon melting analysis with labeled primers: A closed-tube method for differentiating homozygotes and heterozygotes. *Clin Chem*. 2003;49(3):396-406.
73. Montgomery J, Wittwer CT, Palais R, Zhou L. Simultaneous mutation scanning and genotyping by high-resolution DNA melting analysis. *Nat Protoc*. 2007;2(1):59-66. doi:10.1038/nprot.2007.10
74. Aarskog NK, Vedeler CA. Real-time quantitative polymerase chain reaction. A new method that detects both the peripheral myelin protein 22 duplication in Charcot-Marie-Tooth type 1A disease and the peripheral myelin protein 22 deletion in hereditary neuropathy with liability to pressure palsies. *Hum Genet*. 2000;107(5):494-498.
75. Anhuf D, Eggermann T, Rudnik-Schöneborn S, Zerres K. Determination of SMN1 and SMN2 copy number using TaqMan technology. *Hum Mutat*. 2003;22(1):74-78. doi:10.1002/humu.10221
76. Thiel CT, Kraus C, Rauch A, Ekici AB, Rautenstrauss B, Reis A. A new quantitative PCR multiplex assay for rapid analysis of chromosome 17p11.2-12 duplications and deletions leading to HMSN/HNPP. *Eur J Hum Genet*. 2003;11(2):170-178. doi:10.1038/sj.ejhg.5200920
77. Kulka J, Tôkés A-M, Kaposi-Novák P, Udvarhelyi N, Keller A, Schaff Z. Detection of HER-2/neu gene amplification in breast carcinomas using quantitative real-time PCR - a comparison with immunohistochemical and FISH results. *Pathol Oncol Res*. 2006;12(4):197-204. doi:PAOR.2006.12.4.0197
78. Laccone F, Jünemann I, Whatley S, et al. Large deletions of the MECP2 gene detected by gene dosage analysis in patients with Rett syndrome. *Hum Mutat*. 2004;23(3):234-244. doi:10.1002/humu.20004
79. Terribas E, Garcia-Linares C, Lázaro C, Serra E. Probe-based quantitative PCR assay for detecting constitutional and somatic deletions in the NF1 gene: Application to genetic testing and tumor analysis. *Clin Chem*. 2013;59(6):928-937. doi:10.1373/clinchem.2012.194217
80. Heid CA, Stevens J, Livak KJ, Williams PM. Real time quantitative PCR. *Genome Res*. 1996;6(10):986-994.
81. Higuchi R, Fockler C, Dollinger G, Watson R. Kinetic PCR analysis: Real-time monitoring of DNA amplification reactions. *Biotechnology (NY)*. 1993;11(9):1026-1030.
82. Schmittgen TD, Livak KJ. Analyzing real-time PCR data by the comparative C(T) method. *Nat Protoc*. 2008;3(6):1101-1108.





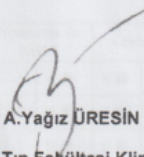
83. Griffiths AJ, Gelbart WM, Miller JH, Lewontin RC. Human Pedigree Analysis. *Modern Genetic Analysis*. 1999. <https://www.ncbi.nlm.nih.gov/books/NBK21257/>. Accessed May 5, 2019.
84. pandas: A Foundational Python Library for Data Analysis and Statistics | R (Programming Language) (47K views). Scribd. <https://www.scribd.com/document/71048089/pandas-a-Foundational-Python-Library-for-Data-Analysis-and-Statistics>. Accessed May 30, 2019.
85. Emrah Yücesan. Ailevi epilepsilerde tüm genom analizleri. 2016.
86. Yapici Z, Akcakaya NH, Tekturk P, Iseri SAU, Ozbek U. A novel gene mutation in PANK2 in a patient with severe jaw-opening dystonia. *Brain Dev*. 2016;38(8):755-758. doi:10.1016/j.braindev.2016.02.010
87. Purcell S, Neale B, Todd-Brown K, et al. PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. *Am J Hum Genet*. 2007;81(3):559-575.
88. O'Connell JR, Weeks DE. PedCheck: A program for identification of genotype incompatibilities in linkage analysis. *Am J Hum Genet*. 1998;63(1):259-266. doi:10.1086/301904
89. Babraham Bioinformatics's FastQC A Quality Control tool for High Throughput Sequence Data. <https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>. Accessed May 7, 2019.
90. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*. 2009;25(14):1754-1760. doi:10.1093/bioinformatics/btp324
91. *A Set of Command Line Tools (in Java) for Manipulating High-Throughput Sequencing (HTS) Data and Formats Such as SAM/BAM/CRAM and VCF.*: Broad Institute/Picard. Broad Institute; 2019. <https://github.com/broadinstitute/picard>. Accessed May 7, 2019.
92. McKenna A, Hanna M, Banks E, et al. The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res*. 2010;20(9):1297-1303. doi:10.1101/gr.107524.110
93. barslmm. *Pointed Interpretation of Clinical Variant Significance: Barslmm/Picus.*; 2019. <https://github.com/barslmm/picus>. Accessed June 13, 2019.
94. Robinson JT, Thorvaldsdóttir H, Winckler W, et al. Integrative genomics viewer. *Nat Biotechnol*. 2011;29(1):24-26. doi:10.1038/nbt.1754
95. Ye J, Coulouris G, Zaretskaya I, Cutcutache I, Rozen S, Madden TL. Primer-BLAST: A tool to design target-specific primers for polymerase chain reaction. *BMC Bioinformatics*. 2012;13:134. doi:10.1186/1471-2105-13-134

96. Kuhn RM, Haussler D, Kent WJ. The UCSC genome browser and associated tools. *Brief Bioinform.* 2013;14(2):144-161. doi:10.1093/bib/bbs038
97. Wildeman M, van Ophuizen E, den Dunnen JT, Taschner PEM. Improving sequence variant descriptions in mutation databases and literature using the Mutalyzer sequence variation nomenclature checker. *Hum Mutat.* 2008;29(1):6-13. doi:10.1002/humu.20654
98. Aziz A, Irfanullah null, Khan S, et al. Novel homozygous mutations in the WNT10B gene underlying autosomal recessive split hand/foot malformation in three consanguineous families. *Gene.* 2014;534(2):265-271. doi:10.1016/j.gene.2013.10.047
99. Blattner A, Huber AR, Röthlisberger B. Homozygous nonsense mutation in WNT10B and sporadic split-hand/foot malformation (SHFM) with autosomal recessive inheritance. *Am J Med Genet A.* 2010;152A(8):2053-2056. doi:10.1002/ajmg.a.33504
100. Brunelle P, Jourdain A-S, Escande F, et al. WNT10B variants in split hand/foot malformation: Report of three novel families and review of the literature. *Am J Med Genet A.* May 2019. doi:10.1002/ajmg.a.61177
101. Kantaputra PN, Kapoor S, Verma P, Intachai W, Ketudat Cairns JR. Split hand-foot malformation and a novel WNT10B mutation. *Eur J Med Genet.* 2018;61(7):372-375. doi:10.1016/j.ejmg.2018.02.001
102. Khan S, Basit S, Zimri FK, et al. A novel homozygous missense mutation in WNT10B in familial split-hand/foot malformation. *Clin Genet.* 2012;82(1):48-55. doi:10.1111/j.1399-0004.2011.01698.x
103. Ugur SA, Tolun A. Homozygous WNT10b mutation and complex inheritance in Split-Hand/Foot Malformation. *Hum Mol Genet.* 2008;17(17):2644-2653. doi:10.1093/hmg/ddn164
104. Ullah A, Gul A, Umair M, et al. Homozygous sequence variants in the WNT10B gene underlie split hand/foot malformation. *Genet Mol Biol.* 2018;41(1):1-8. doi:10.1590/1678-4685-GMB-2016-0162
105. Saleheen D, Natarajan P, Armean IM, et al. Human knockouts and phenotypic analysis in a cohort with a high rate of consanguinity. *Nature.* 2017;544(7649):235-239. doi:10.1038/nature22034
106. Monies D, Abouelhoda M, AlSayed M, et al. The landscape of genetic diseases in Saudi Arabia based on the first 1000 diagnostic panels and exomes. *Hum Genet.* 2017;136(8):921-939. doi:10.1007/s00439-017-1821-8
107. Meng L, Pammi M, Saronwala A, et al. Use of Exome Sequencing for Infants in Intensive Care Units: Ascertainment of Severe Single-Gene Disorders and Effect on Medical Management. *JAMA Pediatr.* 2017;171(12):e173438. doi:10.1001/jamapediatrics.2017.3438

108. Nusse R. Wnt signaling in disease and in development. *Cell Research*. 2005;15(1):28. doi:10.1038/sj.cr.7290260
109. Goessling W, North TE, Loewer S, et al. Genetic interaction of PGE2 and Wnt signaling regulates developmental specification of stem cells and regeneration. *Cell*. 2009;136(6):1136-1147. doi:10.1016/j.cell.2009.01.015
110. Dann CE, Hsieh JC, Rattner A, Sharma D, Nathans J, Leahy DJ. Insights into Wnt binding and signalling from the structures of two Frizzled cysteine-rich domains. *Nature*. 2001;412(6842):86-90. doi:10.1038/35083601
111. Hsieh J-C, Rattner A, Smallwood PM, Nathans J. Biochemical characterization of Wnt-Frizzled interactions using a soluble, biologically active vertebrate Wnt protein. *PNAS*. 1999;96(7):3546-3551. doi:10.1073/pnas.96.7.3546
112. Amit S, Hatzubai A, Birman Y, et al. Axin-mediated CKI phosphorylation of  $\beta$ -catenin at Ser 45: A molecular switch for the Wnt pathway. *Genes Dev*. 2002;16(9):1066-1076. doi:10.1101/gad.230302
113. Liu C, Li Y, Semenov M, et al. Control of  $\beta$ -Catenin Phosphorylation/Degradation by a Dual-Kinase Mechanism. *Cell*. 2002;108(6):837-847. doi:10.1016/S0092-8674(02)00685-2
114. Yost C, Torres M, Miller JR, Huang E, Kimelman D, Moon RT. The axis-inducing activity, stability, and subcellular distribution of beta-catenin is regulated in *Xenopus* embryos by glycogen synthase kinase 3. *Genes Dev*. 1996;10(12):1443-1454. doi:10.1101/gad.10.12.1443
115. Munemitsu S, Albert I, Rubinfeld B, Polakis P. Deletion of an amino-terminal sequence beta-catenin in vivo and promotes hyperphosphorylation of the adenomatous polyposis coli tumor suppressor protein. *Molecular and Cellular Biology*. 1996;16(8):4088-4094. doi:10.1128/MCB.16.8.4088
116. AMINO ACID FREQUENCY.  
<http://www.tiem.utk.edu/~gross/bioed/webmodules/aminoacid.htm>. Accessed June 15, 2019.
117. Katoh M, Katoh M. Comparative genomics on Wnt8a and Wnt8b genes. *Int J Oncol*. 2005;26(4):1129-1133.
118. Summerhurst K, Stark M, Sharpe J, Davidson D, Murphy P. 3D representation of Wnt and Frizzled gene expression patterns in the mouse embryo at embryonic day 11.5 (Ts19). *Gene Expr Patterns*. 2008;8(5):331-348. doi:10.1016/j.gep.2008.01.007
119. Laufer E, Nelson CE, Johnson RL, Morgan BA, Tabin C. Sonic hedgehog and Fgf-4 act through a signaling cascade and feedback loop to integrate growth and patterning of the developing limb bud. *Cell*. 1994;79(6):993-1003.

120. Niswander L, Jeffrey S, Martin GR, Tickle C. A positive feedback loop coordinates growth and patterning in the vertebrate limb. *Nature*. 1994;371(6498):609-612. doi:10.1038/371609a0
121. Kantaputra PN, Carlson BM. Genetic regulatory pathways of split-hand/foot malformation. *Clin Genet*. 2019;95(1):132-139. doi:10.1111/cge.13434
122. Carter TC, Sicko RJ, Kay DM, et al. Copy-number variants and candidate gene mutations in isolated split hand/foot malformation. *J Hum Genet*. 2017;62(10):877-884. doi:10.1038/jhg.2017.56
123. Holder-Espinasse M, Jamsheer A, Escande F, et al. Duplication of 10q24 locus: Broadening the clinical and radiological spectrum. *Eur J Hum Genet*. 2019;27(4):525-534. doi:10.1038/s41431-018-0326-9
124. Roscioli T, Taylor PJ, Bohlken A, et al. The 10q24-linked split hand/split foot syndrome (SHFM3): Narrowing of the critical region and confirmation of the clinical phenotype. *Am J Med Genet A*. 2004;124A(2):136-141. doi:10.1002/ajmg.a.20348
125. Friedli M, Nikolaev S, Lyle R, et al. Characterization of mouse Dactylaplasia mutations: A model for human ectrodactyly SHFM3. *Mamm Genome*. 2008;19(4):272-278. doi:10.1007/s00335-008-9106-0
126. Caubit X, Thangarajah R, Theil T, et al. Mouse Dac, a novel nuclear factor with homology to Drosophila dachshund shows a dynamic expression in the neural crest, the eye, the neocortex, and the limb bud. *Dev Dyn*. 1999;214(1):66-80. doi:10.1002/(SICI)1097-0177(199901)214:1<66::AID-DVDY7>3.0.CO;2-7
127. Gilbert SF, Barresi MJF. *Developmental Biology*. 11 edition. Sinauer Associates is an imprint of Oxford University Press; 2016.
128. Lobo I, Shaw K. Thomas Hunt Morgan, Genetic Recombination, and Gene Mapping. *Nature Education*. 2008;1(1):205.
129. Teer JK, Mullikin JC. Exome sequencing: The sweet spot before whole genomes. *Hum Mol Genet*. 2010;19(R2):R145-151. doi:10.1093/hmg/ddq333

## ETİK KURUL KARARI

	<p>T.C.  <b>İSTANBUL ÜNİVERSİTESİ</b>  <b>İSTANBUL TIP FAKÜLTESİ</b>  <b>KLİNİK ARAŞTIRMALAR ETİK KURULU</b></p>	
Sayı : 492		Tarih : 28.03.2018
Konu: Doç. Dr. Sibel Uğur İŞERİ hk.		
<p>Sayın Doç. Dr. Sibel Uğur İŞERİ          Genetik</p>		
<p><b>İlgi : Genetik Anabilim Dalının 16/03/2018 gün ve 101525 sayılı yazısı</b></p>		
<p>Sorumlu araştırmacılığını üstlendiğiniz ve Sadık Barış SALMAN' ın yürüteceği 2018/372 dosya numaralı "Yarıklı El/Ayak Yapısal Bozukluğunun Genomik Teknikler ve Biyoenformatik Yaklaşımlarla İncelenmesi" başlıklı çalışma kurulumuzun 23/03/2018 tarih ve 06 sayılı toplantısında görüşülerek etik yönden uygun bulunmuş olup, tutanaklar ekte sunulmuştur.</p>		
<p>Bilgilerinizi rica ederim.</p>		
<p>  <b>Prof. Dr. A.Yağız ÜRESİN</b>          İstanbul Tıp Fakültesi Klinik Araştırmalar          Etik Kurul Başkanı</p>		
<p>Eki: İstanbul Tıp Fakültesi Klinik Araştırmaları Etik Kurulu Karar Formu</p>		

## İSTANBUL TIP FAKÜLTESİ KLİNİK ARAŞTIRMALARI ETİK KURULU KARAR FORMU

ARAŞTIRMANIN AÇIK ADI "Yarıklı El/Ayak Yapısal Bozukluğunun Genomik Teknikler ve Biyoenformatik Yaklaşımlarla İncelenmesi"

DEĞERLENDİRİLEN BELGELER	Belge Adı	Tarihi	Versiyon Numarası	Dili		
	ARAŞTIRMA PROTOKOLÜ	19.03.2018		Türkçe <input checked="" type="checkbox"/>	İngilizce <input type="checkbox"/>	Diğer <input type="checkbox"/>
	BİLGİLENDİRİLMİŞ GÖNÜLLÜ OLUR FORMU	<input checked="" type="checkbox"/>		Türkçe <input checked="" type="checkbox"/>	İngilizce <input type="checkbox"/>	Diğer <input type="checkbox"/>
	OLGU RAPOR FORMU	<input type="checkbox"/>		Türkçe <input type="checkbox"/>	İngilizce <input type="checkbox"/>	Diğer <input type="checkbox"/>
	ARAŞTIRMA BROŞÜRÜ	<input type="checkbox"/>		Türkçe <input type="checkbox"/>	İngilizce <input type="checkbox"/>	Diğer <input type="checkbox"/>
DEĞERLENDİRİLEN DİĞER BELGELER	Belge Adı	<input type="checkbox"/>	Açıklama			
	TÜRKÇE ETİKET ÖRNEĞİ	<input type="checkbox"/>				
	SİGORTA	<input type="checkbox"/>				
	ARAŞTIRMA BÜTÇESİ	<input checked="" type="checkbox"/>				
	BIYOLOJİK MATERYEL TRANSFER FORMU	<input type="checkbox"/>				
	HASTA KARTI/GÜNLÜKLERİ	<input type="checkbox"/>				
	İLAN	<input type="checkbox"/>				
	YILLIK BİLDİRİM	<input type="checkbox"/>				
	SONUÇ RAPORU	<input type="checkbox"/>				
	GÜVENLİLİK BİLDİRİMLERİ	<input type="checkbox"/>				
KARAR BELGELERİ	Karar No: 06	Tarih: 23/03/2018				
	İstanbul Üniversitesi Aziz Sancar Deneysel Tıp Araştırma Enstitüsü Genetik Anabilim Dalında görevli Doç. Dr. Sibel Uğur İŞERİ' nin sorumluluğunda ve Sadık Barış SALMAN' ın yürüteceği yukarıda bilgileri verilen araştırma başvuru dosyası ile ilgili belgeler araştırmanın gereke, amaç, yaklaşım ve yöntemleri dikkate alınarak incelenmiş, gerçekleştirilmesinde etik ve bilimsel sakınca bulunmadığına toplantıya katılan Etik Kurul üye tam sayısının salt çoğunluğu ile karar verilmiştir.					

## İSTANBUL TIP FAKÜLTESİ KLİNİK ARAŞTIRMALARI ETİK KURULU

ÇALIŞMA ESASI	19.08.2011 tarihli, 28030 sayılı Resmî Gazetede yayınlanan Klinik Araştırmalar Hakkındaki Yönetmelik						
BAŞKANIN UNVANI / ADI / SOYADI:	Prof. Dr. A. Yağız ÜRESİN						
Unvanı/Adı/Soyadı	Uzmanlık Alanı	Kurumu	Cinsiyet	Araştırma ile ilişki *		Katılım **	İmza
Prof. Dr. A. Yağız ÜRESİN	Farmakoloji ve Klinik Farmakoloji	İstanbul Tıp Fakültesi (Etik Kurul Başkanı)	E <input checked="" type="checkbox"/> K <input type="checkbox"/>	E <input type="checkbox"/> H <input checked="" type="checkbox"/>	E <input checked="" type="checkbox"/> H <input type="checkbox"/>		
Prof. Dr. Berrin UMMAN	Kardiyoloji	İstanbul Tıp Fakültesi (Etik Kurul Başkan Yardımcısı)	E <input type="checkbox"/> K <input checked="" type="checkbox"/>	E <input type="checkbox"/> H <input checked="" type="checkbox"/>	E <input checked="" type="checkbox"/> H <input type="checkbox"/>		
Prof. Dr. Ahmet GÜL	Romatoloji	İstanbul Tıp Fakültesi	E <input checked="" type="checkbox"/> K <input type="checkbox"/>	E <input type="checkbox"/> H <input checked="" type="checkbox"/>	E <input checked="" type="checkbox"/> H <input type="checkbox"/>		
Prof. Dr. Oğuzhan ÇOBAN	Nöroloji	İstanbul Tıp Fakültesi	E <input checked="" type="checkbox"/> K <input type="checkbox"/>	E <input type="checkbox"/> H <input checked="" type="checkbox"/>	E <input checked="" type="checkbox"/> H <input type="checkbox"/>		
Dr. Sevdâ ÖZEL YILDIZ	Biyostatistik	İ.Ü. İstanbul Tıp Fakültesi	E <input type="checkbox"/> K <input checked="" type="checkbox"/>	E <input type="checkbox"/> H <input checked="" type="checkbox"/>	E <input checked="" type="checkbox"/> H <input type="checkbox"/>		

\* -Araştırma ile ilişki  
\*\* -Toplantıda Bulunma

İ.Ü. İstanbul Tıp Fakültesi Klinik araştırmalar Etik kurulu 13.04.2013 tarih, 28617 sayılı Resmî Gazetede yayınlanan Klinik Araştırmalar Hakkında Yönetmelik çerçevesinde kurulmuş ve T.C. Sağlık Bakanlığı Türkiye İlaç ve Tıbbi Cihaz Kurumu tarafından onaylanmıştır. İlgili yönetmelik kapsamında kalan araştırmalar Sağlık Bakanlığından izin almak zorundadır. Yönetmelik kapsamı dışında kalan araştırmalar ise Etik Kurul bünyesinde oluşturulmuş 5 kişilik alt komisyon tarafından değerlendirilmekte olup Sağlık Bakanlığı iznine tabi değildir.

## İSTANBUL TIP FAKÜLTESİ KLİNİK ARAŞTIRMALARI ETİK KURULU KARAR FORMU

ETİK KURULU BİLGİLERİ	ETİK KURULUN ADI	İSTANBUL TIP FAKÜLTESİ KLİNİK ARAŞTIRMALARI ETİK KURULU
	AÇIK ADRESİ:	İ.Ü.İSTANBUL TIP FAKÜLTESİ HULUSİ BEHÇET KÜTÜPHANESİ KAT:3 FATİH/İSTANBUL
	TELEFON	0 (212) 414 21 53
	FAKS	0 (212) 414 21 53
	E-POSTA	itfetikkurul@istanbul.edu.tr.

BAŞVURU BİLGİLERİ	ARAŞTIRMANIN AÇIK ADI	"Yarıklı El/Ayak Yapısal Bozukluğunun Genomik Teknikler ve Biyoenformatik Yaklaşımlarla İncelenmesi"			
	ARAŞTIRMA PROTOKOL KODU	---			
	KOORDİNATÖR/SORUMLU ARAŞTIRMACI UNVANI/ADI/SOYADI	Doç. Dr. Sibel Uğur İŞERİ			
	KOORDİNATÖR/SORUMLU ARAŞTIRMACININ UZMANLIK ALANI	Genetik			
	KOORDİNATÖR/SORUMLU ARAŞTIRMACININ BULUNDUĞU MERKEZ	İstanbul Üniversitesi Aziz Sançar Deneysel Tıp Araştırma Enstitüsü Genetik Anabilim Dalı			
	DESTEKLEYİCİ	İstanbul Üniversitesi Bilimsel Araştırma Projeleri Birimi			
	DESTEKLEYİCİNİN YASAL TEMSİLCİSİ	---			
	ARAŞTIRMANIN FAZI	FAZ 1	<input type="checkbox"/>		
		FAZ 2	<input type="checkbox"/>		
		FAZ 3	<input type="checkbox"/>		
FAZ 4		<input type="checkbox"/>			
ARAŞTIRMANIN TÜRÜ	Yeni Bir Endikasyon	<input type="checkbox"/>			
	Yüksek Doz Araştırması	<input type="checkbox"/>			
	Diğer ise belirtiniz :				
ARAŞTIRMAYA KATILAN MERKEZLER	TEK MERKEZ <input checked="" type="checkbox"/>	ÇOK MERKEZLİ <input type="checkbox"/>	ULUSAL <input checked="" type="checkbox"/>	ULUSLAR ARASI <input type="checkbox"/>	

## İNTİHAL RAPORU İLK SAYFASI

### YARIK ELAYAK YAPISAL BOZUKLUĞUNUN GENOMİK TEKNİKLER VE BİYOENFORMATİK YAKLAŞIMLARLA İNCELENMESİ

#### ORIJINALLIK RAPORU

<b>%5</b>	<b>%3</b>	<b>%0</b>	<b>%5</b>
BENZERLİK ENDEKSİ	İNTERNET KAYNAKLARI	YAYINLAR	ÖĞRENCİ ÖDEVLERİ

#### BİRİNCİL KAYNAKLAR

<b>1</b>	<b>Submitted to Istanbul University</b> Öğrenci Ödevi	<b>%1</b>
<b>2</b>	<b>acikerisim.istanbul.edu.tr</b> İnternet Kaynağı	<b>%1</b>
<b>3</b>	<b>www.docstoc.com</b> İnternet Kaynağı	<b>%1</b>
<b>4</b>	<b>Submitted to Marmara University</b> Öğrenci Ödevi	<b>&lt;%1</b>
<b>5</b>	<b>Submitted to TechKnowledge Turkey</b> Öğrenci Ödevi	<b>&lt;%1</b>
<b>6</b>	<b>Submitted to The Scientific &amp; Technological Research Council of Turkey (TUBİTAK)</b> Öğrenci Ödevi	<b>&lt;%1</b>
<b>7</b>	<b>Submitted to Bursa Teknik Üniversitesi</b> Öğrenci Ödevi	<b>&lt;%1</b>
<b>8</b>	<b>angora.baskent.edu.tr</b>	



## ÖZGEÇMİŞ

### Kişisel Bilgiler

<b>Adı</b>	<b>Sadık Barış</b>	<b>Soyadı</b>	<b>Salman</b>
<b>Doğ.Yeri</b>	<b>Yozgat</b>	<b>Doğ.Tar.</b>	<b>01/12/1993</b>
<b>Email</b>	<b>barslmn@gmail.com</b>	<b>Uyruğu</b>	<b>TC</b>

### Eğitim Düzeyi

	Mezun Olduğu Kurumun Adı	Mez. Yılı
<b>Doktora</b>		
<b>Yük.Lis.</b>		
<b>Lisans</b>	<b>İÜ Fen Fakültesi MBG</b>	<b>2016</b>
<b>Lise</b>	<b>Artvin Anadolu Öğretmen Lisesi</b>	<b>2011</b>

### İş Deneyimi (Sondan geçmişe doğru sıralayın)

	Görevi	Kurum	Süre (Yıl - Yıl)
1.			-
2.			-
3.			-

Yabancı Dilleri	Okuduğunu Anlama*	Konuşma*	Yazma*	KPDS/ÜDS Puanı	(Diğer) Puanı
<b>İngilizce</b>	<b>Çok iyi</b>	<b>Çok iyi</b>	<b>Çok iyi</b>	<b>85</b>	

\*Çok iyi, iyi, orta, zayıf olarak değerlendirin

	Sayısal	Eşit Ağırlık	Sözel
<b>LES Puanı</b>	<b>77</b>	<b>75</b>	<b>70</b>
<b>(Diğer) Puanı</b>			

### Bilgisayar Bilgisi

Program	Kullanma becerisi
<b>Powerpoint</b>	<b>İyi</b>


**Yayımları/Tebliğleri Sertifikaları/Ödülleri**

**Özel İlgi Alanları (Hobileri): Yürüyüş, tempolu yürüyüş, koşu**

