İSTANBUL-2023

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

(DOKTORA TEZİ)

TÜM GENOM DİZİLEME YÖNTEMİNİN GENETİK HASTALIKLARIN TANISINA KATKISI

İSMAİL MUSAB GÜLAÇAR

DANIŞMAN
DOÇ. DR. GÖZDE YEŞİL SAYIN
İKİNCİ DANIŞMAN
PROF. DR. SEHER BAŞARAN

GENETİK ANABİLİM DALI
GENETİK PROGRAMI

İSTANBUL-2023

İTHAF

Evlatlarının amansız hastalıklarına çare bulmak için her fedakarlığı yapan anne ve babalara, bu dünyaya nadir hastalıklarla göz açan ve yüzleri çok az gülen çocuklarımıza ithaf ediyorum.

TEŞEKKÜR

Doktora eğitimimin ilk gününden son gününe kadar beni destekleyen hem derslerinde hem laboratuvarda hem de muhabbetlerimizde kendisinden çok şey öğrendiğim kıymetli hocam Prof. Dr. Seher Başaran'a,

Tez çalışmasına başladığımız günden tezin son günlerine kadar desteğini ve ilgisini esirgemeyen tez danışmanım değerli hocam Doç. Dr. Gözde Yeşil Sayın'a

Doktora ve laboratuvar eğitimimde emeği olan, ne zaman ihtiyaç duysam koşulsuz desteğini sunan değerli hocam Prof. Dr. Birsen Karaman'a,

Engin bilgisi, nezaketi ve bilgi paylaşımında cömertliği ile tanıdığım değerli hocam Prof. Dr. Zehra Oya Uyguner'e,

Tez sürecimde dönem dönem yaşadığım motivasyon kaybımı tekrar kazanmama vesile olan, tez yazım sürecinde düzletmelerle ayrıca katkı sağlayan kıymetli hocam Doç. Dr. Gökhan Görgişen'e,

Protein modelleme ve simülasyon çalışmalarımda yol gösterici olan, değerli bir bilim insanı olan arkadaşım İsmail Erol'a,

Tez izleme komisyonunda yer alan ve katkı sunan değerli hocalarım Prof. Dr. Ayşe Evrim Bayrak ile Prof. Dr. Asuman Gedikbaşı'na ve tez savuma jürimde yer alan sayın hocalarıma,

Laboratuvar çalışmalarında birçok şey öğrendiğim Tıbbı Genetik Anabilim Dalı Sitogenetik Laboratuvarı emektarı olan Nuray, Havva ve Zuhal Hanıma, Moleküler Genetik laboratuvarı emektarlarına, Doç. Dr. Güven Toksoy, Dr. Çağrı Güleç, Dr. Volkan Karaman, Cemal Koç ve Fatma Dağ'a,

Tüm eğitim hayatım boyunca birçok fedakârlık yapan haklarını ödeyemeyeceğim kıymetli Babam ve Annem'e, akademi serüvenimde zamanlarından çokça çaldığım ve her zaman desteklerini gördüğüm Eşim ve göz bebeğim iki oğluma, her zaman ve her şartta yanımda olan kardeşlerime ve ismini sayamadığım üzerimde emeği olan herkese teşekkür ederim

Bu çalışma, İstanbul Üniversitesi Bilimsel Araştırma Projeleri Birimi tarafından desteklenmiştir. Proje No: TDK-2021-37592

İÇİNDEKİLER

BEYAN..... **HATA! YER İŞARETİ TANIMLANMAMIŞ.**
İTHAF..... İİ
TEŞEKKÜR ..... İİİ
İÇİNDEKİLER V
TABLOLAR LİSTESİ..... Vİİİ
ŞEKİLLER LİSTESİ X
SEMBOLLER / KISALTMALAR LİSTESİ ..... Xİİ
ÖZET ..... Xİİİ
ABSTRACT ..... XİV
1. GİRİŞ VE AMAÇ..... 1
2. GENEL BİLGİLER 3
2.1. Nadir Genetik Hastalıklar 3
2.2. Nadir Genetik Hastalıkların Moleküler Tanısındaki Gelişmeler 4
2.3. Nadir Genetik Hastalıklarda Yeni Tanı Yöntemleri 6
2.4. Yeni Nesil Dizileme Yöntemleri 7
2.4.1. Hedefli Dizileme Paneli 7
2.4.2. Tüm Ekzom Dizileme 8
2.4.3. Tüm Genom Dizileme..... 9
2.5. Yeni Nesil Dizileme Yöntemleri ile Saptanan Tesadüfi Bulgular..... 9
2.6. Yeni Nesil Dizileme Yöntemleri ile Saptanan Önemi Bilinmeyen Varyantlar 10
2.7. Tüm Genom Dizilemede Kullanılan Platformlar..... 11
2.8. Tüm Genom Dizilemede Varyant Filtreleme Araçları 12
2.8.1. VEP 14
2.8.2. ANNOVAR..... 15
2.8.3. CADD 16
2.8.4. PolyPhen2 16
2.8.5. GERP 17
2.8.6. MutationTaster..... 17
2.8.7. SIFT 18
2.8.8. REVEL..... 19
2.8.9. DANN 20

2.9. Protein Modelleme ve Simülasyon 20
2.9.1. Moleküler Dinamik Simülasyon 20
2.9.2. Moleküler Dinamik Simülasyonun Sunduğu Faydalar 22
2.9.3. Moleküler Dinamik Simülasyonun Deneysel Çalışmalara Etkisi 22
2.9.4. Moleküler Dinamik Simülasyon için Araç ve Yöntemler 23
2.9.5. Protein Yapı Modellemede Devrimsel Bir Araç: AlphaFold 25
3. GEREÇ VE YÖNTEM 27
3.1. Çalışma Grubunun Oluşturulması 27
3.2. Klinik Değerlendirmeler 27
3.3. Sentromerik İnstabilite Testi 27
3.4. Tüm Genom Dizilemede Moleküler Çalışmalar 28
3.4.1. Yanak Mukozasından Genomik DNA Eldesi 28
3.4.2. Tüm Genom Dizileme Analizi Islak Laboratuvar Çalışmaları 29
3.5. Tüm Genom Dizileme Ham Verilerinin Analizi 41
3.6. Saptanan Varyantların Konfirmasyonu için Primerlerin Tasarlanması 46
3.7. Saptanan Varyant Bölgelerinin Polimeraz Zincir Reaksiyonu (PZR) ile Çoğaltılması 48
3.8. PZR Ürünlerinin Pürifikasyonu 49
3.9. Dizi Analizi PCR Reaksiyonun Kurulması 49
3.10. Dizi Analizi Yapılacak Amplikonların Temizlenmesi 50
3.11. DNA Dizi Analizi 50
3.12. Dizileme Sonuçlarının Değerlendirilmesi ve Biyoinformatik Analizi 51
3.13. Protein Modelleme ve Simülasyon 51
4. BULGULAR 53
4.1. Olgu 1 53
4.1.1. Tüm Genom Dizileme Sonuçları 54
4.1.2. Tüm Genom Dizileme ile CNV Analizi 61
4.2. Olgu 2 63
4.2.1. Tüm Genom Dizileme Sonuçları 64
4.2.2. Tüm Genom Dizileme ile CNV Analizi 70
4.3. Olgu 3 72
4.3.1. Tüm Genom Dizileme Sonuçları 73
4.3.2. Tüm Genom Dizileme ile CNV Analizi 80

4.4. Tüm Olgulara Ait Segragasyon ve Konfirmasyon Çalışması.....81
4.5. ATP8A2, NFU1 ve DNMT3B Genleri Tarafından Sentezlenen Proteinlerin Tüm Sekans Modellenmesi ve Simülasyonu.....86
5. TARTIŞMA.....99
5.1. Aile 1.....100
5.2. Aile 2.....104
5.3. Aile 3.....107
5.4. Sonuç.....110
KAYNAKLAR.....113
HAM VERİLER.....124
FORMLAR.....125
ETİK KURUL KARARI..... **HATA! YER İŞARETİ TANIMLANMAMIŞ.**
İNTİHAL RAPORU İLK SAYFASI.....129
ÖZGEÇMİŞ..... **HATA! YER İŞARETİ TANIMLANMAMIŞ.**

TABLOLAR LİSTESİ

Tablo 3-1: DNBSEQ-T7RS yüksek veri çıktılı dizileme seti (FCL PE100) 31
Tablo 3-2: CPAS Barkod Primer 3 Reaksiyon Kiti 32
Tablo 3-3: Her Akış Hücresi için Önerilen Insert Boyutu ve Teorik Verim* 33
Tablo 3-4: Farklı örnek havuzlama örnekleri 34
Tablo 3-5: FCL SE35/SE50/SE100/PE100 Kitleri için Gerekli ssDNA Hacmi 35
Tablo 3-6: FCLSE35/SE50/SE100/PE100 Make DNB Reaktif Karışımı-1 37
Tablo 3-7: FCLSE35/SE50/SE100/PE100 Primer Hibridizasyon Reaksiyon Koşulları 37
Tablo 3-8: FCLSE35/SE50/SE100/PE100 Make DNB Reaksiyon Karışımı-2 37
Tablo 3-9: FCL SE35/SE50/SE100/PE100 Yuvarlanan Daire Amplifikasyon Koşulu . 38
Tablo 3-10: FCL SE35/SE50/SE100/PE100 için DNB Yükleme Karışımı 40
Tablo 3-11: İnsan ve model organizma çalışmaları veri tabanları 42
Tablo 3-12: Popülasyonlara ve hastalıklara özgü veri tabanları 43
Tablo 3-13: Tez kapsamında kullanılan in siliko varyant tahmin aracı veri tabanları ... 45
Tablo 3-14: Aday genlerin primer dizileri 47
Tablo 3-15: PZR çalışması için kullanılan malzemeler ve final konsantrasyonları 48
Tablo 3-16: PZR çalışma koşulları 49
Tablo 3-17: Sekans Reaksiyon İçeriği 50
Tablo 3-18: Sekans Reaksiyon Sıcaklıkları 50
Tablo 3-19: Kapiller elektroforez yürütme koşulları 51
Tablo 4-1: Olgu 1'e Ait Klinik ve Laboratuvar Bulguları 54
Tablo 4-2: Olgu 1'e Ait TGD Verisinden Biyoinformatik Analiz ile 55
Tablo 4-3: Olgu 1'e Ait TGD Verisinden Biyoinformatik Analiz ile Anote Edilen Varyant Sayıları 56
Tablo 4-4: Olgu 1'e Ait Kalite Kontrol Parametre Sonuçları 57
Tablo 4-5: Olgu-1 Filtreleme Ölçütleri 58
Tablo 4-6: Olgu 1'e Ait TGD verisinin biyoinformatik analizinde gözlenen klinik ilişkili olabilecek varyantlar 59
Tablo 4-7: Olgu 2'ye Ait Klinik ve Laboratuvar Bulguları 64
Tablo 4-8: Olgu 2'ye Ait TGD Verisinden Biyoinformatik Analiz ile Elde Edilen Tüm Varyant Sayıları 64
Tablo 4-9: Olgu 2'ye Ait TGD Verisinden Biyoinformatik Analiz ile Anote Edilen Varyant Sayıları 65
Tablo 4-10: Olgu 2'ye Ait Kalite Kontrol Parametre Sonuçları 66
Tablo 4-11: Olgu-2'nin Filtreleme Ölçütleri 67
Tablo 4-12: Olgu 2'ye Ait TGD verisinin biyoinformatik analizinde gözlenen klinik ilişkili olabilecek varyantlar 68
Tablo 4-13: Olgu 3'e Ait Klinik ve Laboratuvar Bulguları 73
Tablo 4-14: Olgu 3'e Ait TGD Verisinden Biyoinformatik Analiz ile Elde Edilen Tüm Varyant Sayıları 74
Tablo 4-15: Olgu 3'e Ait TGD Verisinden Biyoinformatik Analiz ile Anote Edilen Varyant Sayıları 75
Tablo 4-16: Olgu 3'e Ait Kalite Kontrol Parametre Sonuçları 76
Tablo 4-17: Olgu-3'ün Filtreleme Parametreleri 77
Tablo 4-18: Olgu 3'e Ait TGD verisinin biyoinformatik analizinde gözlenen klinik ilişkili olabilecek varyantlar 78
Tablo 4-19: Tüm Olgulara Ait Sanger Sekans Sonuçları 82

ŞEKİLLER LİSTESİ

Şekil 3-1: NFU1 Geni Ekzon-3 Primer Tasarım Çalışması 47
Şekil 4-1: Olgu-1'e Ait Aile Ağacı 53
Şekil 4-2: Olgu 1'e Ait TGD Verisinden Biyoinformatik Analiz ile Elde Edilen Tüm Varyantlar 56
Şekil 4-3: Olgu 1'e Ait TGD Verisinden Biyoinformatik Analiz ile Anote Edilen Varyantlar 57
Şekil 4-4: Olgu 1'e Ait Okuma Derinliği Grafiği 58
Şekil 4-5: Olgu 1'e Ait TGD verisinin IGV görüntüsü 60
Şekil 4-6: Olgu 1'e Ait TED verisinin IGV görüntüsü 60
Şekil 4-7: Olgu 1 CNV Analizi Anöploidi Sonucu 61
Şekil 4-8: Olgu 1 CNV Analizi Mikrodelesyon/Duplikasyon Sonucu 62
Şekil 4-9: Olgu 1 CNV Analizi Genler Sonucu 62
Şekil 4-10: Olgu 2'ye Ait Aile Ağacı 63
Şekil 4-11: Olgu 2'ye Ait TGD Verisinden Biyoinformatik Analiz ile Elde Edilen Tüm Varyantlar 65
Şekil 4-12: Olgu 2'ye Ait TGD Verisinden Biyoinformatik Analiz ile Anote Edilen Varyantlar 66
Şekil 4-13: Olgu 2'ye Ait Okuma Derinliği Grafiği 67
Şekil 4-14: Olgu 2'e Ait TGD verisinin IGV görüntüsü 69
Şekil 4-15: Olgu 2'ye Ait TED verisinin IGV görüntüsü 69
Şekil 4-16: Olgu 2 CNV Analizi Anöploidi Sonucu 70
Şekil 4-17: Olgu 2 CNV Analizi Mikrodelesyon/Duplikasyon Sonucu 70
Şekil 4-18: Olgu 2 CNV Analizi Genler Sonucu 71
Şekil 4-19: Olgu 3'e Ait Aile Ağacı 72
Şekil 4-20: Olgu 3'e Ait TGD Verisinden Biyoinformatik Analiz ile Elde Edilen Tüm Varyantlar 75
Şekil 4-21: Olgu 3'e Ait TGD Verisinden Biyoinformatik Analiz ile Anote Edilen Varyantlar 76
Şekil 4-22: Olgu 3'e Ait Okuma Derinliği Grafiği 77
Şekil 4-23: Olgu 3'e Ait TGD verisinin IGV görüntüsü 79
Şekil 4-24: Olgu 3 CNV Analizi Anöploidi Sonucu 80
Şekil 4-25: Olgu 3 Analizi Mikrodelesyon/Duplikasyon Sonucu 80
Şekil 4-26: Olgu 3 Analizi Genler Sonucu 81
Şekil 4-27: Aile 1 Sanger Konfirmasyon ve Segregasyon Dizileme Görüntüleri 83
Şekil 4-28: Aile 2 Sanger Konfirmasyon ve Segregasyon Dizileme Görüntüleri 84
Şekil 4-29: Aile 3 Sanger Konfirmasyon ve Segregasyon Dizileme Görüntüleri 85
Şekil 4-30: ATP8A2 için tüm sekans modellemesi 86
Şekil 4-31: Yabanıl Tip (WT) ve Mutant (R467Q) ATP8A2 için RMDS grafiği 87
Şekil 4-32: Yabanıl Tip ve Mutant AT8A2 simülasyonları için RMSF değerleri 88
Şekil 4-33: DNMT3B için tüm sekans modellemesi 89
Şekil 4-34: Yabanıl Tip (WT) ve Mutant (V628L) DMT3B için RMDS grafiği 90
Şekil 4-35: Yabanıl Tip ve Mutant DMT3B simülasyonları için RMSF değerleri 91
Şekil 4-36: NFU1 için tüm sekans modellemesi 92
Şekil 4-37: Yabanıl Tip (WT) ve Mutant (F64S) NFU1 için RMDS grafiği 93
Şekil 4-38: Yabanıl Tip ve Mutant NFU1 simülasyonları için RMSF değerleri 94
Şekil 4-39: NFU1 proteini 64. aminoasit yabanıl tip modeli 95

Şekil 4-40: NFU1 proteini 64. aminoasit mutant modeli 95
Şekil 4-41: ATP8A2 proteini 467. aminoasit yabanıl tip modeli 96
Şekil 4-42: ATP8A2 proteini 467. aminoasit mutant modeli 96
Şekil 4-43: DNMT3B proteini 628. aminoasit yabanıl tip modeli 97
Şekil 4-44: DNMT3B proteini 628. aminoasit mutant modeli 97
Şekil 4-45: NFU1 Proteinin türler arası aminoasit dizisi karşılaştırması 98
Şekil 4-46: ATP8A2 Proteinin türler arası aminoasit dizisi karşılaştırması 98
Şekil 4-47: DNMT3B Proteinin türler arası aminoasit dizisi karşılaştırması 98
Şekil 5-1: Mitokondrideki Demir-Sülfür Küme Sentezi Süreci 101

SEMBOLLER / KISALTMALAR LİSTESİ

ACMG: American College of Medical Genetics and Genomics

bç: Baz çifti

kb: Kilobaz

Tb: Terabaz

CNV: Kopya sayısı değişiklikleri

ddNTP: dideoksi nükleotid trifosfat

DNA: Deoksiribonükleik asit

dNTP: Deoksinükleotid trifosfat

ED: Ekzom dizileme

ESP: Exome Sequencing Project

ExAC: Exome Aggregation Consortium

GD: Genom dizileme

GH: Gebelik haftası

HPO: Human Phenotype Ontology (İnsan Fenotip Ontolojisi)

Indel: İnsersiyon/delesyon

KE: Kapiller elektroforez

KMA: Kromozomal Mikroarray

mRNA: mesajcı RNA

OD: Otozomal dominant

PZR: Polimeraz zincir reaksiyonu

RNA: Ribonükleik asit

SD: Sanger Dizileme

SNV: Tek nükleotid varyantları

TDG: Tüm genom dizileme

TED: Tüm ekzom dizileme

TGK: Tekrarlayan gebelik kayıpları

USG: Ultrasonografi

VUS: Variant uncertain significance/önemi bilinmeyen varyant

WHO: World Health Organization

YND: Yeni nesil dizileme

ÖZET

GÜLAÇAR İM. Tüm Genom Dizileme Yönteminin Genetik Hastalıkların Tanısına Katkısı. İstanbul Üniversitesi Sağlık Bilimleri Enstitüsü, Genetik Anabilim Dalı. Doktora Tezi. İstanbul. 2023.

Amaç: Fenotipik bulguları ile nadir genetik hastalıktan şüphelenilen ancak rutin tanı yöntemleri ile herhangi bir sendrom veya hastalıkla ilişkilendirilmemiş, tanı almamış olgularda Tüm Genom Dizileme yöntemi ile etiyolojiyi aydınlatmak ve olası mutasyonların fonksiyonel etkilerini modellemek amaçlanmıştır.

Yöntem: İstanbul Üniversitesi İstanbul Tıp Fakültesi, Tıbbi Genetik Anabilim Dalı'na başvuran ve daha önce tüm ekzom analizi yapılıp klinikle ilişkili varyant saptanmamış, nöromotor gelişim geriliği olan 3 olgu için tüm genom dizileme işlemi planlandı.

Bulgular: Çalışma kapsamında tüm genom dizileme yöntemiyle 3 olguda da fenotip ile ilişkili birer anlamlı varyant tespit edildi. Olgu-1'de *NFU1* geninde bir varyant, Olgu-2'de *ATP8A2* geninde bir varyant ve Olgu-3'te *DNMT3B* geninde bir varyant saptandı. Varyant tahmin araçları üç varyantı da patojenik olarak saptadı. Üç olgunun ebeveynleri ve kardeşleri için Sanger dizileme yöntemi ile doğrulama ve segregasyon işlemi gerçekleştirildi. Üç gen için tüm sekans protein modelleme ve simülasyon işlemleri gerçekleştirilerek yabanıl tip ve mutant formlar karşılaştırıldı. Varyantların proteinin uzunluğunu değiştirmeyen ancak yapıda dengesizlik oluşturacak değişime sebep olduğu saptandı.

Sonuç: Rutin genetik tanı yöntemleri ile tanısı konulamayan fenotipik bulguları birbirinden farklı olan ve aralarında akrabalık olmayan üç olgu için tüm genom dizileme yapılmış, klinikle ilişkili varyantlar saptanmış ve TGD tekniğinin üstünlüğü gösterilmiştir. Tespit edilen varyantlar ile ilgili modelleme çalışmaları yapılarak literatür karşılaştırılması ile klinikle ilişkileri ortaya konmuştur. İlgili fenotiplerle alakalı ileri çalışmaların yapılmasına zemin hazırlanmıştır.

Anahtar Kelimeler: Nadir Genetik Hastalıklar, Tüm Genom Dizileme, Moleküler Dinamik Simülasyon, Protein Modelleme,

Bu çalışma, İstanbul Üniversitesi Bilimsel Araştırma Projeleri Birimi tarafından desteklenmiştir. Proje No: TDK-2021-37592

ABSTRACT

GÜLAÇAR İM. Contribution of Whole Genome Sequencing Method to Diagnosis of Genetic Diseases. İstanbul University, Institute of Health Science, Department of Genetics. PhD Thesis. İstanbul. 2023.

Purpose: It is aimed to elucidate the etiology by the Whole Genome Sequencing method and to simulate the functional effects of possible mutations in patients with suspected rare genetic diseases whom the routine diagnostic methods failed to reach a definitive diagnosis.

Methods: Whole genome sequencing was performed in 3 cases with neuromotor developmental delay, who referred to Istanbul University Istanbul Faculty of Medicine, Department of Medical Genetics after an indefinite result in whole exome analysis.

Results: Within the scope of the study, a significant variant associated with phenotype was detected in all 3 cases by whole genome sequencing method. In Case-1, a variant was detected in the NFU1 gene, one detected in the ATP8A2 gene in Case-2, and one detected in the DNMT3B gene in Case-3. Variant prediction tools detected all three variants as pathogenic. Confirmation and segregation were performed with the Sanger sequencing method for the parents and siblings of three cases. The wild-type and mutant forms were compared by performing all sequence protein modeling and simulation processes for the three genes. It was determined that the variants caused a change that did not change the length of the protein but would cause an imbalance in the structure.

Conclusion: Whole genome sequencing was performed for three cases in which the previous routine diagnostic tests were inconclusive. Clinically relevant variants detected in all 3 patients showed the superiority of the TGD technique over other routine genetic tests. Computer Modelling of suspected variants was shown to be beneficial when making a prediction of functional effects of the variations. The groundwork has been prepared for further studies on the related phenotypes.

Key Words: Rare Genetic Diseases, Whole Genome Sequencing, Molecular Dynamic Simulation, Protein Modelling, Computer Modelling,

The present work was supported by the Research Fund of Istanbul University. Project No: **TDK-2021-37592**

1. GİRİŞ VE AMAÇ

İleri düzey teknolojik gelişmelere ve asırlar boyu biriken tecrübeye rağmen insanoğlunun sebeplerini tam manasıyla ortaya koyamadığı hala birçok hastalık bulunmaktadır. Bazı genetik hastalıklar, bir toplumda görülme sıklığı bakımından diğer hastalıklara göre daha az görüldüğü için nadir hastalık olarak tanımlanmıştır. Her biri görülme sıklığı bakımından daha az görülse bile binlerce çeşit nadir hastalık, dünya genelinde milyonlarca insanı etkilemektedir [1].

Çok sayıda çeşidi olan nadir hastalıklar heterojen yapıda ve coğrafi olarak da oldukça dağınıktır. Büyük bir kısmı kronik olan ve erken ölümle sonuçlanan nadir hastalıkların az bir kısmının tedavisi bulunmaktadır [2]. Genetik biliminin gelişmesiyle tek gen hastalıkları veya Mendeliyen hastalıklar olarak da tanımlanan nadir hastalıkların büyük bir kısmının genetik kökenli olduğu keşfedilmiştir. Dünya genelinde yaklaşık 400 milyon kişi, 7 binin üzerinde farklı nadir genetik hastalığa sahiptir. Bunlar tek gen hastalığı olsa da bu hastalıkların yaklaşık olarak yarısının genetik temeli tam olarak aydınlatılmamıştır [3].

İnsan genom projesinin tamamlanmasının ardından birçok genetik hastalığın etiyolojisi daha hızlı aydınlatılmaya başlanmıştır. 2009 yılından itibaren Yeni Nesil Dizileme (YND) yöntemlerinin kullanılmasıyla da çok daha hızlı ve verimli bir şekilde birçok genetik hastalığın sebepleri ortaya konmuştur. Her geçen gün maliyetlerin azalması, elde edilen verilerin işlenme hızının artması, yorumlama araçlarının gelişmesi ve verilerin saklanmasındaki gelişmeler; birçok insanın YND yöntemleri ile genetik test hizmeti almasını kolaylaştırmıştır. Hastalıkla ilişkili olduğu bilinen genlerin veya gen gruplarının dizilendiği hedefli dizileme panelleri, daha az maliyetli ve daha hızlı sonuç imkânı verdiği için tercih edilmektedir. Öte taraftan klinikte daha fazla rastlanan yaklaşık 4 binden fazla geni içeren klinik ekzom dizileme testi ile tüm genlerin ekzonlarını kapsayan tüm ekzom dizileme (TED) testleri de gün geçtikçe daha çok tercih edilmektedir. Genetik hastalıkların çok büyük bir kısmına sebep olan varyantlar genlerin ekzon bölgelerinde yer aldığından TED yöntemi en çok tercih edilen yöntemdir. Ancak TED testleri yapıldığı halde klinikle uyumlu herhangi bir varyantın saptanmadığı durumlarda, insan genomunun neredeyse tamamını kapsayan tüm genom dizileme (TGD) tekniği tercih edilmektedir.

TGD tekniği insan genomunun neredeyse tamamını kapsadığından TED tekniğine göre yaklaşık 100 kattan daha fazla veri üretir. Dolayısıyla bu verilerin işlenmesi, yorumlanması ve saklanması görece daha zordur. Ancak TGD yöntemi ile tanısı konmamış, herhangi bir sendrom ile ilişkilendirilmemiş birçok vaka da çözülebilmektedir.

Bu çalışmada genetik bozuklukları klinik bulgularla ortaya çıkan ve rutin genetik testleri yapılarak herhangi bir hastalık veya sendromla ilişkilendirilememiş vakaların, TGD tekniği ile tanısını koymak amaçlandı. Çalışmamızda her birinin klinik bulguları birbirinden farklı olan vakaların TGD tekniği ile analizi, daha önce yapılan TED verilerinin sonuçları ile TGD sonuçları kıyaslanması, ailelerde segregasyon çalışması, klinikle ilişkili olduğu düşünülen varyantların protein modellemesi ve simülasyonunu yapmak hedeflenmiştir.

2. GENEL BİLGİLER

2.1. Nadir Genetik Hastalıklar

Nadir hastalıkların görülme sıklığı diğer hastalıklara göre daha düşük olmasına rağmen dünya genelinde önemli sayıda insanı etkileyen bir hastalık grubudur. Ülkeden ülkeye görülme sıklığı değişiklik gösterse de Avrupa ve Türkiye'de 1/2000 oranında görülen hastalıklar nadir hastalıklar olarak adlandırılmıştır. Bireysel olarak nadir hastalık şeklinde tanımlanmasa da istatistikler nadir hastalıkların tüm nüfusun %10'unu etkilediğini göstermektedir. Akraba evliliği oranının yüksek olduğu toplumlarda nadir hastalıkların çok ciddi sağlık sorunlarına ve sosyal problemlere neden olduğu saptanmıştır [1]. Doğası gereği heterojen bir yapıya sahip olan ve coğrafi olarak farklılıklar gösteren bu hastalıkların sayısı kesin olarak bilinmemektedir. Son yıllarda yapılan çalışmalar sayesinde birçok yeni nadir genetik hastalık tanımlanmaktadır. Büyük bir kısmı kronik olan nadir hastalıkların, çok azı önlenebilmekte veya tedavi edilebilmekte olup genellikle erken ölümle sonuçlanmaktadır [2]. Bu nedenle nadir hastalıklar, sağlık üzerinde önemli bir etkiye sebep olurlar.

37 ülkeden oluşan Avrupa Komisyonu tarafından ortak fonla desteklenen Orphanet (www.orpha.net), nadir hastalığı olan kişilerin tanısı ve bunlara uygulanacak tedavi yöntemlerinin geliştirilmesi için bilgi havuzu oluşturmayı amaçlayan bir ağdır. Orphanet veri tabanında, halka açık epidemiyolojik verilere göre nadir hastalıkların %71,9'u genetik temelli olup, %69,9'u da sadece pediatrik başlangıçlıdır [2]. Çoğu monogenik olan nadir hastalıkların büyük bir bölümü genetik kökenlidir. Monogenik olarak tanımlanmış tüm Mendeliyen hastalıkların yarısından fazlasının ise genetik temeli hala belirsizliğini korumaktadır [3]. Nadir genetik hastalıkların, genetik temelindeki belirsizliğin önemli nedenlerinden biri; genetik ve klinik olarak heterojen bir profile sahip olmalarıdır. Dolayısıyla vakaların tamamen aynı veya benzer fenotiplere sahip olmaları, farklı genetik mekanizmalardan kaynaklanabilmektedir. Bu durum lokus veya alel heterojenitesi ile ilişkili olabilmektedir. Lokus heterojenitesi farklı genlerdeki mutasyonları gösterirken, alel heterojenitesi aynı gendeki farklı alelleri ifade etmektedir [4]. Klinik heterojenite ise aynı gendeki aynı mutasyon veya mutasyonların farklı hastalık fenotiplerine yol açabileceği gerçeğine dayanmaktadır. Ayrıca birçok genetik hastalıkta genetik yatkınlık her zaman tam penetran olmayabilmektedir. Böylece genetik varyasyon fenotipe yansımayabilir ve çeşitli fenotiplerde ifade edilebilir. Genetik hastalıkların klinik

belirtileri genellikle hastalığın tanısına yönelik net bir bilgi vermeyebilir yani patognomonik değildir. Tüm bu durumlar, nadir görülen genetik hastalıkların tanısında ve hasta yönetiminde zorluklar oluşturmaktadır [5].

Globalgenes'in bir raporunda çoğu klinisyenin bu hastalıklar hakkında sınırlı bilgiye sahip olduğu bildirilmiştir. Bu durum, nadir görülen hastalıkların tanısının geç konulmasına neden olmaktadır. Ayrıca, hastanın nadir görülen bir hastalığı olduğundan şüphelenilse bile birçok nadir hastalığın semptomlarının benzer olması nedeniyle yanlış bir tanı konulma olasılığı söz konusudur [2]. Hastalığın klinik olarak tanımlanması zor olduğundan, genetik ve genomik testler tanı için çok büyük önem taşımaktadır. Son 10 yılda, yeni nesil dizileme teknolojileri ile hedefli gen panelleri, ekzom dizileme (ED) ve genom dizileme (GD) gibi testler çok sayıda hastalık geninin aynı anda test edilmesini sağlamaktadır [6].

Son yıllarda gelişen teknoloji ile genetik tanı sistemlerinin kullanılması birçok nadir genetik hastalığın tanısını mümkün kılmıştır. Hastalıkların genetik tanısının ve patogenezinin aydınlatılması, tedavi yaklaşımlarının geliştirilmesine de olanak sağlamaktadır. Genetik tanı sayesinde; ilaç ile tedavi şansı olmasa bile risk taşıyan aileler bilgilendirilerek prenatal hatta preimplantasyon tanı ile sağlıklı çocuk sahibi olma olasılığı yükselmektedir. Ayrıca, geniş ailelerde taşıyıcı bireylerin belirlenmesi hastalıkların görülme riskini de azaltmaktadır.

2.2. Nadir Genetik Hastalıkların Moleküler Tanısındaki Gelişmeler

1980 yılından 2000'li yıllara kadar hastalıkların gelişiminden sorumlu olan genleri tanımlamaya yönelik çalışmalar; nadir, monogenik ve sendromik hastalıklar ile bunların bağlantı analizleri üzerine yoğunlaşmıştır. Hastalığa neden olan alelleri tanımlamak için bağlantı lokuslarındaki genleri haritalamak amacıyla yapılan Sanger dizileme, genetik sinyallerin lokalizasyonuna vesile olmuştur. Elde edilen veriler, hastalığa neden olduğu düşünülen patojenik varyantın fonksiyonel etkilerini belirlemeye yönelik hücre hatları ve hayvan modelleri üzerinde çalışmaların gerçekleştirilmesine olanak sağlamıştır. Genleri ve etkilerini tanımlamak için kat edilen bu süreçler oldukça zahmetliydi. Buna rağmen 2000 yılına kadar Huntington hastalığı ve kistik fibroz gibi önemli biyomedikal etkiye sahip genetik hastalıkların birçoğu da dahil olmak üzere, tahmini sekiz bin tek gen hastalıktan yaklaşık bin tanesi karakterize edilmiştir [11].

2000 yılından sonra, insan genom projesinin tamamlanmasıyla hastalık genlerini haritalamanın önündeki birçok engel kalkmıştır. Buna bağlı olarak nadir tek gen hastalıklarının gelişiminden sorumlu genlerin sayısında önemli miktarda artış olmuştur. Yapısal varyasyonun mikroarray tabanlı tespiti, ekzom ve genom çapında dizileme, in siliko analizler ve keşfedilen genetik varyantların önceliklendirilmesi; hastalık genlerinin tanımlanması sürecinin ilerlemesine büyük katkı sağlamıştır. Yüksek hacimli dizileme teknolojilerinin benimsenmesi, tek mutasyonlardan büyük yapısal yeniden düzenlemelere kadar tüm nedensel genetik varyasyonların tek bir analizle tanımlanmasını sağlamıştır. Bu teknolojiler, araştırma çalışmalarının yanında klinik kullanıma da geçerek genetik hastalıkların daha erken ve daha hızlı teşhisini sağlamıştır [12].

Tıbbi genetik alanının en önemli prensiplerinden biri; tanısı konmamış Mendeliyen hastalıkların genetik etiyolojilerini belirleyebilmesi, önleyici ve tedavi edici müdahalelere kılavuzluk edebilmesidir. Mendeliyen hastalıklar nadir gibi gözükse de milyonlarca kişiyi ve ailelerini etkileyen hastalıklardır [14]. Bugüne kadar 8 binin üzerinde farklı hastalığın genetik temeli belirlenmiş durumdadır [15]. Nadir tek nükleotid varyantların (SNV), küçük insersiyon/delesyon (indel) varyantlarının ve kopya sayısı varyantlarının (CNV) birçok Mendeliyen özelliğin temelinde rol oynadığı gösterilmiştir. Dolayısıyla tanımlanmamış hastalıkların, büyük ölçüde Mendeliyen kalıtım modeliyle kalıtılan tek gen (monogenik) veya tek lokus kaynaklı hastalıklar olduğu düşünülmüştür [16]. Mendeliyen özelliklerinin araştırılması; nadir görülen hastalıkların altında yatan genomik etiyolojilerin ve moleküler mekanizmaların anlaşılabilmesinde önemli bir etkiye sahip olmuştur [17].

Genler veya bireysel genotipler ile klinik bulguları belirlenen fenotipler arasındaki ilişkinin kapsamlı bir şekilde anlaşılması genomik tıbbın klinikte uygulanması olarak adlandırılabilir. Kromozomal mikroarray (KMA) ve ekzom dizileme (ED) gibi genom araştırmaları, hastalık geni keşiflerine vesile olmuştur. Ancak bu ilerlemelere rağmen, protein kodlayan genlerin sadece %20'si (4081/~ 20,000) bir veya daha fazla hastalık özelliği ile ilişkilendirilmiştir. Ayrıca, birden fazla lokustaki varyasyonun, allelik ve lokus heterojenitesinin ve ortak varyantların Mendeliyen özelliklere ne ölçüde katkıda bulunduğu henüz tam olarak anlaşılmamıştır. Dolayısıyla hastalığa sebep olan gen keşfi, sadece hastalık fenotipi ve gendeki varyasyonun birebir uyumundan ibaret değildir [17].

Genom dizileme (GD), yeni nesil dizileme maliyetleri düştüğü için hem araştırma hem de klinik bazda daha kolay erişilebilir hale gelen en son geniş tabanlı test yöntemidir [18]. Tüm genom dizileme yönteminin Mendeliyen hastalıkların teşhisindeki kullanımı ve araştırmalara katkısını incelemek gerekmektedir.

2.3. Nadir Genetik Hastalıklarda Yeni Tanı Yöntemleri

Yüksek verimli dizileme olarak da bilinen Yeni Nesil Dizileme (YND), yüzlerce mega bazdan giga bazlara kadar nükleotid dizisinin üretilebilmesi için deoksiribonükleik asit (DNA) moleküllerinin bir cihazda paralel dizilenmesine olanak sağlar. YND, 2009 yılında ilk platformun ticari olarak piyasaya sürülmesinden bu yana yeni genlerin tanımlanması sürecini önemli ölçüde hızlandırmıştır [19]. YND, genetik hastalık teşhisinde hızlı, yüksek verimli ve uygun maliyetli yaklaşımlarla yeni bir dönem açmasının yanında nadir hastalıkların heterojen genetik profillerinin tanımlanmasında da önemli bir ilerleme sağlamıştır. YND sayesinde birkaç genden yüzlerce gene, tüm ekzomdan tüm genoma eşzamanlı olarak analizler yapılabilmektedir. Böylelikle klinik hipotezin ötesine geçerek aday genlerin araştırılmasına da imkân sağlamaktadır [20]. Ayrıca YND ters fenotipleme olarak bilinen yeni bir sürece yol açmıştır. Bazı durumlarda, YND ve segregasyon analizinin birlikte kullanımı, hastalığa neden olduğu bilinen ancak daha önce farklı bir fenotiple bağlantılı olan bir gendeki patojenik bir mutasyonu tanımlayabilir. Bu gibi durumlarda, hastanın ve aile üyelerinin retrospektif klinik incelemesi, daha önce bilinmeyen ek özelliklerin belirlenmesini sağlayabilir. Bu bağlamda YND, klinisyenlere henüz tam olarak ifade edilmemiş veya henüz ortaya çıkmamış belirtileri araştırmaya başlamalarını sağlayan önemli ipuçlarını ortaya çıkarmak için tanı aşamasının başlangıcında kullanabilecekleri güçlü moleküler araçlar sağlamaktadır [8].

Nadir hastalık araştırmalarında üç ana YND tabanlı test kullanılmaktadır. Bu testler maliyet, analiz kolaylığı ve kapsama göre sıralanabilir: 1) Benzer veya örtüşen fenotipler ile ilişkili gen gruplarının (gen panelleri) kodlama dizilerinin (ekzonlar) paralel dizilimi; 2) İnsan genomunun bilinen tüm kodlama bölgelerinin dizilendiği Tüm Ekzom Dizileme (TED); ve 3) Tüm insan genomunu analiz eden Tüm Genom Dizilime (TGD).

TED, son yıllarda nadir görülen hastalık araştırmalarında yaygın kullanılmaya başlamıştır. Proteine çevrilen insan genomu, genomun sadece %1'ini (~30Mb) kapsar.

Ancak TGD ile karşılaştırıldığında, genomik verileri toplamak ve analiz etmek için önemli ölçüde daha az maliyet ve zaman açısından etkin bir yöntem sunar. Gen panelleri ise TED'e göre çok daha hızlı geri dönüş süreleri, daha az rastlantısal bulgu ve daha yüksek kapsam sunar. Böylece CNV'leri ve somatik mozaik mutasyonları saptama şansını artırır. Fakat, yeni hastalığa neden olan genleri tanımlayamazlar [8].

Hastalıklarla ilişkili yeni genlerin veya yeni fenotiplerle ilişkili bilinen genlerdeki değişimlerin tanımlanması, hastalığın biyolojik temellerini anlamamıza katkı sağlamaktadır. Ayrıca hasta yönetimi için ve hatta bazı durumlarda tedavi edici müdahale için büyük önem taşır. Hastalığa neden olan gen ve değişimleri tanımlamak için iki ana yaklaşım kullanılmıştır: Birincisi, aynı klinik özelliklere sahip bir grup hastanın ekzomunun veya genomunun analizi ve gruptaki herkes için ortak bir gendeki varyantların filtrelenmesidir. İkincisi, izole hastaların ebeveynleri ve/veya aile üyeleri ile birlikte analizi ve varyantların farklı kalıtım biçimlerine (otozomal dominant, resesif, X'e bağlı veya *de novo*) göre filtrelenmesi, böylece hastalığa sebep olan geni tespit etmek için yeterince az sayıda aday varyanta düşürülmesidir. 2012 yılında Goldstein ve çalışma arkadaşları tarafından gerçekleştirilen bir çalışmada, genetik temeli açıklanamayan 12 farklı hastada trio-TED (bir hastanın ve biyolojik ebeveynlerinin tüm ekzom dizilenmesi) uygulanmıştır. Hastaların altısında genetik tanıyı doğrulamayı başarmışlar ve hatta hastaların bazılarına tanıya dayalı potansiyel bir tedavi yaklaşımı önerebilmişlerdir [8].

2.4. Yeni Nesil Dizileme Yöntemleri

Klinik olarak değerlendirilen ve bir Mendeliyen hastalık olduğundan şüphelenilen hastada nadir varyantları belirlemek için yukarıda kısaca değinilen üç ana dizileme yöntemi kullanılmaktadır. Bunlar hedefli dizileme panelleri, tüm ekzom dizileme ve tüm genom dizileme yöntemleridir.

2.4.1. Hedefli Dizileme Paneli

Gen panelleri, belirli bir hastalık veya hastalık grubunun patogenezine sebep olan mutasyonları barındırdığı bilinen genler içinde belirli sayıda spesifik gen veya kodlama bölgesine sahiptir. Hedefli paneller ile TED ve TGD'den daha derin okumalar, daha düşük maliyetle yapılır ve panel dizileme daha kısa zaman diliminde sonuçlanır. Tespit edilen varyantlar, seçilen genlerle sınırlıdır ve daha düşük hacimde veri üretilir. Sonuç olarak

yorumlama için gereken iş yükü daha düşüktür. TED ve TGD'ye kıyasla tesadüfi bulguların ve önemi bilinmeyen varyantların çıkma ihtimali azdır. Ancak, yeni bilgi ve gen keşifleri nedeniyle paneller düzenli güncellemeler gerektirir. Paneller ve TED; yapısal varyantların, tekrarlayan elementlerin ve mitokondriyal genomik varyantların tespitinde sınırlıdır [5]. Gen panelleri başlangıçta bir dizi ilgili hastalıkla ilişkili genleri yakalar, ardından büyük ölçekli paralel dizileme gelir. Paneller, özellikle iyi tanımlanmış hastalıkla ilişkili genlere sahip genetik olarak heterojen bozuklukları incelemek için uygundur [13].

2.4.2. Tüm Ekzom Dizileme

Tüm Ekzom Dizileme (TED), genomun %1-2'sini kapsayan ve hastalığa neden olan mutasyonların yaklaşık %95'ini barındıran genomun protein kodlayan bölgelerini araştırır. TED, bilinen insan genleriyle ilişkili olmayan genlerdeki varyantların tanımlanmasını sağlar. TED, önceden seçilmiş bir panel veya seçilmiş bir gen seti ile birlikte yorumlanabilir. Laboratuvar, hastaların fenotipi ile ilişkili olduğu bilinen genlerin bir listesini içeren biyoinformatik paneli seçebilir. İkinci yaklaşım, laboratuvarın tüm nadir ve potansiyel olarak zarar verici varyantları incelediği ve bu genlerle ilişkili fenotipleri hasta fenotipiyle karşılaştırdığı yaklaşımdır. Bu yaklaşımla henüz keşfedilmemiş genlerdeki varyantların tanımlanması sağlanır. Gen bölgelerinin eksik kapsanması, tekrarlayan elementlerdeki varyasyonu tespit etme yeteneğinin sınırlı olması, somatik mozaiklik durumlarında nedensel varyantlar, yapısal varyantlar ve derin intronik varyantlar TED tekniğini sınırlayan özelliklerdir. Bununla birlikte, teknolojinin ilerlemesiyle bilinen tüm hastalıklara neden olan intronik varyantlar dahil olmak üzere, ekzonların dışındaki bölgelerin daha iyi kapsanması sağlanmıştır [5].

Günümüzde maliyetlerin azalmasıyla, nispeten kolaylaşan örnek toplama prosedürüyle ve geniş genetik tarama potansiyeli sayesinde hastalık tanı algoritmasında daha erken aşamalarda TED uygulamasına eğilim artmıştır. TED'in erken uygulanması sonucunda genetik değişimin tanımlanabildiği durumlar artış göstermiştir. Bununla birlikte, son yıllarda hem dizileme teknolojisinin hem de biyoinformatiğin gelişiminde kaydedilen muazzam ilerlemeye rağmen, araştırmaların çoğunda, TED sonucunda genetik tanı elde edilememesi söz konusudur. Böyle durumlarda, teşhisin kesin olarak doğrulanması için hala başka tanı testlerinin uygulanmasına ihtiyaç duyulmaktadır [13].

2.4.3. Tüm Genom Dizileme

Tüm Genom Dizilime (TGD), insan genomunun büyük bir kısmının dizilenmesini sağlayan bir yöntemdir. TGD; yeni genleri, gen düzenleyici kısımları ve genom dizilimi yoluyla karmaşık kalıtım modellerini anlamamıza yardımcı olan verileri elde etme potansiyeline sahiptir. Bu güçlü araç, genetik nedeni tek bir testle keşfetmeyi mümkün kılar. Böylelikle gelecekte genetik test olarak ilk tercih edilen test olabilir. TGD; tek nükleotid varyantlar (SNV'ler), insersiyon/delesyon (indeller), kopya sayısı varyantları (CNV'ler) ve translokasyonlar gibi yapısal varyantlar dahil olmak üzere diğer dizileme yaklaşımlarından daha geniş bir genetik varyasyon yelpazesini saptayabilir. Ekipman bulunmaması, analiz maliyeti, genomun kompleks yapısı ve kodlamayan genomik varyasyonları yorumlamak için mevcut biyoinformatik programlarının sınırlı olması, TGD yönteminin dezavantajları arasında sayılabilir [5].

TED ve TGD, nadir hastalıkların teşhisinde büyük potansiyele sahiptir. Tek bir testte birçok genin analizini sağlarlar. Aynı zamanda tesadüfi bulgular ve bilinmeyen öneme sahip varyantlar (VUS) da elde edilir. Bu durum klinisyenler ve hastalar için ek zorluklar demektir [5]. TGD; kodlayan ve kodlamayan bölgeler, düzenleyici bölgeler ve kopya sayısı varyasyonlarına (CNV'ler) yol açan SNV'ler dahil olmak üzere tüm genomu tarar. Yeni genler, varyantlar, de novo mutasyonlar ve ilişkili lokuslar için aktif genlerin ve sessiz dizilerin aynı anda incelenmesini kolaylaştırır. TGD'nin diğer bir avantajı ise ekzonik tek nükleotid varyantlarını (SNV'ler) ve ayrıca kodlama yapmayan varyantları, küçük kodlama yapmayan düzenleyici RNA'ları, CNV'leri (delesyon, duplikasyon, ve inversiyon dahil), tekrar genişlemelerini ve kompleks kromozomal yeniden düzenlemeleri de tespit edebilmesidir [13].

2.5. Yeni Nesil Dizileme Yöntemleri ile Saptanan Tesadüfi Bulgular

Tesadüfi bulgular, bir hastada şüphelenilmeyen veya hastanın mevcut fenotipiyle ilgili olmayan bir hastalık için risk altında olduğunu gösterebilir. Bu bulgular, hastalığı önlemek, ertelemek veya hafifletmek için bazı önlemlerin alınabileceği koşullar, genler veya varyantlarla ilgili olabilir [21]. TED ve TGD'den elde edilen verilere bakıldığında, hasta fenotipi ile ilgisi olmayan bir patojenik varyantın tesadüfen keşfedilme olasılığı her zaman vardır. American College of Medical Genetics and Genomics (ACMG) tarafından yayınlanan ve 59 geni içeren bir listede, hastaların birincil semptomlarından bağımsız

olarak bu genlerdeki mutasyonların klinisyenlere bildirilmesi önerilmiştir [22]. Avrupa ve Kanada kılavuzları ise daha muhafazakâr bir yaklaşım benimseyerek tesadüfi bulgu olasılığını azaltmak için TED'in sınırlı kullanımını ve mümkünse hedefli dizileme panelini önermiştir [23].

Tesadüfi bulguların raporlanması; erken teşhise, iyi bir hastalık tahminine ve hastalığı önlemeye imkân vererek tıbbi müdahale için bir fırsat sağlar. Bununla birlikte eksik hastalık penetrasyonu ve öngörülemeyen hastalık sonuçları ile hastalık geninde tamamen iyi huylu olabilecek önemi bilinmeyen varyantların bildirilmesi hususunda endişeler söz konusudur. Ayrıca çocuklarda erişkin başlangıçlı hastalık bulgularının ve tesadüfi bulguların bildirilmesinin hastalar üzerinde oluşturabileceği olumsuz psikolojik etkileri hususunda da endişeler mevcuttur [24]. Büyük ölçekli çalışmalarda, bir sağlık sistemi üzerinde büyük bir mali yük oluşturabilecek, %6'ya varan sıklıkta rastlantısal bulgular rapor edilmiştir [25].

2.6. Yeni Nesil Dizileme Yöntemleri ile Saptanan Önemi Bilinmeyen Varyantlar

YND tarafından üretilen büyük miktarda veri göz önüne alındığında, tüm varyantların klinik önemini belirlemek oldukça zordur. Tanımlanan varyantların çoğu ACMG standartları ve kılavuzlarına göre patojenik, muhtemel patojenik, muhtemel iyi huylu veya iyi huylu olarak sınıflandırılabilir. Varyantların büyük bir kısmı belirsiz öneme sahip varyantlar (VUS) kategorisinde kalır. Varyantların hastanın tanısındaki potansiyel rolü; gnomAD, ClinVar ve HGMD gibi veritabanlarında olası varlığı, varyantın özellikleri, fenotipik özelliklerin bilinen gen varyantlarıyla, sağlıklı veya etkilenmiş aile bireylerindeki varyantlarla kıyaslanmasıyla değerlendirilebilir. Her geçen gün artan bilgi birikimi ve uluslararası veri tabanlarının oluşturulması nedeniyle ACMG sınıflandırmalarına göre sınıf 3 olarak bilinen, belirsiz öneme sahip varyantlar (VUS) artık iyi huylu veya olası patojenik olarak sınıflandırılmamaktadır. Öte yandan laboratuvarlar bir hastanın fenotipi ile ilişkili olup olmadığını belirlemek için bir VUS yerine VUS alt kategorilerini tercih edebilmektedir. Bununla birlikte, uluslararası iş birlikleri sayesinde VUS'lar hakkında bilgi alışverişini kolaylaştırarak bir hastalık geninin tespitinde iyileştirmeler sunmanın yanı sıra bu varyantlar için daha fazla patojenite kanıtı sağlayan ve hızlandıran web tabanlı araçlar bulunmaktadır. VUS'un raporlanmasının gerekip gerekmediğine dair özel bir tavsiye bulunmadığından, tanı laboratuvarları arasında VUS raporlamasıyla ilgili değişken uygulamalar mevcuttur. VUS olarak

sınıflandırılan varyantların durumu muhtemelen zamanla değişecektir. VUS'un ilk analizden 12 ay sonra yeniden yorumlanması tanısal etkinliği artırır. Tanı laboratuvarları için, VUS yeniden sınıflandırılıyorsa tıbbi raporun yeniden düzenlenmesi iyi bir uygulama olarak kabul edilir. Ancak birkaç kılavuza göre, laboratuvarların dizileme verilerini rutin olarak yeniden analiz etmesi zorunlu değildir [26].

VUS prevalansı çalışmalar arasında önemli ölçüde değişmektedir. Oran; hasta kohortuna, tanı yöntemine ve incelenen nadir hastalıkların heterojenliğine bağlıdır [27]. Genetik testin belirsiz sonucu; VUS taşıyıcıları ve aileleri için endişe yaratırken hasta yönetiminden sorumlu klinisyenler için de problem olur. VUS terimini anlamak; hastalar ve genetikçi olmayan klinisyenler için zor olabilir. Öte yandan VUS'a gerektiğinden daha fazla klinik anlamlılık verildiğine dair endişeler de bulunmaktadır [28].

2.7. Tüm Genom Dizilemede Kullanılan Platformlar

Birinci nesil dizileme olarak adlandırılan Sanger dizileme metodu ile yıllar süren çalışmalar sonucu insan genomu dizilenmiştir. Ancak bu dizileme işlemi hem uzun zaman almış hem de ciddi bir maliyete sebep olmuştur.

Yeni nesil dizileme teknolojilerinin gelişimi ile insan ve diğer canlıların genomunu dizilemek daha uygun maliyetlerle ve çok daha kısa zaman diliminde yapılabilmiştir. İkinci nesil dizileme olarak adlandırılan bu yeni teknolojilerden biri Roche firmasının 454 olarak bilinen GS FLX + platformudur. Bu cihaz 700 baz çiftinden 1 milyon okuma üretebilir [29]. Life Technologies firmasının ürettiği platform ise Ion Torrent olarak bilinen Proton cihazıdır. Bu cihaz, bir çip üzerindeki DNA'lardan salınan hidrojen iyonlarını tespit eden katı-durum pH metreli yarı iletken dizileme teknolojisini kullanır. Bu platform, bir Ion P1 çipiyle her bir yürütmede 10Gb diziyi, 200 baz çifti uzunluğunda fragmanları, 60-80 milyon okuma ile üretebilir [30]. Illumina markası HiSeq2500, HiSeq4000, HiSeqX ve NovaSeq6000 gibi modellerle birçok YND platformu sunmuştur. Illumina'nın en yüksek verimli dizileme platformu 10 tane HiSeqX cihazını içeren HiSeq X TEN sistemidir. HiSeq X cihazı, her yürütmede 3 gün içinde, 150 baz çifti uzunluğunda yaklaşık 6 milyar okumadan 1.8 Tb diziyi çıkarır. Bu cihaz, eş zamanlı olarak multiparalel dizileme ve çok yüksek verimli dizileme gerektiren Tüm Genom Dizileme için tasarlanmıştır [31]. Öte yandan Complete Genomics ve MGI Tech tarafından geliştirilen BGISEQ-500 ve DNBSEQ-T7 platformları da sık kullanılan YND

platformları arasındadır. Bu iki platformda, çok büyük ölçekte kısa okumalar üretmek için DNA nanotopu ve kombinatoryal prob çapa sentezi teknolojisi kullanılmaktadır [32].

Üçüncü nesil dizileme teknolojileri olarak adlandırılan ve dizileme işlemini direkt olarak kalıp DNA'dan yapan platformlar Qiagen, Helicos Bioscience ve Pacific Bioscience firmaları tarafından geliştirilmiştir. Üçüncü nesil dizileme teknolojisi alanında lider firma Roche firması denebilir. Bu firma PacBio RS II platformunu Pacific Biosciences şirketi üzerinden piyasaya sunmuştur. PacBio RS II platformu, 20.000 baz çiftinden oluşan binlerce uzun okuma üretebilir. Bu uzun okuma teknolojisi *de novo* genomik yeniden yapılanmaları tanımlamak için ciddi bir avantaja sahiptir. [33]

2.8. Tüm Genom Dizilemede Varyant Filtreleme Araçları

Mendeliyen hastalıkların ve nadir görülen kompleks genetik hastalıkların tanısında Tüm Genom Dizileme (TGD) ve Tüm Ekzom Dizileme (TED) gibi yeni nesil dizileme teknolojilerinin kullanımı yaygınlaşmaya başlamıştır. TGD ve TED uygulamalarındaki maliyetlerin azalması ve testlerin hızlı yapılabilir olması bu teknolojilerin kullanımının artmasına neden olmuştur. Ancak bu işlemlerin sonucunda büyük veriler üretildiğinden, bu verilerin saklanması ile birlikte doğru analizlerin yapılması önem arz etmektedir. Sanger dizilemede belirli bir genin belirli bir bölgesi için elde edilen veriler kilobayt düzeyinde iken tüm genom dizileme ham verilerinin büyüklüğü her bir hasta için 300 gigabayta ulaşmış durumdadır. Bu verilerin işlenip anlamlı hale getirilmesi için çok sayıda analiz programına ihtiyaç vardır. Veribilimcilerin biyoinformatik alanında uzmanlık kazanması ile birlikte son yıllarda çok sayıda analiz programı hizmete sunulmuştur.

Dizileme teknolojileri, referans veri setleri, yazılımlar ve analiz veri hatları önemli ölçüde gelişirken; hastalıklarla ilişkili aday genetik varyantların önceliklendirilmesine dair kılavuz ve standartlar henüz emekleme aşamasındadır. TGD ve TED işlemlerinin ardından elde edilen veriler çok sayıda genetik varyant içermesine rağmen çok büyük bir kısmının hastalıklarla ilişkilendirilebilecek derecede fonksiyonel bir değişime sebep olmadığı görülmektedir. Öte yandan fonksiyonel varyantları tespit etmek için tek bir yol söz konusu değildir. Her filtreleme stratejisi fonksiyonel olabilecek varyantı seçmek için birçok mantıklı tahminde bulunabilirken ilgili aday varyantları erken eleyerek birçok yanlış pozitif sonuç da üretebilmektedir. Ancak herhangi bir bireyin ekzomunda yer alan kalıtsal veya de novo mutasyonlar birçok protein sekansını değiştirse de sadece çok az bir

kısmının proteinin fonksiyonunu değiştirebildiğini unutmamak gerekmektedir. Fonksiyon değiştiren mutasyonların da küçük bir kısmı hastalıkla ilişkili mutasyonlar olabilmektedir. Çünkü fonksiyon değişikliğine sebep olsa da inkomplet penetrans ve penetrans yokluğu sebebiyle hastalık görülmeyebilir [34].

Çok sayıda hesaplamalı algoritmalar, varyantların potansiyel işlevsel etkisini ve ilişkisini tahmin etmek için kullanılmaktadır. Ancak verilerin çok büyük olması sebebiyle yapılan tahmini analizlerin dikkatli bir şekilde değerlendirilmesi gerekmektedir. Miosge ve ark. (2015) yaptıkları fonksiyonel genomik çalışmalar neticesinde, genom dizileme işleminden elde edilen *de novo* ve nadir yanlış anlamlı mutasyonların neredeyse yarısının aslında nötr mutasyonlar olduğunu göstermişlerdir [35].

TGD ve TED işlemlerinin ardından üretilen ham verilerinin işlenmesinde izlenilen süreç boyunca yapılan etkili seçimler, en doğru aday varyantı tespit etme olasılığını artırmaktadır. Bu konuda birçok kılavuz hazırlanmıştır. Bu kılavuzlarda referans transktriptin seçimi, varyant anotasyon araçları, kalıtımsal ve istatistiksel filtrelemeler, varyant sınıflandırma filtrelemeleri, popülasyon frekans filtrelemeleri ve fonksiyonel etki tahminleri gibi adımlar ile ilgili detaylı bilgiler sunulmaktadır.

Herhangi bir varyantın fonksiyonel sonuçları, hangi transkript izoformunun analiz edildiğine bağlı olarak farklılık gösterebileceğinden referans transkript seçimi oldukça önemlidir. McCarthy ve arkadaşlarının yaptığı bir TGD çalışmasında RefSeq ve Ensembl transkript veri setleri karşılaştırılmıştır. 80 milyon varyantın analiz edildiği çalışmada, RefSeq veri seti kullanıldığında çok sayıda fonksiyon kaybı (LoF) varyantı tespit edilememiştir. Bu çalışma varyant anotasyonunda Ensembl veri setlerinin kullanımının daha uygun olduğunu göstermektedir [36].

Hesaplamalı algoritmaların bir kısmı ücretsiz analiz programları olarak kullanılırken bir kısmı da ücretli programlar olarak bulunmaktadır. Bu analiz programları varyant anotasyonlarını, varyant frekanslarını, varyantların fonksiyonel tahminlerini ve hastalıkla ilişkili varyantların tespitini yapabilmektedir [34].

TGD verisinde varyantların frekanslarını tespit eden veritabanları arasında 1000G (1000 Genome Project), Kaviar veritabanı, hrcr1 (Haplotype Refrence Consortium) veritabanı, cg69 (Complete Genomics) veritabanı ve gnomAD veritabanı bulunmaktadır [37].

TED verisinde varyantların frekanslarını tespit eden veritabanları arasında exac03 (Exome Aggregation Consortium) veritabanı, esp6500siv2 (NHLBI-ESP) projesi ve gnomad_exome veritabanı bulunmaktadır [37].

TGD verisinde varyantların fonksiyonel tahminlerini yapan veritabanları arasında 9 milyar mutasyonu skorlayan gerp++ ve cadd (Combined Annotation Dependent Depletion) veritabanları, derin öğrenme ile tahmin skoru üreten DANN veritabanı, kodlayan ve kodlamayan varyantları tahmin edebilen fathmm ve eigen veritabanları, çeşitli genomik ve epigenomik anotasyonları birleştirerek genom boyunca tahmin yapabilen gwava veritabanı bulunmaktadır [37].

TED verisinde varyantların fonksiyonel tahminlerini yapan veritabanları arasında dbnsfp30a veritabanı bulunmaktadır. Bu veritabanı SIFT, PolyPhen2 HDIV, PolyPhen2 HVAR, LRT, MutationTaster, MutationAssessor, FATHMM, MetaSVM, MetaLR, VEST, CADD, GERP++, DANN, fitCons, PhyloP ve SiPhy veri setlerini içermektedir. Ancak bu veritabanında sadece kodlayan varyantlar bulunmaktadır [37].

Hastalıklara ilişkin varyantları tahmin eden veritabanları arasında ClinVar veritabanı, cosmic70 veritabanı, icgc21 (International Cancer Genome Consortium) veritabanı ve nci60 (NCI-60 insan tümör hücre hattı panel ekzom dizileme) veritabanı bulunmaktadır. Varyant tanımlamaları için kullanılan veritabanları arasında snp142 (dbSNP) ve ANNOVAR geliştiricileri tarafından geliştirilen avsnp142 veritabanları bulunmaktadır [37].

Varyant anotasyonu için birçok hesaplamalı araç geliştirilmiştir. SnpEff, ANNOVAR ve Variant Effect Predictor (VEP) programları en çok kullanılan araçlar arasındadır. Üç program da Ensembl transkriptlerine dayanarak anotasyon gerçekleştirmektedir. Yine McCarthy ve arkadaşlarının yaptığı çalışmada ANNOVAR programının diğerlerine göre birçok avantajı olduğu gösterilmiştir. Özellikle çerçeve kayması, anlamsız ve katlanma bölgesi varyantlarının anotasyonunda ANNOVAR ön plana çıkmaktadır. Ancak Dashti kılavuzunda VEP programının kullanımı tavsiye edilmektedir [34].

2.8.1. VEP

VEP hem ticari amaçlı hem de ticari olmayan çalışmalar için ücretsiz hizmet veren varyant analiz programıdır. Apache 2.0 lisansıyla (www.ensembl.org/Tools/VEP) web

adresinden veya bir API aracılığı ile erişilebilir durumdadır. VEP ek analizler için algoritmik uzantıları mümkün kılan esnek bir eklenti mimarisine sahiptir. Örneğin, birçok transkript için dokular arasında ekspresyon seviyelerini belirtmek üzere deneysel bir eklenti geliştirilmiştir. Bu eklenti ile ilgilenilen bir hastalığın hedeflediği dokularda olan varyantlar önceliklendirilebilir. VEP hem birçok standart formatta rapor üretirken hem de özel formatlarda rapor üretebilmektedir. Ücretsiz veritabanlarındaki alel frekanslarla SNV ve indel varyantları; dbNSFP veritabanına dayanarak gen/transkript etkisi, evrimsel korunma skorlarını ve tahmini fonksiyonel etki skorlarını anote edebilmektedir. Zigosite, gen sembolü, genomik lokasyon ile eş anlamlı, yanlış anlamlı ve çerçeve kayması varyantlar gibi sınıflandırmalar filtrelemede kolaylık sağlamaktadır. Tüm versiyonları varyant çağırma araçlarının ürettiği VCF formatı ile uyumludur. En çok kullanılan varyant çağırma aracı olan Genome Analysis Toolkit ile de uyum göstermektedir [38].

2.8.2. ANNOVAR

ANNOVAR en çok kullanılan varyant anotasyon araçlarından biridir. Annovar'a (http://annovar.openbioinformatics.org/) adresinden erişim sağlanabilmektedir. Annovar, Perl programlama dilinde yazılmış bir komut satırı aracıdır ve Perl yorumlayıcısının kurulu olduğu çeşitli işletim sistemlerinde çalıştırılabilir. Girdi olarak en yaygın kullanılan VCF formatı da dahil olmak üzere çeşitli girdi formatlarını alabilir. Birkaç farklı formatta ise (anote edilmiş VCF dosyası, sekmeyle ayrılmış metin dosyası veya virgülle ayrılmış metin dosyası gibi) girdi dosyasındaki her bir varyant için anote edilmiş bir varyant çıktı dosyası üretebilir. Örneğin, girdi dosyası '1 881627 881627 G A' olarak bir satır içerir; bu, kromozom 1'de, konum 881627'de G'den A'ya nükleotid değişikliği olan bir varyant olduğu anlamına gelir. Sekmeyle ayrılmış çıktı dosyası, orijinal girdi bilgilerinin yanında ek anotasyon bilgileri içermektedir. Örneğin 'ekzonik' (genomik fonksiyon), 'NOC2L' (etkilenen gen), 'NM_015658' (etkilenen transkript), 'eş anlamlı SNV' (kodlayan varyantın fonksiyonel rolü), 'c.2334G>C' (transkriptteki nükleotid değişikliği) ve 'p.A328G' (proteindeki amino asit değişikliği) gibi ek bilgileri içermektedir [38].

Genel olarak ANNOVAR, gen tabanlı anotasyon, bölge tabanlı anotasyon ve filtre tabanlı anotasyon olmak üzere üç farklı türde anotasyon sağlamaktadır. Gen tabanlı anotasyon ile bir varyantın bilinen genlerdeki etkisi birçok bilgi ile sunulmaktadır. Öncelikle varyantın konumu verilir. Ekzonik, intronik, katlanma bölgesi, 3'UTR, 5'UTR

ve genler arası bölgelerdeki yeri saptanır. Daha sonra eğer varyant ekzonik ise protein kodlamadaki fonksiyonu; eş anlamlı, anlamsız, yanlış anlamlı, çerçeve kayması ve eklenme/delesyon gibi etkiler şeklinde tespit edilerek etkilenen transkript ve ilgili aminoasitteki değişim belirlenir. Bölge bazlı anotasyon; varyantın özellikle korunmuş genomik elementler, sitogenetik bantlar, mikroRNA hedef bölgeleri ve ENCODE bölgelerine denk gelip gelmediğini anlamak için kullanışlıdır. Filtre tabanlı anotasyon ise 1000 Genom Projesi ile ABD Ulusal Sağlık Enstitüleri-Ulusal Kalp, Akciğer ve Kan Enstitüsü (NIH-NHLBI) ESP6500 ekzom dizileme projesinde elde edilen varyantların alel frekanslarını, anlamsız varyantların SIFT ve PolyPhen puanlarını bulmak ve dbSNP veri tabanında bir varyantın varlığının veya yokluğunun belirlenmesi için kullanılabilmektedir [38].

2.8.3. CADD

Fonksiyonel tahmin araçlarından biri olan CADD, zararlı ve potansiyel olarak patojenik olma olasılığı en yüksek olan varyantları belirlemek için referans genom boyunca varyantları skorlayan bir veritabanıdır. Bu nedenle CADD'ın ana hedefi, milyonlarca aday arasından varyantların önceliklendirilmesidir. Bu işlem sırasında CADD; ailesel çalışmalarda elde edilen de novo, baskın ve çekinik varyantlar ile popülasyona dayalı çalışmalarda tanımlanan varyantları tespit edebilmektedir. 2014 yılında piyasaya sürülmesinden bu yana CADD, insan genetik varyasyonunu değerlendirmek için en yaygın kullanılan araçlardan biri haline gelmiştir. Diğer araçlar da kendi performanslarını ölçmek amacıyla sıklıkla CADD aracını tercih etmektedir. Ayrıca, CADD, evrimsel değişikliklerin yorumlanmasından insan popülasyonlarındaki değişken uygunluk etkilerinin teorik olarak araştırılmasına kadar değişen evrimsel çalışmalarda da kullanılabilmektedir [39].

2.8.4. PolyPhen2

İnsanlar arasında fenotipik farklılığı sağlayan en fazla genetik varyasyonun tek nükleotid değişimlerinden (SNP) kaynaklı olduğu bilinmektedir. Protein dizisinde yer alan ve aminoasitlerde değişime sebep olan SNP'lere eş anlamlı olmayan SNP'ler denir. Bu SNP'lerin sayısı oldukça fazladır ve aminoasitlerde alelik değişimlerin proteinin yapısını ve fonksiyonunu değiştirdiğini gösteren birçok çalışma bulunmaktadır. Bu değişimler çoklu dizi hizalaması ve üç boyutlu (3D) protein analizleri ile tahmin edilebilirler. PolyPhen-2 de bir aminoasit değişiminin proteinin yapısına ve fonksiyonuna

etkisini tahmin eden araçlardan biridir. İnsan genetiği araştırmalarında ortaya çıkan çok büyük sayıda nadir genetik varyantın yorumlanması için en temel araçlardan biri konumundadır. Bu otomatik aracın yaptığı tahmin; değişimi belirleyen birkaç sekansa, filogenetik ve yapısal özelliklere dayanmaktadır. Bir proteindeki belirli bir aminoasit değişimi için PolyPhen2, çeşitli sekansları ve değişim bölgesinin yapısal özelliklerini ortaya çıkararak bu değişimi olasılıklar ile sınıflandırmaktadır [40].

2.8.5. GERP

Tahmin araçları arasında sıkça kullanılan araçlardan bir diğeri ise Genomik Evrimsel Derece Belirleme (GERP) aracıdır. GERP skoru, nötr beklentiye kıyasla çoklu tür sekans hizalamada değişim sayısında azalma ile tarif edilmektedir. Örneğin GERP skorunun 4 olması, belirli bir bölgede, filogeni boyunca evrimin nötr derecesine dayalı olarak beklenenden 4 değişim daha az olması manasına gelir. Yani GERP skoru birçok türün arasında sekans korunmasının ölçümü demektir. Ancak GERP skoru genellikle evrimsel genomik çalışmalarda, bir tür içinde segrege olan mutasyonlar üzerinde seçilimin gücünü ölçmek için kullanılır. Bu uygulamalarda, birçok türde yüksek oranda korunan bölgelerde ortaya çıkan mutasyonların zararlı olduğu dolayısıyla bir tür içindeki genetik yüke katkıda bulunduğu varsayılır. Niceliksel olarak bir türdeki her segrege olan bölge için GERP skoru, o bölgedeki segrege olan mutasyonlara denk gelir. Örneğin Schubert ve arkadaşları tarafından yapılan bir çalışmada, yabani ve evcil atlarda zararlı mutasyonların modeli incelenmiştir. Evcilleştirilmiş atlarda GERP skor yükünde bir artış olduğunu gözlemleyerek evcilleştirmenin zararlı mutasyonlarda bir artışa neden olduğunu ileri sürmüşlerdir [41]. Henn ve arkadaşları ise, GERP aracını kullanarak insanlardaki aminoasit değişimlerine sebep olan mutasyonların uyumluluk etkisini incelemişler ve 4-6 arası GERP skorlarını çok zararlı etkiler olarak tanımlamışlardır. Öte yandan bu zararlı alellerin Afrikalı olmayan topluluklarda daha fazla olduğunu gözlemlemişlerdir [42].

2.8.6. MutationTaster

İnsan genetiği çalışmalarının ana amaçlarından biri kalıtsal hastalığa sahip hastalardaki patojenik DNA dizi değişimlerini belirlemektir. Yeni nesil dizileme teknolojisi yüzlerce aday genin dizilenmesine, bağlantı analizlerine ve tüm ekzom çalışmalarına olanak sağlamaktadır. Bu da potansiyel olarak birçok hastalık yapıcı değişimin belirlenmesine yol açmaktadır. Otomatik varyant analizleri zaman ve kaynak

tasarrufu sağlayarak derinlemesine analiz yapma imkanı sunmaktadır. Ancak mevcut değerlendirme araçları sadece aminoasit değişimlerinin sonucunu tahmin ederken binlerce sorguyu makul bir sürede yapamamaktadır. Yüksek çıktılı dizileme verilerini incelemenin zorluğunu aşmak adına geliştirilen araçlardan biri de MutationTaster aracıdır. Bu araç potansiyel hastalık yapıcı DNA dizi değişimlerini hızlı bir şekilde tahmin eden web tabanlı bir uygulamadır. MutationTaster farklı biyomedikal veritabanlarından bilgileri birleştirerek mevcut analiz araçlarını kullanmaktadır. Analizler; evrimsel korunmayı, katlanma bölgesi değişimlerini, protein özellik kaybını ve mRNA miktarını etkileyen değişimleri içermektedir. Test sonuçları, hastalık potansiyelini tahmin eden Bayer sınıflandırıcısı tarafından değerlendirilmektedir. Tipik bir sorgu 0.3 saniyeden daha az bir sürede tamamlanır. Değişimin doğasına bağlı olarak MutationTaster, sessiz eş anlamlı veya intronik değişimdeki model, tek bir aminoasidi etkileyen değişimlerdeki model ya da aminoasit dizisinde kompleks farklılıklara sebep olan değişimdeki model olmak üzere üç farklı modelden bir tanesini seçmektedir [43].

2.8.7. SIFT

Tek nükleotid değişimleri (SNP'ler), insan genomunda hastalıkla ilişkili bir bölgenin tespiti için kullanılan belirteçlerdir. Kodlayan ve düzenleyici bölgelerdeki SNP'lerin hastalıkla ilişkili olduğu tahmin edilmektedir. Bir proteindeki aminoasit değişikliğine sebep olan eş anlamlı olmayan SNP'ler özellikle ön plana çıkmaktadır. Çünkü kalıtsal hastalıklara neden olan gen problemlerinin birçoğu aminoasit değişikliğinden kaynaklanmaktadır. SIFT (tolere edilenden tolere edilmeyene gruplama), bir aminoasit değişiminin protein fonksiyonunu, dolayısıyla fenotipi etkileyip etkilemediğini tahmin eden bir programdır. SIFT insan varyant veritabanlarını kullanarak nötr bir polimorfizm ile hastalığa sebep olan mutasyonları ayırabilir. SIFT programı bir protein ailesindeki önemli aminoasitlerin korunması gerektiğini varsayarak iyi korunmuş pozisyonlarda gerçekleşen değişimleri zararlı mutasyonlar olarak belirlemektedir. Örneğin bir protein sekansında sadece izolösin aminoasidi varsa, SIFT izolösinin protein fonksiyonu için önemli olduğunu varsayarak bu noktadaki herhangi bir aminoasit değişikliğini zararlı olarak bildirmektedir. Öte yandan bir protein dizisindeki bir pozisyonda hidrofobik olan izolösin, valin ve lösin aminoasitleri varsa, SIFT bu pozisyon için hidrofobik özelliği önemli olarak varsayarak bu pozisyonda meydana gelen değişim sonucunda herhangi bir hidrofobik aminoasit gelmesi durumunda SIFT bunu tolere edilen

bir değişim olarak görmektedir. Ancak yüklü veya polar bir aminoasit gelmesi durumunda bu değişimi proteinin fonksiyonunu etkileyecek bir değişiklik olarak raporlayacaktır. Bir aminoasit değişiminin proteinin fonksiyounu etkileyip etkilemediğini tahmin etmek için SIFT, değişimin olduğu pozisyonu ve yeni gelen aminoasit tipini incelemektedir. Bir protein dizisi verildiğinde SIFT, ilişkili proteinleri seçer ve bu proteinlerin dizilerini hizalar. Hizalamadaki her aminoasit pozisyonuna dayanarak SIFT, bir pozisyonda yer alan aminoasidin tolere edilip edilmeyeceğini her pozisyonda en sık var olan aminoaside bakarak bir hesaplama yapar. Eğer normalize edilen değer eşik değerin altında ise değişim zararlı olarak değerlendirilir [44].

2.8.8. REVEL

REVEL, nadir yanlış anlamlı varyantların patojenitesini tahmin etmeye yönelik bir sistemdir. Nadir varyantlar, dizileme çalışmalarında keşfedilen önemi bilinmeyen varyantların çok büyük bir kısmını kapsamaktadır. REVEL mevcut yöntemlerle kıyaslandığında; özellikle %3'ün altında alel frekansına sahip, nadir görülen nötr yanlış anlamlı varyantlardan hastalık yapıcı mutasyonları ayırdetmek için en iyi performansa sahip araçtır. Klinisyenler ve araştırmacılar tarafından kullanımı kolaylaştırmak için, insan genomundaki tüm potansiyel eşanlamlı olmayan tek nükleotid varyantlarını barındıran dbNSFP v2.7 veritabanındaki tüm yanlış anlamlı varyantların REVEL skorları hesaplanmıştır. Bundan ötürü REVEL, nadir genetik varyantların yüksek doğrulukta yorumlanması için gerekli olan bir patojenite tahmin araç ihtiyacını karşılamaktadır. REVEL yönteminin birkaç güçlü yönü bulunmaktadır. İlk olarak REVEL, gelecekte yeni nesil dizileme çalışmalarıyla keşfedilecek yeni varyantlara benzeyen hastalıklar ve nötr varyantların belirlenmesi için test edilmiş ve geliştirilmiştir. Bu çalışmalar düşük alel frekansına sahip varyantları ve önceden keşfedilmiş varyantlara nispeten daha az etkiye sahip varyantları içermektedir. REVEL'deki nötr varyantların alel frekansları, özellikle nadir varyantları yorumlarken performansı geliştirmek için %0.1 ile %1 arasında sınırlandırılmıştır. İkinci olarak REVEL; MutPred ve VEST gibi önemli özellikleri olan bireysel tahmin araçlarını kapsamaktadır. Daha önce MutPred skorları kullanılabilir halde olmamasına rağmen günümüzde dbNSFP v2.7 veritabanındaki tüm yanlış anlamlı mutasyonlar için hesaplanmış durumdadır [45].

2.8.9. DANN

Hastalıklara sebep olan varyantları belirlemek oldukça zor olabilmektedir. Varyantların çoğu, genomun kodlanan bölgelerinde rolü olduğu düşünülen ancak tam manasıyla bilinmeyen kodlanmayan bölgelerdeki varyantlardır. Birçok anotasyon metodu sadece protein kodlayan bölgelerdeki varyantları anote edebilmektedir. Dolayısıyla genomun %98'lik kısmı anote edilmemiş durumdadır. CADD metodunun genomun hem kodlanan hem de kodlanmayan kısımlarını anote edebildiği belirtilmişti. CADD, simüle genetik varyantlardan gözlemlenen genetik varyantları ayırt etmek için bir doğrusal çekirdek SVM (Destek Vektör Makinesi) kullanmaktadır. Gözlemlenen genetik varyantlar insan genomları arasındaki farklılıklardan türetilmiştir. Bu varyantlar insan şempanze ortak atalarından çıkarılmıştır. Doğal seçilim etkisi nedeniyle gözlemlenen varyantlar zararlı mutasyonları içermemektedir. Ancak simüle varyantlar bu zararlı varyantları fazlasıyla içermektedir. CADD'ın SVM'si verinin sadece doğrusal gösterimini öğrenebildiği için performansı sınırlanmaktadır. Bu problemi çözmek için DANN (Nöral Ağı kullanan genetik varyantların Zararlı Anotasyonu) adı verilen bir DNN algoritması geliştirilmiştir. Bir DNN, girdi ve çıktı katmanları arasında birimlerin birçok gizli katmanıyla oluşturulmuş yapay sinir ağıdır [46].

2.9. Protein Modelleme ve Simülasyon

Genetik çalışmalar sonucunda elde edilen varyantların herhangi bir fenotipik değişime sebep olup olmadığını anlamak için belirli yollar izlenir. Elde edilen varyantların bir proteinin fonksiyonel durumunu değiştirme ihtimalini ölçebilen birçok web tabanlı in silico araçlar bulunmaktadır. Bu araçlarla tek nükleotid değişiminin lokasyonuna göre patojen bir etkiye sahip olup olmadığı veya iyi huylu denilebilecek bir varyant olup olmadığını tespit edilebilir. Ancak bu araçların potansiyel olarak zararlı diye sınıflandırdığı bir varyantın proteinin yapısına nasıl etki ettiğini göreceğimiz tahmin araçları protein modelleme ve simülasyon araçları ile daha iyi anlaşılabilir.

2.9.1. Moleküler Dinamik Simülasyon

Bir biyomolekülün atomik seviyedeki yapısını ve fonksiyonunu bilmek oldukça değerlidir. Modelleme çalışmaları atomik seviyede bir biyomolekülün yapısına ve fonksiyonuna dair bilgi edinmemize imkan vermektedir. Aynı zamanda biyomoleküller atomik seviyede hareket halindedir. Bu düzeyde bir biyomolekülün hareketine dair bilgi edinmek için ise simülasyon çalışmaları önem arz etmektedir. Bir proteinin veya başka

bir biyomolekülün her atomunun zaman içinde nasıl hareket edeceğini, Newton'un hareket yasaları uyarınca hesaplayan ve atomlar arası etkileşimi bir fizik modeline dayalı olarak gerçekleştiren hesaplamalı tekniğe Moleküler Dinamik (MD) simülasyonları denir [47]. Bu hesaplamalı yöntem, zaman içinde etkileşen parçacıklardan oluşan büyük sistemlerin davranışını modellemek ve incelemek için klasik mekaniğin ilkelerini kullanır.

Moleküler dinamik simülasyonlar; proteinin yapısındaki konformasyonel değişim, proteine ligand bağlanması ve protein katlanması dahil olmak üzere birçok moleküler süreci simüle edebilir. Ayrıca bu simülasyonlar, bir protein yapısındaki tüm atomlara ait pozisyonları, femtosaniye (10^{-15} saniye) zamansal çözünürlükte ortaya koyabilir. Öte yandan bu simülasyonlar; biyomoleküllerin mutasyon, fosforilasyon, protonasyon veya bir ligandın eklenmesi ile çıkarılması gibi değişimlere atomik düzeyde nasıl tepki vereceğini tahmin eder. MD simülasyonları genellikle x-ışını kristalografisi, kriyo-elektron mikroskobu (kriyo-EM), nükleer manyetik rezonans (NMR), elektron paramanyetik rezonans (EPR) ve Förster rezonansı enerji transferi (FRET) dahil olmak üzere çok çeşitli deneysel yapısal biyoloji teknikleriyle birlikte kullanılır [48].

Son yıllarda bilimsel çalışmalarda MD'ye ilgi oldukça artmıştır. Bu ilginin artmasına sebep olan birçok gelişme olmakla birlikte özellikle sinirbilim alanında yapılan çalışmalarda birçok molekülün deneysel olarak yapıları çok sayıda araştırmacı tarafından incelenmektedir. Örneğin iyon kanalları, nörotrasmitter taşıyıcıları ve G-proteine bağlı reseptörler (GPCR) gibi zar proteinleri en çok çalışılan moleküllerdir. Zar proteinlerinin yapılarını tayin etmede kristalografi metodu yeterli değildi. Ancak son zamanlarda kristalografideki gelişmelerle birçok yapı aydınlatılmıştır. Kriyo-EM tekniğindeki gelişmeler de ciddi katkılar sunmuştur. Aynı zamanda bu gelişmeler 2017 yılında Nobel Ödülüne layık görülmüştür [49]. Deneysel yapıların aydınlatılması, MD simülasyonları için bir başlangıç noktası olmuştur. Simülasyonlar birçok yapısal soruya cevap vermek için oldukça iyi bir zemin hazırlamıştır. Proteinlerin fonksiyonun nasıl gerçekleştiği, belirli koşullarda proteinlerin neden patolojik rol edindiği, yapı temelli en iyi ilacın nasıl dizayn edileceği ve protein fonksiyon çalışmaları için en iyi protein dizaynının nasıl yapılacağı gibi sorulara net bir cevap bulunmasını, simülasyonlar sağlayabilir [48].

2.9.2. Moleküler Dinamik Simülasyonun Sunduğu Faydalar

MD simülasyonlara son yıllarda artan ilginin sebeplerinden biri simülasyonların son birkaç yılda daha fazla erişlebilir ve daha güçlü hale gelmesidir. Yakın zamana kadar MD simülasyonu için süper bilgisayarlar gerekliydi. Ancak günümüzde iyi bir grafik işlem birimine (GPU) sahip lokal bilgisayarlarla daha uygun maliyetlere de simülasyonlar yapılabilir hale geldi [50]. Yeterli uzmanlığı olmayan kullanıcılar için geliştirilen yazılımlarla simülasyon adımlarını tamamlamak da mümkün oldu. Her ne kadar simülasyonlar fiziksel modellere göre yaklaşık modeller olarak kabul edilse de son yıllardaki gelişmelerle simülasyonlar çok iyi seviyede doğruluk değerine ulaştı [48].

MD simülasyonlarının çalışma mekanizması; bir biyomolekül sisteminde yer alan (su ortamı veya çift lipid tabakası içindeki bir protein) her bir atoma diğer atomlar tarafından uygulanan kuvvetlerin zamana bağlı olarak tahmin edilmesine dayanmaktadır. Simülasyon, belirli bir zaman aralığında sistemin atomik düzeyde konum ve hızın konfigrasyonunu gösteren üç boyutlu bir film gibidir [48].

2.9.3. Moleküler Dinamik Simülasyonun Deneysel Çalışmalara Etkisi

MD simülasyonların deneysel uygulamaya göre birçok avantajı bulunmaktadır. Örneğin simülasyonda her atomun zamana göre her noktadaki konumu ve hareketi yakalanabilir. Deneysel yapıda bu mümkün değildir. Ayrıca simülasyonun koşulları kesin olarak bilinebilir ve kontrol edilebilir. Bir proteine hangi ligandların bağlandığı, proteinin herhangi bir mutasyona sahip olup olmadığı, post translasyonel modifikasyonun olup olmadığı, çevresindeki başka moleküllerin ilk konformasyonu, protonasyon durumu, sıcaklık ve voltaj etkisi gibi etkenler karşılaştırılarak çeşitli moleküler bozulmaların etkileri belirlenebilir [48].

MD simülasyonların deneysel yapı belirleme yöntemlerine karşı başka bir üstünlüğü de bir biyomolekülün çeşitli bölgelerinin hareketini ve esnekliğini gösterebilmesidir. X-ray kristalografisi veya kryo-EM gibi deneysel yapı belirleme teknikleri genelde ortalama bir yapı ortaya koyabilir. Ancak simülasyon ile bir molekülün çeşitli bölgelerinin dengede ne kadar hareket ettiği ve ne tür yapısal dalgalanmalara maruz kaldığı da ölçülebilir. Öte yandan simülasyonlar, bir proteinin fonksiyonu için ve bir ligandın bağlanması için gerekli olan su moleküllerinin veya tuz iyonlarının dinamik davranışını da ortaya koyabilir. Yani simülasyon ile sadece proteinin yapısı değil aynı

zamanda proteinin yer aldığı ortamın da maruz kalınan değişikliklere nasıl cevap vereceği belirlenebilir [51,52].

MD simülasyon ile modellenen bir yapının doğruluğu test edilebilir ve bu yapının iyileşitirilmesi de sağlanabilir. Bir kristal yapı, kristal örgü paketlenmesinden ötürü veya bir zar proteini, çift lipid tabakasının olmayışından ötürü artefaktlardan etkilenebilir. Ancak bu artefaktların düzeltilmesi için öncelikle kristal yapının uygun bir çözücü ortamda yer alarak yapısının daha uygun bir konformasyona dönüşmesi simüle edilebilir [53].

MD simülasyonlar ile bir biyomolekül sisteminin çeşitli bozulmalara nasıl tepki vereceği belirlenebilmektedir. Örneğin yapısı deneysel olarak belirlenen bir proteinin, bu proteine bağlı bir ligandın çıkarılması durumunda proteininin konformasyonun nasıl değiştiği simüle edilebilir. Ayrıca bağlı ligandın çıkarılıp deneysel yapıda yer almayan başka bir ligandın eklenmesinin konformasyon değişikliğine olan etkisi gözlenebilir [54]. Öte yandan daha fazla nadir genetik hastalık çalışmalarında yaygın olarak görülen bir veya birkaç aminoasidin mutasyona uğraması ile mutasyon ya da mutasyonların işlevsel etkisi ve yabanıl tip ile arasındaki fonksiyon farkı simülasyon ile incelenebilir [55]. Bir aminoasidin fosforillenmesi veya farklı tipte bir translasyon sonrası modifikasyonu simüle edilmesiyle tepki gözlenebilir [56]. Yine başka bir simülasyon ile atomlara dış kuvvet uygulanarak transmembran voltaj ve mekanik zorlamanın etkisi incelenebilir [57].

MD simülasyonlar deneysel çalışmalara göre bazı faydalar üretirken öte taraftan deneysel çalışmaların daha verimli planlanmasına da katkıda bulunmaktadır. Örneğin bir protein veya ligand tasarımında çok fazla aday molekül bulunurken simülasyon çalışmaları ile deneysel çalışmada test edilecek aday sayısı oldukça az sayıya indirilebilir. Dolayısıyla simülasyon çalışmaları, deneysel çalışmalar için bir filtre görevi görür [58].

2.9.4. Moleküler Dinamik Simülasyon için Araç ve Yöntemler

MD simülasyonları için gereken donanım ve yazılımlar her geçen gün daha verimli, daha ulaşılabilir ve daha az maliyetli hale gelmektedir. Daha önce sadece süper bilgisayarlarda MD simülasyonlar yapmak mümkünken daha sonra geleneksel merkezi işlem birimlerinde (CPU) ve daha az maliyetli grafik işlem birimlerinde (GPU) yapılmaya başlanmıştır. Üç donanım türü de günümüzde kullanılırken herbirinin diğerlerine göre üstünlüğü bulunmaktadır. Örneğin süper bilgisayarlar diğer donanım türlerine göre daha

hızlı işlem yapabilmektedir. GPUlar ise daha uygun maliyetle hızlı sonuçlar verdiği için tercih edilen donanım türlerinden biridir. CPUlar ise en yaygın bulunan donanım türü olduğundan birçok kullanıcı tarafından tercih edilmektedir. Simülasyon için gerekli araçlardan biri de kuvvet alanı platformlarıdır. En çok kullanılan kuvvet alanları CHARMM, AMBER ve OPLS olarak sıralanabilir. [59].

MD simülasyonları için en çok kullanılan üç kuvvet alanı benzer işlevleri yerine getirse de herbirinin diğerine göre avantajları ve dezavantajları bulunmaktadır. Örneğin CHARMM36m ve tamamlayıcısı olan CHARMM Genel Kuvvet Alanı (CGenFF); proteinler, lipidler ve ilaç benzeri ligandlar için optimize edilmiş ve doğrulanmış parametrelere sahiptir. OPLS kuvvet alanı ise muhtemelen ligand parametrelerini en kapsamlı şekilde optimize eden programdır [59].

MD simülasyonları için gerekli yazılım programları için de GROMACS, NAMD, AMBER, CHARMM, Desmond ve OpenMM yazılımları bulunmaktadır. AMBER ve CHARMM yazılımları, kuvvet alanı olan AMBER ve CHARMM programları ile karıştırılmamalıdır. Tüm simülasyon yazılımları birçok kuvvet alanını desteklemektedir. Bu yazılımlar benzer hesaplamaları yapsa da en verimli hangi donanıma uygun oldukları ve belli başlı desteklenen özelliklerde en verimli hangisinde çalıştıkları noktasında farklılık gösterirler [60].

MD simülasyonu başlatmadan önce moleküler sistemi hazırlamak için birkaç adımın atılması gerekmektedir. Örneğin kristal yapılarda genellikle konumları tam olarak belirlenemeyen, hidrojen atomları eksiktir. Hidrojen atomlarını sisteme dahil etmek gerekir. Simulasyon ortamının pH sına göre, öncelikle yapıda yer alan amino asitlerin protonasyon dereceleri tespit edilir, ardından gerekli protonlar proteine eklenir. Simulasyon kutusuna yerleştirilen protein, eğer sitoplazmik protein ise su eklenerek çözülür. Eğer bir zar proteini ise model membran lipitleri ortama eklenir. Sistemin net yükü bulunuyorsa, fizyolojik tuz konsantrasyonunda (örneğin 0.15 M NaCl) ortama iyonlar eklenerek, ortam nötralleştirilir. Çok kullanılan simülasyon yazılımlarının birçoğu bu hazırlık aşamalarını barındırmaktadır [61].

MD simülasyon yazılımlarını tercih ederken önemli etkenlerden biri de sonuçları analiz etme yöntemidir. MD simülasyonlarının analizi birçok sebepten ötürü oldukça zorlayıcı olabilir. Bu simülasyonlar oldukça fazla miktarda veri üretmektedir. Örneğin tipik bir simülasyonda 100.000 atomun konumu ve hızı, milyarlarca zaman aralığına bağlı

olarak takip edilmektedir. Oldukça büyük bir veriyi barındıran bu simülasyonda önemli ve en ilgili biyolojik kısmı tespit etmek bir hayli zordur. Dolayısıyla bu adımları en verimli şekilde yapabilmek için tercih edilen yazılım programı önem arz etmektedir [48].

MD simülasyonlarından en yararlı ve verimli verileri çıkarmanın yolu ancak deneysel uygulamalardan elde edilen bilgiler ışığında mümkündür. Analiz süreci, genelikle moleküler düzeyde canlandırma ve niceliksel analiz metotolarının kombinasyonunu gerektirmektedir [60].

MD simülasyonları her ne kadar protein modellemede ileri bir seviyeye geldiyse de bu simülasyonları sınırlayan durumlar da göz ardı edilmemelidir. Örneğin, MD'de kullanılan kuvvet alanlar her ne kadar iyi düzeyde geliştirilmiş olsa da nihayetinde yaklaşık değerlerdir [62]. Ayrıca MD simülasyonlar yapıldığı sırada kovalent bağlar kırılmaz ya da yeniden oluşmaz. Yani titre edilebilen amino asit rezidülerin protonasyon durumu sabittir. Bu durum simülasyonun başında dikkate alınarak düzenlenmelidir. Sabit pH'lı simülasyonların yapıldığı durumlarda bu durum geçerli değildir [63]. Öte yandan doğru bir simülasyon, doğru bir deneysel protein yapısına veya iyi bir homolog modele bağlıdır. Dolayısıyla doğru bir simülasyon çalışmasının, deneysel çalışmaların varlığından oldukça etkilendiği söylenebilir.

2.9.5. Protein Yapı Modellemede Devrimsel Bir Araç: AlphaFold

İnsan genom projesinin tamamlanmasıyla yeni çalışmaların önü açılmış oldu. Protein kodlayan genlerden yola çıkarak proteinlerin yapısını keşfetmek de araştırmacılar için oldukça meraklı bir serüven oldu. Nihayetinde binlerce bireysel çalışma ile büyük projeler kapsamında yapılan genomik ve proteomik çalışmalar yaklaşık 50 bin insan proteinin yapısını ortaya koydu. Bu çalışmalar sayesinde *Homo Sapeiens*, Protein Data Banka (PDB) platformunda en iyi gösterilen türdür. Bu kadar yoğun çalışmaya rağmen insan proteinlerinin yalnızca %35'i PDB'de haritalanmıştır [64]. Bunların çoğu da sadece proteinin bir bölgesini içeren yapılardır. On yıllarca yapılan birçok çalışmanın neticesinde insan protein dizilerindeki rezidülerin yalnızca %17'si deneysel olarak belirlenebilmiştir [65]. Öte yandan protein yapısı tahmin araçları; protein aileleri, spesifik fonksiyonel sınıflar, proteomlar içindeki domainler veya tüm zinciri ortaya koymada oldukça yararlı olmuştur. SWISS-MODEL veritabanı, Genome3D ve ModBase projeleri çok sayıda protein yapısının ortaya konmasına olanak sağlamış, aynı zamanda birçok araştırmacının ücretsiz olarak kullanımı için imkan sunmuştur [66-68]. Protein yapılarını ortaya koyan

bu tahmin araçlarının haricinde protein dizaynı, fonksiyon anotasyonu, hastalık tahmini ve domain tanımlanması konularında da biyoinformatik araçlar geliştirilmiştir. Ancak protein yapılarını hassas bir ölçüde tahim eden ve büyük ölçekte tahmin eden araştırmalar son yıllarda dikkat çekici bir seviyeye gelmiştir.

Protein yapı tahimininde önemli bir çalışma olan Critical Assessment of protein Structure Prediction (CASP), yani Protein Yapı Tahmininde Önemli Değerlendirme çalışması son versiyonu ile ciddi bir mesafe kat etmiştir. CASP14 versiyonu, AlphaFold2 adıyla tanıtıldı. Bu versiyon bir önceki sürüm olan CASP13'e göre oldukça farklı bir model kullanmaktadır. Özellikle yüksek doğruluk hassasiyeti açısından önceki sürümlere göre bir üstünlüğü bulunmaktadır. En gelişmiş makine öğrenme metodu ile çalışan AlphaFold insan proteoumunun neredeyse tamamının yapısını (%98,5) hassas bir ölçüde tahmin edebilmektedir. Tüm tahmin modellerine (https://alphafold.ebi.ac.uk/) internet adresinden ulaşılabilir. AlphaFold platformu, Google tarafından 2014 yılında satın alınan DeepMind yapay zeka geliştirme firması tarafından geliştirilmiştir. Bu platformun şuan internet üzerinden sağlayıcısı (hosting) Avrupa Biyoinformatik Enstitüsü'dür [69].

İki yapı arasındaki benzerliği ölçmek için yaygın olarak kullanılan yöntemlerden biri ortalama kare sapma (RMSD) değeridir. AlphaFold modelleri ile en iyi performans gösteren yöntem ile kıyas edildiğinde, 2.80 Å (angstrom) RMSD değerine kıyasla Alphafold modelleri 0.96 Å RMSD'lik bir medyan omurga doğruluğuna sahiptir. Öte yandan, AlphaFold modelleri protein omurgası tahmini doğru olduğunda kalıntı yan zincirlerin konumunu tahmin etmede yüksek bir doğruluğa sahiptir [70].

AlphaFold veri tabanı, 2700 kalıntıdan daha uzun proteinler için modeller içermemektedir. Bu nedenle, konjenital kardiyomiyopati [71] ve kas distrofisinden sorumlu ana genler olan Titin ve Distrofin tarafından kodlananlar gibi 207 büyük (kalıntı aralığı 2701-34350), biyolojik olarak önemli insan proteinleri için hiçbir model mevcut değildir [72].

3. GEREÇ VE YÖNTEM

3.1. Çalışma Grubunun Oluşturulması

Çalışmamıza İstanbul Üniversitesi İstanbul Tıp Fakültesi, Tıbbi Genetik Anabilim Dalı'na genetik etiyolojinin aydınlatılması amaçlı başvuran, daha önce tüm ekzom analizi yapılıp klinikle ilişkili varyant saptanmamış nöromotor gelişim geriliği olan ve onam formu alınan 3 aile dahil edildi. Etkilenmiş ve sağlıklı kardeşlerin yanı sıra ebeveynler de çalışmaya dahil edildi. Böylece çalışmamıza 3 indeks olgu, 2 etkilenmiş ve 12 sağlıklı birey olmak üzere toplam 17 olgu dahil edildi. 3 indeks olgunun yanak mukozasından (buccal smear) örnek alınarak DNA elde edildi.

Çalışmanın yapılması için etik kurul belgeleri hazırlanarak İstanbul Üniversitesi İstanbul Tıp Fakültesi Klinik Araştırmalar Etik Kurulu'na başvuruldu. Kurul 18/12/2020 tarih ve 31 sayılı toplantısında görüşerek çalışmamızı etik yönden uygun buldu.

Etik kurul onayının ardından İstanbul Üniversitesi Bilimsel Araştırma Projeleri Koordinasyon Birimi'ne doktora tez projesi başvurusu yapıldı. 23.02.2021 tarihinde BAP Komisyonu tarafından tez projesi onaylandı ve TDK-2021-37592 kodlu, "Nadir genetik sendromlu bir olgu ve ailesinde tüm ekzom ve/veya tüm genom analizleri ile genetik etiyolojinin belirlenmesi" başlıklı proje olarak yürürlüğe girdi.

3.2. Klinik Değerlendirmeler

Olguların fiziki muayenesi yapıldı ve ayrıntılı anamnez alınarak en az üç kuşak olacak şekilde aile ağaçları çizildi. Olgulara ait biyokimya ve enzim parametreleri, beyin bilgisayarlı tomografi ve manyetik rezonans sonuçları, karyotip analizi ve tüm ekzom dizileme sonuçları kayıt altına alındı. Tüm olgulara ait tüm ekzom dizileme verileri tekrar incelenip değerlendirildi.

3.3. Sentromerik İnstabilite Testi

İstanbul Üniversitesi İstanbul Tıp Fakültesi, Tıbbi Genetik Anabilim Dalı Sitogenetik Laboratuvar'ında olgularımızdan birine sentromerik instabilite testi yapmak için kan örneği alınarak lenfosit kültürü yapıldı. Ancak üreme olmadığından diğer adımlara geçilemedi. Ayrıca aynı olgudan tekrar örnek alınarak kromozom kırıklarını tespit etmek amacıyla DEB testi yapılması planlandı. Fakat alınan örnekten lenfosit kültürü üretememesinden ötürü ileri adımlara geçilemedi.

3.4. Tüm Genom Dizilemede Moleküler Çalışmalar

3.4.1. Yanak Mukozasından Genomik DNA Eldesi

Tüm genom analizi yapılacak hastaların yanak mukozalarından Nebula Genomics firmasının sağladığı pamuklu çubukla (swab) örnekler alındı. Koruyucu sıvının bulunduğu tüpe koyularak oda sıcaklığında muhafaza edildi. Genomik DNA izolasyonu Genomik DNA Purifikasyon Kiti (Monarch®-NEB) aracılığıyla gerçekleştirildi. Genomik DNA izolasyonu, kit içeriğinde bulunan protokole göre aşağıda belirtilen basamaklarla gerçekleştirildi.

Başlamadan önce:

- RNase A ve Proteinase K -20°C'de saklandı.
- Monarch gDNA Wash Buffer konsantresine şişe etiketinde belirtildiği gibi etanol (≥ %95) eklendi.
- Soğuk PBS kullanıldı.
- Mikrosantrifüj önceden 4°C'ye soğutuldu.
- Numune lizizi için bir termal karıştırıcı ve bir ısıtma bloğu 56°C'ye ayarlandı.
- Elüsyona hazırlanmak için bir ısıtma bloğu 60°C'ye ayarlandı. Uygun hacimde elüsyon tamponu 60°C'ye (numune başına 35–100 µl) önceden ısıtıldı. Sıcaklıklar genellikle cihazda belirtilenden daha düşük olduğundan sıcaklık doğrulandı.
- Bir liziz numunesiyle bir kolon bir kereden fazla yüklenmedi. Matrisin liziz numunesine aşırı maruz kalması, zarın genişlemesine ve yerinden çıkmasına neden olabilir.

Protokol :

1. 1.5 ml reaksiyon tüpüne 1 ml soğuk PBS eklendi.
2. Bukkal swab hücrelerini içeren toplama parçasını kesildi ve PBS ile içeren tüpe eklendi.
3. Hücre materyalini 10 saniye vortekslendi, ardından swab PBS' den çıkarıldı.
4. Hücre materyali 4 °C'de 30 saniye 2000xg santrifüjlendi.

5. Süpernatan tüp içerisinde yaklaşık 100 µl kalana kadar toplandı ve atıldı.
6. Karışım vorteks ile süspanse edildi.
7. Karışım üzerine 10 µl Proteinaz K ve 3 µl RNase A eklendi ve vortekslendi.
8. Tüp içerisine 100 µl Cell Lysis Buffer eklendi, kısa süre tam hızda vortekslendi ve 30 dakika 56°C'de inkübe edildi.
9. İnkübasyon sonrasında örnek üzerine 400 µl gDNA Binding Buffer eklendi ve 5-10 saniye vorteks ile iyice karıştırıldı.
10. 600 µl lizat/bağlama tamponu karışımı spin kolona yüklendi.
11. Spin kolon öncelikle 1.000 x g' de 3 dakika sonra maksimum hızda (> 12.000 x g) 1 dakika santrifüj edildi.
12. Kolon, yeni bir toplama tüpüne aktarıldı ve üzerine 500 µl Yıkama Tamponu eklendi.
13. Tüp, 1 dakika maksimum hızda (12.000 x g) santrifüj edildi. Toplama tüpündeki süpernatan uzaklaştırıldı.
14. Kolona tekrar 500 µl Yıkama Tamponu eklendi ve örnek 1 dakika maksimum hızda (>12.000 x g) santrifüjlendi.
15. Kolon yeni 1,5 ml'lik bir mikrofüj tüpüne aktarıldı. Üzerine 100 µl önceden ısıtılmış (60°C) gDNA Elüsyon Tamponu eklendi ve maksimum hızda (> 12.000 x g) 1 dakika santrifüjlendi.

3.4.2. Tüm Genom Dizileme Analizi Islak Laboratuvar Çalışmaları

Hastalardan mukozal swab yöntemi ile alınan örneklere Nebula Genomics tarafından MGI DNBSEQ-T7RS platformunda çalışma prensipleri aşağıda detaylı olarak verilen tüm genom dizileme işlemi uygulandı [73].

DNBSEQ-T7RS Yüksek Verimli Dizileme Seti, DNBSEQ-T7RS üzerinde DNA veya RNA dizilemesi için özel olarak tasarlanmıştır. Bu dizileme seti yalnızca bilimsel araştırmalar için kullanılmak üzere tasarlanmıştır.

Bu dizileme platformu, DNBSEQTM teknolojisini kullanır. Bir dizileme çalışması, bir DNA bölgesinin hibridizasyonu ile başlar. Ardından kombinatoryal prob çengel dizileme (cPAS) kimyası kullanılarak DNA Nanoball'a (DNB) bir floresan prob

eklenir. Son olarak, yüksek çözünürlüklü görüntüleme sistemi floresan sinyalini yakalar. Optik sinyalin dijital olarak işlenmesinden sonra dizileme platformu yüksek kaliteli ve yüksek doğrulukta dizileme bilgisi üretir.

Dizileme çalışması sırasında, kontrol yazılımı otomatik olarak temel arama analiz yazılımını çalıştırır ve ikincil analiz için ham dizileme veri çıktıları sağlar.

Nadir bir varyant arandığında dizileme çalışmasında, dizileme döngülerinin sayısı dizileme okuma uzunluğuna bağlıdır. Örneğin, bir PE100 döngü çalışması, toplam 200 (2 × 100) döngü için her uçtan 100 döngü okuması gerçekleştirir. Ek dizileme çalışmasının sonunda, gerekirse ekstra 10 döngü indeks okuması gerçekleştirilebilir.

Tablo 3.1'de 100X derinlikte bir tüm genom dizileme işlemi için gerekli olan kitlerin bilgisi, kullanım miktarı, taşıma sıcaklığı ve saklama sıcaklığı bilgileri yer almaktadır. Bu çalışma için DNBSEQ-T7RS platformunda kullanılacak kit FCL PE100 kitidir. Problar için kullanılacak kit olan CPAS barkod reaksiyon kitinin bilgileri Tablo 3.2'de yer almaktadır.

Tablo 3-1: DNBSEQ-T7RS yüksek veri çıktılı dizileme seti (FCL PE100)

Sekanslama Kiti	İçerik	Miktar	Taşıma Sıcaklığı	Stok Sıcaklığı
DNBSEQ-T7RS Sequencing Flow Cell Katalog numarası: 1000016269	Dizileme Akış Hücresi (T7 FCL)	1 EA	0℃~30℃	0℃~30℃
DNBSEQ DNB Make Reagent Kit Katalog Numarası: 1000016115	Düşük TE Tamponu	960 μL	-15℃ ve altı sıcaklıklarda	-25℃~-15℃
	DNB Yapma Tamponu	400 μL		
	DNB Yapma Enzim Mix I	800 μL		
	DNB Yapma Enzim Mix II(LC)	80 μL		
	DNB Durdurma Reaks. Tamponu	400 μL		
DNBSEQ-T7RS DNB Load Reagent Kit Katalog Numarası: 1000016114	DNB Yükleme Tamponu I	300 μL	ve altı sıcaklıklarda	-25℃~-15℃
	DNB Yükleme Tamponu II	150 μL.		
	Boş mikrotüpler 0.5mL	1tüp		
	Yükleme Sonrası Plate	1 EA		
DNBSEQ-T7RS High-throughput Sequencing Kit (FCL PE100) Katalog Numarası: 1000016111	dNTPs Miks II	4.50 mL.	-15℃ ve altı sıcaklıklarda	-25℃~-15℃
	dNTPs Miks IV	5.00 mL.		
	Sekans Enzim Miks	4.75 mL		
	MDA Reaktifi	4.20 mL		
	MDA Enzim Miks	0.60 mL		
	Sekans Reaktif Kartuşu	1 EA		
	Şeffaf Kaplama Filmi	2 adet		
DNBSEQ-T7RS Cleaning Reagent Kit（FCL PE100） Katalog Numarası: 1000016120	Yıkama Kartuşu	1 EA	40℃ ve altı sıcaklıklarda	0℃~30℃

Tablo 3-2: CPAS Barkod Primer 3 Reaksiyon Kiti

Ürün	Reaksiyon Kiti	İçerik	Miktar	Taşıma Sıcaklığı	Stok Sıcaklığı
Primer for double barcode sequencing (Pair End Sequencing use only)	CPAS Barcode Primer 3 Reagent Kit Katalog numarası: 1000020834	1μMAD153 Barcode Primer 3 Katalog numarası: 1000020833	3.50 mL.1tüp	-15℃ altında	-25℃~-15℃

Sekanslama İş Akışı

I. **DNB Hazırlama:** DNB Hazırlama Kit'inde yer alan reaktifler kullanılır ve DNB yapmak için DNA kütüphanesi hazırlanır.

II. **DNB Yükeme:** DNB Yükleme Reaktif Kit'inde yer alan reaktifleri kullanarak DNB, MGIDL-T7RS yükleyicisindeki akış hücrelerine yüklenir.

III. **Sekanslama Reaksiyonunu Hazırlama:** Reaksiyon kiti kontrol edildikten sonra çözülür ve gerekli reaktifler eklenir. Aynı zamanda saf su deposu doldurulur.

IV. **Sekanslama:** Dizilemeyi başlatmak için son kontroller yapılır ve parametreler girildikten sonra dizileme başlatılır.

I. DNB Hazırlama

A. Kütüphane Insert Boyutu Önerisi

Bu dizileme setleri, MGI Library Prep Kitleri kullanılarak hazırlanan kütüphanelerle uyumludur. Genel amaç için kütüphane tek zincirli dairesel DNA'yı (ssDNA) ifade eder. En iyi dizileme kalitesi için, kütüphanenin insert büyüklüğü 50-500 nükleotid (nt) arasında olmalıdır. MGIEasy stLFR Library Prep Kit ile hazırlanan stLFR kütüphanesi için kütüphane dairesel dsDNA'dır. Kütüphanenin insert boyutunun 200-1500 nt arasında olması tavsiye edilir. Bu çalışmada FCL PE 100 kiti kullanıldı. Diğer çalışmalar ve bu çalışmada kullanılan kitlerdeki insert boyutlarına ait bilgi Tablo 3.3'te yer almaktadır.

Tablo 3-3: Her Akış Hücresi için Önerilen Insert Boyutu ve Teorik Verim*

Sekanslama Kit	Önerilen Insert boyutu (nt)	Uygulama Alanı	Haritalanabilen Okuma（M)* *	Data Çıktısı (Gb)* *
FCL SE35	50-230	NIPT	5000	175
FCL SE50	50-230	NIPT，PMSEQ	5000	250
FCL SE100	200-400	PMSEQ	5000	500
FCL PE100	200-400	WGS，WES，RNAseq，Single Cell	5000	1000
FCL PE150	300-500	WGS，WES，RNAseq	5000	1500
App-A FCL PE100	200-400	WGS，WES，RNAseq	5000	1000
App-A FCL PE150	300-500	WGS，WES，RNAseq	5000	1500
stLFR FCL PE100	200-1500	stLFR	4000	800

*Bir dizileme kiti seçerken insert boyutunu ve gerekli veri çıktısını göz önünde bulundurulur.

**Gerçek veri çıktısı, farklı kütüphane türleri ve uygulamalarına göre değişiklik gösterecektir.

B. Kütüphane Konsantrasyonu

- Kütüphane gereksinimi, ilgili kütüphane hazırlama kiti kullanım kılavuzuna tabidir.
- Tipik dairesel tek sarmallı DNA (ssDNA) kütüphanesi için, ssDNA kütüphane konsantrasyonu ≥ 2 fmol/μL olmalıdır.
- PCR içermeyen kütüphane için ssDNA kütüphane konsantrasyonu ≥ 3,75 fmol/μL olmalıdır.

  Not: Kütüphane konsantrasyonu bilinmiyorsa, QubitTM ssDNA Test Kiti ve QubitTM Florometre kullanılarak ssDNA kütüphane miktar tayininin (ng/μL) gerçekleştirilmesi önerilir. ssDNA kütüphane konsantrasyonunu ng μL'den fmol/μL'ye dönüştürmek için aşağıdaki denklemi kullanın:

  *Konsantrasyon (fmol/μL) = 3030 * Konsantrasyon (ng/μL) / N*

  N, nükleotidlerin sayısını temsil eder (adaptör dahil toplam kütüphane uzunluğu).

C. Kütüphane Havuzu Oluşturma

1. Birlikte havuz oluşturulabilecek numune sayısı

DNBSEQ-T7RS sekanslama platformu aynı anda 4 akış hücresinin dizilenmesini gerçekleştirebilir ve her akış hücresi teorik olarak 5000M okuma üretebilir. PE100 dizileme için, bir akış hücresi teoride 1 terabayt (Tb) veri üretebilir. Her akış hücresi için birlikte havuz oluşturulacak örneklerin sayısı; gerekli veri çıktısına, okuma uzunluğuna ve spesifik uygulamaya bağlıdır. DNB havuzu oluşturmadan önce DNB'nin ölçülmesi önerilir. Bir kılavuz olarak, havuzlamadaki varyasyondan ve tüm barkodların aynı miktarda DNB'den aynı miktarda veri çıkışı üretmeyecek olması nedeniyle Tablo 3.4'te açıklandığı gibi teorik veri çıktısının %90'ından daha fazla toplam veri çıktısı olan daha fazla numuneyi havuzda birleştirmeyin.

Örnek 1: İnsan Tüm Genom Dizilemesi (WGS).

PE100 sekanslama kitini kullanırken, her akış hücresinde 9 numune önerilir.

Örnek 2: stLFR örneği. PE100 dizileme kiti kullanılırken, gerekli dizileme derinliği 40X ise, her bir akış hücresi için 6 örneğin havuzlanması önerilir.

Örnek 3: Çeşitli uygulamalarla örnekleri havuzlama

PE150 dizileme kiti kullanılırken, dizilenecek örnekler WGS (100G/numune) ve RNASeq (50G/numune) içeriyorsa, her akış hücresi için 4 WGS örneğinin ve 18 RNASeq örneğinin havuzlanması önerilir .

Tablo 3-4: Farklı örnek havuzlama örnekleri

İndeks	Okuma Uzunluğu	Her Örnek için Minimum Veri	*Pooling örnek numarası	Her örnek için teorik veri çıktı aralığı
1	PE100	100Gb	9	112Gb-137Gb
2	stLFR PE100	120Gb	6	120Gb-146Gb
3	PE150	50Gb	18 RNAseq	52Gb-64Gb
		100Gb	4 WGS	102Gb-122Gb

*Havuzlama varyasyonunun ± %10 dahilinde olduğu varsayılır.

2. Barkod için baz dengesini doğrulama

Her dizileme döngüsünde dengeli bir baz bileşimi, yüksek dizileme kalitesi için çok önemlidir. Barkoddaki her pozisyon için minimum A, C, G, T baz bileşiminin %12,5'ten az olmaması özellikle tavsiye edilir. Belirli bir numune havuzu için, barkod içindeki A, C, G, T'nin minimum baz bileşimi %5-12,5 arasındaysa, barkod bölme hızı tehlikeye girebilir. Barkodun herhangi bir konumunda A, C, G, T'nin minimum baz bileşimi %5'ten az ise, barkodda daha dengeli bir baz kompozisyonuna sahip olmak için havuzlama stratejisinin yeniden tasarlanması kuvvetle önerilir. Aynı barkoda sahip iki veya daha fazla numunenin birlikte havuzlanmaması gerektiğine dikkat etmek de önemlidir. Aksi takdirde okumayı doğru şekilde yapmak imkansızdır.

D. DNB'yi Oluşturma

1. FCL SE35, FCL SE50, FCL SE100 ve FCL PE100 için DNB hazırlama

- **Gerekli ssDNA Kütüphane Miktarını Hesaplama**

FCL SE35, FCL SE50, FCL SE100 ve FCL PE100 için bir akış hücresini yüklemek için 270μL DNB gereklidir. Bir DNB oluşturma reaksiyonu, 100μL veya 50μL DNB yapabilir. Gerekli veri çıktısına ve DNA kütüphanesinin türlerine bağlı olarak 100μL veya 50μL'lik DNB'yi seçin. Örneğin, normal ssDNA kütüphanesi için sırasıyla 100μL ve 50μL DNB yapmak için 40fmol ve 20fmol ssDNA kütüphanesi gereklidir. PCR içermeyen ssDNA kütüphanesi için sırasıyla 100μL ve 50μL DNB yapmak için 75fmol ve 37.5fmol ssDNA kütüphaneleri gereklidir. Tablo 3.5, 100μL veya 50μL DNB yapmak için gereken gerekli ssDNA kütüphane hacmini gösteriyor.

Tablo 3-5: FCL SE35/SE50/SE100/PE100 Kitleri için Gerekli ssDNA Hacmi

Numune tipi	Gerekli ssDNA hacmi: V(μL)	
	100 μLDNB Reaksiyonu	50 μLDNB Reaksiyonu
Normal kütüphane	40fmol/C	20fmol/C
PCR free	75fmol/C	37.5fmol/C

Not : C, ssDNA kütüphanesinin (fmol/μL) konsantrasyonunu ifade eder.

Belirli bir A örneği için, "a" milyon baz veri çıktısı gerektiriyorsa ve bu akış hücresi için toplam teorik beklenen veri çıktısı "b" milyon baz ise, o zaman örnek A için havuzda gerekli DNB hacmi (V) şu şekildedir:

$$V = a/b * 270 (\mu L).$$

Havuzlanan toplam numune sayısı < 6 ise, 100 µL DNB reaksiyon hacminin seçilmesi önerilir ve 100 µL DNB reaksiyonu yapan reaksiyon sayısının yuvarlama (V/100) + 1'e eşit olması önerilir.

(Örneğin: V = 80 ise, bir adet 100 µL DNB reaksiyonu gerektirir; V = 120 ise, iki adet 100 µL DNB reaksiyonu gerektirir.)

Havuzlanan toplam numune sayısı ≥ 6 ise, 50 µL DNB reaksiyon hacminin seçilmesi önerilir ve 50 µL DNB reaksiyonu yapan reaksiyon sayısının yuvarlama (V/50) + 1'e eşit olması önerilir.

Not: Tüm numuneler potansiyel olarak bulaşıcı kabul edilmeli ve ilgili ulusal düzenlemelere uygun olarak işlenmelidir.

- **Reaktifleri DNB Yapımı İçin Hazırlama**

Kütüphane kullanılana kadar buz üzerine yerleştirildi. DNB Hazırlama Tamponu, Düşük TE Tamponu ve DNB Reaksiyon Durdurma Tamponu, DNB Make Reaktif Kitinden çıkarıldı ve reaktifleri oda sıcaklığında eritildi. Tablo 3.6'da gerekli hacimler gösterildi. DNB Enzim Miks I'i buz üzerinde çözüldü. Çözüldükten sonra reaktifleri bir vorteks karıştırıcı kullanarak 5 saniye karıştırıldı, kısaca santrifüjlendi ve kullanılana kadar buz üzerinde tutuldu.

- **DNB Oluşturma**

Reaksiyon karışımı 0,2 mL'lik 8'li strip tüp içerisinde aşağıdaki tabloya göre buz üzerinde hazırlandı.

Tablo 3-6: FCLSE35/SE50/SE100/PE100 Make DNB Reaktif Karışımı-1

İçerik	100 µL DNB reaksiyon için Hacim (µL)	50 µL DNB reaksiyon için Hacim (µL)
ssDNA kütüphanesi	V *	V **
Düşük TE Tamponu	20-V	10-V
DNB Hazırlama Tamponu	20	10
Toplam Hacim	40	20

* ve ** : V, sırasıyla 100 µL ve 50 µL DNB reaksiyonu için Tablo 3.5' ten belirlenen ssDNA kitaplığının hacmidir.

DNB hazırlama reaksiyon karışımı 1'i vorteksle iyice karıştırıldı ve mini santrifüj kullanarak 5 saniye santrifüjlendi. Karışım Tablo 3.7' de belirtilen koşullarda primer hibridizasyon reaksiyonuna tabi tutuldu.

Tablo 3-7: FCLSE35/SE50/SE100/PE100 Primer Hibridizasyon Reaksiyon Koşulları

Sıcaklık	Zaman
Kapak Sıcaklığı（105℃）	On
95℃	1 dk
65℃	1 dk
40℃	1 dk
4℃	Devam

DNB Hazırlama Enzim Miks II (LC) buz üzerine yerleştirildi. Yaklaşık 5 saniye santrifüjlendi ve kullanılana kadar tekrar buz üzerinde bekletildi.

Tablo 3-8: FCLSE35/SE50/SE100/PE100 Make DNB Reaksiyon Karışımı-2

İçerik	100 µL DNB reaksiyon için Hacim (µL)	50 µL DNB reaksiyon için Hacim (µL)
DNB Hazırlama Enzim Karışımı-I	40	20
DNB Hazırlama Enzim Karışımı-II	4	2

- DNB Hazırlama Miks1 reaksiyon sonucunda DNB Hazırlama Enzim Miks II (LC) ile Tablo 3.8'de belirtildiği şekilde 6-8 kez hafifçe pipetleyerek karıştırıldı.
- Karışım vortekslendi ve mini santrifüj kullanarak 5 saniye santrifüjlendi. Karışım Tablo 3.9'da belirtilen koşullarda yuvarlanan daire amplifikasyonu reaksiyonuna tabi tutuldu.

Tablo 3-9: FCL SE35/SE50/SE100/PE100 Yuvarlanan Daire Amplifikasyon Koşulu

Sıcaklık	Zaman
Kapak sıcaklığı（35℃）	On
30℃	25 dk
4℃	Devam

- Reaksiyon sonunda, tüp içerisine 100 μL DNB reaksiyonu için 20 μL Stop DNB Reaksiyonu Durdurma Tamponu eklendi. 1000 μL' lik pipet kullanarak 5-8 kez pipetleyerek hafifçe karıştırıldı.

E. DNB'yi Ölçme ve Havuz Oluşturma

1. DNB'yi Ölçmek

DNB oluşturma işlemi tamamlandıktan sonra DNB'yi ölçmek için QubitTM ssDNA Test Kiti ve QubitTM Florometre cihazı kullanıldı. Bunun nedeni dizileme, DNB konsantrasyonunun 8 ng/μL'nin üzerinde olmasını ya da 40 ng/μL'nin altında olması gerekir. Konsantrasyon 8 ng/μL'den düşükse yeni bir DNB yapılması, konsantrasyon 40 ng/μL'yi yüksek ise DNB'nin yüklemeden önce 20 ng/μL'ye seyreltilmesi gerekmektedir.

2. DNB Havuzu Oluşturma

Havuzdaki her numune için gereken DNB (μL) miktarı, bu numunenin nispi miktarına ve belirli tipte dizileme kiti tarafından tanımlanan bir akış hücresini yüklemek için gereken toplam DNB miktarına bağlıdır. Bu nedenle nispi miktar hesaplaması aşağıda belirtildiği şekilde hesaplanmıştır.

Her numune için nispi miktarı hesaplama

Havuzda 8 numune (A'dan H'ye) olduğu varsayıldığında, her numune için nispi miktar şu şekilde tanımlanır:

A numunesinin nispi miktarı (A1) = numune A için gerekli veri çıkışı / numune A için DNB konsantrasyonu;

B numunesinin nispi miktarı (B1) = numune B için gereken veri çıkışı / numune B için DNB konsantrasyonu;

...

H numunesinin bağıl miktarı (H1) = numune H için gereken veri çıkışı / numune H için DNB konsantrasyonu;

Tüm numune için toplam bağıl miktarı (V) hesaplama

V = A1 + B1 + … + H1

Her numune için gereken DNB hacmini hesaplama

SE35/SE50/SE100/PE100, App-A PE100 ve 270μL DNB gerektiren stLFR PE100 için kullanılan her bir FCL akış hücresi için havuzlama için DNB hacmi şu şekilde hesaplanır:

Örnek A için DNB hacmi, A2 = 270 * A1 / V

B numunesi için DNB hacmi, B2 = 270 * B1 / V

…

H numunesi için DNB hacmi, H2 = 270 * H1 / V

400μL DNB gerektiren PE150 ve App-A PE150 için kullanılan her FCL akış hücresi için havuzlama için DNB hacmi şu şekilde hesaplanır:

Örnek A için DNB hacmi, A2 = 400 * A1 / V

B numunesi için DNB hacmi, B2 = 400 * B1 / V

…

H numunesi için DNB hacmi, H2 = 400 * H1 / V

Not: Her DNB'nin gerekli hacmini aspire etmek için normal pipet uçları kullanıldı ve karıştırmak için geniş çaplı uçlar kullanıldı.

II. DNB Yükleme

A. Yükleme sonrası plakayı ve tampon çözeltileri hazırlama

Yükleme Sonrası Plakayı Çözme:

Yükleme Sonrası Plakası DNB Yükleme Reaktif Kitinden çıkarıldı ve oda sıcaklığında 2 saat su banyosunda eritildi. Yükleme Plakası tamamen çözüldükten sonra, kullanana kadar 2-8°C'lik bir buzdolabına koyuldu. Ayrıca Yükleme Plakası gece boyunca 2-8 °C buzdolabında da çözülebilir. 5 kez karıştırmak için

Yükleme Sonrası Plakasını nazikçe ters çevrildi ve ardından kullanmadan önce 1 dakika santrifüj yapıldı.

B. DNB Yükleme Reaktiflerini Hazırlama

DNB Yükleme Tamponu II, DNBSEQ-T7RS DNB Yükleme Reaktif Kitinden çıkarıldı. Reaktifler oda sıcaklığında yaklaşık yarım saat su banyosunda eritildi. Çözüldükten sonra reaktifler bir vorteksle 5 saniye karıştırıldı. Kısa bir süre santrifüj yapıldı ve kullanılana kadar buz üzerinde tutuldu. DNB Yük Tamponu II'de kristal çökeltisi bulunduğundan, kullanımdan önce çökeltiyi yeniden çözmek için reaktif 1-2 dakika sürekli vorteks ile kuvvetlice karıştırıldı.

C. 0.1M NaOH Reaktifini Hazırlama

Her Yükleme Sonrası plakası en az 4 mL 0.1M NaOH gerektirdiği için uygun hacimde 0.1M NaOH, hazırlandı.

D. DNB Yükleme Karışımını Hazırlama

0,5 mL'lik mikrosantrifüj tüpü içinde Tablo 3.10' de belirtilen hacimlerde DNB Yükleme karışımı hazırlandı.

Tablo 3-10: FCL SE35/SE50/SE100/PE100 için DNB Yükleme Karışımı

Yükleme Sırası	İçerik	Hacim (µL)
1	DNB*	270
2	DNB Yükleme Tamponu II	90
3	DNB Hazırlama Enzim Miks II（LC）	1

Dizileme Akış Hücresini Hazırlama

- Akış hücresini DNBSEQ-T7RS dizileme akış hücresi kutusundan çıkarıldı.
- Dış paketi açıldı.
- Toz giderici kullanarak akış hücresinin arkası temizlendi.

III. Sekans Reaksiyonunu Hazırlama

Sekans reaksiyonu aşağıda belirtilen basamaklar şeklinde hazırlandı.

- High-throughput Sequencing Kit içinden sekans kartuşu çıkarıldı. 4-5 saat oda sıcaklığında su banyosunda bekletildi.
- Kuyulara uygun şekilde yükleme yapmak için sekans kartuşu hazır hale getirildi.
- dNTP Miks IV, dNTP Miks II ve Enzim Miks 9 ve 10 numaralı kuyulara eklendi.

- MDA Enzim Miks ve MDA Reaktifi 8 numaralı kuyuya eklendi.
- Gerekli adımlar tamamlandıktan sonra yıkama kartuşuna saf su haznesi dolduruldu ve kullanılan kite göre ölçüsü ayarlandı.

IV. Sekanslama

Sekanslama için uygulanan basamaklar aşağıda belirtilmiştir

- Reaksiyon kartuşu yerleştirildi.
- Dizileme arayüzü girildi.
- Akış hücresi yüklendi.
- Dizileme parametreleri ayarlandı.
- Parametreler kontrol edildikten sonra dizileme başlatıldı.

3.5. Tüm Genom Dizileme Ham Verilerinin Analizi

Tüm Genom Dizileme işlemi DNBSEQ-T7RS platformunda gerçekleştirildikten sonra FASTQ formatında ham veri üretilir. Üretilen bu verinin işlenmesi ve yorumlanması için birçok platform bulunmaktadır. TGD verisi büyük bir veri olduğundan ve ürettiği varyant miktarı 700 binden fazla olduğundan ötürü, veri analizini kolaylaştıran ve daha verimli hale getiren birçok platform geliştirilmiştir. Yapay zeka destekli algoritmalar kullanılarak klinikle uyumlu varyant belirleme platformlarından biri olan NGS Cloud Platformu TGD verisinin analizi için tercih edilmiştir [74]. Aşağıda genel olarak TGD verisinin işlenmesi aşamasında takip edilen adımlar ve kullanılan veritabanları anlatılmıştır.

- Dizileme reaksiyonundan sonra platformun ürettiği FASTQ formatındaki veri alındı.
- Alınan veriler, ticari olarak hizmet veren Pairend firmasına ait bir yapay zeka destekli genetik veri analiz platformu olan NGS Cloud platformu kullanılarak işlendi.
- Alınan ham verilerin insan referans genomuna göre hizalanması için Bowtie, Burrows-Wheeler Aligner (BWA), Samtools, Picard Tools araçları kullanıldı. Ham veriler (.fastq > .bam/.bai) hizalanmış formata dönüştürüldü.
- Olguya ait ham veriler kalite kontrolden geçirildi.

- Hizalanan veriler içinden GATK 4.2.3.0 platformu kullanılarak varyantlar çağırıldı. Hizalanmış veriler (.bam/.bai > vcf) varyantların çağırıldığı formata dönüştürüldü.
- ANNOVAR ve VEP aracılığı ile çağırılan varyantların anotasyonları yapıldı (vcf > anotasyon).
- Varyant anotasyonlarının ardından filtreleme ölçütleri belirlenerek anlamlı varyantlar filtrelendi.
- Aday varyantlar içinde klinikle uyumlu ve fenotipi belirleme potansiyeli güçlü olan varyantlar, değişim özelliğine göre in silico araçlar kullanılarak patojenite tahminleri yapıldı.
- Uygun varyantlar için segregasyon ve konfirmasyon çalışması planlandı.
- Tablo 3-11'de referans genom, gen dizileri, primer sentezi, varyant analizi ve hastalık ilişkisi gibi bilgilere ulaşmak için kullanılan tüm veritabanları paylaşılmıştır.

Tablo 3-11: İnsan ve model organizma çalışmaları veri tabanları

PubMed	https://www.ncbi.nlm.nih.gov/pubmed/
Rat Genome Database (*Rattus norvegicus*)	https://rgd.mcw.edu/
Mouse Genome Informatics (*Mus musculus*)	http://www.informatics.jax.org/
KEGG PATHWAY	https://www.genome.jp/kegg/pathway.html
Reactome	https://reactome.org/
Gene Expression Atlas	https://www.ebi.ac.uk/gxa/home
The Human Protein Atlas	https://www.proteinatlas.org/

Genetik dizileme platformlarından elde edilen ham verilerin referans genoma hizalanması, hizalanmış verilerden varyantların çağırılması ve varyantların anotasyonunun ardından filtreleme basamaklarının tamamlanmasına kadar olan süreçlerde açık kaynaklı veya ücretli veri tabanları kullanılmaktadır. Popülasyonlara özgü veri tabanları, hastalıklara özgü veri tabanları ve dizilere özgü veritabanlarından en sık

kullanılan veri tabanları Tablo 3.12'de gösterilmiştir. Bu veritabanları kullanılarak elde edilen varyantların görülme sıklığı gibi bilgiler elde edilebilmektedir.

Tablo 3-12: Popülasyonlara ve hastalıklara özgü veri tabanları

Popülasyonlara özgü veri tabanları		
Exome Aggregation Consortium	Populasyon çalışmaları ya da projelerden elde edilmiş ekzom verilerini içerir	gnomAD veri tabanına eklenmiştir
The Genome Aggregation Database	Geniş kapsamlı ekzom (125.748) ve genom (71.702) dizileme çalışmalarından elde edilmiş verileri içerir	https://gnomad.broadinstitute.org/
Exome Variant Server	Avrupa ve Afro-Amerikan kökenli bireylere ait varyantları içerir	https://evs.gs.washington.edu/EVS/
1000 Genomes Project	2008-2015 yılları arasında yürütülen, 26 farklı populasyona ait <%1 sıklıktaki varyantların tespit edilmesini hedefleyen projedir	Ensembl GRCh38 veri tabanına eklenmiştir
dbSNP	Farklı kaynaklarda bildirilmiş ≤50 bç'lik varyantlara dair bilgileri içerir	https://www.ncbi.nlm.nih.gov/snp/
dbVar	Farklı kaynaklarda bildirilmiş ≥50 bç'lik varyantlara dair bilgileri içerir	https://www.ncbi.nlm.nih.gov/dbvar
Hastalıklara özgü veri tabanları		
ClinVar	Varyantların hastalıklar ve fenotipler ile ilişkisine dair bilgileri içerir	https://www.ncbi.nlm.nih.gov/clinvar/

OMIM	Genler ve hastalık ilişkili gen varyantlarını içeren geniş kapsamlı bir veri tabanıdır	https://omim.org
Human Gene Mutation Database	Lüteratürde yayınlanmış hastalık ilişkili varyant bilgilerini içerir	http://www.hgmd.cf.ac.uk/ac/index.php
Human Genome Variation Society	Fenotip ve populasyonlar ile ilişkili varyantlar hakkında birçok farklı veri tabanını içerir	https://www.hgvs.org/
Leiden Open Variation Database	YND teknikleri ile elde edilmiş, genlere dair verileri içerir	https://www.lovd.nl/
DECIPHER	Klinisyenler ve araştırmacılara yönelik moleküler sitogenetik teknikler ile elde edilmiş genomik varyant verilerini içerir	https://decipher.sanger.ac.uk/
Dizilere özgü veri tabanları		
NCBI Genome	İnsan genomu referans dizilerini içerir	https://www.ncbi.nlm.nih.gov/genome
RefSeq	Medikal, fonksiyonel ve araştırmaya yönelik çalışmalardan elde edilmiş referans dizileri içerir	https://www.ncbi.nlm.nih.gov/refseq/
Locus Reference Genomic (LRG)	Klinik çalışmalar ve yayınlardan elde edilmiş gen, transkript ve protein dizilerini içerir	https://www.lrg-sequence.org/

MitoMap	İnsan mitokondriyal DNA'sı için Cambridge referans genomunu içerir	https://www.mitomap.org/MITOMAP
Ensembl	Omurgalılar için karşılaştırmalı, fonksiyonel ve evrimsel genomik bilgiyi, genlere ait varyant ve protein bilgilerini içerir	https://www.ensembl.org/index.html
UniProt	Protein dizileri ve fonksiyonları hakkında karşılaştırmalı bilgiler içerir	https://www.uniprot.org/
UCSC	Genomik bilgiyi ve gen dizilerini içerir	https://genome-euro.ucsc.edu/

Elde edilen varyantların gen veya protein dizisine nasıl etkide bulunduğunu tahmin eden birçok in siliko varyant tahmin aracı bulunmaktadır. Tablo 3.13'de yanlış anlamlı varyantların etkisini tahmin eden araçların yanında kırpılma bölgesinde meydana gelen varyantların etkisini ölçmek için en çok kullanılan araçlar bulunmaktadır.Ayrıca varyantın nükleotid korunmuşluğu ile ilgili analizini yapan araçlar da tabloda yer almaktadır.

Tablo 3-13: Tez kapsamında kullanılan in siliko varyant tahmin aracı veri tabanları

Yanlış anlamlı değişimleri için	
MutationTaster	http://www.mutationtaster.org/
PROVEAN	http://provean.jcvi.org/index.php
PolyPhen-2	http://genetics.bwh.harvard.edu/pph2/
ConSurf	https://consurf.tau.ac.il/
FATHMM	http://fathmm.biocompute.org.uk/index.html
MutationAssessor	http://mutationassessor.org/r3/
PANTHER	http://pantherdb.org/
PhD-SNP	http://snps.biofold.org/phd-snp/phd-snp.html
MutPred2	http://mutpred.mutdb.org/#qform
SNPs&GO	https://snps-and-go.biocomp.unibo.it/snps-and-go/

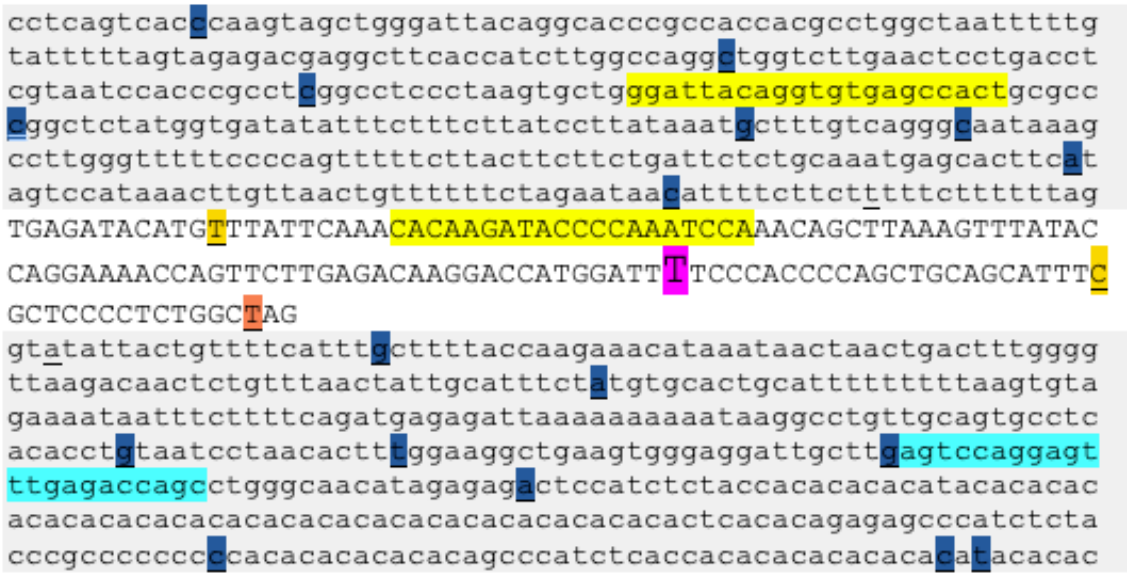
nsSNPAnalyzer	http://snpanalyzer.uthsc.edu/
Kırpılma bölgesi için	
GeneSplicer	https://www.cbcb.umd.edu/software/GeneSplicer/gene_spl.shtml
Human Splicing Finder	http://www.umd.be/HSF/index.html
MaxEntScan	http://hollywood.mit.edu/burgelab/maxent/Xmaxentscan_scoreseq.html
NetGene2	http://www.cbs.dtu.dk/services/NetGene2/
Nükleotid Korunmuşluğu için	
GERP	http://mendel.stanford.edu/SidowLab/downloads/gerp/
PhastCons	http://compgen.cshl.edu/phast/
PhyloP	http://compgen.cshl.edu/phast/help-pages/phyloP.txt

3.6. Saptanan Varyantların Konfirmasyonu için Primerlerin Tasarlanması

Tüm genom dizileme işleminden sonra anlamlı olarak filtrelenen varyantların konfirmasyonu ve ailesel segregasyonu yapmak amacıyla ilgili gen bölgelerine spesifik primerler tasarlandı. Tasarlama aşamasında, primerlerdeki oligonükleotidlerin tekrarlayan dizi bölgesinde olmamasına, herhangi bir polimorfizm içermemesine ve içermesi kaçınılmaz ise toplum sıklığının %1'in altında olmasına, primerlerin DNA'ya bağlanma ısılarının birbirine yakın olmasına (~2-5˚C farklılık), ararlarında çift bağ yapısı bulunan A-T nükleotitlerinin sayısının üçlü bağ yapısı bulunan C-G sayısında olabildiğince eşit olmasına önem verildi. Primer3 suite online in siliko aracı kullanılarak primerler tasarlandı. Primerlerin ilgili bölgelere spesifik olup olmadığı NCBI Blast programı kullanılarak teyit edildi. Primerlerin hairpin, self dimer ve hetero dimer yapıp yapmadığına IDT (Integrated DNA Technologies) Oligo Analyzer online platformu kullanılarak bakıldı. Tez kapsamında tasarlanan primer dizilerine ait bilgiler Tablo 3.14'de yer almaktadır.

Tablo 3-14: Aday genlerin primer dizileri

Gen ve Ekzon	Primer Dizisi	Primer Bağlanma Sıcaklığı (°C)	PZR Ürün Boyu
ATP8A2-16F	5'- TGAAGGCAGATTCTTTCCAGG-3'	57.92	166 bç
ATP8A2-16R	5'- AGCACTTACATGGCGATCCT-3'	59.17	166 bç
DNMT3B-17F	5'- TGCTTCTGTGTCTCTCTGGC-3'	59.68	194 bç
DNMT3B-17R	5'-AAGGTACAGACTGCCCTCAC-3'	59.02	194 bç
NFU1-3F	5'-ACCCCAAATCCAAACAGCTT-3'	57.90	177 bç
NFU1-3R	5'-GTTGTCTTAACCCCAAAGTCAGT-3'	58.80	177 bç
NFU1-3F YENİ	5'-GGATTACAGGTGTGAGCCACT-3'	56.50	627 bç



Şekil 3-1: NFU1 Geni Ekzon-3 Primer Tasarım Çalışması

NFU1 geni 3.ekzon için primer tasarlanmış ancak anne-baba sanger sonuçları çocuklarla uyuşmadığından allel drop out ihtimali düşünülerek yeniden primer tasarlanmıştır. Primer tasarlanan bölge GC bakımından oldukça zengindir. Ayrıca lacivert renkle renklendirilen bölgeler polimorfik bölgeleri göstermektedir.

3.7. Saptanan Varyant Bölgelerinin Polimeraz Zincir Reaksiyonu (PZR) ile Çoğaltılması

İlgili gen bölgelerini çoğaltmak amacıyla olgulara ait DNA'lar ve tasarlanan primerler ile termal döngü cihazı kullanılarak PZR çalışması yapıldı. Her bir çalışma için kullanılan malzemeler, final konsantrasyonları ve PZR çalışma koşulları Tablo 3.15 ve 3.16'de verildi.

Tablo 3-15: PZR çalışması için kullanılan malzemeler ve final konsantrasyonları

Malzeme Adı	Final Konsantrasyon
10X (NH4)2SO4 tampon solüsyonu (Fermantas)	1 X
25 mM MgCl2 (Fermantas)	2 mM
25 µM dNTP (Fermantas)	0.2 mM
2 µM İleri ve geri primerler (Sentebiolab)	0.2 µM
500 U (5 U/µl) Taq polimeraz enzimi (Fermantas)	0.1 µl
Distile Su	25 µl'ye tamamlanacak şekilde eklenir
DNA	100-200 ng/g
Final Hacim	25 µl

Tablo 3-16: PZR çalışma koşulları

Aşama	Sıcaklık	Süre	Döngü
Ön Denatürasyon	94°C	10 dakika	x1 döngü
Denatürasyon	94°C	30 saniye	x35 döngü
Bağlanma	Primer Bağlanma Isısı		
Uzama	72°C		
Son Uzama	72°C		x1 döngü
Bekleme	4°C		x1 döngü

PZR uygulamasından sonra çoğalan gen bölgelerinin ilgili gen bölgeleri olduğunu göstermek ve bant boylarının hedeflenen bç ile aynı olduğunu göstermek amacıyla, 5 µl PZR ürünü ile 2 µl Gel Red boyası karıştırılarak %1.5' lik agaroz jel, içerisinde 1xTBE tamponu bulunan elektroforez tankına yerleştirildi. PZR ürünü ve 50bç' lik marker yükleme tamponu kullanılarak kuyucuklara mikropipet yardımıyla yüklendi. Elektroforez tankına bağlı güç kaynağı ile 120 voltta 30 dakika yürütüldü. Süre sonunda örnekler UV ışık veren transilluminatör (Syngene DigiGenius) yardımıyla incelendi.

3.8. PZR Ürünlerinin Pürifikasyonu

PZR sonrası, tüp içerisindeki ürün dışı PZR bileşenlerinin uzaklaştırılmasını sağlayan pürifikasyon basamağıdır. Bu basamakta PZR pürifikasyon kitlerinden ExoSap kullanıldı. 5 µl PZR ürününe 1 µl Exosap eklendi. Karışım 37 °C 15 dk, 80 °C de 15' dk bekletildi.

3.9. Dizi Analizi PCR Reaksiyonun Kurulması

Dizi analizi reaksiyonunda Sanger sonlandırma metodu temelli florasan işaretli dideoksi nükleotidlerin kullanıldığı Big Dye Terminator V3.1 (ABI) kiti kullanıldı. Dizi analizi reaksiyonu; tek yönlü primer, 5X tampon, Big Dye Terminator enzim, distile su ve hedef DNA içeriğinden oluşmaktadır. Kurulan reaksiyon ısı döngüsü cihazında

(SimpliAmp; ABI) çoğaltıldı. Sekans reaksiyon içeriği ve PZR koşulları Tablo 3.17 ve 3.18' de verilmiştir.

Tablo 3-17: Sekans Reaksiyon İçeriği

	1X
dH_2O	1.8 µl
5X Buffer	2 µl
PrF (1uM)	3,2 µl
BigDye	1 µl
DNA	2 µl
Total	**10 µl**

Tablo 3-18: Sekans Reaksiyon Sıcaklıkları

Denatürasyon	96°C	1 dk	1x döngü
Denatürasyon	96°C	10 sn	25x döngü
Bağlanma	50°C	5 sn	
Uzama	60°C	4 dk	
Sonlanma	4°C	∞	

3.10. Dizi Analizi Yapılacak Amplikonların Temizlenmesi

Dizi analizi reaksiyonu sonrası, ürün dışı bileşenlerin ortamdan uzaklaştırılmasını sağlayan pürifikasyon basamağıdır. Bu basamakta birbirinden farklı pürifikasyon metotları kullanılsa da bu çalışmada Sephadex (GML) metodu kullanıldı. 50 ml' lik falkon tüpünde, 1 g Sephadex üzerine 14 ml steril su eklenerek karıştırıldı ve 1-2 saat 4°C' de bekletildi. Ara ara çalkalandı ve kullanmadan hemen önce yine iyice karıştırıldı. Spin kolonların (GML) her birine 700 µl olacak şekilde dağıtıp 5000 rpm'de 2 dk santrifüj edildi. Kolon tüpünün alt tarafına süzülen su atıldı. Daha sonra sekans ürünü kolona değdirilmeden ortasına bırakıldı ve 5000 rpm' de 2 dk santrifüj edildi. Süzülen kısım sekans cihazında yürütülmek üzere hazırlandı.

3.11. DNA Dizi Analizi

Kapiller elektroforez (KE) sistemi ile çalışan ABI-3500 cihazı 8 kapillerli dizi sistemi ve polimeri içerir. DNA sekanslama ve/veya fragman analizi imkanı sunar. Elektrokinetik bir ayırım yaparak, yüksek voltaj ile akım verilir, moleküllere elektrik yüklenerek hareketlenmeleri sağlanır ve anot ucunda yerleştirilmiş hassas sensörler ile izlenirler. Çok az miktarda ürün kullanılarak hassas ve otomatize bir sistem ile diziler okunur.

Plak cihaza yerleştirildikten sonra bilgisayar aracılığı ile daha önceden kalibre edilmiş boya seti seçildi ve Tablo 3.17'de belirtilen koşullarda yürütme yapıldı.

Tablo 3-19: Kapiller elektroforez yürütme koşulları

Yürütme voltajı	13,4 kVolt
Ön yürütme voltajı	15 kVolt
Enjeksiyon voltajı	1,6 kVolt
Yürütme süresi	2,52 saniye
Ön yürütme süresi	180 saniye
Enjeksiyon zamanı	8 saniye
Veri gecikmesi	250 saniye

3.12. Dizileme Sonuçlarının Değerlendirilmesi ve Biyoinformatik Analizi

Elektroforez sonucunda cihaz tarafından verilen elektrokaryogram verileri biyoinformatik analize tabi tutuldu. Elde edilen karyogram Finch TV, 4peaks vb. yazılım programlarıyla insan genom projesiyle elde edilen genomik veri tabanları ile karşılaştırılarak, gen üzerindeki değişimler belirlendi.

3.13. Protein Modelleme ve Simulasyon

Tüm genom dizileme sonrası elde edilen aday varyantların ailesel segregasyonu ve konfirmasyonu ile elde edilen, klinikle uyumlu varyantların protein modellemesi yapıldı.

- Genlere ait amino asit dizilerinin tüm sekans dizisi (https://genome.ucsc.edu/) adresinden indirildi.
- AlphaFold colab platformu kullanılarak tüm sekans modelleri oluşturuldu.
- İndirilen tüm aminoasit sekans dizisi AlphaFold2_advanced arayüzünde yer alan sekans kısmına girildi ve modeller oluşturuldu.
- N-terminal'den C-terminal kısmına kadar aminoasit dizileri bölgesel olarak renklendirildi.
- Alphafold modelleri kaba taneli moleküler dinamik simülasyona tabi tutuldu.
- Modellenen proteinin moleküler dinamik simülasyon sisteminin hazırlanması için CHARMM-GUI platformu kullanıldı.
- CHARMM-GUI platformunda Martini 2.2 ve Martini 2.0 FF modeli seçilerek kaba taneli protein modeli oluşturuldu.

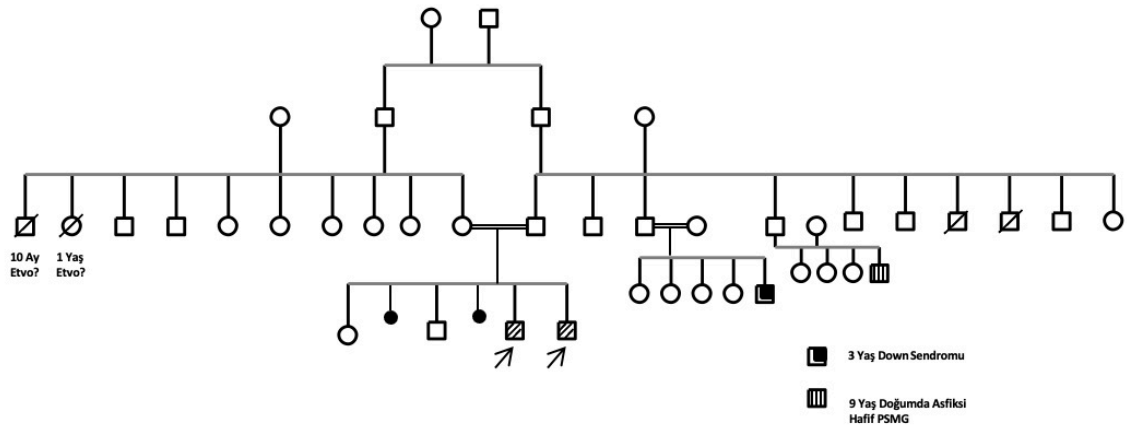
- Moleküler dinamik simülasyon için GROMACS platformu kullanıldı.
- Kaba taneli modeller daha sonra all-atom olarak 100 nano saniye simulasyonundan modellendi.
- Gmx cluster tool platformu küme çerçeveleri oluşturmak için kullanıldı.
- Her bir protein modeli için en iyi kümeler kullanılarak yeni sistemler oluşturuldu.
- Sistemler oluşturulduktan sonra her bir proteinin yabanıl tipi ile mutant tipi modellendi.
- Yeni sistemler iki kez 500 nano saniye simülasyona tabi tutuldu ve her bir deneyde 5000 çerçeve elde edildi.
- Analiz için gmx cluster tool platformu, 4.0 cut-off değeri ile çerçeveler bir temsili yapıya dönüştürüldü.
- Protein modellerinin merkezi atomlarının kök ortalama kare sapma (RMSD) değerleri hesaplandı, grafikler oluşturuldu.
- Ayrıca protein modellerinin karbon alfa atomlarının kök ortalama kare salınım (RMSF) değerleri hesaplandı, grafikler oluşturuldu.
- Simülasyonlar Tübitak'ın TRUBA sunucularında gerçekleştirildi.

4. BULGULAR

4.1. Olgu 1

Ebeveynleri birinci derece kuzen olan 6 yaşındaki Olgu-1, İ.Ü. Tıbbi Genetik Polikliniği'ne yürüme bozukluğu ile başvurdu. Alt ekstremitede artmış Derin Tendon Refklesleri (DTR) ve azalmış kas gücü bulunmaktaydı. Son muayenesinde 6 yaşında olan Olgu-1, 18 aylıkken gribal enfeksiyon geçirmiş, ailenin ifadesine göre enfeksiyon sonrası yürümesi bozulmuş, düşmeye ve yorulmaya başlamıştır. Enfeksiyon geçtikten sonra parmak ucunda yürümesi başlamıştır. Olgu-1'in bir tane etkilenmiş kardeşi ve iki tane de sağlıklı kardeşi bulunmaktadır. Etkilenmiş kardeşte de benzer bulgular bulunmaktadır. Etkilenmiş kardeş, 2 yaşında fizik tedavi alarak yürüyebilmiştir. Çocuk nöroloji polikliniğinde değerlendirilen olgunun kraniyal MR bulguları normal olarak kaydedilmiştir. Enfeksiyon geçirdiklerinde her iki kardeşte de yürüme bozukluğu ve spastisite artmaktadır. Her iki kardeşte de nöbet öyküsü bulunmamaktadır. Fizik muayenesinde 22.5 kg, 118cm boyunda olduğu ve baş çevresinin 52cm olduğu ve kas gücünün ¼ şeklinde olduğu tespit edilmiştir. Olgu-1'e mevcut klinik bulguları ile herediter spastik parapleji ön tanısı konmuştur. Tablo 4-1'de Olgu-1'e ait bulgular gösterilmiştir.

Olguya ait aile ağacı Şekil 4.1'de gösterilmiştir. Birinci derece kuzen evliliği yapan Olgu-1'in ebeynleri, ebeveynlerin anne-babası ve kardeşleri aile ağacında gösterilmiştir. Olgu-1'in bir etkilenmiş kardeşi ve iki de sağlıklı kardeşi bulunmaktadır. Ayrıca iki kardeş de düşük olarak bildirilmiştir.



Şekil 4-1: Olgu-1'e Ait Aile Ağacı

Tablo 4-1: Olgu 1'e Ait Klinik ve Laboratuvar Bulguları

Olgunun Yaşı/Cinsiyeti	7 yaş/Erkek
Ebeveyn Akrabalık Durumu	1.Derece Kuzen
Anne Gebelik Öyküsü/Olgu Doğum	G6P4Ab2 / G3
Ailede Benzer Öykü	Etkilenmiş Kardeş
Fetal USG Bulgusu	Antenatal II.düzey USG yapılmış patoloji saptanmamış
Perinatal Hikaye	Miadında Normal Doğum ile 3500 g doğmuş
Ön tanı	Herediter Spastik Parapleji

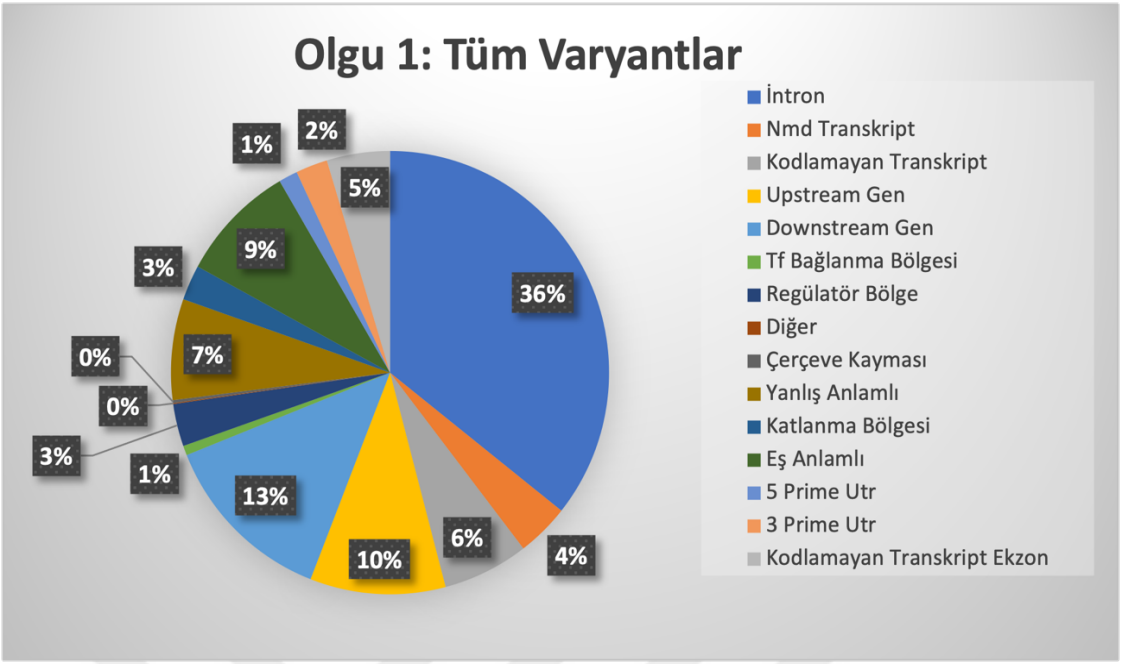
4.1.1. Tüm Genom Dizileme Sonuçları

Olgu 1'in yanak mukozasından alınan örnekten DNA izolasyonu yapıldıktan sonra ortalama 100X okuma derinliği olan kit ile tüm genom dizileme işlemi sonucunda elde edilen toplam varyant sayıları Tablo 4-2'de ve Şekil 4-2'de gösterilmiştir.

Tablo 4-2: Olgu 1'e Ait TGD Verisinden Biyoinformatik Analiz ile Elde Edilen Tüm Varyant Sayıları

Toplam Varyant Sayısı	754.780
İntron	269.473
Nmd Transkript	29.312
Kodlamayan Transkript	47.738
Upstream Gen	75.478
Downstream Gen	98.302
Tf Bağlanma Bölgesi	5.263
Regülatör Bölge	23.414
Diğer	514
Çerçeve Kayması	1.625
Yanlış Anlamlı	55.530
Katlanma Bölgesi	19.394
Eş Anlamlı	65.335
5 Prime UTR	10.511
3 Prime UTR	17.765
Kodlamayan Transkript Ekzon	35.126

Biyoinformatik analiz sonucunda toplam 754.780 varyant elde edilmiştir. Bu varyantların %36' lık kısmını intronik, %13' lük kısmını downstream gen ve %10' luk kısmını ise upstream gen varyantları oluşturmaktadır.



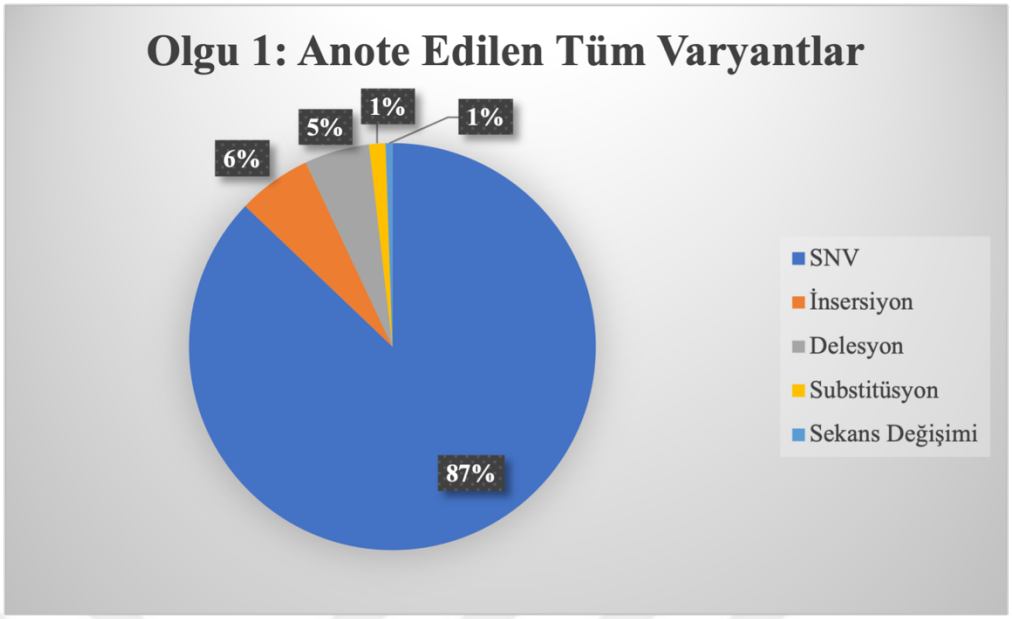
Şekil 4-2: Olgu 1'e Ait TGD Verisinden Biyoinformatik Analiz ile Elde Edilen Tüm Varyantlar

TGD verisinden biyoinformatik analizler ile anote edilen varyant sayısı 60.243 olarak belirlenmiştir. Bunlardan 52.467' sini (%87) SNV, 3531' ini (%6) insersiyon, 3114' ünü (%5) delesyon, 784' ünü (%1) yer değiştirme, 323' ünü (%1) sekans değişimi oluşturmaktadır. Diğer anote edilen varyantlar ise 23 tanedir (Tablo 4-3 ve Şekil 4-3).

Tablo 4-3: Olgu 1'e Ait TGD Verisinden Biyoinformatik Analiz ile Anote Edilen Varyant Sayıları

Anote Edilen Toplam Varyant Sayısı	60.243
SNV	52.467
İnsersiyon	3.531
Delesyon	3.115
Substitüsyon	784
Sekans Değişimi	323
Diğer	23

Sekans değişimi tam olarak değişimin türü belirlenememiş nükleotid değişimleridir. UPD, delesyon, delins, insersiyon, inversiyon, yapısal değişimler ve sekans uzunluk değişimleri bu değişimler arasında yer alabilir.



Şekil 4-3: Olgu 1'e Ait TGD Verisinden Biyoinformatik Analiz ile Anote Edilen Varyantlar

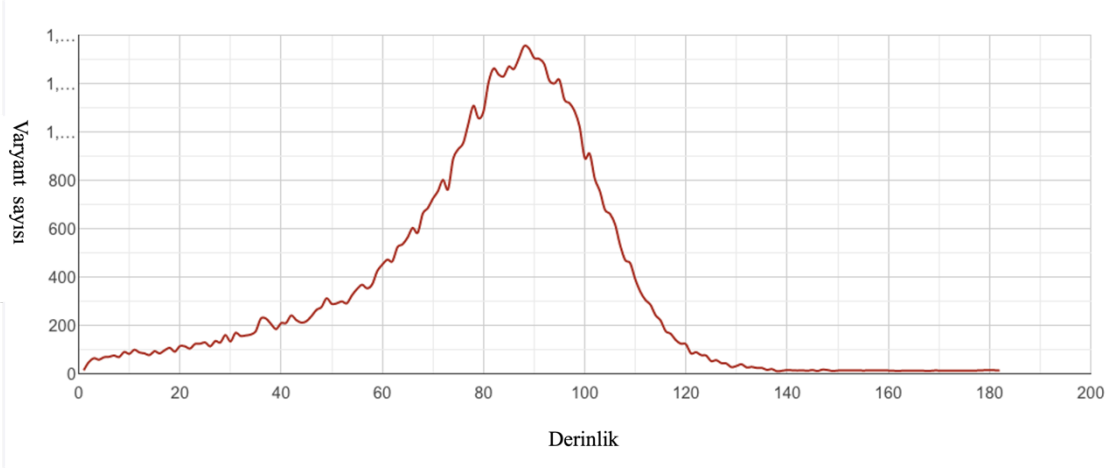
TGD çalışmasından elde edilen Olgu-1'e ait ham verilerin kalite kontrol ile ilişkili parametre sonuçları Tablo 4-4'te gösterilmiştir. Indel varyantların tek nükleotid değişimlerine (SNV) oranı 0,148'tir. Heterozigot varyantların homozigot varyantlara oranı 1,48'dir. Allelik imbalans oranı %5,53'tür. Allelik imbalans, alel fraksiyonu 0,3'ten küçük veya buna eşit ve 0,7'den büyük veya ona eşit olan heterozigot varyantların toplam yüzdesidir.

Tablo 4-4: Olgu 1'e Ait Kalite Kontrol Parametre Sonuçları

Indel/SNV Oranı	0,148*
Transisyon/Transversiyon Oranı	5,04
Heterozigot/Homozigot Oranı	1,48**
Düşük Okuma Derinliği	%0,66
Düşük Genotip Kalitesi	%1,24
Muhtemel Homozigot	%5,35
Muhtemel Heterozigot	%0,02
Alelik İmbalans	%5,53
Ortalama Derinlik	83,94x
Medyan Derinlik	84x
Maksimum Derinlik	6879x

* Referans değer 1/8'dir (0,125) [75].

**Bu değerin 1.4 ile 2.0 arasında olması beklenmektedir [76].



Şekil 4-4: Olgu 1'e Ait Okuma Derinliği Grafiği

Olgu-1'in TGD işlemi için ortalama 100X derinliğinde okuma yapam kit kullanılmıştır. Şekil 4-4'te görüldüğü üzere Olgu-1'in ortalama derinliği 90X civarındadır.

Tablo 4-5: Olgu-1 Filtreleme Ölçütleri

İnsan Fenotip Ontolojisi (HPO) terimleri	Spastik parapleji, alt ekstremitelerde hiperfleksi
Minör Alel Frekansı (MAF)	<%1
In-house Frekansı	<%1
Kalıtım Modeli	Otozomal resesif
Zigosite	Homozigot

Filtreleme ölçütleri (Tablo 4-5) girildikten sonra on binlerce varyant arasından 4 varyant filtrelenmiştir. *NFU1* geni dışında *NUS1, SNX7* ve *TRIM71* genlerinde varyant bulunmuştur. Filtreleme ölçütünde homozigot varyantlar filtrelense de klinikle ilişkili olabilecek heterozigot varyantlar da değerlendirildi. Tablo 4-6'da genlerin transkript bilgileri, varyantın konumu, yerleşim yeri, kalıtım modeli, görülme sıklığı ve varyantın olası etkisi gösterilmiştir. Varyantların kalıtım modeli ve fenotiple ilişkisi değerlendirilmiş ve anlamlı varyant ile konfirmasyon ve segregasyon çalışmalarına devam edilmiştir.

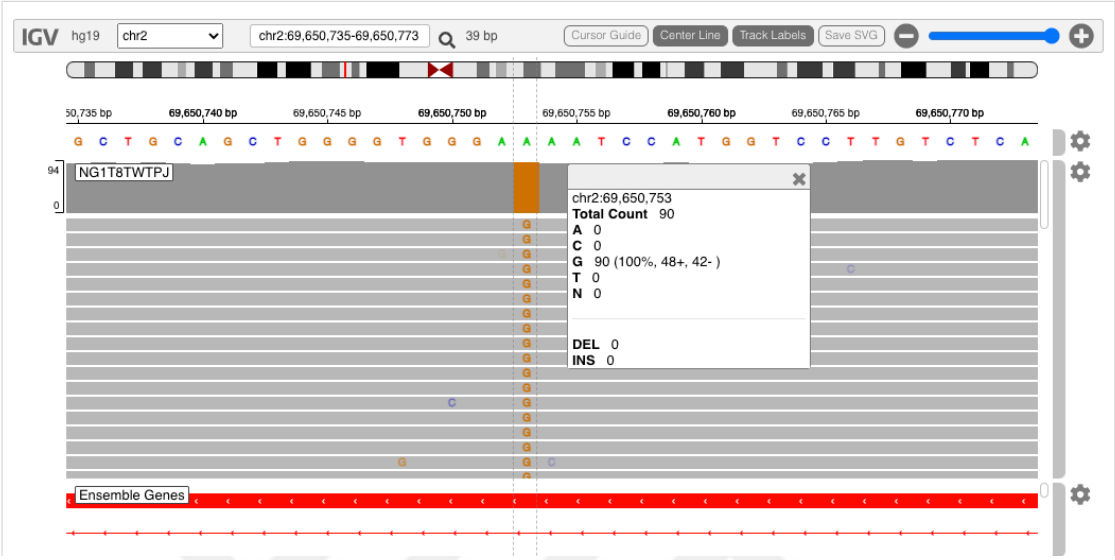
Tablo 4-6: Olgu 1'e Ait TGD verisinin biyoinformatik analizinde gözlenen klinik ilişkili olabilecek varyantlar

Gen (Transkript)	Varyant[1]	Yerleşim	Zigosite	MAF (%)[2]	Sınıflama[3]
NFU1 (ENST00000303698)	c.191T>C p.(Phe64Ser) 2:69650753	Ekzon 3/8	Homozigot	0	Önemi Belirsiz
NUS1 (ENST00000368494)	c.331G>C p.(Glu111Gln)	Ekzon 1/5	Heterozigot	0	Önemi Belirsiz
SNX7 (ENST00000306121)	c.410G>A p.(Arg137Gln)	Ekzon 3//9	Homozigot	0.0009	Önemi Belirsiz
TRIM71 (ENST00000383763)	c.1166T>C p.(Ile389Thr)	Ekzon 4/4	Heterozigot	0	Muhtemel Patojen

(1)HGSV (Human Genome Variation Society) tavsiyelerine göre isimlendirilmiştir (Dunnen ve ark 2016). (2)gnomAD'da (The Genome Aggregation Database) yer alan ekzom ve genom verilerindeki varyantların total allel frekansını (TAF) yansıtmaktadır. (3)ACMG (American College of Medical Genetics) rehberlerine (Richards ve ark. 2015) ve ACGS (Association for Clinical Genomics Science) rehberlerine (V.4.01) uygun şekilde sınıflandırıldı.

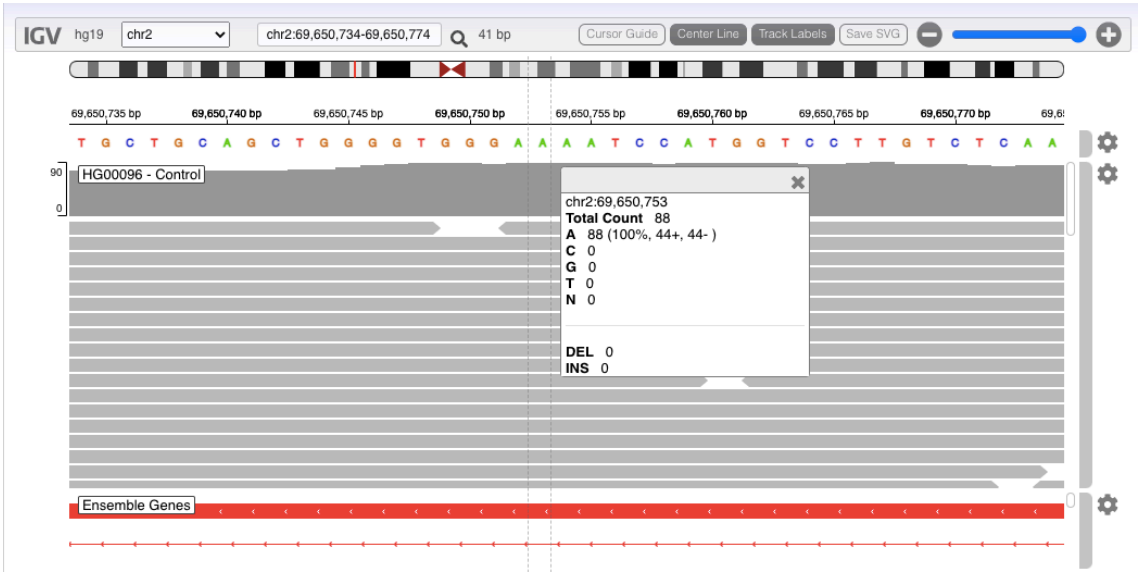
Sonuç: Olgu-1 için yapılan TGD sonucunda *NFU1* geninin 3.ekzonunda c.191T>C değişimi tespit edilmiştir. Bu değişim homozigot olarak tespit edilmiştir. 90 defa okunan bölgede tüm dizilerde okunan nükleotid Guanin nükleotididir. Referans nükleotid ise Adenin nükleotididir. Total allel frekansı (MAF) popülasyon veritabanlarında bildirilmemiştir. Bu varyant ENST00000303698 transkriptine göre yanlış anlamlı varyant şeklinde bir değişim olup, bir veya daha fazla bazı değiştiren, farklı amino asit dizisine neden olan fakat uzunluğun korunduğu bir dizi varyantı biçiminde tanımlanmaktadır. DANN, Eigen, fathmm-MKL, M-Cap, MetaLR, MetaSVM, MutPred, MutationAssesor, Mutation Taster, Provean, Polyphen2 ve SIFT gibi *in silico* araçlarda yapılan değerlendirmeye göre bu değişim; gen veya gen ürününde zarar verici bir etkiyi desteklemektedir (PP3). Bu değişim ClinVar veritabanında yer almamaktadır. Mevcut bilgiler ışığında bu varyant önemi belirsiz olarak sınıflandırılmıştır. Bu varyant incelediğimiz kadarıyla bilimsel literatürde şimdiye kadar yayımlanmamıştı. Ancak tez kapsamında elde edilen bulgularla birlikte hastamızın da dahil olduğu bir kohort çalışması tez süresi devam ederken yayınlandı. Bu çalışmada *NFU1* varyantlarının fenotip ilişkisi

ortaya konmuştur [77]. *NFU1* genindeki varyantlar, çoklu mitokondriyal disfonksiyon sendromu 1 (MMDS1) ile ilişkilendirilmiştir. Bu varyant fenotiple uyumlu olabileceğinden ailede segregasyon çalışması yapılmıştır.



Şekil 4-5: Olgu 1'e Ait TGD verisinin IGV görüntüsü

Olgu 1'e ait NFU1 geninin 3.ekzonunda yer alan ilgili varyantın Tüm Genom Dizileme verisinin Integrative Genomics Viewer (IGV) uygulamasındaki görüntüsü Şekil 4-5'te yer almaktadır. 90 defa okunan ilgili bölgede referans nükleotid Adenin (A) iken, indeks olgunun ilgili noktası 90 defa Guanin (G) olarak okunmuştur.

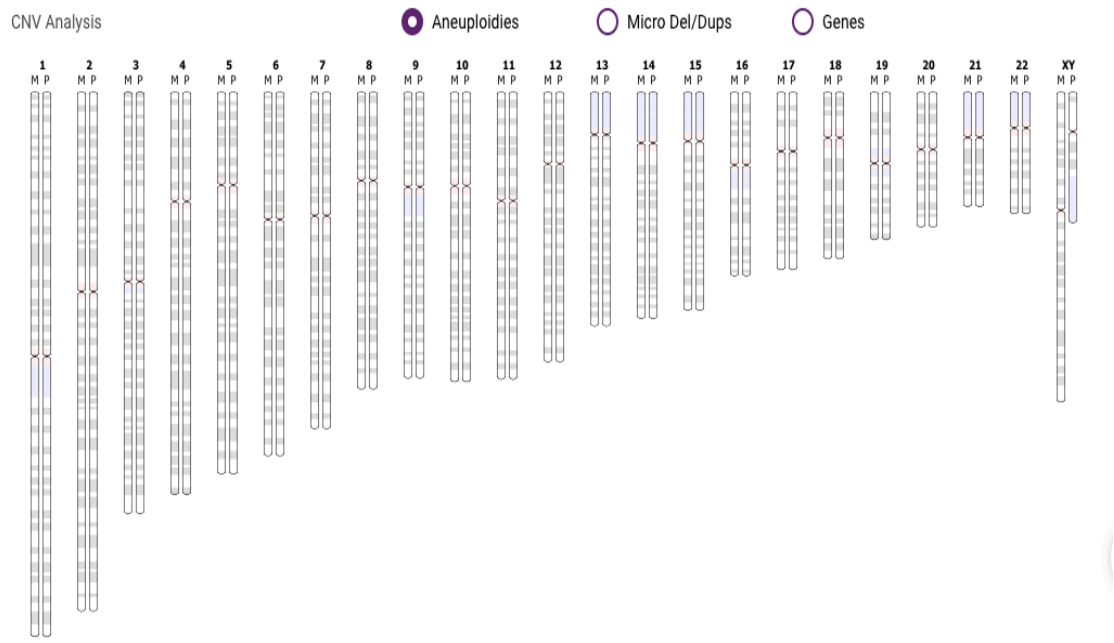


Şekil 4-6: Olgu 1'e Ait TED verisinin IGV görüntüsü

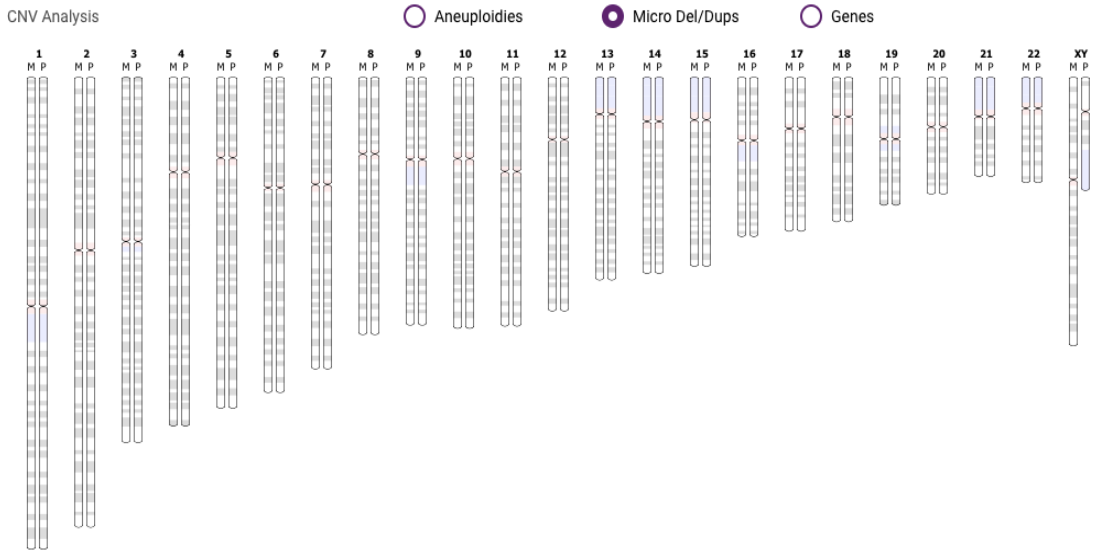
Daha önce yapılan Tüm Ekzom Dizileme verisinin Integrative Genomics Viewer (IGV) uygulamasındaki görüntüsü Şekil 4-6'da yer almaktadır. Periferik kan örneğinden DNA elde edilerek TED işlemi gerçekleştirilmiştir. 88 defa okunan ilgili bölgede referans nükleotid Adenin (A) iken, indeks olgunun ilgili noktası 88 defa Adenin (A) olarak okunmuştur. *NFU1* haricindeki varyantlar kıyas edildiğinde TED işleminde kapsanmayan bölgelerin olduğu tespit edildi. Kapsanan bölgelerde ise TED ve TGD verilerinde aynı sonuçlar tespit edildi.

4.1.2. Tüm Genom Dizileme ile CNV Analizi

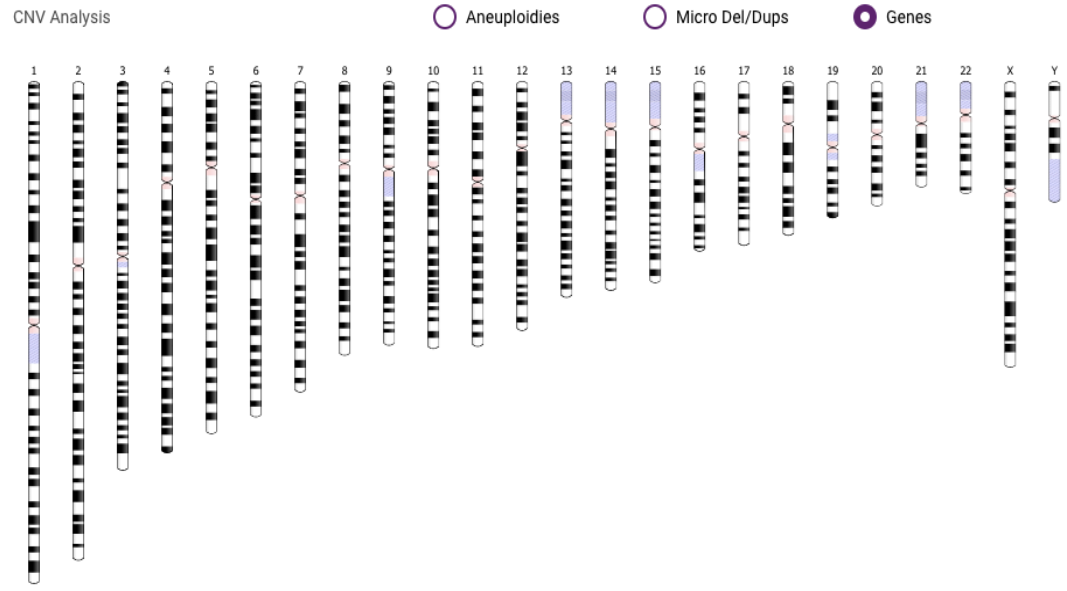
Olgu-1 ile ilgili CNV analizi yapılmış; anöploidi, mikro delesyon/duplikasyon ve gen bazında CNV'lere rastlanmamıştır (Şekiller 4-7, 4-8, 4-9).



Şekil 4-7: Olgu 1 CNV Analizi Anöploidi Sonucu



Şekil 4-8: Olgu 1 CNV Analizi Mikrodelesyon/Duplikasyon Sonucu

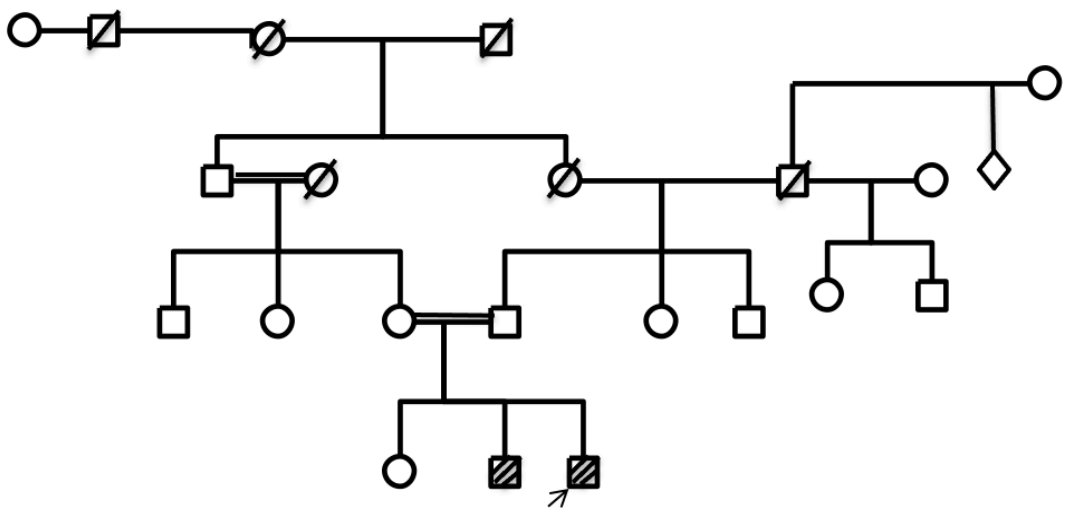


Şekil 4-9: Olgu 1 CNV Analizi Genler Sonucu

Sonuç: Yapılan CNV analizlerinde Olgu 1'e ait anöploidi, mikrodelesyon veya duplikasyon ve genler düzeyinde bir kopya sayısı değişikliğine rastlanmamıştır.

4.2. Olgu 2

Olguya ait aile ağacı Şekil 4-10'da gösterilmiştir. Birinci derece kuzen evliliği yapan Olgu-2'nin ebeynleri, ebeveynlerin anne-babası, kardeşleri ve dede-ninelerine ait bilgiler aile ağacında gösterilmiştir. Olgu-2'nin bir etkilenmiş kardeşi ve bir de sağlıklı kardeşi bulunmaktadır.



Şekil 4-10: Olgu 2'ye Ait Aile Ağacı

Ebeveynleri birinci derece kuzen olan 8 yaşındaki erkek Olgu-2, multipl konjenital anomalileri ve mental retardasyonu nedeniyle genetik incelemeye alınmıştır. Aile, 6-7 aylıkken ateşli hastalık ve gribal enfeksiyon geçiren olgunun, 8 aylıkken baba kelimesini kullanmış olmasına rağmen daha sonra hiçbir kelime söylemediğini ifade etmiştir. Bir yaşından itibaren nöbet geçiren olgumuzun enfeksiyon sonrası emeklemesinin durduğu ve kasılmaların başladığı ailesi tarafından bildirilmiştir. Fizik muayenesinde baş çevresi 47cm, DTR'leri hafif artmış, babinski negatif olarak belirlendi. Üriner ve fekal inkontinans mevcut olan olguda serebral palsi ön tanısı düşünülmüştür. Olgu-2'nin benzer bulguları olan etkilenmiş bir kardeşi ve bir de sağlıklı kardeşi bulunmaktadır. Olgunun bulguları Tablo 4-7'de yer almaktadır.

Tablo 4-7: Olgu 2'ye Ait Klinik ve Laboratuvar Bulguları

Olgunun Yaşı/Cinsiyeti	8 yaş/Erkek
Ebeveyn Akrabalık Durumu	1.Derece Kuzem
Anne Gebelik Öyküsü/Olgu Doğum	G3P3AO / G3
Ailede Benzer Öykü	Etkilenmiş Kardeş/Epliepsi PSMG
Fetal USG Bulgusu	-
Perinatal Hikaye	Seksio mükerrer, 3300 g, Küvöz bakımı yok
Ön tanı	Serebral Palsi

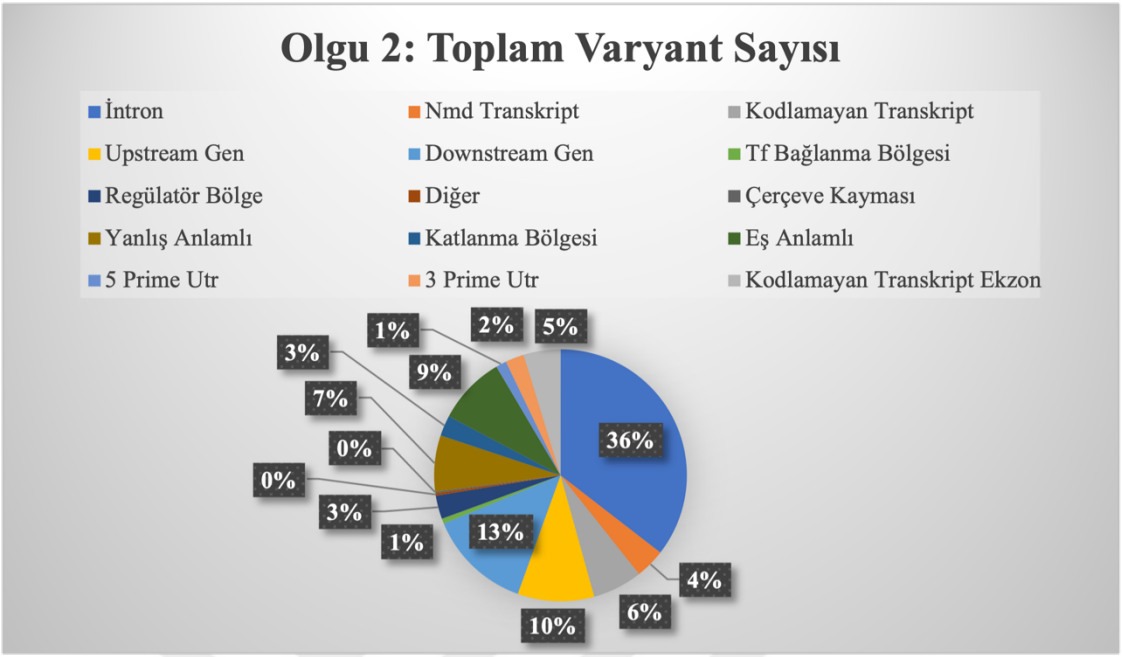
4.2.1. Tüm Genom Dizileme Sonuçları

Olgu 2'nin yanak mukozasından alınan örnekten DNA izolasyonu yapıldıktan sonra ortalama 100X okuma derinliği olan kit ile tüm genom dizileme işlemi yapılmıştır. Dizilime sonucunda elde edilen toplam varyant sayıları Tablo 4-8'de ve Şekil 4-11'de gösterilmiştir.

Tablo 4-8: Olgu 2'ye Ait TGD Verisinden Biyoinformatik Analiz ile Elde Edilen Tüm Varyant Sayıları

Toplam Varyant Sayısı	754.443
İntron	267.892
Nmd Transkript	28.845
Kodlamayan Transkript	47.976
Upstream Gen	73.962
Downstream Gen	99.881
Tf Bağlanma Bölgesi	5.008
Regülatör Bölge	22.699
Diğer	2.509
Çerçeve Kayması	1.501
Yanlış Anlamlı	54.215
Katlanma Bölgesi	19.895
Eş Anlamlı	66.190
5 Prime Utr	10.249
3 Prime Utr	18.077

Biyoinformatik analiz sonucunda toplam 754.443 varyant elde edilmiştir. Bu varyantların 267.892 tanesi intronik, 99.881 tanesi downstream gen ve 73.962 tanesi ise upstream gen varyantları oluşturmaktadır.

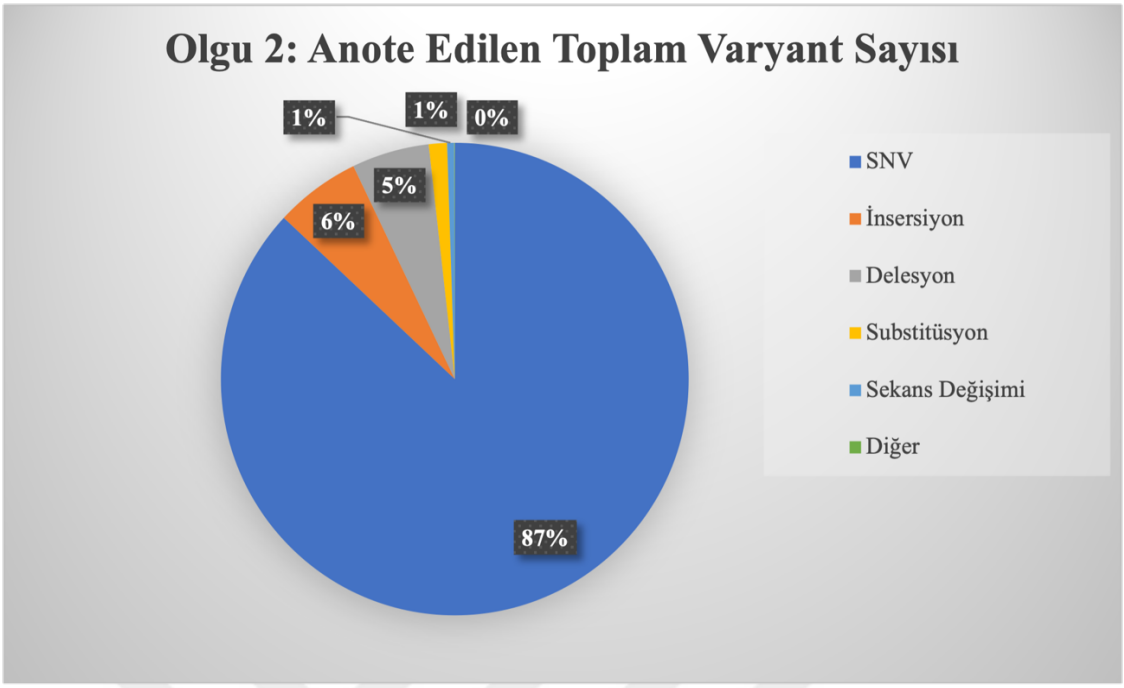


Şekil 4-11: Olgu 2'ye Ait TGD Verisinden Biyoinformatik Analiz ile Elde Edilen Tüm Varyantlar

TGD verisinden biyoinformatik analizler ile anote edilen varyant sayısı 59.841 olarak belirlenmiştir. Bunlardan 52.035'ini (%87) SNV, 3537'ini (%6) insersiyon, 3200'ünü (%5) delesyon, 753'ünü (%1) yer değiştirme, 291'ini (%1) sekans değişimi oluşturmaktadır. Diğer anote edilen varyantlar ise 25 tanedir (Tablo 4-9 ve Şekil 4-12).

Tablo 4-9: Olgu 2'ye Ait TGD Verisinden Biyoinformatik Analiz ile Anote Edilen Varyant Sayıları

Anote Edimiş Toplam Varyant Sayısı	59.841
SNV	52.035
İnsersiyon	3.537
Delesyon	3.200
Substitüsyon	753
Sekans Değişimi	291
Diğer	25



Şekil 4-12: Olgu 2'ye Ait TGD Verisinden Biyoinformatik Analiz ile Anote Edilen Varyantlar

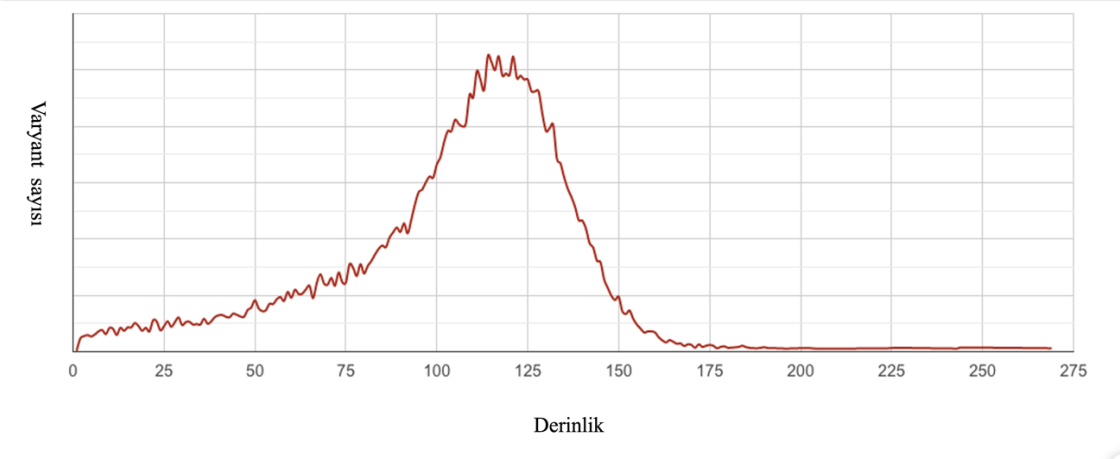
TGD çalışmasından elde edilen Olgu-2'ye ait ham verilerin kalite kontrol ile ilişkili parametre sonuçları Tablo 4-10'da gösterilmiştir. Indel varyantların tek nükleotid değişimlerine (SNV) oranı 0,150'dir. Heterozigot varyantların homozigot varyantlara oranı 1,49'dur. Alelik imbalans oranı %4,95'tir.

Tablo 4-10: Olgu 2'ye Ait Kalite Kontrol Parametre Sonuçları

Indel/SNV Oranı	0.150*
Transisyon/Transversiyon Oranı	5.02
Heterozigot/Homozigot Oranı	1.49**
Düşük Okuma Derinliği	%0.58
Düşük Genotip Kalitesi	%1.01
Muhtemel Homozigot	%5.16
Muhtemel Heterozigot	%0.02
Alelik İmbalans	%4.95
Ortalama Derinlik	109.57x
Medyan Derinlik	111x
Maksimum Derinlik	7114x

* Referans değer 1/8'dir (0,125)[75].

**Bu değerin 1.4 ile 2.0 arasında olması beklenmektedir [76].



Şekil 4-13: Olgu 2'ye Ait Okuma Derinliği Grafiği

Olgu-2'nin TGD işlemi için ortalama 100X derinliğinde okuma yapam kit kullanılmıştır. Şekil 4-13'te görüldüğü üzere Olgu-2'nin ortalama derinliği 120X civarındadır.

Tablo 4-11: Olgu-2'nin Filtreleme Ölçütleri

İnsan Fenotip Ontolojisi (HPO) terimleri	Mental retardasyon, motor gecikme, nöbetler, konuşma yokluğu, uzun yüz, makrotiya, yürüyememe, serebral palsi
Minör Alel Frekansı (MAF)	<%1
In-house Frekansı	<%1
Kalıtım Modeli	Otozomal resesif
Zigosite	Homozigot

Olgu-2 için yapılan TGD sonucunda filtreleme ölçütleri (Tablo 4-11) girildikten sonra on binlerce varyant arasından 5 varyant filtrelenmiştir. *ATP8A2* geni dışında *ARID1A, PIK3R1, ATP5F1A* ve *ARX* genlerinde varyant bulunmuştur. Tablo 4-12'da genlerin transkript bilgileri, varyantın konumu, yerleşim yeri, kalıtım modeli, görülme sıklığı ve varyantın olası etkisi gösterilmiştir. Varyantların kalıtım modeli ve fenotiple ilişkisi değerlendirilmiş ve anlamlı varyant ile konfirmasyon ve segregasyon çalışmalarına devam edilmiştir.

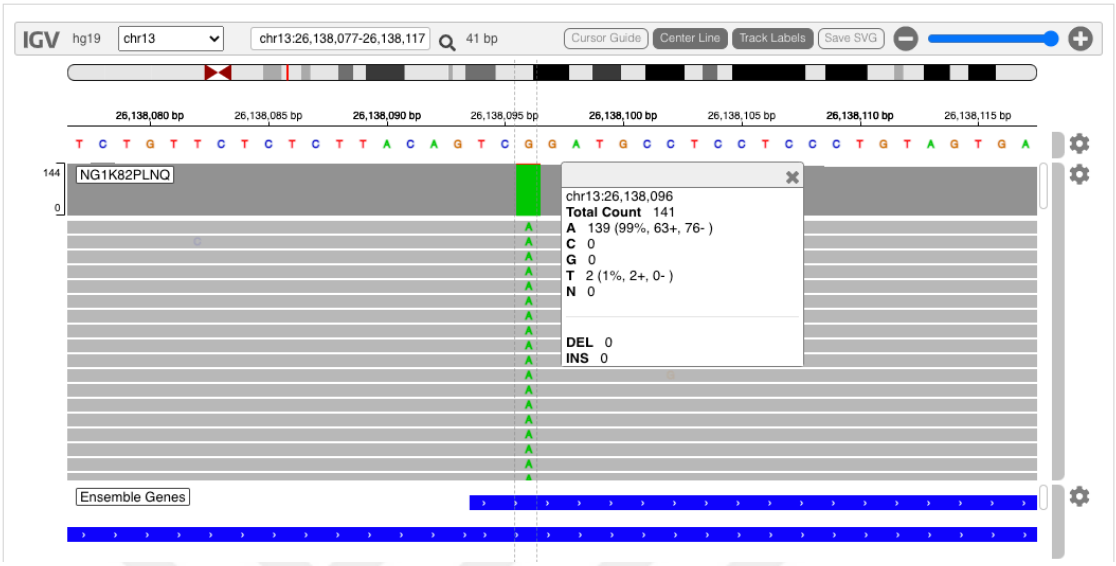
Tablo 4-12: Olgu 2'ye Ait TGD verisinin biyoinformatik analizinde gözlenen klinik ilişkili olabilecek varyantlar

Gen (Transkript)	Varyant[1]	Yerleşim	Zigosite	MAF (%)[2]	Sınıflama[3]
ATP8A2 (ENST00000381655)	c.1400G>A p.(Arg467Gln) 13:26138096	Ekzon 16/37	Homozigot	0,01653	Önemi Belirsiz
ARID1A (ENST00000324856)	c.448T>C p.(Phe150Leu) 1:27023342	Ekzon 1/20	Heterozigot	0	Önemi Belirsiz
PIK3R1 (ENST00000274335)	c.814C>T p.(Leu272Phe) 5:67576535	Ekzon 5/15	Heterozigot	0	Önemi Belirsiz
ATP5F1A (ENST00000282050)	c.595_596insA p.(Leu199TyrfsTer10) 18:43669586	Ekzon 6/13	Heterozigot	0	Patojenik
ARX (ENST00000379044)	c.434C>G p.(Ala145Gly) X:25031678	Ekzon 2/5	Hemizigot	0	Önemi Belirsiz

(1)HGSV (Human Genome Variation Society) tavsiyelerine göre isimlendirilmiştir (Dunnen ve ark. 2016). (2)gnomAD'da (The Genome Aggregation Database) yer alan ekzom ve genom verilerindeki varyantların total allel frekansını (TAF) yansıtmaktadır. (3)ACMG (American College of Medical Genetics) rehberlerine (Richards ve ark. 2015) ve ACGS (Association for Clinical Genomics Science) rehberlerine (V.4.01) uygun şekilde sınıflandırıldı.

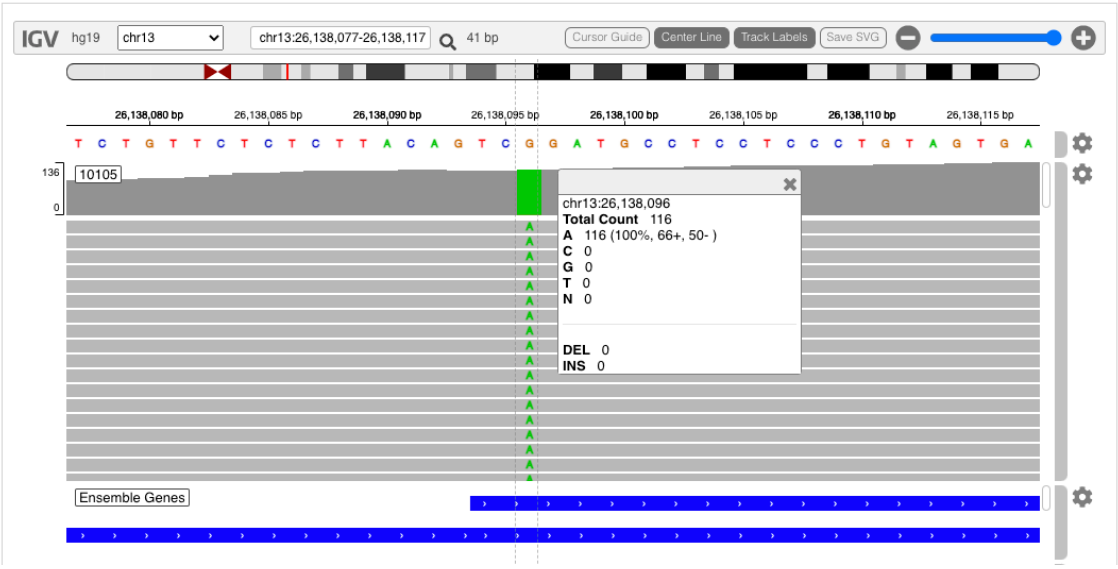
Sonuç: *ATP8A2* geninin 16.ekzonunda c.1400G>A değişimi tespit edilmiştir. Bu değişim homozigot olarak tespit edilmiştir. Okunan bölgede neredeyse tüm dizilerde okunan nükleotid Adenin nükleotididir. Referans nükleotid ise Guanin nükleotididir. Total allel frekansı (MAF) popülasyon veritabanlarında %0.016 olarak bildirilmiştir. Bu varyant ENST00000381655 transkriptine göre yanlış anlamlı varyant şeklinde bir değişim olup, bir veya daha fazla bazı değiştiren, farklı amino asit dizisine neden olan fakat uzunluğun korunduğu bir dizi varyantı biçiminde tanımlanmaktadır. Tespit edilen varyant, ACMG kriterlerine göre önemi belirsiz (VUS), MutationTaster tahmin aracına göre patojenik olan homozigot varyant olarak saptanmıştır. Bu varyant fenotiple uyumlu olabileceğinden ailede segregasyon çalışması yapılmıştır. *ATP8A2* genindeki varyantlar, germline patojenik varyantları otozomal resesif kalıtım gösteren serebellar ataksi, zeka

geriliği ve denge kuramama sendromu 4 olarak adlandırılan (CAMRQ4; OMIM: 615268) bir hastalıkla ilişkilendirilmiştir.



Şekil 4-14: Olgu 2'e Ait TGD verisinin IGV görüntüsü

Olgu 2'ye ait *ATP8A2* geninin 16.ekzonunda yer alan ilgili varyantın Tüm Genom Dizileme verisinin Integrative Genomics Viewer (IGV) uygulamasındaki görüntüsü Şekil 4-14'te yer almaktadır. 141 defa okunan ilgili bölgede referans nükleotid Guanin (G) iken, indeks olgunun ilgili noktası 139 defa Adenin (A) olarak okunmuştur.



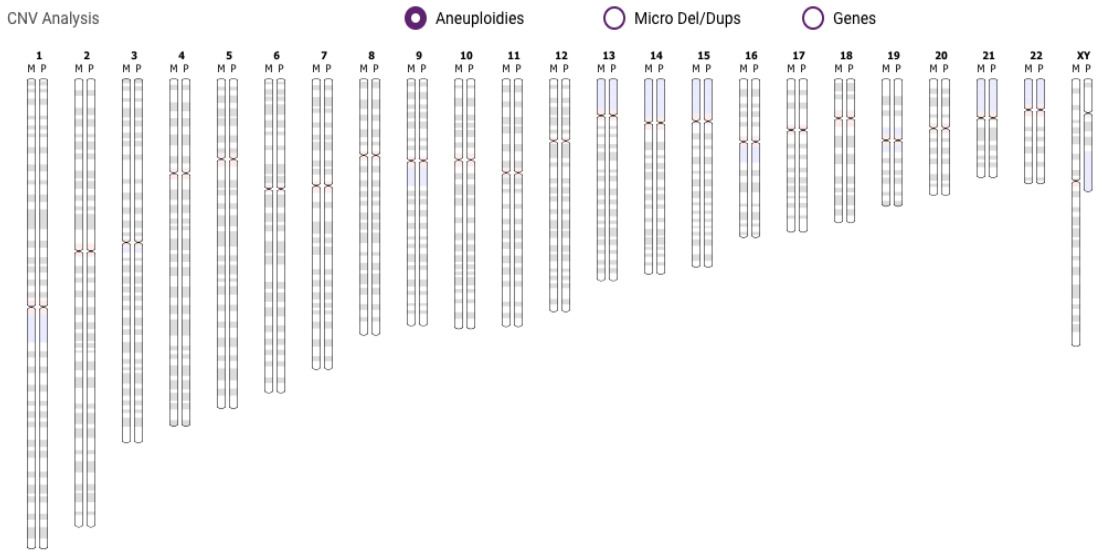
Şekil 4-15: Olgu 2'ye Ait TED verisinin IGV görüntüsü

Olgu 2'ye ait *ATP8A2* geninin 16.ekzonunda yer alan ilgili varyantın Tüm Ekzom Dizileme verisinin Integrative Genomics Viewer (IGV) uygulamasındaki görüntüsü Şekil

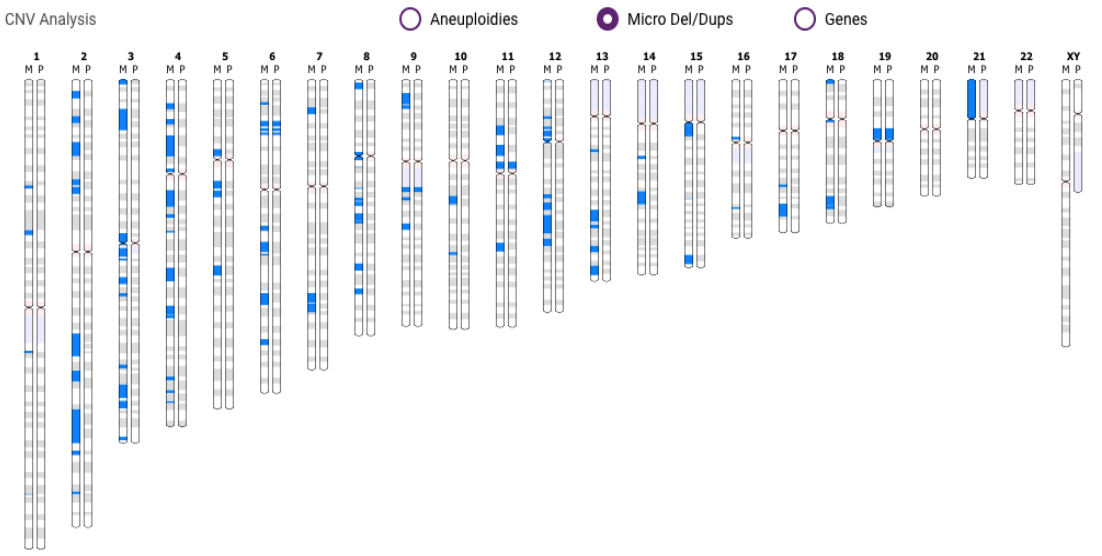
4-15'te yer almaktadır. 116 defa okunan ilgili bölgede referans nükleotid Guanin (G) iken, indeks olgunun ilgili noktası 116 defa Adenin (A) olarak okunmuştur.

4.2.2. Tüm Genom Dizileme ile CNV Analizi

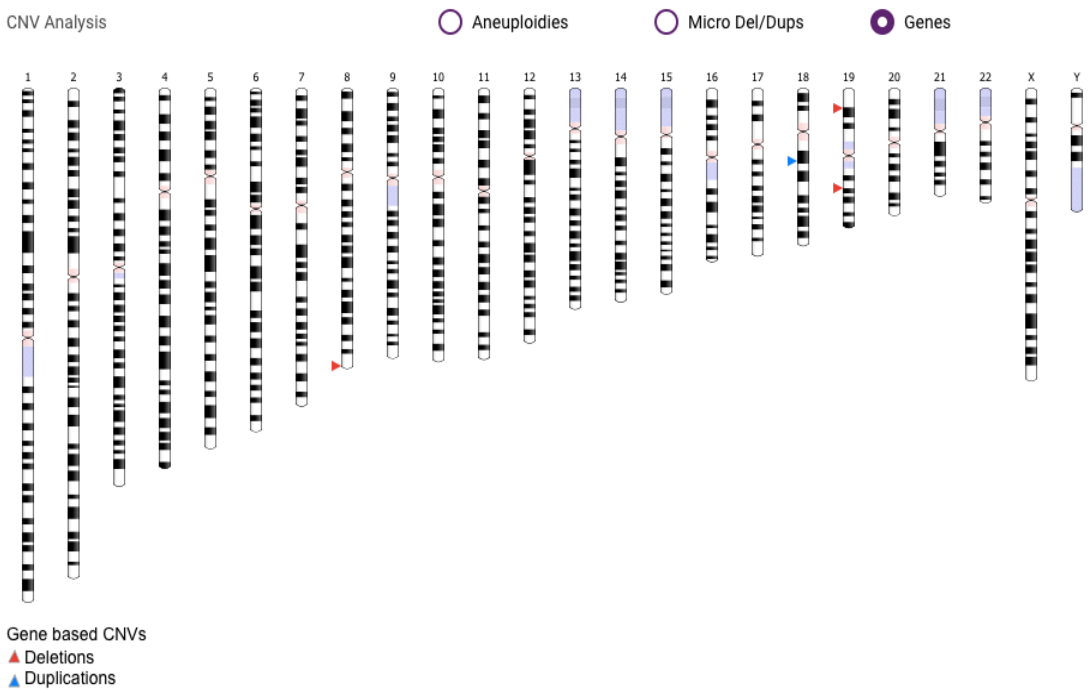
Olgumuzla ilgili CNV analizi yapılmış; anöploidi, mikro delesyon/duplikasyon ve gen bazında CNV'lere rastlanmamıştır. (Şekil 4-16, 4-17 ve 4-18)



Şekil 4-16: Olgu 2 CNV Analizi Anöploidi Sonucu



Şekil 4-17: Olgu 2 CNV Analizi Mikrodelesyon/Duplikasyon Sonucu

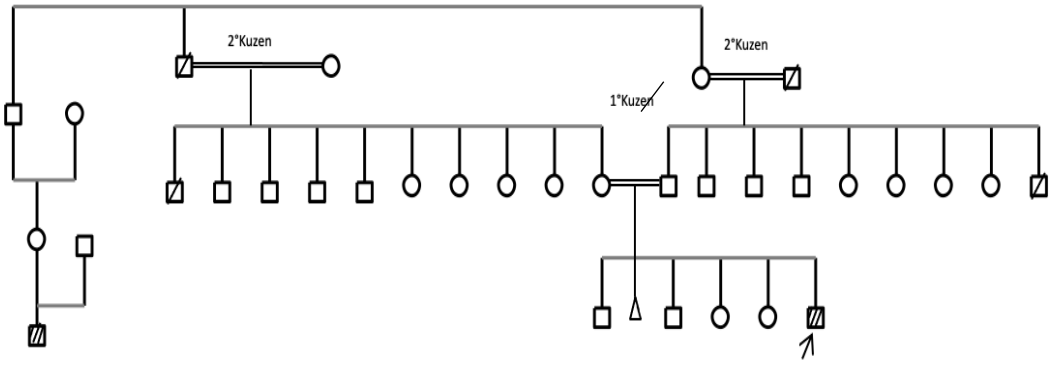


Şekil 4-18: Olgu 2 CNV Analizi Genler Sonucu

Sonuç: Yapılan CNV analizlerinde Olgu 1'e ait anöploidi, mikrodelesyon veya duplikasyon ve genler düzeyinde bir kopya sayısı değişikliğine rastlanmamıştır.

4.3. Olgu 3

Olguya ait aile ağacı Şekil 4-19'da gösterilmiştir. Birinci derece kuzen evliliği yapan Olgu-3'ün ebeveynleri, ebeveynlerin anne-babası ve kardeşleri aile ağacında gösterilmiştir. Olgu-3'ün üç sağlıklı kardeşi, bir vefat eden kardeşi bir de düşük bulunmaktadır.



Şekil 4-19: Olgu 3'e Ait Aile Ağacı

3 yaşında ve erkek olan Olgu-3'ün gelişme geriliği ve dismorfik görünümü bulunmaktadır. Fizik muayenesinde 8.6 kg, boyu 84 cm, baş çevresi 42cm, her iki testis inguinal kanal yerleşimli olarak tespit edilmiştir. Trigonosefali, bilateral epikantus, medialde seyrek kaşlar, bifid burun ucu, sağ kulak belirgin, ince üst dudak, çenede vertikal çizgilenme ve dimple; dismorfik bulgular arasındadır. EEG'sinde normal uyku zemininde sağ temporo-parieto-oksipitalde birkaç kez ortaya çıkan düşük amplitüdlü yavaş diken dalga deşarjı ve sağ temporooksipitoparietal bölgede seyrek olarak ortaya çıkan fokal epileptiform anomalisi tespit edilmiştir. Karyotip analizi XY'dir. Array-CGH testinde klinik bulgularla ilişkilendirilmiş bir kopya sayısı değişikliği tespit edilememiştir. Konjenital mikrosefalisi olan Olgu-3'ün bilişsel gelişim geriliği ve kriptoorşidizmi mevcuttur. Ayrıca eklem laksitesi, parmak pozisyon anomalisi aksiyal hipotonisite, ekstremitelerde hipertonisite, bruksizm ve konvülziyon öyküsü vardır. Üç sağlıklı kardeşi olan olgu için daha önce yapılan genetik tetkiklerden herhangi anlamlı bir varyant tespit edilememiştir. Olgu-3'e ait klinik ve laboratuvar bulgular Tablo 4-13'te gösterilmiştir.

Tablo 4-13: Olgu 3'e Ait Klinik ve Laboratuvar Bulguları

Olgunun Yaşı/Cinsiyeti	2 yaş 10 ay/Erkek
Ebeveyn Akrabalık Durumu	1.Derece Kuzen
Anne Gebelik Öyküsü/Olgu Doğum	G7P5Ab2 / G7
Ailede Benzer Öykü	G1, 3yaşında konvülziyon sonrası eks
Fetal USG Bulgusu	Antenatal II.düzey USG yapılmış, sorun izlenmemiş
Perinatal Hikaye	Miadında, Sezaryen, 3830 g, 52cm boy, 34cm Baş çevresi
Ön tanı	Primer Mikrosefali

Olgu 3 için lenfosit kültürü yapılarak metafazda kromozom incelenmesi planlanmıştı. Ancak iki ayrı zamanda olgudan alınan kan ile kültür çalışması yapılmış fakat lenfosit üremesi gözlenmemiştir. Dolayısıyla kromozom analizi yapılamamıştır. Olgunun immün yetmezliği olduğundan olgu devamlı enfeksiyonel hastalık geçirmekte ve çok sık olarak antibiyotik tedavisi almaktadır. İki ayrı dönemde kan örneği alınmasına rağmen lenfosit kültürünün üretememesine sebep olarak antibiyotik kullanımı düşünülmüştür. Çalışma süresi, olgudan iki kez kan örneği alınmasından ötürü ailenin hassasiyeti ve proje bütçesi göz önünde bulundurulduğunda tekrar örnek alınması mümkün olmamıştır.

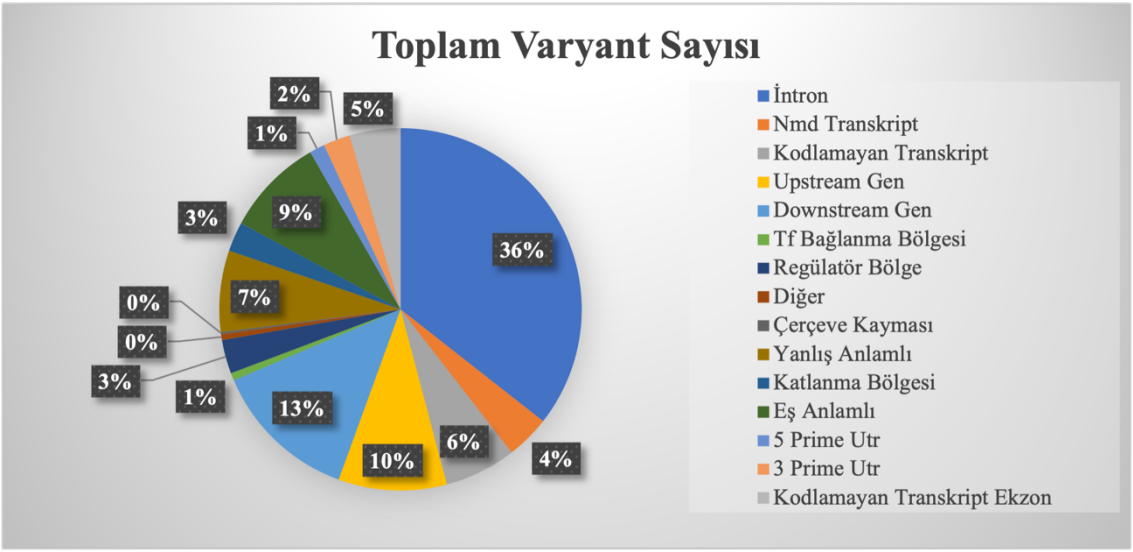
4.3.1. Tüm Genom Dizileme Sonuçları

Olgu 3'ün yanak mukozasından alınan örnekten DNA izolasyonu yapıldıktan sonra ortalama 100X okuma derinliği olan kit ile tüm genom dizileme işlemi yapılmıştır. Kalite kontrol basamağından geçmeyen örneğimizden ötürü yeni kit alınmıştır. Alınan kit ortalama 30X okuma derinliğinde olduğundan okuma sonuçları diğer olgulara göre farklılık göstermiştir. Dizilime sonucunda elde edilen toplam varyant sayıları Tablo 4-14'de ve Şekil 4-20'de gösterilmiştir.

Tablo 4-14: Olgu 3'e Ait TGD Verisinden Biyoinformatik Analiz ile Elde Edilen Tüm Varyant Sayıları

Toplam Varyant Sayısı	734.252
İntron	261.319
Nmd Transkript	28.706
Kodlamayan Transkript	46.864
Upstream Gen	70.980
Downstream Gen	95.805
Tf Bağlanma Bölgesi	4.897
Regülatör Bölge	22.159
Diğer	3.784
Çerçeve Kayması	1.406
Yanlış Anlamlı	53.339
Katlanma Bölgesi	19.358
Eş Anlamlı	64.727
5 Prime Utr	9.981
3 Prime Utr	17.628
Kodlamayan Transkript Ekzon	33.299

Olgu-3'ün TGD verisine ait biyoinformatik analiz sonucunda toplam 734.252 varyant elde edilmiştir. Bu varyantların %36' lık kısmını (261.319) intronik, %13' lük kısmını (95.805) downstream gen ve %10' luk kısmını (70.980) ise upstream gen varyantları oluşturmaktadır.

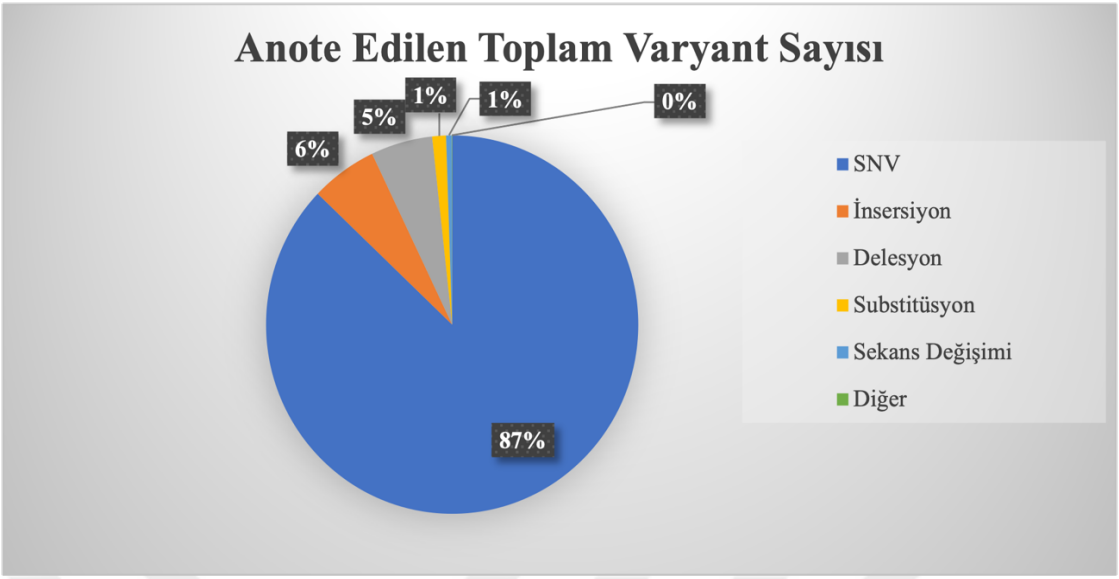


Şekil 4-20: Olgu 3'e Ait TGD Verisinden Biyoinformatik Analiz ile Elde Edilen Tüm Varyantlar

TGD verisinden biyoinformatik analizler ile anote edilen varyant sayısı 58.074 olarak belirlenmiştir. Bunlardan 50.612'ini (%87) SNV, 3353'ünü (%6) insersiyon, 3111'ini (%5) delesyon, 695'ini (%1) yer değiştirme, 279'unu (%1) sekans değişimi oluşturmaktadır. Diğer anote edilen varyantlar ise 24 tanedir.

Tablo 4-15: Olgu 3'e Ait TGD Verisinden Biyoinformatik Analiz ile Anote Edilen Varyant Sayıları

Anote Edimiş Toplam Varyant Sayısı	58.074
SNV	50.612
İnsersiyon	3.353
Delesyon	3.111
Substitüsyon	695
Sekans Değişimi	279
Diğer	24



Şekil 4-21: Olgu 3'e Ait TGD Verisinden Biyoinformatik Analiz ile Anote Edilen Varyantlar

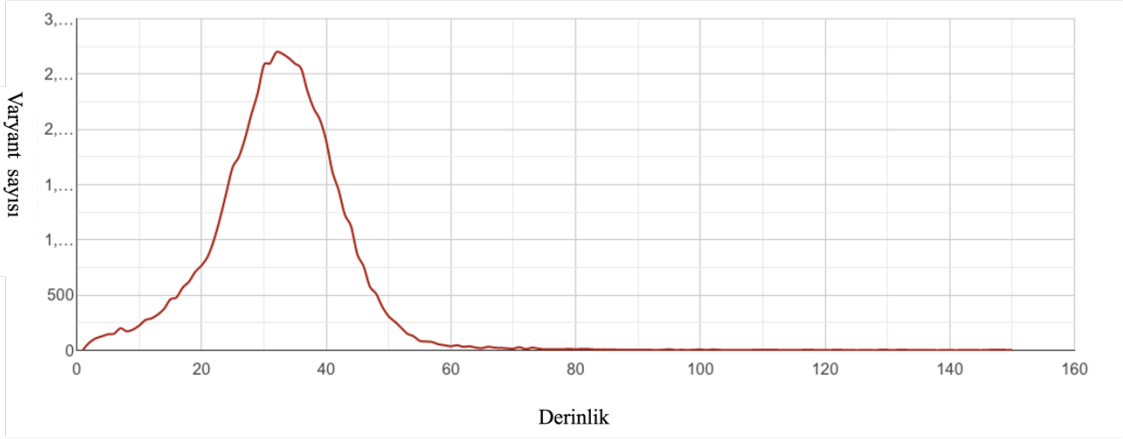
TGD çalışmasından elde edilen Olgu-3'e ait ham verilerin kalite kontrol ile ilişkili parametre sonuçları Tablo 4.14'te gösterilmiştir. Indel varyantların tek nükleotid değişimlerine (SNV) oranı 0,147'dir. Heterozigot varyantların homozigot varyantlara oranı 1,43'tür. Alelik imbalans oranı %5,11'dir.

Tablo 4-16: Olgu 3'e Ait Kalite Kontrol Parametre Sonuçları

Indel/SNV Oranı	0.147*
Transisyon/Transversiyon Oranı	5.05
Heterozigot/Homozigot Oranı	1.43**
Düşük Okuma Derinliği	%1.37
Düşük Genotip Kalitesi	%2.8
Muhtemel Homozigot	%1.97
Muhtemel Heterozigot	%0.02
Alelik İmbalans	%5.11
Ortalama Derinlik	34.01x
Medyan Derinlik	33x
Maksimum Derinlik	4384x

* Referans değer 1/8'dir (0,125) [75].

**Bu değerin 1.4 ile 2.0 arasında olması beklenmektedir [76].



Şekil 4-22: Olgu 3'e Ait Okuma Derinliği Grafiği

Olgu-3'ün TGD işlemi için ortalama 30X derinliğinde okuma yapam kit kullanılmıştır. Şekil 4.22'de görüldüğü üzere Olgu-2'nin ortalama derinliği 35X civarındadır.

Tablo 4-17: Olgu-3'ün Filtreleme Parametreleri

İnsan Fenotip Ontolojisi (HPO) terimleri	Gelişme geriliği, bilişsel gelişim geriliği, konjenital mikrosefali, kriptoorşidizmi dismorfik yüz bulguları
Minör Alel Frekansı (MAF)	<%1
In-house Frekansı	<%1
Kalıtım Modeli	Otozomal resesif
Zigosite	Homozigot

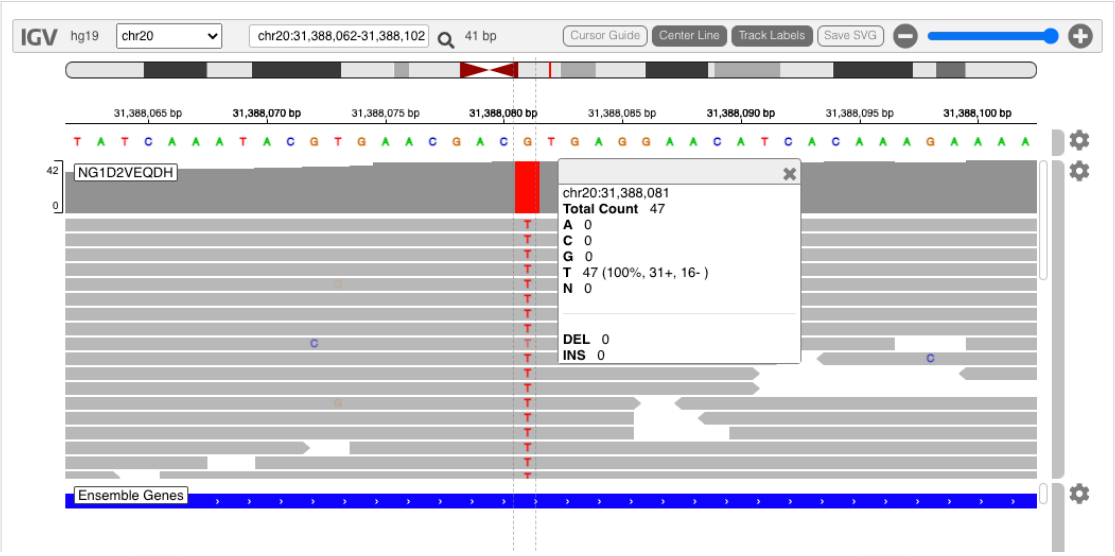
Olgu 3 için yapılan TGD işleminde filtreleme ölçütleri (Tablo 4-17) girildikten sonra on binlerce varyant arasından 4 varyant filtrelenmiştir. *DNMT3B* geni dışında *HSPA9, SOX18* ve *FIG4* genlerinde varyant bulunmuştur. Tablo 4-18'de genlerin transkript bilgileri, varyantın konumu, yerleşim yeri, kalıtım modeli, görülme sıklığı ve varyantın olası etkisi gösterilmiştir. Varyantların kalıtım modeli ve fenotiple ilişkisi değerlendirilmiş ve anlamlı varyant ile konfirmasyon ve segregasyon çalışmalarına devam edilmiştir.

Tablo 4-18: Olgu 3'e Ait TGD verisinin biyoinformatik analizinde gözlenen klinik ilişkili olabilecek varyantlar

Gen (Transkript)	Varyant[1]	Yerleşim	Zigosite	MAF (%)[2]	Sınıflama[3]
DNMT3B (ENST00000328111)	c.1882G>T p.(Val628Leu) 20:31388081	Ekzon 17/23	Homozigot	0	Önemi Belirsiz
HSPA9 (ENST00000297185)	c.1535C>G p.(Pro512Arg) 5:137893656	Ekzon 13/17	Heterozigot	0.0144	Önemi Belirsiz
SOX18 (ENST00000340356)	c.438G>C p.(Leu146Phe) 20:62680236	Ekzon 2/2	Heterozigot	0	Önemi Belirsiz
FIG4 (ENST00000230124)	c.834A>T p.(Lys278Asn) 6:110062705	Ekzon 8/23	Heterozigot	0.3977	Önemi Belirsiz

(1)HGSV (Human Genome Variation Society) tavsiyelerine göre isimlendirilmiştir (Dunnen ve ark. 2016). (2)gnomAD'da (The Genome Aggregation Database) yer alan ekzom ve genom verilerindeki varyantların total allel frekansını (TAF) yansıtmaktadır. (3)ACMG (American College of Medical Genetics) rehberlerine (Richards ve ark. 2015) ve ACGS (Association for Clinical Genomics Science) rehberlerine (V.4.01) uygun şekilde sınıflandırıldı.

Sonuç: Olgu-3 için yapılan TGD sonucunda *DNMT3B* geninin 17.ekzonunda c.1882G>T p.(Val628Leu) değişimi tespit edilmiştir. Bu değişim homozigot olarak tespit edilmiştir. 47 defa okunan bölgede tüm dizilerde okunan nükleotid Timin nükleotididir. Referans nükleotid ise Guanin nükleotididir. Bu varyant ENST00000328111 transkriptine göre yanlış anlamlı varyant şeklinde bir değişim olup, bir veya daha fazla bazı değiştiren, farklı bir amino asit dizisine neden olan fakar uzunluğunun korunduğu bir dizi varyantı biçiminde tanımlanmaktadır. Bu varyant ACMG kriterlerine göre önemi belirsiz (VUS), 6 farklı tahmin aracına göre patojenik olan homozigot bir varyant olarak saptanmıştır. *DNMT3B* genindeki patojenik varyantlar, otozomal resesif kalıtım gösteren immünyetmezlik sentromik instabilite-yüz anomalileri sendromu 1 (ICF1; OMIM: 242860) ile ilişkilendirilmiştir.



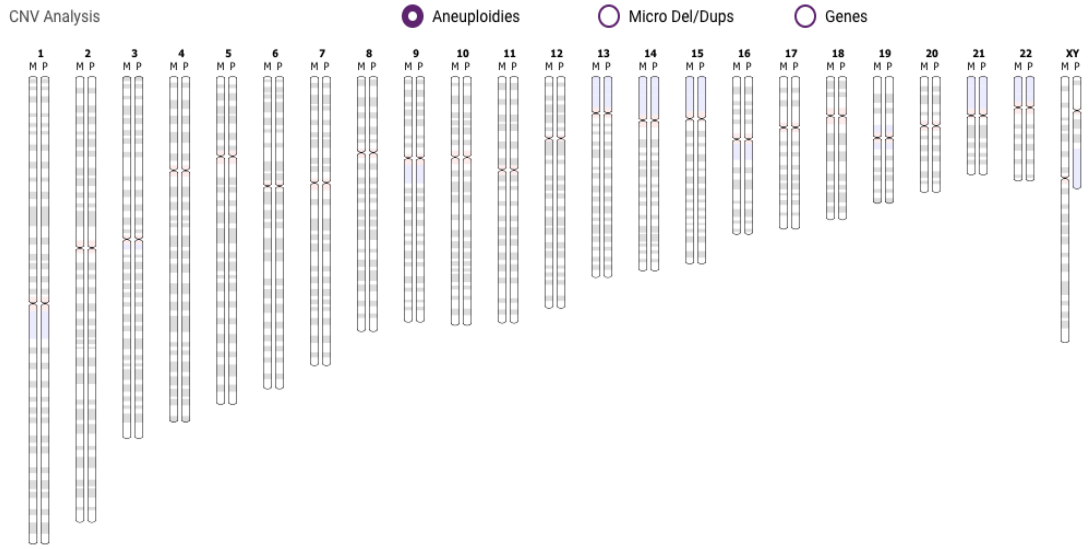
Şekil 4-23: Olgu 3'e ait TGD verisinin IGV görüntüsü

Olgu 3'e ait *DNMT3B* geninin 17.ekzonunda yer alan ilgili varyantın Tüm Genom Dizileme verisinin Integrative Genomics Viewer (IGV) uygulamasındaki görüntüsü Şekil 4.23'de yer almaktadır. 47 defa okunan ilgili bölgede referans nükleotid Guanin (G) iken, indeks olgunun ilgili noktası 47 defa Timin (T) olarak okunmuştur.

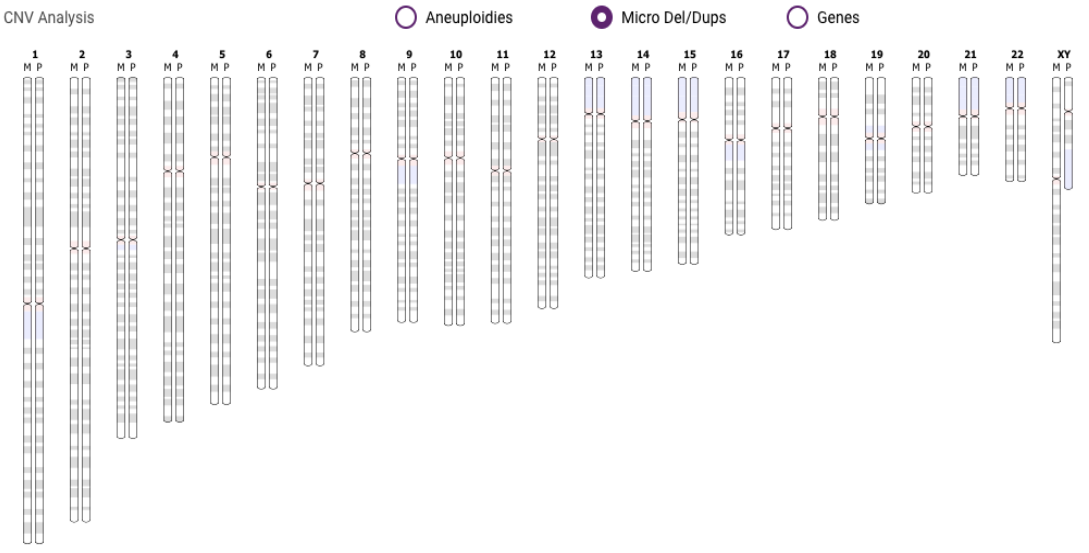
Olgu-3'e ait başka bir merkezde (Biruni Üniversitesi) tapılan TED raporuna ulaşılmıştır. Periferik kandan DNA izole edilerek yapılan çalışma TWIST 36Mb capture kiti ile MGI DNBSEQ-G400 cihazında gerçekleştirilmiştir. 169,45 ortlama derinlik ile 62,654,344 temiz okuma alınmıştır. Hastanın klinik bulguları ie ilşikilendirilmiş bir genetik değişimin saptanmadığı bildirilmiştir. TED işlemine ait ham veriler ilgili merkez ve aileden talep edilmiş ancak temin edilememiştir. Bu nedenle hastanın TED ham verisinden tekrar analiz yapılamamıştır.

4.3.2. Tüm Genom Dizileme ile CNV Analizi

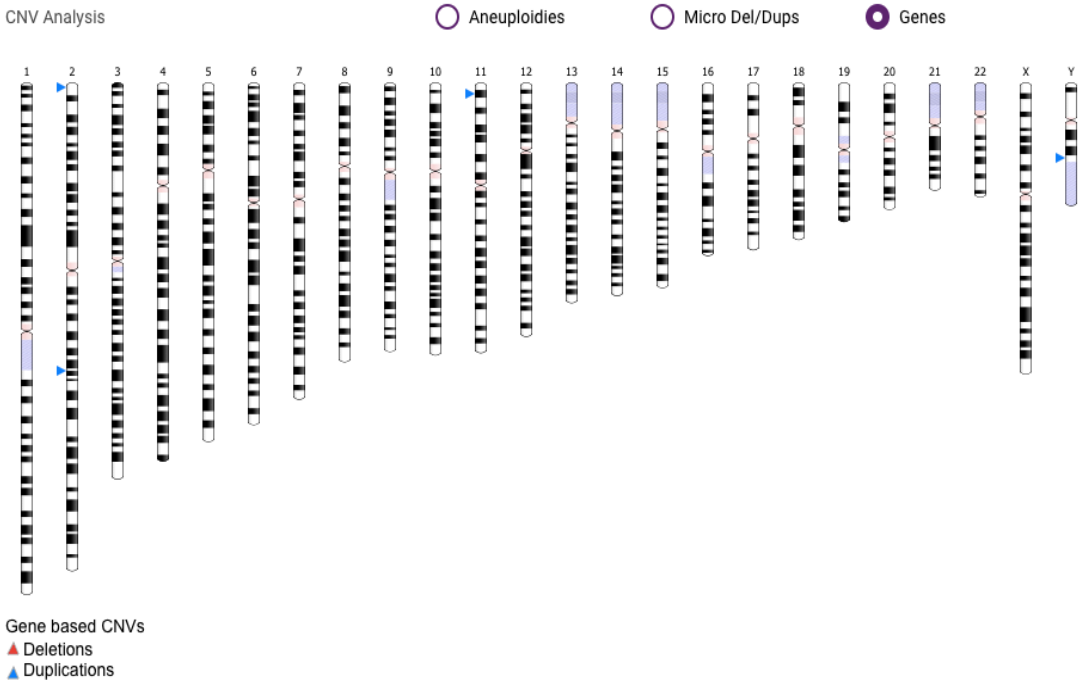
Olgumuzla ilgili CNV analizi yapılmış; anöploidi, mikro delesyon/duplikasyon ve gen bazında CNV'lere rastlanmamıştır. (Şekiller 4.24, 4.25 ve 4.26)



Şekil 4-24: Olgu 3 CNV Analizi Anöploidi Sonucu



Şekil 4-25: Olgu 3 Analizi Mikrodelesyon/Duplikasyon Sonucu



Şekil 4-26: Olgu 3 Analizi Genler Sonucu

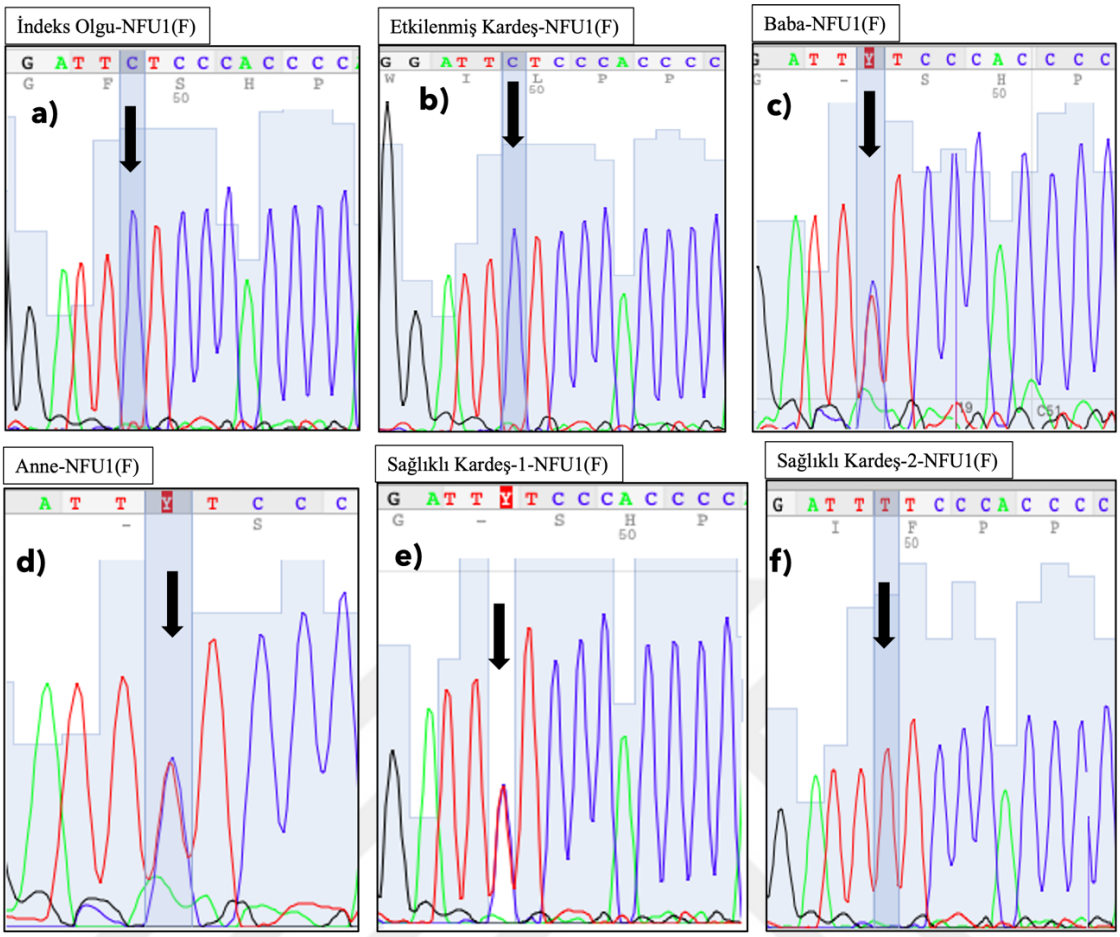
Sonuç: Yapılan CNV analizlerinde Olgu 3'e ait anöploidi, mikrodelesyon veya duplikasyon ve genler düzeyinde anlamlı bir kopya sayısı değişikliğine rastlanmamıştır.

4.4. Tüm Olgulara Ait Segragasyon ve Konfirmasyon Çalışması

3 farklı ailede 3 indeks olgunun Tüm Genom Dizileme sonucunda elde edilen varyasyonların doğrulanması amacıyla ve diğer aile bireylerinde ilgili varyantın segregasyonu amacıyla tüm olgulardan elde edilen DNA örnekleri ile birlikte öncelikle ilgili genlerde PZR aşaması gerçekleştirilmiş ardından Sanger Sekanslama metodu ile dizileme işlemi gerçekleştirilmiştir. Sanger sekans analiz sonuçları Tablo 4.16'da yer almaktadır. Ayrıca Şekiller 4.27, 4.28 ve 4.29'da sanger sekanslamaya ait kromatogram görselleri yer almaktadır.

Tablo 4-19: Tüm Olgulara Ait Sanger Sekans Sonuçları

Aile	Olgu	Gen	Varyant	Zigosite
Aile-1	İndeks Olgu	NFU1	c.191T>C	Homozigot Mutant
	Etkilenmiş Kardeş	NFU1	c.191T>C	Homozigot Mutant
	Anne	NFU1	c.191T/C	Heterozigot
	Baba	NFU1	c.191T/C	Heterozigot
	Sağlıklı Kardeş-1	NFU1	c.191T/C	Heterozigot
	Sağlıklı Kardeş-2	NFU1	c.191 T	Homozigot Yabanıl Tip
Aile-2	İndeks Olgu	ATP8A2	c.1400G>A	Homozigot Mutant
	Etkilenmiş Kardeş	ATP8A2	c.1400G>A	Homozigot Mutant
	Anne	ATP8A2	c.1400G>A	Homozigot Mutant
	Baba	ATP8A2	c.1400 A/G	Heterozigot
	Sağlıklı Kardeş	ATP8A2	c.1400 A/G	Heterozigot
Aile-3	İndeks Olgu	DNMT3B	c.1882G>T	Homozigot Mutant
	Anne	DNMT3B	c.1882 G/T	Heterozigot
	Baba	DNMT3B	c.1882 G/T	Heterozigot
	Sağlıklı Kardeş-1	DNMT3B	c.1882 G/T	Heterozigot
	Sağlıklı Kardeş-2	DNMT3B	c.1882 G	Homozigot Yabanıl Tip
	Sağlıklı Kardeş-3	DNMT3B	c.1882 G/T	Heterozigot

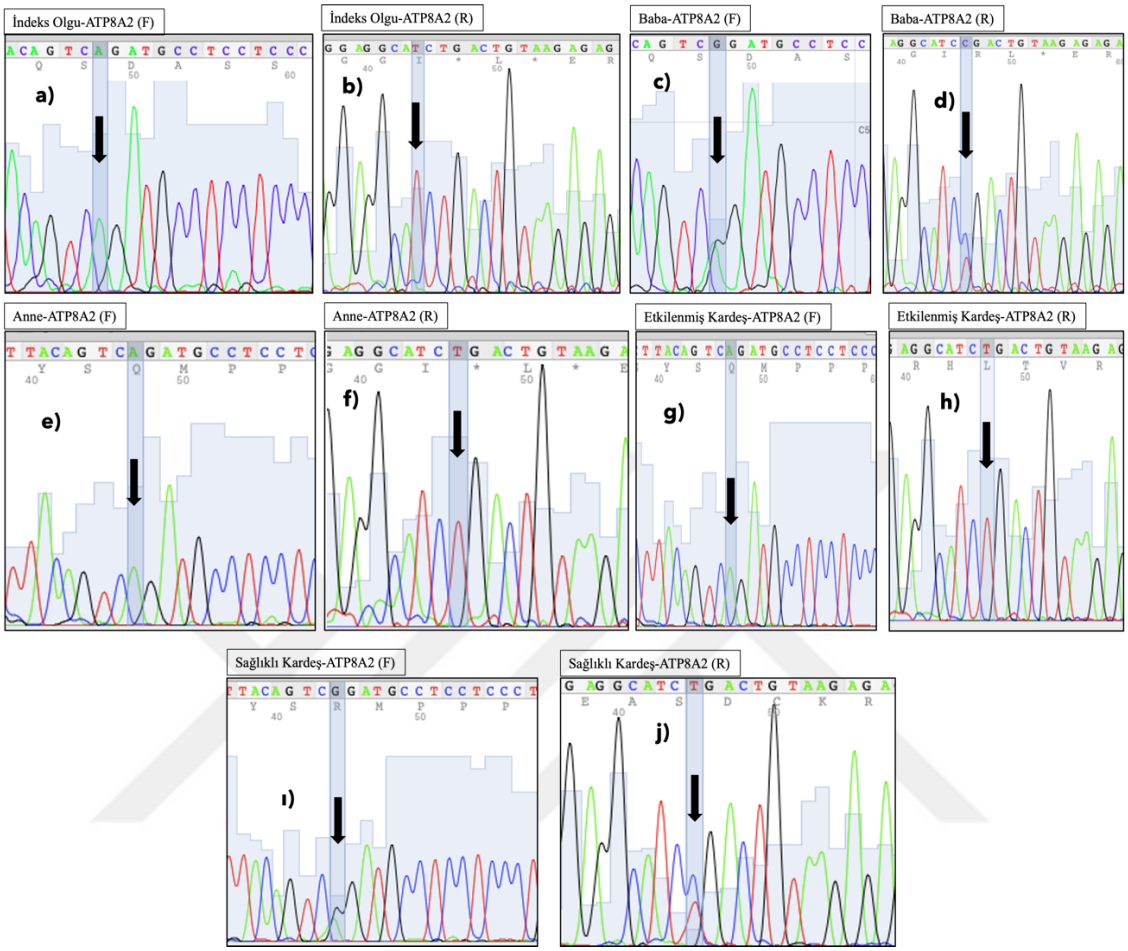


Şekil 4-27: Aile 1 Sanger Konfirmasyon ve Segregasyon Dizileme Görüntüleri

a) İndeks Olgu Homozigot Mutant (c.191T>C) **b)** Etkilenmiş Kardeş Homozigot Mutant (c.191T>C) **c)** Baba Heterozigot (c.191T/C) **d)** Anne Heterozigot (c.191T/C) **e)** Sağlıklı Kardeş-1 Heterozigot (c.191T/C) **f)** Sağlıklı Kardeş-2 Homozigot Yabanıl Tip (c.191T)

Olgu-1'de TGD aracılığı ile elde edilen anlamlı *NFU1* varyantının konfirmasyonu ve ailesel segregasyonu, Sanger dizileme metodu ile yapılmıştır. İlk bulgularda Olgu-1 ve etkilenmiş kardeşinde ilgili varyant homozigot mutant olarak tespit edilmiştir (Şekil 4-27a ve 4-27b). Sağlıklı kardeşlerin birinde homozigot yabanıl tip olarak tespit edilmiştir (Şekil 4-27f). Diğer sağlıklı kardeşte ise heterozigot olarak varyant tespit edilmiştir (Şekil 4-27e). Ancak anne ve babadaki nokta, homozigot mutant olarak tespit edilmiştir. Sağlıklı kardeşlerden birinin heterozigot olması, diğerinin de homozigot yabanıl tip olması, anne-babanın sonucuyla çelişmektekdir. Dolayısıyla reaksiyon tekrarlanmıştır. Tekrar sonucunda aynı sonuçların elde edilmesi nedeniyle yeniden primer dizaynı yapılmıştır. Guanin ve Sitozin (GC) oranı oldukça yüksek bir bölge olduğu için alel drop out ihtimali göz önünde bulundurularak primer dizaynı bu ihtimal dahilinde yapılmıştır. Yeni dizayn edilen primerler ile anne babanın ilgili noktası heterozigot olarak tespit edilmiştir (4-27c

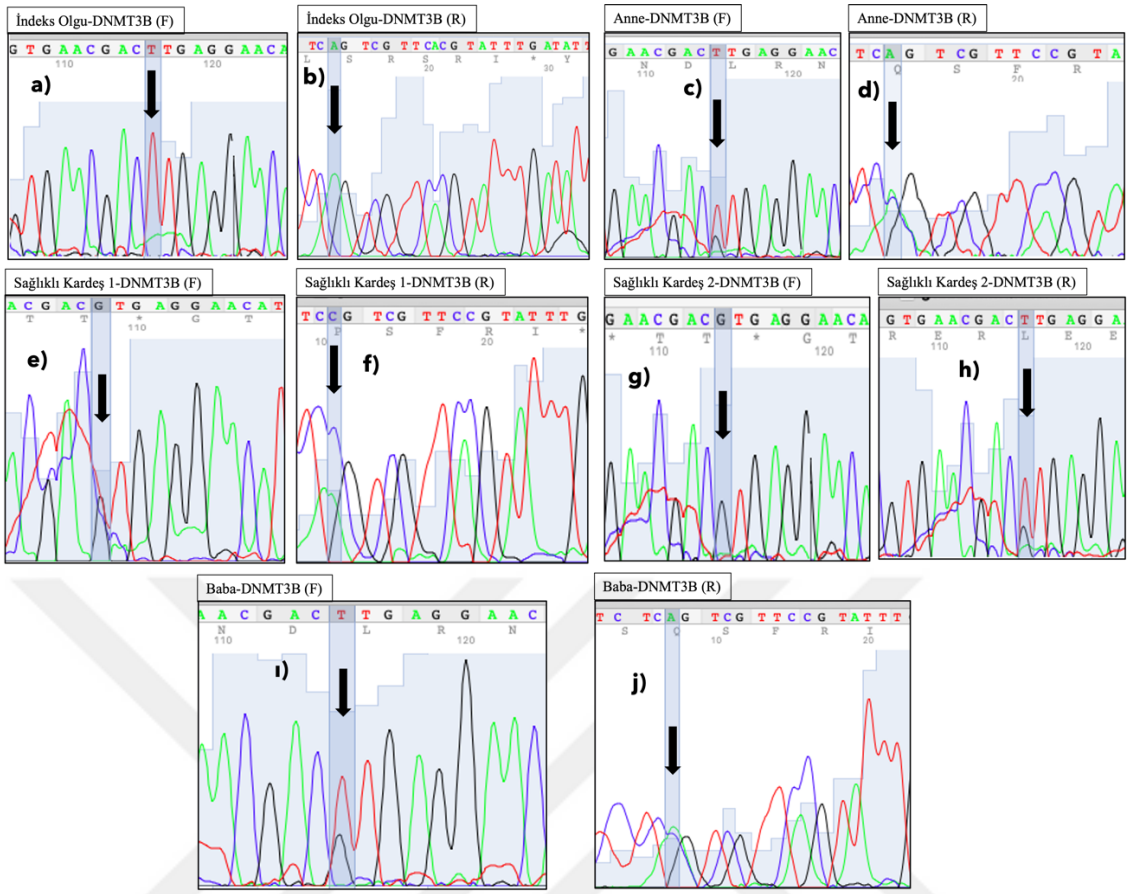
ve 4-27d). Dolayısıyla Olgu-1'deki c.191T>C değişimi segrege olmuştur. Akraba evliliği olan bir aile için *NFU1* genindeki bu değişim kalıtım modeliyle de uygunluk göstermiştir.



Şekil 4-28: Aile 2 Sanger Konfirmasyon ve Segregasyon Dizileme Görüntüleri

a) İndeks Olgu Homozigot Mutant c.1400G>A **b)** İndeks Olgu Ters Zincir **c)** Baba Heterozigot(c.1400 A/G) **d)** Baba Ters Zincir **e)** Anne Homozigot Mutant (c.1400G>A) **f)** Anne Ters Zincir **g)** Etkilenmiş Kardeş Homozigot Mutant (c.1400G>A) **h)** Etkilenmiş Kardeş Ters Zincir **ı)** Sağlıklı Kardeş Heterozigot (c.1400G/T) **j)** Sağlıklı Kardeş Ters Zincir

Olgu-2'de TGD aracılığı ile elde edilen anlamlı *ATP8A2* varyantının sanger konfirmasyonu ve ailesel segregasyonu Şekil 4-28'te gösterilmiştir. İlk bulgularda Olgu-2 ve etkilenmiş kardeşinde ilgili varyant homozigot mutant olarak tespit edilmiştir (Şekil 4-28a,b ve Şekil 4-28g,h). Baba ve sağlıklı kardeşte c.1400A/G noktası heterozigot olarak tespit edilmiştir (Şekil 4-28c,d ve Şekil 4-28ı,j). Ancak anne homozigot mutant olarak tespit edilmiştir. Bu durum tekar değerlendirilmiş ve anne için sanger dizileme yeniden yapılmıştır. Fakat sonuç yine homozigot mutant olarak tespit edilmiştir (Şekil 4-28e,f).



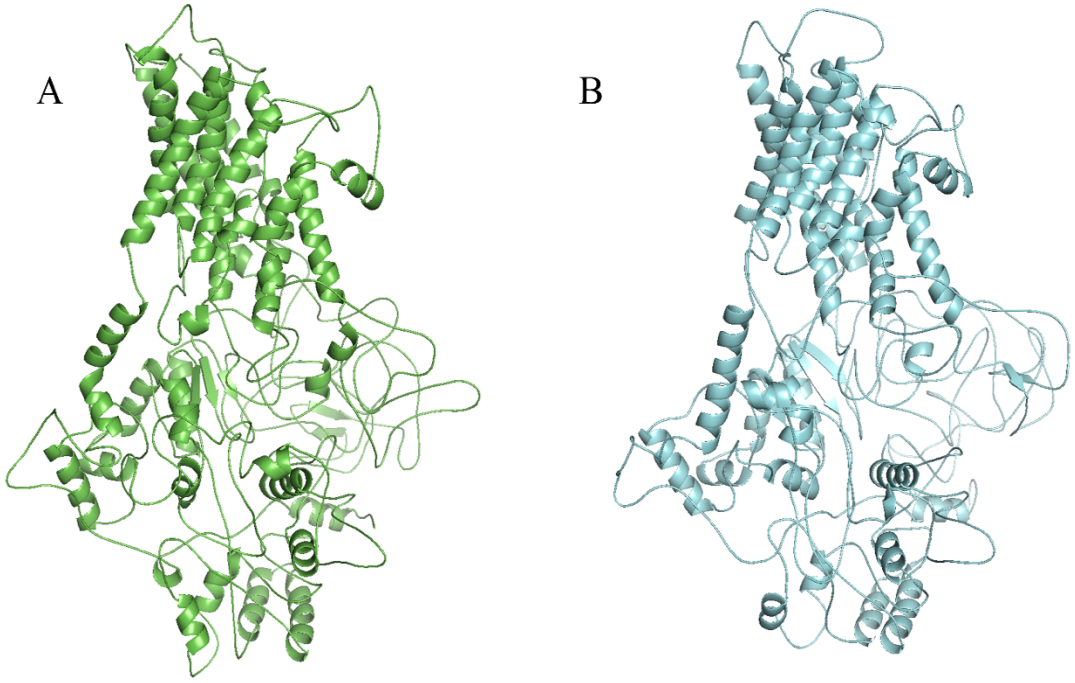
Şekil 4-29: Aile 3 Sanger Konfirmasyon ve Segregasyon Dizileme Görüntüleri

a) İndeks Olgu Homozigot Mutant (c.1882G>T) **b)** İndeks Olgu Ters Zincir **c)** Anne Heterozigot (c.1882 G/T) **d)** Anne Ters Zincir **e)** Sağlıklı Kardeş-1 Heterozigot (c.1882 G/T) **f)** Sağlıklı Kardeş-1 Ters Zincir **g)** Sağlıklı Kardeş-2 Homozigot Yabanıl Tip (c.1882 G) **h)** Sağlıklı Kardeş-3 Heterozigot (c.1882 G/T) **ı)** Baba Heterozigot (c.1882 G/T) **j)** Baba Ters Zincir

Olgu-3'de TGD aracılığı ile elde edilen anlamlı *DNMT3B* varyantının sanger konfirmasyonu ve ailesel segregasyonu Şekil 4-29'da gösterilmiştir. İlk bulgularda Olgu-3'teki ilgili varyant homozigot mutant olarak tespit edilmiştir (Şekil 4-29a,b). Anne ve sağlıklı kardeş-1'de c.1882G/T noktası heterozigot olarak tespit edilmiştir (Şekil 4-29c,d ve Şekil 4-29e,f). Sağlıklı kardeş-2 homozigot yabanıl tip olarak tespit edilmiştir (Şekil 4-29g). Sağlıklı kardeş-3 ve baba ise heterozigot olarak tespit edilmiştir (Şekil 4-29h ve Şekil 4-29ı,j). Dolayısıyla Olgu-3'deki c.1882G>T değişimi segrege olmuştur. Akraba evliliği olan bir aile için *DNMT3B* genindeki bu değişim kalıtım modeliyle de uygunluk göstermiştir.

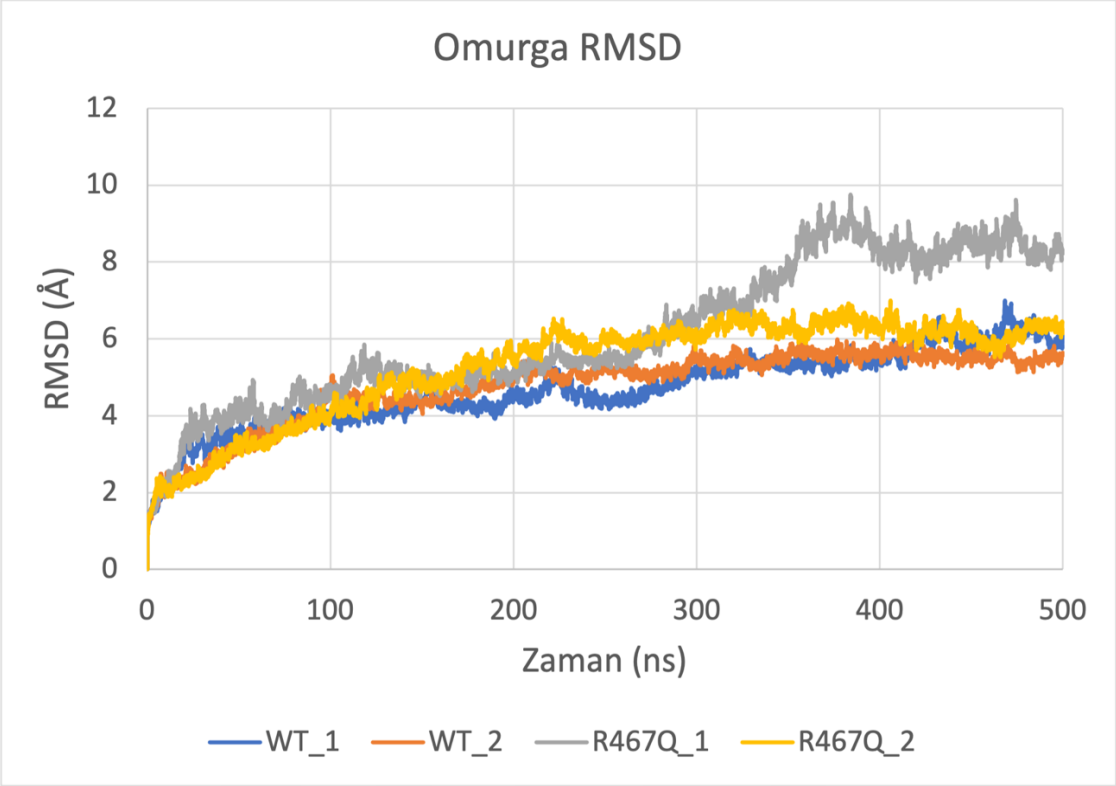
4.5. ATP8A2, NFU1 ve DNMT3B Genleri Tarafından Sentezlenen Proteinlerin Tüm Sekans Modellenmesi ve Simülasyonu

Çalışmamıza dahil edilen 3 farklı ailenin indeks olgularında varyant tespit edilen *ATP8A2* (ENST00000381655.7 transkripti 1188 aminoasit), *NFU1* (ENST00000303698.7 transkripti 230 amino asit) ve *DNMT3B* (ENST00000328111.6 transkripti 853 aminoasit) genleri tarafından sentezlenen proteinlerin tüm sekans modellemesi gerçekleştirildi. Varyant içeren sekansları da modellenerek karşılaştırma yapıldı.



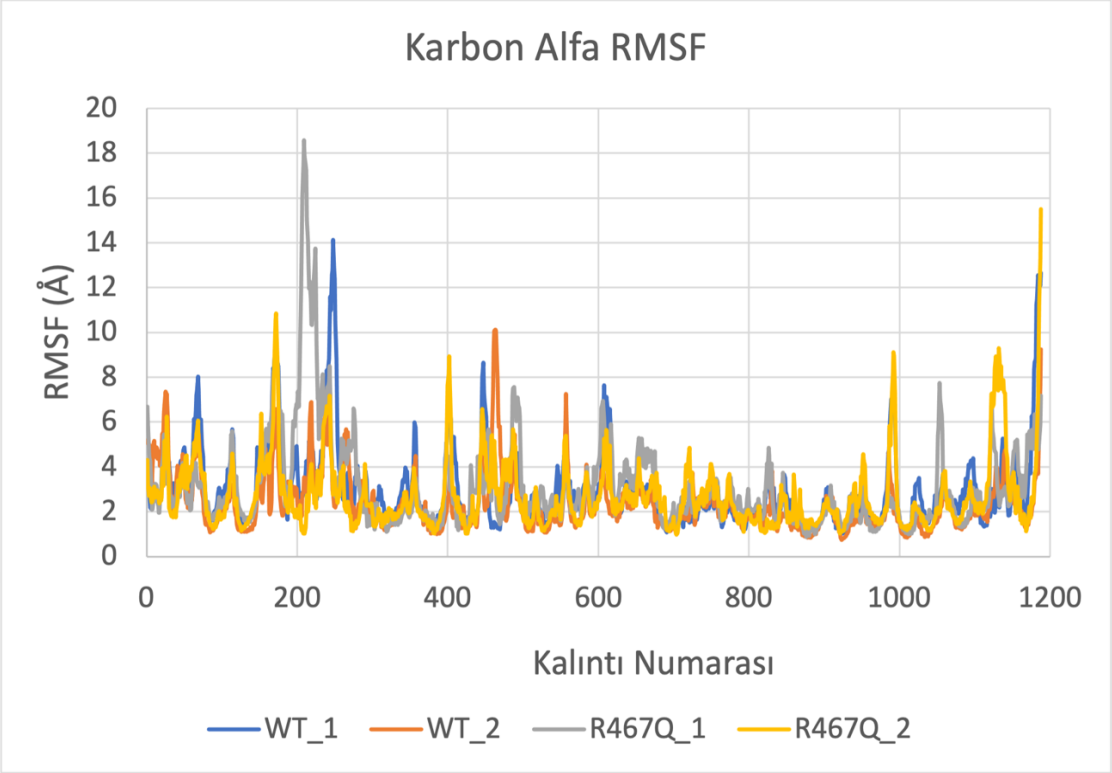
Şekil 4-30: ATP8A2 için tüm sekans modellemesi A) Yabanıl Tip (WT) için tüm sekans modellemesi B) Mutant (R467Q) için tüm sekans modellemesi

Çalışmamızda tespit edilen *ATP8A2* genindeki Arg467Gln değişimi proteinin katlanma bölgesinde yer almaktadır. Bu değişimin proteinin yapısına ve fonksiyonuna etkisini araştırmak için protein modelleme ve simülasyon çalışması yapılmıştır. ATP8A2 proteinin tüm sekans modellemesi yapılmıştır. Ayrıca mutant aminoasiti içeren yapının modellemesi de yapılmıştır. Şekil 4-30'da, yabanıl tip ve mutant ATP8A2 için temsili yapıları göstermektedir. Yabanıl tip için temsili yapı, tüm çerçevelerin %38'ini temsil eder (bkz. A), mutant (R467Q) için ilk küme, tüm çerçevelerin %27'sini temsil eder (bkz. B). Yabanıl tip ve R467Q simülasyonlarını oluşturan her iki sistem için 12 küme oluşturulmuştur.



Şekil 4-31: Yabanıl Tip (WT) ve Mutant (R467Q) ATP8A2 için RMDS grafiği

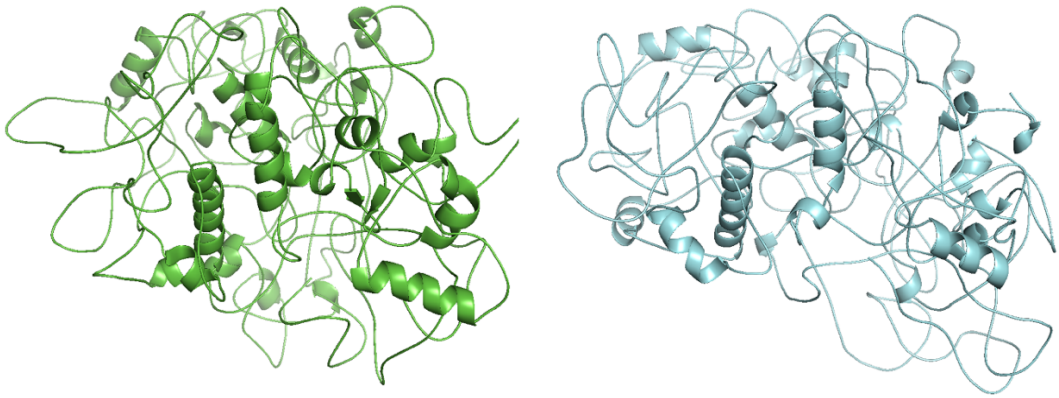
Modellenen proteinlerin moleküler dinamik simülasyonu yapıldı. Proteinlerin zamana göre hareketleri nicel olarak analiz edildi. Atomlar hareket ederken bir zaman adımının alınması ile omurgadaki protein alfa karbonlarının başlangıç koordinatlarına göre hesaplanması için kök ortalama kare sapması (RMSD) değerleri omurga atomlarına dayalı olarak hesaplanmıştır. Şekil 4-31'de, ATP8A2'deki moleküler dinamik (MD) simülasyonları boyunca RMSD değişikliklerini göstermektedir. Yabanıl tip ve mutant simülasyonlarının RMSD ortalamaları; WT-1, WT-2, R467Q-1 ve R467Q-2 için sırasıyla 4.69, 4.77, 6.12 ve 5.27'dir. RMSD ortalamalarına bakıldığında, 4.73 ve 5.70 Å değerleri elde edilir. Mutant simülasyonlarının değerleri, yabanıl tip simülasyonlarına kıyasla daha yüksektir. Bu da mutantın genel protein yapısını dengesizleştirdiğini gösteriyor.



Şekil 4-32: Yabanıl Tip ve Mutant AT8A2 simülasyonları için RMSF değerleri

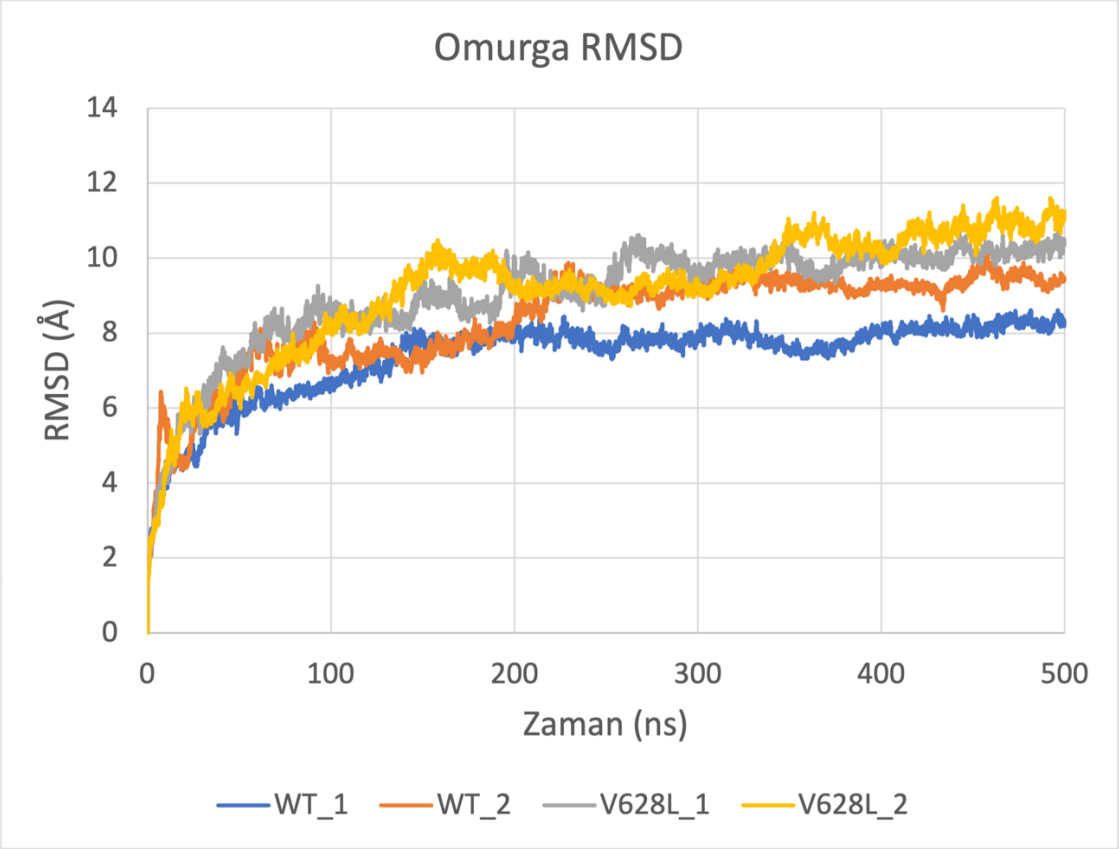
Proteini oluşturan kalıntıların (rezidü) dalgalanmasını incelemek için moleküler dinamik simülasyonların ikinci analiz yöntemi olan kök ortalama kare dalgalanması (RMSF) değerleri hesaplandı. Şekil 4-32'de ATP8A2'nin yabanıl tip ve mutant formlarının RMSF değerleri gösterildi. RMSF değerlerindeki en büyük farklılıklar, diğerlerine kıyasla R467Q-1 simülasyonlarında 200-226 kalıntı aralığında gözlendi.

A B



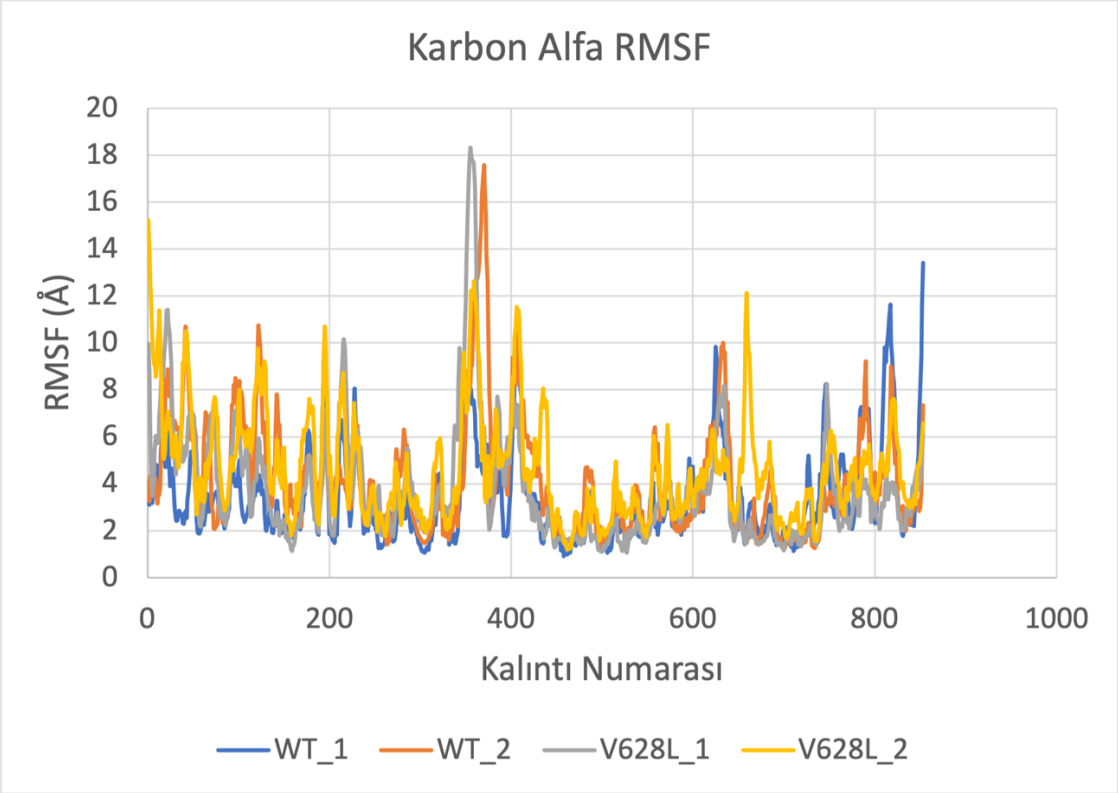
Şekil 4-33: DNMT3B için tüm sekans modellemesi A) Yabanıl Tip (WT) için tüm sekans modellemesi B) Mutant (V628L) için tüm sekans modellemesi

Çalışmamızda tespit edilen DNMT3B Val628Leu değişimi proteinin 628. kodon SAM-bağımlı metiltransferaz C5-tipi domaininde ve S-adenozil-L-metiyonin bağlanma bölgesinde yer almaktadır (PM1). Bu değişimin proteinin yapısına ve fonksiyonuna etkisini araştırmak için protein modelleme ve simülasyon çalışması yapılmıştır. DNMT3B proteinin tüm sekans modellemesi gerçekleştirilmiştir. Ayrıca mutant aminoasiti içeren yapının modellemesi de yapılmıştır. Şekil 4-33, yabanıl tip ve mutant DNMT3B için temsili yapıları gösterir. WT için temsili yapı, tüm çerçevelerin %10'unu temsil eder (bkz. A), mutant (V628L) için ilk küme, tüm çerçevelerin %9'unu temsil eder (bkz. B). WT ve V628L simülasyonları için sırasıyla 40 ve 55 küme bulundu.



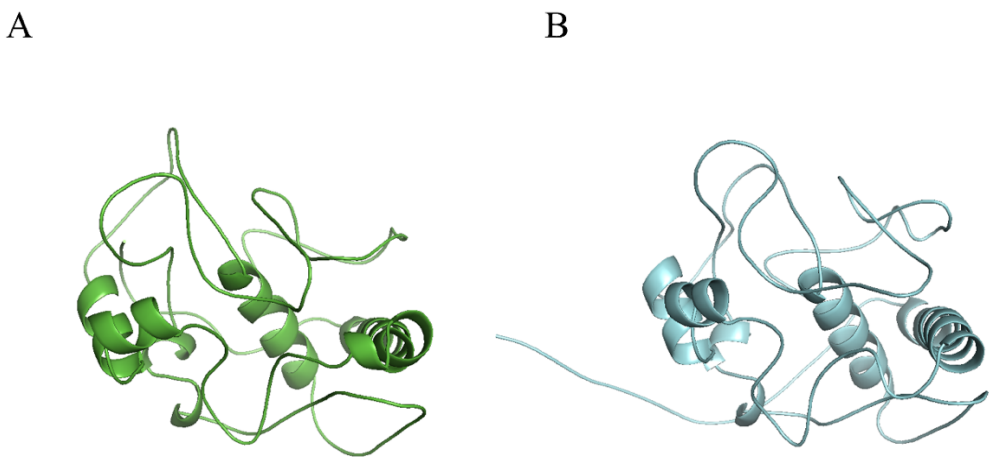
Şekil 4-34: Yabanıl Tip (WT) ve Mutant (V628L) DMT3B için RMDS grafiği

Modellenen proteinlerin moleküler dinamik simülasyonu yapıldı. Proteinlerin zamana göre hareketleri nicel olarak analiz edildi. Atomlar hareket ederken bir zaman adımının alınması ile omurgadaki protein alfa karbonlarının başlangıç koordinatlarına göre hesaplanması için kök ortalama kare sapması (RMSD) değerleri omurga atomlarına dayalı olarak hesaplanmıştır. Şekil 4-34, DNMT3B'deki MD simülasyonları boyunca RMSD değişikliklerini göstermektedir. WT ve mutant simülasyonlarının RMSD ortalamaları; WT-1, WT-2, V628L-1 ve V628L-2 için sırasıyla 7.36, 8.35, 9.03 ve 9.10'dur.



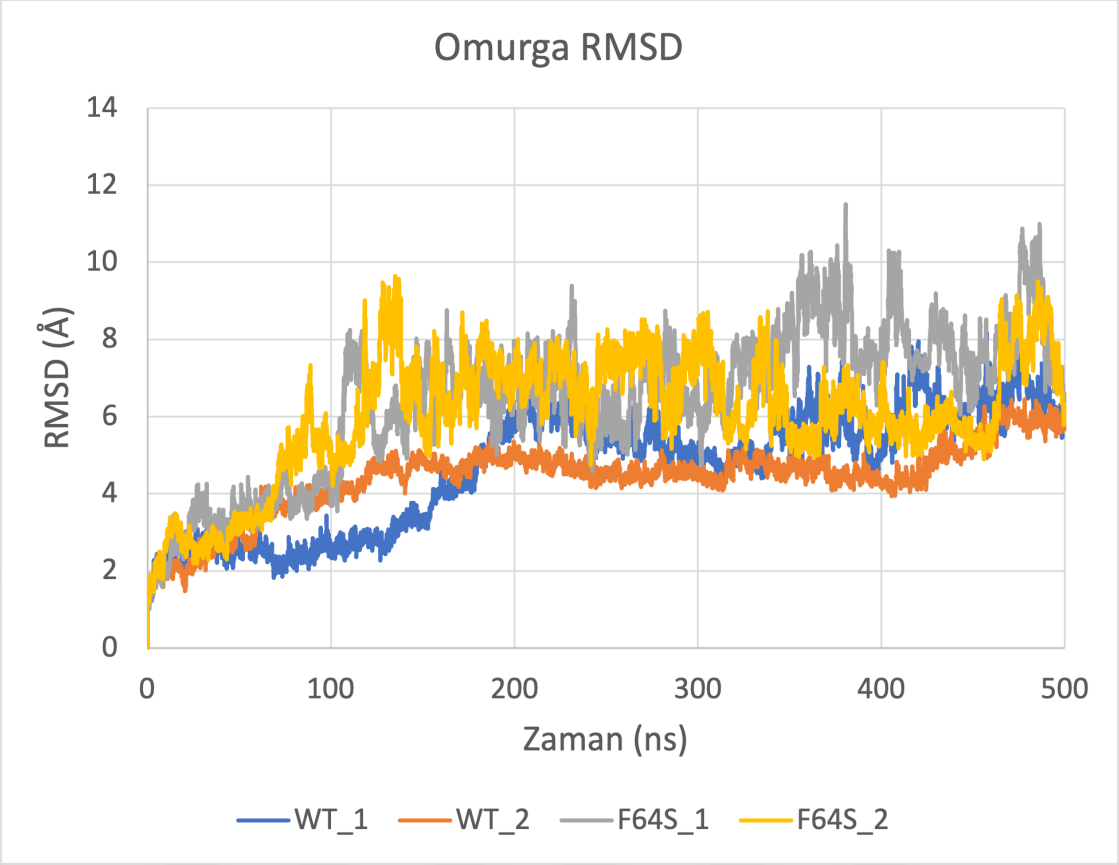
Şekil 4-35: Yabanıl Tip ve Mutant DNMT3B simülasyonları için RMSF değerleri

Proteini oluşturan kalıntıların (rezidü) dalgalanmasını incelemek için moleküler dinamik simülasyonların ikinci analiz yöntemi olan kök ortalama kare dalgalanması (RMSF) değerleri hesaplandı. Şekil 4-35, DNMT3B için yabanıl tip ve mutant sistemler için RMSF grafiğini göstermektedir. RMSF değerlerindeki en büyük farklılıklar, diğerlerine kıyasla yabanıl tip ve V628L-1 simülasyonlarında 357-375 kalıntı aralığında gözlendi.



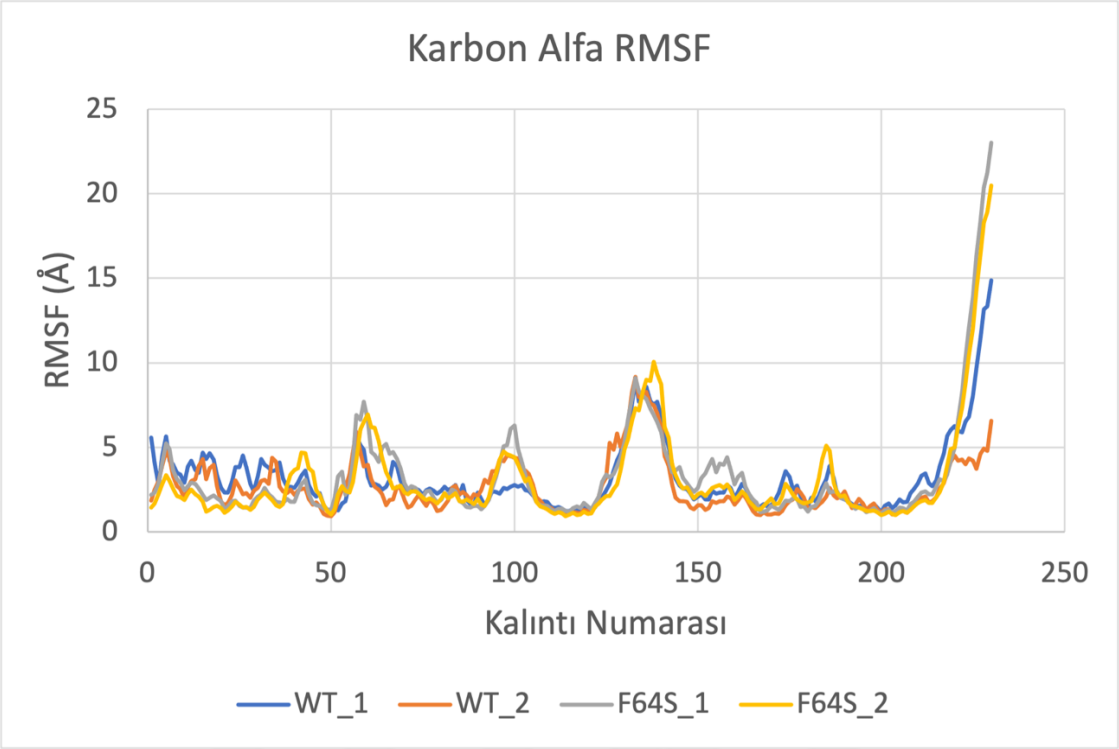
Şekil 4-36: NFU1 için tüm sekans modellemesi A) Yabanıl Tip (WT) için tüm sekans modellemesi B) Mutant (F64S) için tüm sekans modellemesi

Çalışmamızda tespit edilen NFU1 Phe64Ser değişimi proteinin N-terminal kısmında yer almaktadır. Bu değişimin proteinin yapısına ve fonksiyonuna etkisini araştırmak için protein modelleme ve simülasyon çalışması yapılmıştır. NFU1 proteinin tüm sekans modellemesi gerçekleştirilmiştir. Ayrıca mutant aminoasiti içeren yapının modellemesi de yapılmıştır. Şekil 4-36'da yer alan yabanıl tip (WT) ve mutant modeller karşılaştırıldığında; WT için temsili yapı, tüm çerçevelerin %27'sini temsil eder (bkz. A). Mutant (F64S) için ilk küme, tüm çerçevelerin %12'sini temsil eder (bkz. B). WT ve F64S simülasyonları için sırasıyla 33 ve 104 küme oluşmuştur.



Şekil 4-37: Yabanıl Tip (WT) ve Mutant (F64S) NFU1 için RMDS grafiği

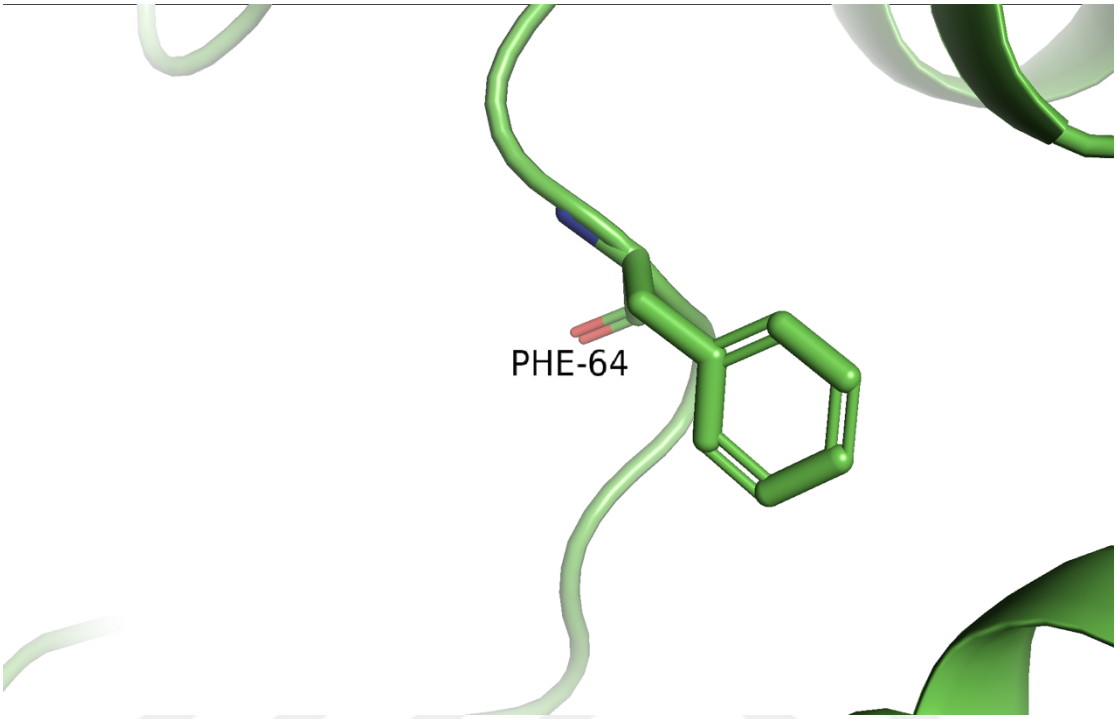
Modellenen proteinlerin moleküler dinamik simülasyonu yapılarak proteinlerin zamana göre hareketleri nicel olarak analiz edildi. Atomlar hareket ederken bir zaman adımının alınması ile omurgadaki protein alfa karbonlarının başlangıç koordinatlarına göre hesaplanması için kök ortalama kare sapması (RMSD) değerleri hesaplandı. RMSD, omurga atomlarına dayalı olarak hesaplanmıştır. Şekil 4-37'de, *NFU1*'deki MD simülasyonları boyunca RMSD değişikliklerini göstermektedir. WT ve mutant simülasyonlarının RMSD ortalamaları; WT-1, WT-2, F64S-1 ve F64S-2 için sırasıyla 4.71, 4.41, 6.47 ve 6.04'tür. Yabanıl tip ve mutant simulasyonlarının omurga RMSD değerleri ortalamaları karşılaştırıldığında sırasyla 4.56 ve 6.26 Å değerleri elde edilmiştir. Bu değerler göz önüne alınarak, F64S mutantının yapısal olarak yabanıl tip ile karşılatırıldığında daha az stabilite gösterdiği söylenebilir. Bu durumda F64S, NFU1 proteinin yapısal stabilitesini bozan bir mutasyondur çıkarımı yapılabilir. Mutant simülasyonlarının değerleri, WT simülasyonlarına kıyasla daha yüksektir. Bu da mutantın genel protein yapısını dengesizleştirdiğini göstermektedir. İlginç bir şekilde, WT-2 tekrar simülasyonu N-terminalinde yüksek değerler göstermez.



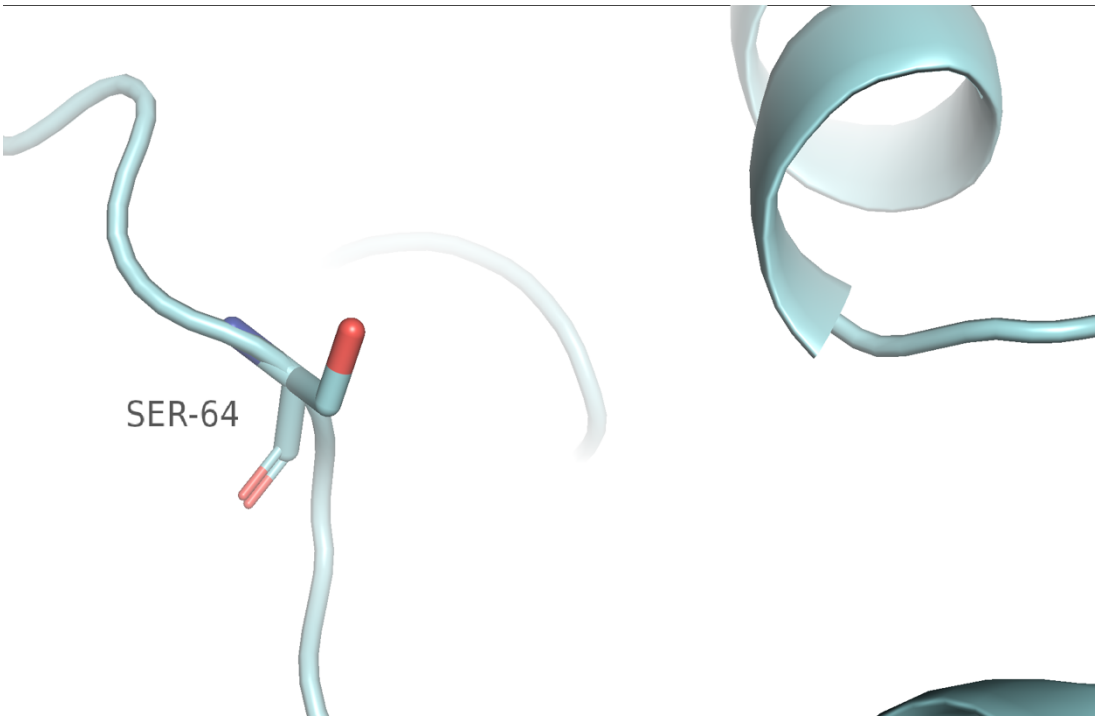
Şekil 4-38: Yabanıl Tip ve Mutant NFU1 simülasyonları için RMSF değerleri

Proteini oluşturan kalıntıların (rezidü) dalgalanmasını incelemek için moleküler dinamik simülasyonların ikinci analiz yöntemi olan kök ortalama kare dalgalanması (RMSF) değerleri hesaplandı. Şekil 4-38, NFU1 yabanıl tip ve mutant sistemler için RMSF grafiğini göstermektedir. RMSF'deki en yüksek değişiklikler, WT-1'in N-terminal kuyruğunda ve diğer iki mutant simülasyonunda gözlendi.

NFU1 proteininde 64. sırada yer alan aminoasite ait yabanıl tip ve mutant modeli aşağıda gösterilmiştir. Şekil 4-39 ve 4-40'ta yabanıl tipte fenilalanin olan aminoasit mutant tipte serin aminoasidine dönüşmüştür.

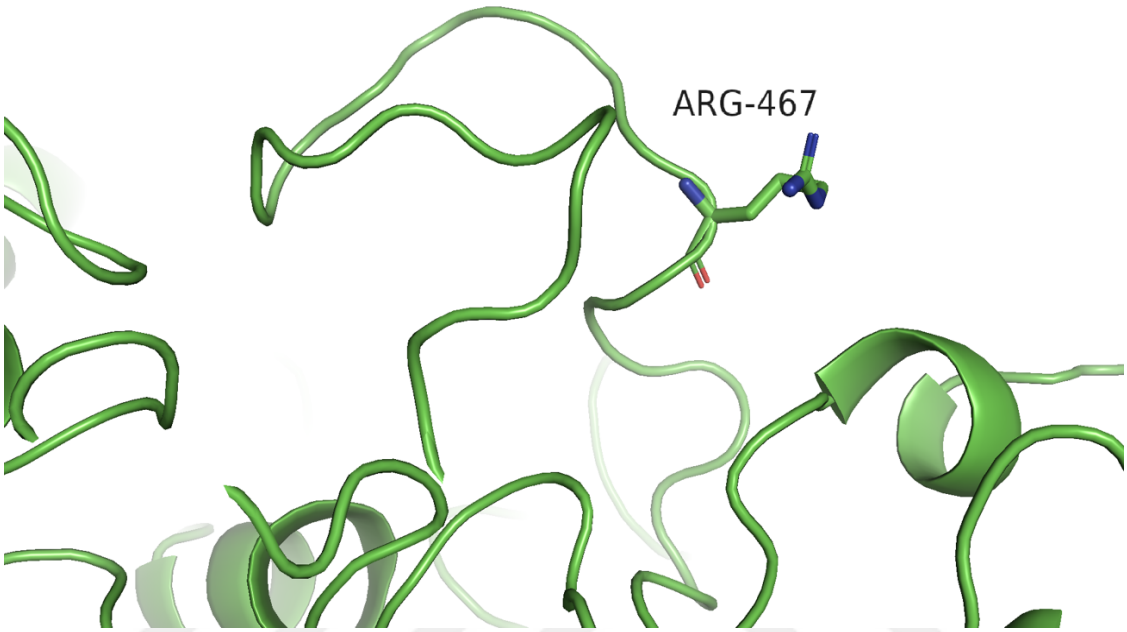


Şekil 4-39: NFU1 proteini 64. aminoasit yabanıl tip modeli

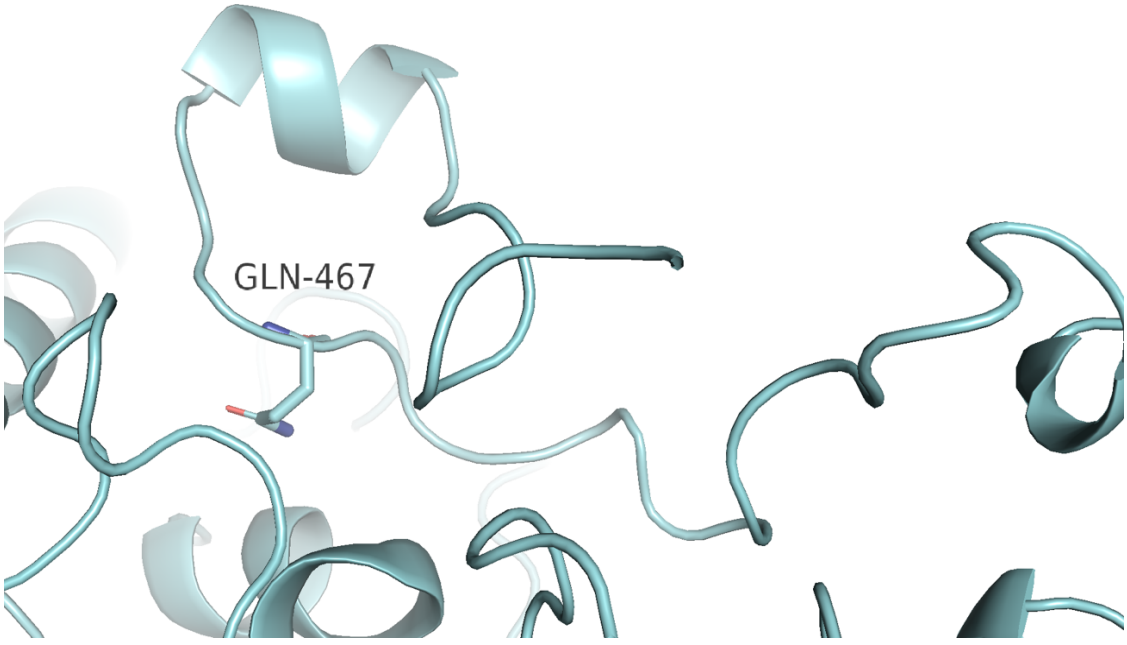


Şekil 4-40: NFU1 proteini 64. aminoasit mutant modeli

ATP8A2 proteininde 467. sırada yer alan aminoasite ait yabanıl tip ve mutant modeli aşağıda gösterilmiştir. Şekil 4-41 ve 4-42'de yabanıl tipte arjinin olan aminoasit mutant tipte glutamin aminoasidine dönüşmüştür.

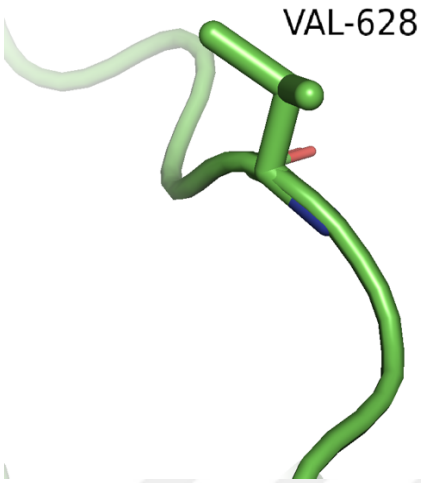


Şekil 4-41: ATP8A2 proteini 467. aminoasit yabanıl tip modeli

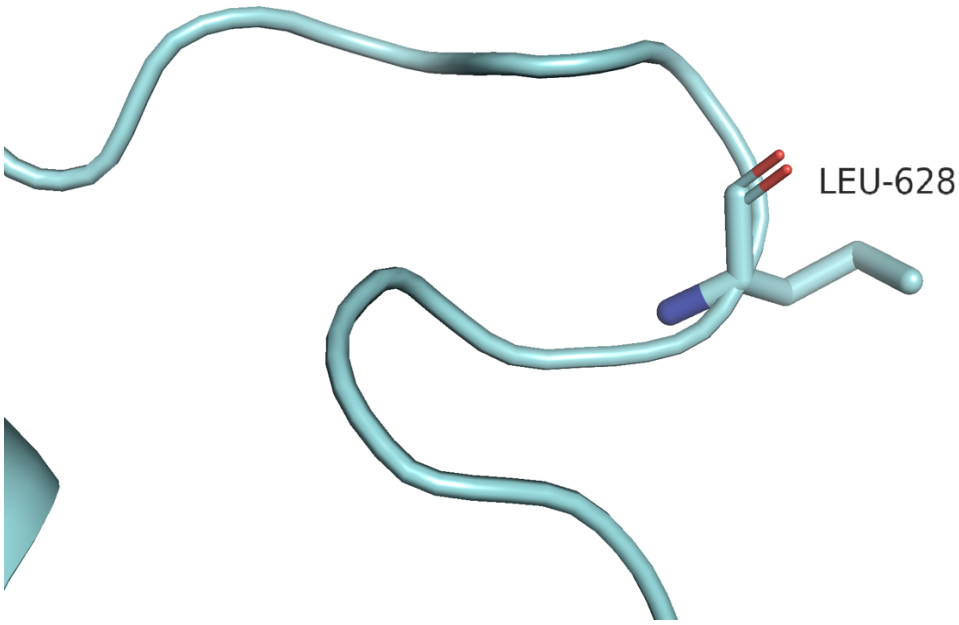


Şekil 4-42: ATP8A2 proteini 467. aminoasit mutant modeli

DNMT3B proteininde 628. sırada yer alan aminoasite ait yabanıl tip ve mutant modeli aşağıda gösterilmiştir. Şekil 4-43 ve 4-44'te yabanıl tipte valin olan aminoasit mutant tipte lösin aminoasidine dönüşmüştür.



Şekil 4-43: DNMT3B proteini 628. aminoasit yabanıl tip modeli



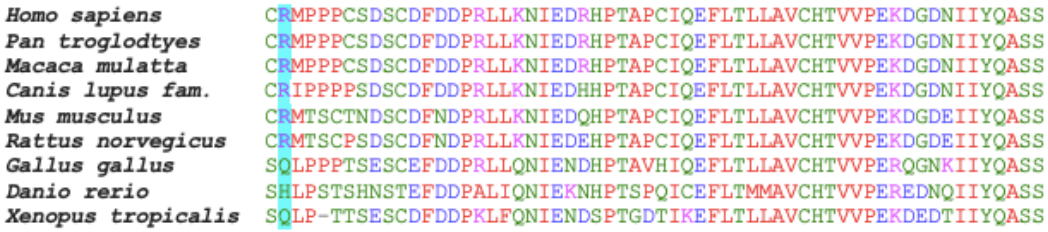
Şekil 4-44: DNMT3B proteini 628. aminoasit mutant modeli

ATP8A2, DNMT3B ve NFU1 proteinlerindeki aminoasit değişiminin türler arası protein dizileri karşılaştırılarak tespit edilen varyantın korunmuş bölgede olup olmadığı gösterilmiştir. Aşağıdaki şekillerde türler arası aminoasit dizileri karşılaştırılmış ve varyant noktaları gösterilmiştir.



Şekil 4-45: NFU1 Proteinin türler arası aminoasit dizisi karşılaştırması

Şekil 4-45'te insan NFU1 proteinin 36. aminoasitinden 96. aminoasitine kadar olan aminoasit dizisi diğer canlı türlerindeki ortolog protein dizisi ile karşılaştırılmıştır. Karşılaştırılan tüm canlılarda F64 korunmuş bölgede bulunan bir aminoasittir.



Şekil 4-46: ATP8A2 Proteinin türler arası aminoasit dizisi karşılaştırması

Şekil 4-46'da insan ATP8A2 proteinin 466. aminoasitinden 526. aminoasitine kadar olan aminoasit dizisi diğer canlı türlerindeki ortolog protein dizisi ile karşılaştırılmıştır. Yakın canlılar da R467 korunsa da Gallus gallus'tan itibaren aminoasit korunmamıştır.



Şekil 4-47: DNMT3B Proteinin türler arası aminoasit dizisi karşılaştırması

Şekil 4-47'de insan DNMT3B proteinin 586. aminoasitinden 646. aminoasitine kadar olan aminoasit dizisi diğer canlı türlerindeki ortolog protein dizisi ile karşılaştırılmıştır. Karşılaştırılan tüm canlılarda V628 korunmuş bölgede bulunan bir aminoasittir.

5. TARTIŞMA

Genetik tanı için Yeni Nesil Dizileme teknikleri son yıllarda oldukça önem kazanmış ve yaygın kullanılmaya başlanmıştır. En yaygın kullanılan YND tekniği, genomda bilinen genlerin tüm ekzonlarını ve ekzon/intron sınırlarını kapsayan Tüm Ekzom Dizileme tekniğidir. İnsan genomundaki tüm genlere ait ekzon sayısı 180 bin civarındadır. 30 Mb büyüklüğünde olan tüm ekzonlar, insan genomunun yaklaşık %1'ini teşkil eder. Hastalık ilişkili mutasyonların %85'inin insan genomunun protein kodlayan bölgelerinde olduğu tahmin edilmektedir [78]. Dolayısıyla TED, tüm tek gen hastalıklarının teşhisinde en önemli tekniklerden biri olarak öne çıkmaktadır.

Tüm Genom Dizileme tekniği ise insan genomunun protein kodlayan ve kodlamayan tüm bölgelerinin dizilenmesidir. TGD uygulaması; TED'e göre daha yüksek maliyetli olması, çok daha büyük bir veri ürettiğinden analiz aşamasının daha uzun zaman alması ve daha fazla sayıda polimorfik değişimler saptanmasının yol açtığı doğrulama çalışmaları gibi sebeplerden ötürü hastalıkların tanısında daha az tercih edilmektedir. Bununla birlikte yeni çalışmalar, nörodejenaratif hastalıkların patogenezinde kodlamayan bölgelerin/genlerin de rol oynadığını göstermiş ve TED ile sonuç alınamayan vakalar için TGD tekniğinin bir sonraki test olarak kullanılması gerektiği önerilmiştir [79]. Test maliyetlerinin düşmesi ve artan deneyimlerle üretilen verinin daha verimli ve hızlı bir şekilde işlenmesi ile TGD kullanımı daha da yaygınlaşacaktır.

YND teknolojisi, birçok hastalığın genetik tanısına katkı sağladığı gibi nörodejeneratif hastalıkların tanısında da önemli katkılar sağlamaktadır. Bu teknikler, yeni mutasyonların/genlerin hastalığın patogenezine nasıl katkıda bulunduğunu da göstermiştir. Örneğin CoQ10 yetmezliğine sebep olan *ADCK3* geni ve diğer genlerdeki yeni mutasyonlar, hastalıkla ilişkili bilinen genlerdeki mutasyon frekanslarını değiştirmiştir [80]. Bilinen tek gen hastalıkları ile ilişkili yeni genler de keşfedilmektedir. Örneğin *KCND3* geni ve spinoserebellar ataksia (SCA22/SCA19) hastalığı arasında ilişki olduğu ortaya konmuştur. *KCND3* gen mutasyonlarının, serebellar dejenerasyonun patogenezinde yer aldığı ve nöronal uyarılmanın kilit düzenleyicisi olarak iyon kanalları işlevleri üzerinde etkin olduğu gösterilmiştir [81].

Öte yandan YND teknikleri sayesinde bilinen genlerin yeni fenotiplerle ilişkili olduğu da gösterilmiştir. Yani fenotipik olarak farklı hastalıklarda ortak yolakların varlığı ortaya konmuştur. Örneğin *KIF5A* geninin spastik parapleji tip 10 hastalığına neden olduğu bilinmekteydi. Yeni yapılan bir çalışmada bu gendeki 5 yanlış anlamlı mutasyonun, spastik paraplejiden aksonal nöropatiye kadar, hem merkezi hem de çevresel aksonopatilerin patogenezinde, hücrelerarası trafik ve aksonal iletimde muhtemel ortak yolaklarda etkisi olduğu gösterilmiştir [82]. Tüm bu örnekler YND tekniklerinin hastalıkların veya yeni genlerin sadece tanısında değil aynı zamanda patogenezininin aydınlatılmasında da katkı sağlayabildiğini göstermektedir [83].

Çalışmamızda fenotiple uyumlu olabilecek varyantlar bir YND tekniği olan TGD yöntemi ile tespit edilmiştir. TGD reaksiyonları DNBSEQ-T7RS platformunda gerçekleştirildikten sonra ham veriler NGS Cloud veritabanında analiz edilmiştir. Yapay zeka destekli bir algoritma ile klinikle ilgili varyantları kullanıcıya sunan platformda daha önce yapılan 37 örneklik bir validasyon çalışmasında, platformun kliniği açıklama oranı %94,4 olarak tespit edilmiştir. Vakaların %44.1'inde kliniği açıklayan varyant ilk sırada, %73,5'inde ise ilk üç sırada raporlanmıştır. Yüksek başarısından ötürü bu platform tercih edilmiştir [73].

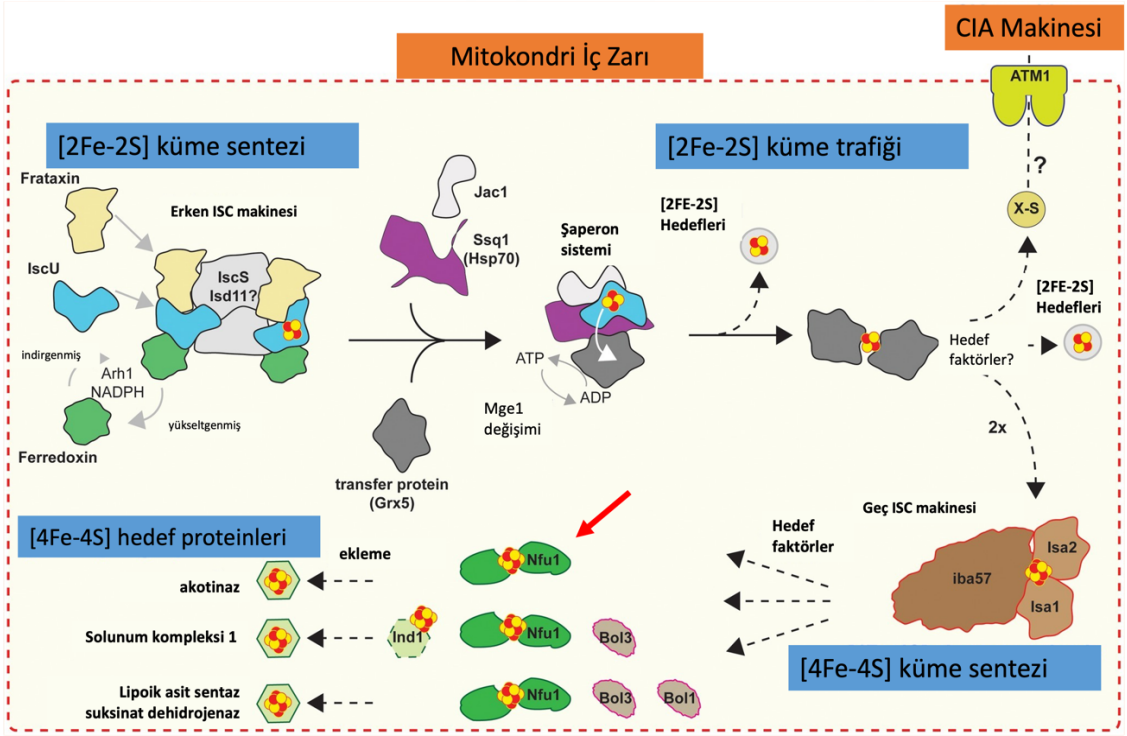
Bu tez kapsamındaki çalışmamızda TED analizinde klinik ile ilişkili olabilecek bir varyant saptanmayan olgularda TGD yönteminin uygulanmasının tanıya katkısı araştırılmış ve TGD analizinde saptanan aday varyantların validasyon ve segregasyon çalışmaları yapıldıktan sonra protein modelleme ve simülasyon işlemleri de yapılmıştır. Bu işlemlerin sonuçları, her bir indeks olgumuz ve ailesi için ayrı ayrı tartışılacaktır.

5.1. Aile 1

Ebeveynleri birinci derece kuzen olan olguda yürüme bozukluğu, alt ekstremitede artmış Derin Tendon Refklesleri (DTR) ile azalmış kas gücü saptanmış ve olgu spastik parapleji tanısı ile değerlendirilmiştir. Daha önce yapılan TED analizinde herhangi bir anlamlı varyant bulunmadığından TGD analizi yapılarak on binlerce varyant arasından *NFU1* geni ile birlikte *NUS1, SNX7* ve *TRIM71* genlerinde varyant bulunmuştur. *NFU1* genindeki varyant (c.191T>C), fenotip ve kalıtım modeli ile uyumlu olmasından ötürü aday varyant olarak değerlendirilmiştir.

NFU1 geni, lipoat içeren 2-oksait dehidrogenazların normal olgunlaşması ve mitokondriyal solunum zinciri kompleksinin birleştirilmesi için gereken demir-sülfür (Fe-S) kümelerinin üretiminde temel bir rol oynamaktadır [85].

Demir-Sülfür kümeleri; mitokondriyel solunum zincirindeki elektron iletimi, sitrik asit döngüsü, hem biyosentezi, DNA replikasyonu ile tamiri ve demir homeostazı gibi çeşitli hücresel fonksiyonlarda önemli kofaktörler olarak görev alırlar [86]. Demir-Sülfür Küme Yapısı (NFU1); [4Fe-4S] kümelerinin olgunlaşması ve bunların hedef proteinlere iletimi sürecinde rol alan bir Fe-S küme biyosentez faktörüdür. Hedef proteinleri arasında lipoik asit sentetaz, mitokondriyel akonitaz ve solunum zinciri kompleksi I ve II proteinleri bulunmaktadır [87].



Şekil 5-1: Mitokondrideki Demir-Sülfür Küme Sentezi Süreci*
* Tsaousis (2019) çalışmasındaki şekil yeniden düzenlenmiştir.

Şekil 5.1'de gösterilen model, mitokondrideki Demir-Sülfür protein kümelerinin sürecini anlatmaktadır. Demir-Sülfür kümesi (ISC) makinesi olarak adlandırılan süreçte 18 protein yer almaktadır. Bu proteinlerin tamamı [2Fe-2S] ve [4Fe-4S] kümelerinin sentezi ile bunların uygun apoproteinlere iletiminden sorumludur. Bu çok elemanlı sistemde NFU1 proteini önemli bir adımda rol almaktadır [88].

NFU1 genindeki patojenik varyantların otozomal resesif kalıtım gösteren çoklu mitokondriyel disfonksiyon sendromu 1'e (MMDS1; OMIM: 605711) neden olduğu bilinmektedir. MMDS1; halsizlik, solunum yetersizliği, nörolojik gelişim eksikliği, laktik asidoz ve erken ölüm ile sonlanan sistemik bir enerji metabolizması bozukluğu olan ciddi bir hastalıktır [85,89]. Ancak Tonduti ve ark. (2015) tarafından yayınlanan yeni bir spastik parapleji fenotipi, *NFU1* genindeki c.146delC (p.Pro49LeufsX8) ve c.565G > A (p.Gly189Arg) patojenik varyantları ile ilişkilendirilmiştir. Bildirdikleri vaka daha hafif, geç başlangıçlı, tekrarlayan ve ara veren epizotlar ile giden, yavaş ilerleyen spastik paraparesis, hafif kognitif bozukluk ve uzun yaşam süresi ile karakterize olup bugüne kadar bilinen tipik fenotip dışında olarak tanımlanmıştır. Erken dönemde MRG'de gözlenen beyaz cevher değişiklikleri, üçüncü dekatta karışık tipte duysal motor nöropatisi ile kombine hale gelmiştir. Yazarlar sonuç olarak, yüksek glisin konsantrasyonları ile birlikte giden yavaş ilerleyen lökoensefalopati olgularında *NFU1* geni patojenik varyantlarından şüphelenilmesi gerektiğini ifade etmişlerdir [90].

Tez kapsamında elde edilen *NFU1* varyantının da dahil olduğu ve yakın zamanda yayınlanan bir kohort çalışma ile, 10 farklı aileden 19 etkilenmiş bireyin *NFU1* geninde, bi-allelik yanlış anlamlı varyantlar olduğu ve bu varyantların, nörogelişimsel gecikme ve ciddi hipotoni ile karakterize, erken başlangıçlı saf herediter spastik paraplejiden kompleks olanına kadar çok farklı fenotiplerle ilişkili olabildiği gösterilmiştir. Ayrıca, kohort grubunun çoğunluğunda ateşli bir hastalıktan sonra dönüşümsüz veya geri dönüşümlü nörolojik dekompensasyon olduğuna dikkat çekilmiştir. Bireyler arasında MR bulgularındaki ak madde anormalliklerinde değişkenlikler olduğu bildirilmiştir. Çalışmada MMDS1 ve HSP kliniklerinin *NFU1* geni ilişkili fenotipler olduğu ileri sürülmüştür. Çalışmada *NFU1* geninde saptanan tüm varyantlar TED ve TGD teknikleri ile gösterilmiştir. Bulunan tüm varyantların popülasyon frekansı ya çok çok nadirdir veya hiç bildirilmemiştir. Protein modelleme çalışmaları ile bu varyantların tamamının protein bölgesinde yer aldığı gösterilmiştir. Bu varyantlardan 6 tanesi NifU N-terminal kısmında iken, 3 tanesi NifU C-terminal (Demir/Sülfür) bağlanma bölgesinde bulunmaktadır [77].

Çalışmamızda tespit edilen varyantın proteinde yol açtığı Phe64Ser değişimi proteinin N-terminal kısmında yer almaktadır. Bu çalışma kapsamında fonksiyon analizi yapılamadığından, varyantın proteinin yapısına ve fonksiyonuna etkisini araştırmak için protein modelleme ve simülasyon çalışmaları yapılmıştır. Mutant ve yabanıl tip

proteinlere ait simülasyon değerleri (RMSD ve RMSF değerleri), mutasyonun protein yapısını dengesizleştirdiğini göstermektedir. Tespit edilen varyant döngü (loop) bölgesinde olduğundan, ikincil yapıda olan bir değişime göre protein fonksiyonunu daha az değiştireceği beklenmektedir. Öte yandan bu varyant aromatik bir amonio asidi (Fenilalanin) polar bir aminoaside (Serin) dönüştürdüğünden aminoasit çevresindeki etkileşimleri değiştirecek potansiyeldedir. Ayrıca türler arası korunmuşluk analizine bakıldığında (Şekil 4-45) tüm türlerde aynı pozisyonda fenilalanin aminoasidi bulunmaktadır. Aminoasit korunmuş bir bölgede yer aldığından bu değişimin protein fonksiyonuna etki edebileceği de ileri sürülebilir.

Olgu-1'in *NFU1* genindeki varyantın bulunduğu bölge GC nükleotitleri olarak zengin bir bölge olduğundan daha önce yapılan TED işleminde yanlış okuma olması muhtemeldir. Literatürde TED işleminin kısıtlarından bahseden çalışmalar bulunmaktadır. Yapısal yeniden düzenlemeler, kopya sayısı değişimleri, tekrar eden dizi genişlemeleri ve GC ile AT zengin bölgelerin okunması TED işleminin kısıtları arasında yer almaktadır [84]. Bu durumun olabileceğini gösteren diğer kanıt da ilk yapılan Sanger sekanslamada anne ve babada yanlış sonuç alınmasıdır. Anne ve baba için iki defa Sanger dizileme yapılmasına rağmen sonuç homozigot mutant (c.191T>C) olarak saptanmıştır. Ancak sağlıklı çocuklardan birinin sonucu heterozigot diğerinin homozigot yabanıl tip olmasından dolayı anne-babadaki varyant sonucunun GC zengin bölgedeki polimorfizmlerden kaynaklı olduğu düşünülerek yeniden primer dizaynı yapılmış ve yeni okumada anne-baba heterozigot olarak tespit edilmiştir. Sanger dizilemede yaşanan süreç daha önce yapılan TED sonucunun yanlış olmasının sebebini açıklayan ek bir bulgu olarak düşünülebilir. Öte yandan TGD işleminde primerlerin daha geriden bölgeye bağlanması ve amplikonların boyutunun uzun olması, dizileme işleminin doğru sonuç vermesini sağlamış olması muhtemeldir.

Olgu-1 ve ailesi için gerçekleştirilen TGD işlemi, biyoinformatik analizler, *NFU-1* geninde bulunan varyantın mutasyonel analizi, Sanger konfirmasyonu ve segregasyon çalışması, protein modelleme ve simülasyon çalışmaları, literatürde bulunan benzer çalışmalar göstermektedir ki *NFU1* geninde bulunan c.191T>C değişimi Olgu-1'in fenotipi için oldukça kuvvetli bir adaydır.

5.2. Aile 2

Ebeveynleri birinci derece kuzen olan 8 yaşındaki erkek olgunun multipl konjenital anomalisi ve mental retardasyonu bulunmaktadır. Serebral palsi ön tanısı düşünülen olgunun ailenin ifadesine göre enfeksiyon sonrası emeklemesi durmuş ve kasılmalar başlamıştır. Etkilenmiş kardeşte de aynı bulgular bulunmaktadır. Daha önce TED analizi yapılmış ve herhangi bir anlamlı varyant olmadığı raporlanmıştı. Ancak TED ham verisi bu çalışma kapsamında tekrar analiz edildiğinde *ATP8A2* geninde anlamlı bir varyant saptanmıştır. Genomun başka bölgelerinde klinikle ilişklili varyant bulunma ihtimaline binaen TGD analizi yapıldığında *ATP8A2* geni ile birlikte *ARID1A, PIK3R1, ATP5F1A* ve *ARX* genlerinde varyant bulunmuştur. Ancak yapılan incelemelerde sadece *ATP8A2* genindeki c.1400A>G varyantı, fenotip ve kalıtım modeli ile uyumlu bulunmuştur.

ATP8A2 geni, hücre zarlarında yer alan aminofosfolipidlerin translokasyonundan sorumlu 14 üyeli P-ATPaz ailesinin sınıf 1a üyesi olan ATP8A2 proteinini kodlamaktadır. Beyin, retina, testis ve spinal kord, ATP8A2'nin en fazla eksprese olduğu dokulardır. Fosfatidilserin ve fosfatidiletanolamin olarak bilinen fosfolipitler ATP8A2'nin substratlarıdır. ATP8A2, bu fosfolipitleri hücre zarı boyunca çevirerek fosfolipit asimetrisini sağlamaktadır [91,92].

ATP8A2 proteinin; ATP bağlanma bölgesi olan N-domaini, fosforilasyonun olduğu P-domaini, fosforile durumda olan ara ürünlerin defosforile edildiği A-domaini ve lipitlerin zar boyunca translokasyonunu sağlayan 10 adet transmembran segmenti içeren M-domaini bulunmaktadır. ATP8A2 ile fonksiyonel bir heterodimer kompleks oluşturan CDC50 (TMEM30A olarak da bilinir) de ek bir alt birimdir [93,94].

ATP8A2 genindeki germline patojenik varyantların otozomal resesif kalıtım gösteren serebellar ataksi, zeka geriliği ve denge kuramama sendromu 4 olarak adlandırılan (CAMRQ4; OMIM: 615268) bir hastalığa neden olduğu bilinmektedir [95].

Onat ve ark. (2013) tarafından akraba evliliği olan bir Türk ailede *ATP8A2* geni ve insan fenotip ilişkisi tanımlanmıştır. Etkilenmiş bireylerde zeka geriliği, disartri ve dört ayak (quadrupedal) duruşun eşlik ettiği/etmediği gövdesel ataksi ile karakterize ciddi bir nörolojik hastalık gözlenmiştir. *ATP8A2* geninin 12. ekzonunda yer alan c.1128C>G varyantı, 376. pozisyondaki izolosin aminoasidini metiyonin aminoasidine dönüştürmüştür. Varyant, ATP8A2 proteinin E1 E2 ATPaz domaininde yer aldığından

ve diğer canlı türlerinde de korunmuş bir pozisyonda yer aldığından olgunun fenotipinden sorumlu mutasyon olarak düşünülmüştür [96].

McMillian ve ark. (2018) tarafından bildirilen 11 hastada, biallelik *ATP8A2* mutasyonlarına eşlik eden gelişme geriliği, ağır hipotoni, hareket problemleri, koreoatetozis, distoni, yüz kasları diskinezisi, optik atrofi, beslenme güçlükleri, bir kısmında kortikal atrofi ve geç miyelinlenme gibi bulgular gösterilmiştir. Optik atrofi neredeyse tüm olgularda ilk yaşlarda saptandığından oftalmolojik değerlendirmelerin tekrar edilmesi gerektiği ifade edilmiştir. Esas bulguların erken bir şekilde bilinmesi, bu kompleks nörolojik hastalığın tanısını da kolaylaştıracağı dile getirilmiştir [97].

Heidari ve ark. (2021) tarafından bildirilen üç hasta ile fenotipik spektrumun genişlediği ve bilişsel gerilik, global gelişme geriliği, yürümenin gelişmemesi, konuşmanın gelişmemesi ya da geç başlaması, nöbetler, hipotoni, kas güçsüzlüğü, koreoatetozis, tremori bas titübasyonu, üst ve alt ekstremitelerde azalmış derin tendon refleksleri, distoni, yüz kasları diskinezisi, mikrosefali, oftalmopleji, optik atrofi, beslenme güçlükleri, diş anomalileri, diş eti hiperplazisi ve eklem katılığı gibi bulguların da hastalık tablosuna dahil olabileceği bildirilmiştir. Hastaların birinde katlanma bölgesinde c.1868-2A>G varyantı, ikincisinde c.2473G>C; (p.Asp825His) varyantı ve üçüncü hastada c.1312A>G; (p.Met438Val) varyantı bildirilmiştir. Tahmin araçlarında tüm varyantların zararlı olduğu ve oldukça korunmuş bölgelerde yer aldığı bildirilmiştir. Asp825 aminoasidi P-domaininde Mg^{+2} iyonunu koordine etmekte, Met438 aminoasidi ise N-domainin P-domainine kıvrıldığı pozisyonda yer almaktadır. Hücre kültürü çalışması yapılarak Asp825His varyantının ekspresyon seviyesine bakılmış, Asp825His varyantının yanlış katlandığı ve ATP8A2-CDC50 kompleksini oluşturamadığı gösterilmiştir [98].

Literatürde bildirilen olguların klinik bulguları gibi çalışmamıza dahil edilen Olgu-2'nin de mental retardasyonu, konjenital anomalileri, gelişme geriliği, hipotoni ve diş anomalileri gibi bulguları bulunmaktadır. TGD analizinde tespit edilen *ATP8A2* genindeki varyant, Olgu-2 ve etkilenmiş kardeşinin Sanger dizileme sonucunda homozigot mutant ve baba ile sağlıklı kardeşte heterozigot olarak tespit edilmiştir. Ancak, annenin varyantı homozigot olarak taşıdığı gözlenmiştir. Bu nedenle annede tekrarlanan Sanger dizilemede yine aynı sonuç (homozigot mutant) alınmıştır. Bu teknik bulgular ile saptanan varyantın ailedeki fenotiple ilişkili olmadığı düşünülebilir. Ancak literatürdeki

çalışmalarda *ATP8A2* genindeki yanlış anlamlı varyantların benzer fenotiplere katkı sunduğu ileri sürüldüğünden, çalışmamızda tespit ettiğimiz varyantın klinik tabloya eşlik edebileceği düşünülmüş ve ileri çalışmalar yapmanın gerekli olduğu kanaatine varılmıştır [96, 98]. Annenin homozigot mutant bir genotipte olmasına rağmen fenotipik bulgusunun olmaması; epigenetik faktörler, gen düzenleyici bölgelerdeki etkileşim ve multigenik kalıtım gibi faktörlerin fenotipik farklılıklara yol açabildiği durumlarla açıklanabilir. Öte yandan sanger sekanslama hatası olarak allel drop olasılığında annedeki c.1400 noktası yanlış okunmuş olabilir. Sanger sekanslama iki ayrı koşulda tekrar edilmesine rağmen aynı sonuç elde edilmiş, ancak bu sonuç yanlış okuma ihtimalini dışlamamıştır. Tez bütçesi ve zaman problemi olduğundan sadece sanger dizileme adımı tekrarlandı. Bölgeye özel yeniden primer tasarımı ve allel drop out ihtimalinin dışlanması ileri çalışmaların planlarına dahil edilebilir.

Daha önce başka bir merkezde yapılan TED analizi sonunda oluşturulan raporda *ATP8A2* genindeki c.1400G>A varyantına değinilmemesi, filtreleme yapılırken ilgili bu genin incelemeye/filtremeye alınmaması olabilir. Bu durum TED veri analizinde gerçekleştirilen filtreleme basamaklarının bilinmeyen varyantların tanımlanması üzerinde oldukça önemli bir role sahip olduğunun göstergesidir.

Çalışmamızda tespit edilen *ATP8A2* genindeki c.1400G>A (p.Arg467Gln) değişimi proteinin katlanma bölgesinde yer almaktadır. Bu değişimin proteinin yapısına ve fonksiyonuna etkisini araştırmak için protein modelleme ve simülasyon çalışması, proteinin tüm sekans ve mutant aminoasiti içeren yapının modellemesi de yapılmıştır. Mutant ve yabanıl tip proteinlere ait simülasyon değerleri (RMSD ve RMSF değerleri), mutasyonun protein yapısını dengesizleştirdiğini göstermektedir. 467. pozisyondaki arjinin aminoasidi pozitif yüklü bir aminoasittir. Glutamin aminoasidi ise polar bir aminoasittir. Arjinin aminoasidi glutamin aminoasidine dönüştüğünden fonksiyonel bir değişim beklenebilir. Öte yandan türler arası korunmuşluk analizleri incelendiğinde (Şekil 4-47) Gallus ve Xenopusta aynı pozisyonda glutamin aminoasidi bulunmaktadır. Dolayısıyla ilgili varyantın, proteinde büyük bir fonksiyon değişikliğine sebep olmayacağı da ileri sürülebilir.

Olgu-2 ve ailesi için gerçekleştirilen TGD işlemi, biyoinformatik analizler, *ATP8A2* geninde bulunan varyantın mutasyonel analizi, Sanger konfirmasyon ve segregasyon çalışması, protein modelleme ve simülasyon çalışmaları, literatürde bulunan

benzer çalışmalar göstermektedir ki *ATP8A2* geninde bulunan değişim Olgu-2'nin fenotipi için bir aday olabilir.

5.3. Aile 3

Ebeveynleri birinci derece kuzen olan olgunun gelişme geriliği, bilişsel gelişim geriliği, konjenital mikrosefali ile kriptoorşidizmi gibi malformasyonları ve dismorfik yüz bulguları bulunmaktadır. Olgunun TGD analizi sonucunda *DNMT3B* geni dışında *HSPA9, SOX18* ve *FIG4* genlerinde varyant bulunmuş ancak otozomal resesif kalıtım modeline uygunluk gösteren ve patojenik varyantları immünyetmezlik sentromik instabilite-yüz anomalileri sendromu 1 (ICF1; OMIM: 242860) ile ilişkilendirilen *DNMT3B* genindeki varyant, fenotipik bulgular için en güçlü aday varyant olarak düşünülmüştür.

ICF sendromu; dismorfik yüz bulguları ve immünoglobulin eksikliğine bağlı immun yetmezlik ile karakterize nadir genetik bir hastalıktır. Fitohemagglütinin (PHA) ile uyarılmış lenfosit kültürlerinde elde edilen metafazlarda 1., 9. ve 16. kromozomlarda dallanma yapısının varlığı, ICF sendromu için özgün sitogenetik bir bulgudur [99]. ICF1 sendromlu olguların bir kısmında, *DNMT3B* genindeki patojenik varyantlar ile açıklanabilecek DNA hipometilasyonunun genomun bir kısmında gözlendiği bildirilmiştir [100]. Özellikle 1., 9. ve 16. kromozomların satelit 2 ile 3 tekrar kısımlarında DNA metilasyonun düştüğü gösterilmiştir [101, 102]. *DNMT3B* genindeki A603T/STP807ins and V699G/R54X varyantların DNA hipometilasyona sebep olduğu fonksiyonel bir çalışmada gösterilirken, aynı çalışmada *BOLL, SYCP2* ve *NCRNA00221* genlerinin de DNA hipometilasyonuna sebep olabileceği de gösterilmiştir [103].

İmmunglobulin eksikliği ICF sendromunda saptanan bir bulgu olduğundan vakalara ICF sendromu erken tanısı konulduktan sonra immunglobulin takviyesi ile hastalığın seyrinin değişebileceği ifade edilmiştir. Ayrıca özellikle ağır enfeksiyon semptomları olan hastalar için allojenik kök hücre naklinin terapötik bir seçenek olması gerektiğine de dikkat çekilmiştir [104].

Weemas ve ark. (2013), ICF sendromuna *DNMT3B* ve *ZBTB24* genlerindeki mutasyonların sebep olduğunu göstermişlerdir. Bu sendromun klinik özelliklerini karakterize etmek ve genotip-fenotip korelasyonunu belirlemek için 44 ICF hastasının klinik ve genetik verileri karşılaştırılmıştır. Hastaların 23 tanesinde *DNMT3B* genindeki

mutasyonlar saptanarak bu gendeki mutasyonların olduğu grup, ICF1 sendromu şeklinde sınıflandırılmıştır. ICF2 sendromu olarak tanımlanan hastaların 13 tanesinde ise *ZBTB24* geninde mutasyonlar saptanmıştır. Herhangi bir gende varyant tespit edilemeyen 8 hasta ise ICFX olarak sınıflandırılmıştır. Hastaların tümü benzer immunolojik, morfolojik ve epigenetik bulgular gösterse de sistematik değerlendirmede bazı farklılıkların olduğu tespit edilmiştir. ICF1 hastalarında daha çok hümoral immün yetmezlik bulunurken, ICF2 hastalarında daha çok bilişsel gerilik ön planda yer almakta, konjenital malformasyonlar ise her iki hasta grubunda da görülmektedir [105]. Bu tespitler hastalığın klinik teşhisinde ve mutasyon analizinde oldukça önemli yönlendirmeler sağlamaktadır.

Kellner ve ark. (2019); dismorfik yüz bulguları olmayan, gelişme geriliği saptanan, rinovirüs ve enterovirüs gibi viral enfeksiyonlar geçiren, normal sayıda T ve B hücreli, agamaglobulinemisi olan 14 aylık bir bebeğin mutasyon analizinde *DNMT3B* geninde c.2177T>G ve c.2281G>C değişimlerini heterozigot olarak saptamıştır. Parental analizler ise anne ve babanın bu varyantlar için bileşik heterozigot olduğunu göstermiştir. 726. pozisyondaki valini glisine çeviren c.2177T>G varyantı C-5 metiltransferaz bölgesinin C-terminal kısmına yakın pozisyonda yer almaktadır [106]. Bu varyantların saptandığı ICF1 sendromlu 3 vaka daha bildirilmiştir [107,108]. Diğer varyant olan c.2281G>C varyantı 21. ekzonda yer almakta ve 761. pozisyondaki glutamik asidi glutamine çevirmektedir. Bu değişim C-5 metiltransferaz bölgesinin C-terminal kısmına yakın evrimsel açıdan oldukça korunmuş bir bölgede yer almaktadır. Tahmin araçları da bu değişimi patojenik olarak sınıflamaktadır. Vakada *DNMT3B* geninde biallelik mutasyonlar görüldüğünden kromozom analizi yapılmış ve ICF1 sendromu için karakterize 1., 9. ve 16. kromozomlarda sentromerik yoğunlaşma, tüm kol delesyonu, translokasyon, izokromozom ve çoklu radyal konfigrasyonlar gibi perisentrik kromozom bozuklukları saptanmıştır. Dolayısıyla sitogenetik bulgular da *DNMT3B* mutasyonlarını doğrulamıştır [106]. Bu sendrom için *DNMT3B* mutasyonlarının yanında sitogenetik bulguların varlığı da tanı için kritik bir öneme sahiptir. Dolayısıyla genetik danışma ve tedavi yönlendirme için de sitogenetik bulgular önem arz etmektedir.

DNMT1, DNMT3A ve DNMT3B enzimleri; kromatin yapısı ve transkripsiyon regülasyonunda temel bir adım olan DNA metilasyonunun başlamasından ve devam ettirilmesinden sorumlu olan üç DNA metiltransferaz enzimidir. DNMT3L katalitik bir inaktif kofaktör olarak DNMT3A ve DNMT3B enzimlerini aktive eder [109,110].

DNMT3B fonksiyonunda meydana gelen problemlerin karsinogenezde rol aldığı ve bazı genetik hastalıklarla ilişkili olduğu bildirilmiştir. DNMT3B proteinin farklı dokularda alternatif kırpılma ile birçok izoformunun olduğu gösterilmiştir [111,112]. Emburg ve Robertson (2011), *DNMT3B* geninde meydana gelen mutasyonların ve kırpılmaların; DNA'ya bağlanma yeteneğini, diğer DNMT3'lerle etkileşimini ve DNA'yı metilleme fonksiyonunu değiştirdiğini göstermiştir. Bu çalışmada özellikle DNMT3B'nin katalitik bölgesi olan C-terminal bölgesinin proteinin fonksiyonu için ne kadar önemli olduğu saptanmıştır. Ayrıca alternatif kırpılmaların ve protein domain kaybının DNMT3B'nin hücresel lokalizasyonunu etkilediği de gösterilmiştir. Sonuç olarak inaktif DNMT3B izoformlarının genomik metilasyon modelini etkilediği öne sürülmüştür [113].

Moarefi ve Chedin, ICF sendromu mutasyonlarının, DNMT3B aracılığı ile gerçekleşen DNA metilasyonunda birçok biyokimyasal bozukluğa sebep olduğunu göstermişlerdir. Bölgesel yönlendirilmiş mutasyon (site-directed mutagenesis) yöntemi ile *DNMT3B* geninde varyantlar oluşturulmuştur. Bu yöntemle DNMT3B fonksiyonunun birçok yeni ve önemli yönleri ortaya konmuş, DNMT3B proteinin S-adenozil L-metiyonine (SAM) sabit bir şekilde bağlanması için DNA kofaktörüne ihtiyaç duyduğu gösterilmiştir. *DNMT3B* geni mutasyonlarının; homo-oligomerizasyon, SAM kullanımı ve DNA bağlanması gibi fonksiyonlarda bozulmaya yol açtığı gösterilmiştir. V606A, C651S, A766P, H814R, R840Q ve R823G varyantlarının proteinin katalitik özelliklerini değiştirdikleri gösterilmiştir [114].

Çalışmamızda *DNMT3B* geninde homozigot olarak saptanan varyantın (c.1882G>T) neden olduğu amino asit değişikliğinin yer aldığı 628. kodon SAM-bağımlı metiltransferaz C5-tipi domaininde ve S-adenozil-L-metiyonin bağlanma bölgesinde yer almaktadır (PM1). ACMG tarafından "önemi bilinmeyen varyant" olarak sınıflandırılmış bu varyant, total allel frekansı popülasyon veritabanlarında bildirilmemiş ve bildiğimiz kadarıyla ClinVar veri tabanında da yer almayan bu varyant bilimsel literatürde şimdiye kadar yayımlanmamıştır. Varyant analizi yapan veritabanları bu varyant için gen veya gen ürününde "zarar verici" bir etkiyi desteklemektedir (PP3).

Çalışmamızda saptanan varyant, DNMT3B proteininin 628. pozisyondaki valin aminoasidini lösin aminoasidine dönüştürmektedir. İki aminoasit de hidrofobik karakterde olmasına karşın yan zincirler göz önünde bulundurulduğunda CH_2 grubunun eklenmesiyle mikro çevrede aminoasitler arası etkileşimlerin bozulabileceği göz önünde

bulundurulmalıdır. Bu çalışma kapsamında fonksiyon analizi yapılamadığından, varyantın proteinin yapısına ve fonksiyonuna etkisini araştırmak için protein modelleme ve simülasyon çalışmaları yapılmıştır. Mutant ve yabanıl tip proteinlere ait simülasyon değerleri (RMSD ve RMSF değerleri), mutasyonun protein yapısını dengesizleştirdiğini göstermektedir. Ayrıca türler arası korunmuşluk analizine bakıldığında (Şekil 4-47) tüm türlerde aynı pozisyonda valin aminoasidi bulunmaktadır. Korunmuş bir bölgede yer aldığından bu değişimin protein fonksiyonuna etki ettiği de ileri sürülebilir. Elde edilen tüm veriler, olgumuzda saptanan varyantın patojenik olma olasılığını güçlendirmektedir.

Literatürde bildirilen vakalarla örtüşen bir şekilde olgumuzda da bilişsel gerilik, gelişim geriliği, dismorfik bulgular ve immünyetmezlik ile ilişkilendirilebilecek kronik enfeksiyon hikayesi bulunmaktadır. Enfeksiyon nedeniyle aldığı antibiyotik tedavisine bağlı olarak lenfosit kültüründe üreme sağlanamayan olgumuzun sitogenetik bulguları tanıda katkı sağlayamamıştır. Tespit edilen varyant SAM bölgesinde bulunan bir aminoasidi değiştirdiğinden bu varyantın DNA metilasyon profilini değiştirebileceği öngörülmektedir. Bu çalışma kapsamında imkan olmadığından ileri çalışmalarda olgunun metilasyon profilinin incelenmesi de planlanabilir. Ayrıca sık sık enfeksiyon geçirdiği için immunoloji polikinliğine yönlendirilen olguda immunglobulin eksikliği gösterilirse saptanan varyantın immunglobulin düzeyine etkisi de araştırılabilir. Fonksiyonel çalışmalarla ilgili varyantın etkisi daha detaylı olarak incelenebilir. Sonuç olarak elde edilen varyant; yapılan çalışmalar ve literatür kıyaslamasıyla fenotipten sorumlu güçlü bir aday varyant olarak ileri sürülebilir.

5.4. Sonuç

Yeni Nesil Dizileme araçlarından TGD ve TED'in tercih edilmesi için farklı gerekçeler olsa da literatürde iki tekniği karşılaştıran çalışmalar, hangi tekniğin kullanılması gerektiği ile ilgili yönlendirici olmaktadır. Belkadi ve ark. (2015), TGD tekniğinin ekzom varyantlarını saptama bakımından TED tekniğine göre birçok parametre ile daha güçlü bir yöntem olduğunu bildirmişlerdir. Akraba olmayan altı bireyin TGD ve TED verileri karşılaştırıldığında; iki yöntem ile ekzom bölgelerinden elde edilen SNV ve indel sayılarının birbirine yakın olduğu, yüksek kalitede elde edilen varyant sayılarının da yakın olduğu, yüksek kalitede yalnızca TED ve yalnızca TGD tarafından saptanan varyant sayısının TGD'de fazla olduğu, Sanger dizileme ile rastgele seçilen varyantlar arasında "yanlış pozitif varyant" oranının TED yöntemiyle elde edilen

varyantlarda çok daha fazla olduğu, TGD ile tespit edilip TED tarafından tespit edilemeyen varyant sayısının (656), TED tarafından tespit edilip TGD tarafından tespit edilemeyen varyant sayısından (26) çok daha fazla olduğu ve kopya sayısı varyantlarını tespit etmede TED'in güvenilir olmadığı saptanmıştır. Daha maliyetli olmasına rağmen tüm bu veriler, TGD'nin TED'e göre daha güçlü bir yöntem olduğunu göstermektedir [115].

Meienber ve ark. (2016), klinik dizilemede özellikle ekzonların GC zengin bölgelerini TED yönteminin yeterince kapsamadığı, TGD yönteminin ise bu bölgeleri de kapsayarak daha doğru sonuçlar vermesi nedeniyle Mendeliyen hastalıkların tanısı için PCR gerektirmeyen TGD yönteminin tercih edilmesi gerektiğini önermişlerdir [116]. Lelieveld ve ark. (2015) da TGD yönteminin TED'e kıyasla genomun kodlayan kısımlarını daha derin okumalarla kapsadığını ve daha nadir varyantların saptanması için TGD'nin daha üstün olduğuna dikkat çekmişlerdir [117].

Genel olarak TGD'nin TED'e karşı üstünlüğü sıralandığında öncelikle TGD genomun kodlayan ve kodlamayan bölgelerinde SNV, indel, SV ve CNV'leri tespit edebilirken, TED ile yalnızca kodlayan bölgeler ile düzenleyici bölgelerde değişimleri tespit edebilir [116]. TGD daha güvenilir dizileme kapsamına sahipken TED'in yakalama problarının hibridizasyon verimliliğindeki farklılıklar, genomda hiç kapsamadığı veya çok az kapsadığı yerlerin olmasına sebep olabilir. TGD'nin kapsama homojenliği TED'e göre üstündür. Genomun birbirine benzeyen sekans bölgelerine yönelik verimli TED yakalama problarını üretmek güç olduğundan TED ile hedef dışı bölgeleri yakalama olasılığı yüksektir. TGD yönteminde kütüphane hazırlığı için özellikle GC zengin bölgelerde hatalı okumalara sebep olan PCR amplifikasyonu gerekli değildir. TED yönteminde yakalama için en az 1ug DNA gerekli olduğundan PCR amplifikasyonu gereklidir. TGD yönteminde sekans okuma uzunluğunun bir limiti yoktur. Çok uzun amplikonlar okunabilir. Bu da hatalı okumaların önüne geçmek için iyi bir avantajdır. Ancak TED yönteminde genellikle 120 bazdan daha kısa problar tasarlanır. Bu da okuma uzunluğunun küçük olmasına neden olur. TGD yönteminde TED yönetimdeki kapsama seviyesine ulaşmak için çok daha az ortalama okuma derinliği yeterlidir. TGD yönteminde referans genoma eşleşmede yaşanan hatalar minimum düzeydedir. Ancak TED yönteminde problar heterozigot noktalarda daha çok referans alleli çoğaltmaya meyilli olduğundan yanlış negatif SNV tespit edilir [118]. Ayrıca TGD'nin, TED

yönteminde tespit edilemeyen Mendel hastalıklarına dahil olduğu bilinen bazı genler dahil olmak üzere, yaklaşık %16'sı homozigot olan, yüzlerce potansiyel olarak zararlı kodlama SNV'sini tespit edebildiği gösterilmiştir [115].

Bu tez çalışmasında fenotipik bulguları ile nadir genetik hastalıktan şüphelenilen ancak rutin tanı yöntemleri ve TED analizi ile herhangi bir sendrom veya hastalıkla ilişkilendirilmeyen üç olgunun, TGD yöntemi ile etiyolojisi aydınlatılmış ve protein modelleme-simülasyon çalışmaları ile saptanan mutasyonların protein yapılarını dengesizleştirdiği gösterilmiştir. Üç olguda tespit edilen varyantlar ilk kez saptanmasına rağmen literatürdeki benzer fenotipik bulguları olan vakaların aynı genin benzer domainlerinde mutasyonlarının olması, bu varyantların da fenotipten sorumlu olabileceği görüşünü desteklemiştir. Olgu-1'in TGD analizi sonucunda bildirilen varyantın daha önce yapılan TED analizinde olmadığı gösterilmiştir. GC zengin bir bölgede yer alan varyantların TED analizinde tespit edilememe sorunu literatürde de daha önce bildirilmişti. Dolayısıyla bu çalışma ile varyantın tespitinde TED'e karşı TGD'nin üstünlüğü bir kez daha gösterilmiştir. Çalışmanın sonucunda ileri çalışmalar için yönlendirici sonuçlar elde edilmiştir. Fonksiyonel çalışmaların yapılması için literatürdeki çalışmalara da bakılarak bir yol haritası oluşturulabilir. TGD testinin kapsamı dışında kalan kromozomal yeniden düzenlemeler ve epigenetik mekanizmaların fenotipi etkileyebileceği TGD ile tanı konulamayan olgularda göz ardı edilmemelidir.

KAYNAKLAR

1. Pogue ER, ve ark. Rare genetic diseases: update on diagnosis, treatment and online resources. *Dru Dis T.* 2018; 23: 187-195. doi: 10.1016/j.drudis.2017.11.002
2. Wakap SN, ve ark. Estimating cumulative point prevalence of rare diseases: analysis of the Orphanet database. *Eur J Hum Genet*. 2020; 28(2): 165–173. doi: 10.1038/s41431-019-0508-0
3. Rahit KMTH, Graovac MT. Genetic Modifiers and Rare Mendelian Disease. *Genes.* 2020; 11(3): 239. doi: 10.3390/genes11030239
4. Deng Y, Pan W. Significance testing for allelic heterogeneity. *Genetics* 2018; 210:25–32. https://doi.org/10.1534/genetics.118.301111
5. Vinkšel M. Improving diagnostics of rare genetic diseases with NGS approaches. *J Community Genet.* 2021; 12(2): 247–256. doi: 10.1007/s12687-020-00500-5
6. Bick D, ve ark. Case for genome sequencing in infants and children with rare, undiagnosed or genetic diseases. *J Med Genet.* 2019; 56(12): 783–791. doi: 10.1136/jmedgenet-2019-106111
7. Hirschhorn JN. Genomewide association studies--illuminating biologic pathways. *N Engl J Med*. 2009; 360(17):1699-701. doi: 10.1056/NEJMp0808934.
8. Fernández-Marmiesse A. NGS Technologies as a Turning Point in Rare Disease Research, Diagnosis and Treatment. *Curr Med Chem.* 2018 Jan; 25(3): 404–432. doi: 10.2174/0929867324666170718101946
9. Faviez C, ve ark. Diagnosis support systems for rare diseases: a scoping review. *Orphanet J Rare Dis.* 2020 Apr 16; 15(1):94. doi: 10.1186/s13023-020-01374-z.
10. Posey JE. Genome sequencing and implications for rare disorders. *Orphanet J Rare Dis.* 2019; 14: 153. doi: 10.1186/s13023-019-1127-0
11. Claussnitzer M, ve ark. A brief history of human disease genetics. *Nature.* 2020; 577(7789): 179–189. doi:10.1038/s41586-019-1879-7.
12. Bick D, Jones M, Taylor SL, Taft RJ, Belmont J. Case for genome sequencing in infants and children with rare, undiagnosed or genetic diseases. *J Med Genet*. 2019; 56(12):783-791. doi: 10.1136/jmedgenet-2019-106111.
13. Sun H, ve ark. Next-Generation Sequencing Technologies and Neurogenetic Diseases. *Life.* 2021 Apr; 11(4): 361. doi: 10.3390/life11040361

14. Baird PA, Anderson TW, Newcombe HB, Lowry RB. Genetic disorders in children and young adults: a population study. *Am J Hum Genet*. 1988; 42(5):677–93.
15. McKusick VA. Mendelian inheritance in man and its online version, OMIM. Am *J Hum Genet*. 2007;80(4):588–604.
16. Lupski JR. Genomic disorders: structural features of the genome can lead to DNA rearrangements and human disease traits. *Trends Genet*. 1998;14(10):417–22.
17. Posey JE, O'Donnell-Luria AH, Chong JX, Harel T, Jhangiani SN, Coban Akdemir ZH, ve ark. Insights into genetics, human biology and disease gleaned from family based genomic studies. *Genet Med*. 2019;21(4): 798–812.
18. Wetterstrand KA. *DNA sequencing costs: data from the NHGRI genome sequencing program (GSP),* Erişim 5.02.2019, www.genome.gov/sequencingcostsdata
19. Hirschhorn JN. Genomewide association studies--illuminating biologic pathways. *N Engl J Med*. 2009; 360(17): 1699-1701.
20. Jamuar SS, Tan EC. Clinical application of next-generation sequencing for Mendelian diseases. *Hum Genomics*. 2015; 9:10
21. Delanne J, Nambot S, Chassagne A, ve ark. Secondary findings from whole-exome/genome sequencing evaluating stakeholder perspectives. *Eur J Med Genet.* 2019; 62:103529
22. Kalia SS, Adelman K, Bale SJ, ve ark. Recommendations for reporting of secondary findings in clinical exome and genome sequencing. *Genet Med.* 2019 19:249–255. https://doi.org/10.1038/gim.2016.190
23. Claustres M, Kožich V, Dequeker E, ve ark. Recommendations for reporting results of diagnostic genetic testing (biochemical, cytogenetic and molecular genetic). *Eur J Hum Genet.* 2014; 22:160–170. https://doi.org/10.1038/ejhg.2013.125
24. Hehir-Kwa JY, Claustres M, Hastings RJ, ve ark. Towards a European consensus for reporting incidental findings during clinical NGS testing. *European Journal of Human Genetics*. 2015; 1:1601–1606.
25. Van El CG, Cornel MC, Borry P, ve ark. Whole-genome sequencing in health care. *Eur J Hum Genet*. 2103; 21:580–584. https://doi.org/10.1038/ejhg.2013.46

26. Richards S, Aziz N, Bale S, ve ark. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med.* 2015; 17:405–424. https://doi.org/10.1038/gim.2015.30
27. Trujillano D, Bertoli-Avella AM, Kumar Kandaswamy K, Weiss MER, Köster J, Marais A, ve ark. Clinical exome sequencing: results from 2819 samples reflecting 1000 families. *Eur J Hum Genet*. 2017; 25:176–182. https://doi.org/10.1038/ejhg.2016.146
28. Vears DF, Sénécal K, Borry P. Reporting practices for variants of uncertain significance from next generation sequencing technologies. *Eur J Med Genet.* 2017; 60:553–558. https://doi.org/10.1016/j.ejmg.2017.07.016
29. Branford CT. *GS FLX+ System* [Internet], c1996-2016, Erişim 01.10.2022, Roche Diagnostics Co.: http://454.com/products/gs-flx-system/
30. Waltham MA. *Ion Proton System Specifications* [Internet]. Erişim 1.10.2022, https://www.thermofisher.com/us/en/home/life-science/sequencing/next generation-sequencing/ion-torrent-next-generation-sequencing-workflow/ion torrent-next-generation-sequencing-run-sequence/ion-proton-system-for-next generation-sequencing/ion-proton-system-specifications.html.
31. HiSeq X System Specifications [Internet]. Illumina Inc.; Erişim 01.10.2022, http://www.illumina.com/systems/hiseq-x-sequencing-system/performance-specifications.html.
32. Drmanac R, Sparks AB, Callow MJ, ve ark. Human genome sequencing using unchained base reads on self-assembling DNA nanoarrays. *Science*. 2010; 327(5961):78–81.
33. The original long-read sequencer [Internet]. Pacific Biosciences; Erişim 1.10.2022, http://www.pacb.com/products-and-services/pacbio-systems/rsii/.
34. Dashti MJS, Gamieldien J. A practical guide to filtering and prioritizing genetic variants. *Bio Tech*. 2017; 62:18-30
35. Miosge LA, Field MA, Sontani Y, Cho VS, Johnson A, Palkova B, Balakishnan R, ve ark. Comparison of predicted and actual consequences of missense mutations. *Proc Natl Acad Sci*. 2015; 112:E5189-E5198.

36. McCarthy DJ, Humburg P, Kanapin A, Rivas MA, Gaulton K, Cazier JB, Donnelly P. Choice of transcripts and software has a large effect on variant annotation. *Genome Med.* 2014; 6:26.
37. Annovar Filter-based Annotation, Erişim 18.12.2022, https://annovar.openbioinformatics.org/en/latest/user-guide/filter/
38. Yang H, Wang K. Genomic variant annotation and prioritization with ANNOVAR and wANNOVAR. *Nat Protoc.* 2015; 10(10): 1556–1566. doi: 10.1038/nprot.2015.105
39. Rentzsch P, ve ark. CADD: predicting the deleteriousness of variants throughout the human genome. *Nucleic Acids Res.* 2019; 47(Database issue): D886–D894. doi: 10.1093/nar/gky1016
40. Adzhubei I, ve ark. Predicting functional effect of human missense mutations using PolyPhen-2. *Curr Protoc Hum Genet*. 2013; 7:7.20. doi: 10.1002/0471142905.hg0720s76.
41. Schubert M, Jo´nsson H, Chang D, Der Sarkissian C, Ermini L, Ginolhac A, ve ark. Prehistoric genomes reveal the genetic foundation and cost of horse domestication. *Proc Natl Acad Sci.* 2014; 111: E5661–9. https://doi.org/10.1073/pnas.1416991111
42. Huber CD, ve ark. Population genetic models of GERP scores suggest pervasive turnover of constrained sites across mammalian evolution. *PLoS Genet.* 2020 May; 16(5): e1008827. doi: 10.1371/journal.pgen.1008827
43. Schwarz JM, ve ark. MutationTaster evaluates disease-causing potential of sequence alterations. *Nat Methods.* 2010; 7(8):575-6. doi: 10.1038/nmeth0810-575.
44. Pauline C, Henikoff S. SIFT: predicting amino acid changes that affect protein function. *Nucleic Acids Res.* 2003; 31(13): 3812–3814. doi: 10.1093/nar/gkg509
45. Ioannidis NM, ve ark. REVEL: An Ensemble Method for Predicting the Pathogenicity of Rare Missense Variants. *Am J Hum Genet*. 2016; 99(4):877-885. doi: 10.1016/j.ajhg.2016.08.016
46. Quang D, ve ark. DANN: a deep learning approach for annotating the pathogenicity of genetic variants. *Bioinformatics*. 2015; 31(5): 761–763. doi: 10.1093/bioinformatics/btu703

47. Karplus M, McCammon JA. Molecular dynamics simulations of biomolecules. *Nat Struct Biol*. 2022; 9: 646–652.

48. Scott AH, Dror RO. Molecular dynamics simulation for all. *Neuron Review*. 2018; 99: 1129-1143.

49. Fernandez-Leiro R, Scheres SH. Unravelling biological macromolecules with cryo-electron microscopy. *Nature*. 2016; 537: 339–346.

50. Salomon-Ferrer R, Gotz AW, Poole D, Le Grand S, Walker RC. Routine microsecond molecular dynamics simulations with AMBER on GPUs. *J. Chem. Theory Comput.* 2013; 9: 3878–3888.

51. Berneche S, Roux B. Energetics of ion conduction through the K+ channel. *Nature*. 2011; 414: 73–77.

52. Li J, Shaikh SA, Enkavi G, Wen PC, Huang Z, Tajkhorshid E. Transient formation of water-conducting states in membrane transporters. *Proc. Natl. Acad. Sci*. 2013; 110: 7696–7701.

53. Burg JS, Ingram JR, Venkatakrishnan AJ, Jude KM, Dukkipati A, Feinberg EN, ve ark. Structural basis for chemokine recognition and activation of a viral G protein-coupled receptor. *Science*. 2015; 347: 1113–1117.

54. Wacker D, Wang S, McCorvy JD, Betz RM, Venkatakrishnan AJ, Levit A,ve ark. Crystal structure of an LSD-bound human serotonin receptor. *Cell.* 2017; 168: 377–389.e12.

55. Cordero-Morales JF, Jogini V, Lewis A, Vasquez V, Cortes DM, Roux B, Perozo E. Molecular driving forces determining potassium channel slow inactivation. *Nat. Struct. Mol. Biol*. 2007; 14: 1062–1069.

56. Fields JB, Nemeth-Cahalan KL, Freites JA, Vorontsova I, Hall JE, Tobias DJ. Calmodulin gates aquaporin 0 permeability through a positively charged cytoplasmic loop. *J Biol Chem*. 2017; 292: 185–195.

57. Delemotte L, Tarek M, Klein ML, Amaral C, Treptow W. Intermediate states of the Kv1.2 voltage sensor from atomistic molecular dynamics simulations. Proc. *Natl. Acad. Sci*. 2011; 108: 6109–6114.

58. Chevalier A, Silva DA, Rocklin GJ, Hicks DR, Vergara R, Murapa P, ve ark. Massively parallel de novo protein design for targeted therapeutics. *Nature*. 2017; 550: 74–79.
59. Harder E, Damm W, Maple J, Wu C, Reboul M, Xiang JY, ve ark. OPLS3: a force field providing broad coverage of drug-like small molecules and proteins. *J. Chem. Theory Comput.* 2016; 12: 281–296.
60. Abraham MJ, Murtola T, Schulz R, Pall S, Smith JC, Hess B, Lindahl E. GROMACS: high performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX.* 2015; 1: 19–25.
61. Betz RM. Dabble. 2017; http://doi.org/10.5281/zenodo.836914 (StanfordUniversity).
62. Lindorff-Larsen K, Maragakis P, Piana S, Eastwood MP, Dror RO, Shaw DE. Systematic validation of protein force fields against experimental data. *PLoS ONE 7*. 2017; e32131.
63. Goh GB, Hulbert BS, Zhou H, Brooks CL. Constant pH molecular dynamics of proteins in explicit solvent with proton tautomerism. *Proteins.* 2014; 82: 1319–1331.
64. wwPDB Consortium. Protein Data Bank: the single global archive for 3D macromolecular structure data. *Nucleic Acids Res*. 2018; 47: D520–D528
65. SWISS-MODEL. Homo sapiens (human). 2021; https://swissmodel.expasy.org/repository/ species/9606
66. Waterhouse A, ve ark. SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res.* 2018; 46: W296–W303
67. Sillitoe I, ve ark. Genome3D: integrating a collaborative data pipeline to expand the depth and breadth of consensus protein structure annotation. *Nucleic Acids Res.* 2020; 48: D314–D319
68. Pieper U, ve ark. ModBase, a database of annotated comparative protein structure models and associated resources. *Nucleic Acids Res.* 2014; 42: D336–D346
69. Tunyasuvunakool K, Adler J, Wu Z, Green T, Zielinski M, Zidek A, ve ark. Highly accurate protein structure prediction for the human proteome. *Nature*. 2021; 596: 590–596. https://doi.org/10.1038/ s41586-021-03828-1.

70. Jumper J, Evans R, Pritzel A, Green T, Figurnov M,bRonneberger O, ve ark. Highly accurate protein structure prediction with AlphaFold. *Nature*. 2021; 596: 583–589.
71. Jungbluth H, Treves S, Zorzato F, Sarkozy A, Ochala J, Sewry C, ve ark. Congenital myopathies: disorders of excitation-contraction coupling and muscle contraction. *Nature Rev. Neurol.* 2018; 14:151–167.
72. Nowak KJ, Davies KE. Duchenne muscular dystrophy and dystrophin: pathogenesis and opportunities for treatment. *EMBO Rep.* 2004; 5: 872–876.
73. DNBSEQ-T7RS User Manual, MGI Tech Co., 2020, Erişim Tarihi 01.09.2022, https://en.mgi-tech.com/download/files/yi_id/1/type_id/1
74. Yüksel Z, Esen E, Gerdan S, Yılmaz Aİ, Gerdan ÖF. Klinik Karar Destek Sistemi Olarak: NGS Cloud. *14.Ulusal Tıbbi Genetik Kongresi*. 2020; 8.
75. Shigemizu D, ve ark. A practical method to detect SNVs and indels from whole genome and exome sequencing data. *Sci Rep*. 2013; 3: 2161-2167.
76. Samuels DC, ve ark. Heterozygosity Ratio, a Robust Global Genomic Measure of Autozygosity and Its Association with Height and Disease Risk. *Genetics*. 2016; 204 (3) : 893-904.
77. Kaiyrzhanov R, Zaki MS, Lau T, Sen S, Azizimalamiri R, Zamani M, ve ark. Phenotypic continuum of NFU1-related disorders. *Ann Clin Transk Neurol*. 2022; 9(12): 2025-2035.
78. Choi M, Scholl UI, Ji W, Liu T, Tikhonova IR, Zumbo P, ve ark. Genetic diagnosis by whole exome capture and massively parallel DNA sequencing. *Proc Natl Acad Sci USA*. 2009; 106(45): 19096-19101.
79. MacArthur DG, Manolio TA, Dimmock DP, Rehm HL, Shendure J, Abecasis GR, ve ark. Guidelines for investigating causality of sequence variants in human disease. *Nature*. 2014; 508(7497): 469-476.
80. Liu YT, Hersheson J, Plagnol V, Fawcett K, Duberley KE, Preza E, ve ark. Autosomal-recessive cerebellar ataxia caused by a novel ADCK3 mutation that elongates the protein: clinical, genetic and biochemical characterisation. *J Neurol Neurosurg Psychiatry*. 2013; 85(5): 493-498.

81. Lee YC, Durr A, Majczenko K, Huang YH, Liu YC, Lien CC, ve ark. Mutations in KCND3 cause spinocerebellar ataxia type 22. *Ann Neurol*. 2012; 72(6): 859-869.
82. Liu YT, Laura M, Hersheson J, Horga A, Jaunmuktane Z, Brandner S, etal. Extended phenotypic spectrum of KIF5A mutations: From spastic paraplegia to axonal neuropathy. *Neurology*. 2014; 83(7): 612-619.
83. Shevchenko Y, Bale S. Clinical Versus Research Sequencing. *Cold Spring Harb Perspect Med*. 2016; 6(11): 1-10.
84. Curry PDK, Broda KL, Caroll CJ. The role of RNA sequencing as a new genetic diagnosis tool. *Current Genetic Medicine Reports*. 2021; 9: 13-21.
85. Cameron JM, Janer A, Levandovskiy V, Mackay N, Rouault TA, Tong WH, ve ark. Mutations in iron-sulfur cluster scaffold genes NFU1 and BOLA3 cause a fatal deficiency of multiple respiratory chain and 2-oxoacid dehydrogenase enzymes. *Am J Hum Genet*. 2011; 89:486–95.
86. Yoon T, Cowan JA. Frataxin-mediated iron delivery to ferrochelatase in the final step of heme biosynthesis. *J Biol Chem*. 2004; 279(25):25943-25946.
87. Lill R, Dutkiewicz R, Freibert SA, ve ark. The role of mitochondria and the CIA machinery in the maturation of cytosolic and nuclear iron-sulfur proteins. *Eur J Cell Biol.* 2015;94(7–9):280-291.
88. Tsaousis AD. On the orgin of iron/sulfur cluster biosynthesis in eukaryotes. *Front Microbiol.* 2019; 10(2478): 1-10.
89. Seyda A, Newbold RF, Hudson TJ, Verner A, MacKay N, Winter S, ve ark. A novel syndrome affecting multiple mitochondrial functions, located by microcell-mediated microcell-mediated transfer to chromosome 2p14-2p13. *Am J Hum Genet*. 2001; 68:386–96.
90. Tonduti D, Dorboz I, Imbard A, ve ark. New spastic paraplegia phenotype associated to mutation of NFU1. *Orphanet J Rare Dis*. 2015; 10:13.
91. Andersen JP, Vestergaard AL, Mikkelsen SA, Mogensen LS, Chalat M, Molday RS. P4-ATPases as phospholipid flippases—Structure, function, and enigmas. *Frontiers in Physiology*. 2016; 7: 1–23. https://doi.org/10.3389/fphys.2016.00275
92. Coleman JA, ve ark. Localization, purification, and functional reconstitution of the P4-ATPase Atp8a2, a phosphatidylserine flippase in photoreceptor disc membranes. *J Biol Chem*. 2009; 284(47): 32670-32679.

93. Wang J, Molday LL, Hii T, Coleman JA, Wen T, Andersen JP, Molday RS. Proteomic analysis and functional characterization of P4-ATPase phospholipid flippases from murine tissues. *Scientific Reports.* 2018; 8(1): 10795. https://doi.org/10.1038/s41598-018-29108-z

94. van der Velden LM, Wichers CGK, van Breevoort AED, Coleman JA, Molday RS, Berger R, ve ark. Heteromeric interactions required for abundance and subcellular localization of human CDC50 proteins and class 1 P4-ATPases. *Journal of Biological Chemistry*. 2010; 285(51): 40088–40096. https://doi.org/10.1074/jbc.M110.139006

95. Cacciagli P, Haddad MR, Mignon-Ravix C, El-Waly B, Moncla A, Missirian C, Villard L. Disruption of the ATP8A2 gene in a patient with a t(10;13) de novo balanced translocation and a severe neurological phenotype. *European Journal of Human Genetics*. 2010; 18(12):1360–1363. https://doi.org/10.1038/ejhg.2010.126

96. Onat OE, Gulsuner S, Bilguvar K, Basak AN, Topaloglu H, Tan M, ve ark. Missense mutation in the ATPase, aminophospholipid transporter protein ATP8A2 is associated with cerebellar atrophy and quadrupedal locomotion. Eur *J Hum Genet*. 2013; 21(3) : 281-285.

97. McMillan HJ, Telegrafi A, Singleton A, Cho MT, Lelli D, Lynn FC, Yoon G. Recessive mutations in ATP8A2 cause severe hypotonia, cognitive impairment, hyperkinetic movement disorders and progressive optic atrophy. Orphanet Journal of Rare Diseases, 2018; 13(1): 86. https://doi.org/10.1186/s13023-018-0825-3

98. Heidari E, Harrison AN, Jafarina E, Tavasoli AR, Almadani N, Molday RS, Garshasbi M. Novel variants in critical domains of ATP8A2 and expansion of clinical spectrum. *Hum Mutat*. 2021; 42(5) : 491-497

99. Tiepolo L, Maraschio P, Gimelli G, Cuoco C, Gargani GF, Romano C. Multibranched chromosomes 1, 9 and 16 in a patient with IgA and IgE deficiency. *Hum Genet.* 1979; 51: 127–37.

100. Hansen RS, Wijmenga C, Luo P, Stanek AM, Canfield TK, Weemaes CMR, Gartler SM. The DNMT3B DNA methyltransferase gene is mutated in the ICF immunodeficiency syndrome. *Proc Natl Acad Sci.* 1999; 96: 14412–7.

101. Kondo T, ve ark. Whole-genome methylation scan in ICF syndrome: hypomethylation of non-satellite DNA repeats D4Z4 and NBL2. *Hum Mol Genet*. 2000; 9: 597–604.

102. Jeanpierre M, ve ark. An embryonic-like methylation pattern of classical satellite DNA is observed in ICF syndrome. *Hum Mol Genet.* 1993; 2: 731–5.

103. Simo-Riudalbas L, ve ark. Genome-Wide DNA Methylation Analysis Identifies Novel Hypomethylated Non-Pericentromeric Genes with Potential Clinical Implications in ICF Syndrome. *PLoS One*. 2015 ; 10(7): e0132517. doi: 10.1371/journal.pone.0132517.

104. Hagleitner MM, Lankester A, Maraschio P, Hultén M, Fryns JP, Schuetz C, Gimelli G, ve ark. Clinical spectrum of immunodeficiency, centromeric instability and facial dysmorphism (ICF syndrome). *J Med Genet*. 2008; 45(2):93-9.

105. Weemaes CM, van TolMJ, Wang J, van Ostaijen-ten DamMM, van Eggermond MC, Thijssen PE, ve ark. Heterogeneous clinical presentation in ICF syndrome: correlation with underlying gene defects. *Eur J Hum Genet*. 2013; 21(11): 1219–25. https://doi.org/10.1038/ejhg.2013.40.

106. Kellner ES, ve ark. The Value of Chromosome Analysis to Interrogate Variants in DNMT3B Causing Immunodeficiency, Centromeric Instability, and Facial Anomaly Syndrome Type I (ICF1). Journal of Clinical Immunology. volume 2019; 39: 857–859. https://doi.org/10.1007/s10875-019-00704-6

107. Hansen RS, Wijmenga C, Luo P, Stanek AM, Canfield TK, Weemaes CM, ve ark. The DNMT3B DNA methyltransferase gene is mutated in the ICF immunodeficiency syndrome. *Proc Natl Acad Sci*. 1999; 96(25): 14412–7. https://doi.org/10.1073/pnas.96.25.14412.

108. Xu GL, Bestor TH, Bourc'his D,Hsieh CL, TommerupN, Bugge M, ve ark. Chromosome instability and immunodeficiency syndrome caused by mutations in a DNA methyltransferase gene. *Nature*. 1999; 402(6758): 187–91. https://doi.org/10.1038/46052.

109. Kareta MS, Botello ZM, Ennis JJ, Chou C, Chedin F. Reconstitution and mechanism of the stimulation of de novo methylation by human DNMT3L. *J Biol Chem*. 2006; 281: 25893–25902.

110. Ooi SKT, ve ark. DNMT3L connects unmethylated lysine 4 of histone H3 to de novo methylation of DNA. *Nature.* 2007; 448: 714–717.

111. Ostler KR, ve ark. Cancer cells express aberrant DNMT3B transcripts encoding truncated proteins. *Oncogene*. 2007; 26: 5553–5563.

112. Weisenberger DJ, Velicescu M, Cheng JC, Gonzales FA, Liang G, Jones PA. Role of the DNA methyltransferase variant DNMT3b3 in DNA methylation. *Mol Cancer Res*. 2004; 2: 62–72.

113. Emburg BOV, Robertson KD. Modulation of Dnmt3b function in vitro by interactions with Dnmt3L, Dnmt3a and Dnmt3b splice variants. *Nucleic Acids Researchs*. 2011; 39(12): 4984-5002. doi:10.1093/nar/gkr116

114. Moarefi AH, Chedin F. ICF Syndrome Mutations Cause a Broad Spectrum of Biochemical Defects in DNMT3B-Mediated De Novo DNA Methylation. *J. Mol. Bio.* 2011; 409: 758-772. doi:10.1016/j.jmb.2011.04.050

115. Belkadi A, ve ark. Whole-genome sequencing is more powerful than whole-exome sequencing for detecting exome variants. *Proc Natl Acad*. 2015; 112(17) :5473-8. doi: 10.1073/pnas.1418631112. Epub 2015 Mar 31.

116. Meienberg J, Bruggmann R, Oexle K, Matyas G. Clinical sequencing: is WGS the better WES?. *Hum Genet.* 2016; 135(3): 359-62.

117. Lelieveld SH, Speilman M, Mundlos S, Veltman JA, Glissen C. Comparison of Exome and Genome Sequencing Technologies for the Complete Capture of Protein-Coding Regions. *Hum Mutat*. 2015; 36(8): 815-22. doi: 10.1002/humu.22813. Epub 2015 Jun 11.

118. Kim K, Seong MW, Chung WH, Park SS, Leem S, Park W ve ark. Effect of Next-Generation Exome Sequencing Depth for Discovery of Diagnostic Variants. *Genomics Inform.* 2015; 13(2):31-39.

HAM VERİLER

Olgu 1, 2 ve 3 için yapılan Tüm Genom Dizileme işleminde üretilen FastQ formatında ham veriler, harici bellekte saklanmaktadır.

FORMLAR

T.C. İSTANBUL ÜNİVERSİTESİ
İSTANBUL TIP FAKÜLTESİ
TIBBİ GENETİK ANABİLİM DALI

BİLGİLENDİRİLMİŞ GÖNÜLLÜ OLUR FORMU (BGOF)
GENETİK ARAŞTIRMALAR İÇİN BİLGİLENDİRİLMİŞ OLUR FORMU
(0-11 Yaş Çocuklar İçin)

AÇIKLAMA:
Bu Bilgilendirilmiş Gönüllü Olur Formu İstanbul Üniversitesi İstanbul Tıp Fakültesi Tıbbi Genetik Anabilim Dalı kliniğine başvuran kadın ve erkekleri ***"Nadir genetik sendromlu bir olgu ve ailesinde tüm ekzom ve/veya tüm genom analizleri ile genetik etiyolojinin belirlenmesi "*** araştırmasına davet etmek üzere hazırlanmıştır.

Sorumlu Araştırmacının Adı, Soyadı	: Doç. Dr. Gözde YEŞİL SAYIN
Araştırmayı yürütecek kuruluşun adı	: İstanbul Tıp Fakültesi Tıbbi Genetik ABD.
Destekleyici kuruluş	: İstanbul Üniversitesi, BAP
Araştırmanın adı	: ***"Nadir genetik sendromlu bir olgu ve ailesinde tüm ekzom ve/veya tüm genom analizleri ile genetik etiyolojinin belirlenmesi "***

Birinci Bölüm: Araştırma Hakkında Bilgi

"**Nadir genetik sendromlu bir olgu ve ailesinde tüm ekzom ve/veya tüm genom analizleri ile genetik etiyolojinin belirlenmesi "** adını verdiğimiz çalışmaya katılmak üzere davet edilmiş bulunmaktasınız. Bu çalışmada yer almayı kabul etmeden önce çalışmanın ne amaçla ve nasıl yapılacağını anlamanız ve katılıp katılmama doğrultusundaki kararınızı bu bilgilendirme sonrası özgürce vermeniz gerekmektedir. Araştırma hakkında sözlü olarak size aktaracağım bilgiler yazılı olarak da size bir sonraki bölümde sunulacaktır. Size özel hazırlanmış bu bilgilendirmeyi lütfen dikkatlice okuyunuz, sorularınıza açık yanıtlar isteyiniz. Bu belgedeki son bölüm onay işlemleri ile ilgilidir. Araştırmaya katılmayı kabul ederseniz lütfen bu bölümü imzalayınız. Okuma ve yazma konusunda engelleriniz olduğu takdirde bir tanığın gözetiminde bu belgeyi onaylamanız istenecek ve gerektiğinde parmak iziniz alınacaktır.

Araştırmanın konusu ve amacı:
Avrupa ve Türkiye'de 2000 kişide 1 kişi sıklığında görülen hastalıklar nadir hastalıklar olarak adlandırılmıştır. Nadir hastalıkların çoğu genetik kökenlidir [1]. Bu hastalıkların birçoğunun tanı ve tedavisi daha zordur. Son yıllarda gelişen teknoloji ile genetik tanı sistemleri geliştiği için birçok nadir genetik hastalığın tanısı mümkün olmuştur. Hastalıkların genetik tanısının aydınlatılması tedavisinin geliştirilmesine olanak sağlamaktadır.

Bu çalışmada akraba evliliği yapmış ailelerdeki bilinen herhangi bir sendromla ilişkilendirilememiş hasta olan bireyde ve ailelerinde ileri moleküler tekniklerle (tüm genom analizi/tüm ekzom analizi) tanının konulması hedeflenmektedir. Ayrıca geniş aile taramaları ile taşıyıcıların belirlenecek ve ailelere genetik danışma verilerek olası riskler ve sağlıklı çocuk sahibi olmalarına yardımcı teknikler anlatılacaktır. Varsa hastalığın tedavisi konusunda bilgilendirilme yapılacak ve geniş saha çalışması ile halkın genetik hastalıklarla ilgili riskleri konusunda aydınlatılması sağlanacaktır.

Yapılacak işlemin tanımı ve izlenecek yöntem:

Çalışmaya, İstanbul Tıp Fakültesi, Tıbbi Genetik Anabilim Dalına yönlendirilen ve klinik olarak belirli bir sendromla ilişkilendirilmeyen, olası genetik etyoloji düşündüren olgular ile bu olguların anne-babaları, sağlıklı ve etkilenmiş kardeşleri dahil edilmiştir.

Hastanın çalışma konusunda bilgilendirmesi ve onayının alınmasından sonra öyküsü, aile ağacı çizimi ve fizik muayene yapılacaktır. Daha sonra hastalardan steril EDTA'lı tüpe 6 ml periferik kan örneği alınacaktır. Alınan kandan ticari kit yardımı ile DNA izolasyonu yapılacaktır. Ayrıca bazı hastalarda ticari kitlerin örnek toplama yönetiminden ötürü iç yanaktan sürüntü alınacaktır. İzolasyon sonrası çalışmaya alınacak DNA örneklerinin kalitesi ölçülecektir. Kalite yeterli ise tekrar kan alınmayacaktır. Olgulardan elde edilecek DNA örnekleri uygun koşullarda (-20°C buzdolabında) çalışma için saklanacaktır. Mevcut hastalığa neden olan genetik değişimlerin aydınlatılması için bu örneklerde yeni nesil dizileme teknolojisi (Tüm Genom Dizileme/Tüm Ekzom Dizileme) ile inceleme analizleri uygulanacaktır. Veri tabanları kullanılarak genetik değişimler saptanacak ve fenotip ile ilişkili aday mutasyonların sınıflandırılması sağlanacaktır. İlişkili olduğu düşünülen ve daha önce tanımlanmamış olan değişimler, Sanger dizileme yöntemi ile konfirme edilecektir.

Olası riskler ve faydalar:

Kan alınması sırasında oluşabilecek riskler:

(1) İğne-batmasına bağlı olarak az bir acı duyulması.

(2) İğne batması sonrasında çok nadiren enfeksiyon gelişmesi durumu.

Bu olası risklerin giderilmesi için kan örneğini almadan önce deri dezenfeksiyonu yapan etanol gibi kimyasallar, gazlı pedler, yara bandı ve tek kullanımlık şırıngalar hazır bulundurulacaktır.

İç yanaktan sürüntü alınma sırasında oluşabilecek riskler:

(1) Sürüntü alınırken hafif bir acı duyulması

(2) Nadiren iç yanakta tahriş oluşması.

Bu olası risklerin giderilmesi için tecrübeli sağlık personeli örnek alımını yapacaktır.

Olası faydalar:

Genetik sebebin aydınlatılması hasta olgunun yanı sıra tüm ailede diğer riskli kişilerin belirlenmesi ve bu bireylerin evliliklerinde genetik danışmaya yönelik imkanlar sunacak, olası aday genlerin keşfi ile hastalığın temelinde rol alan moleküler mekanizmaların anlaşılmasına ve ileride yeni tedavi yöntemlerinin tasarlanmasına olanak sağlayacaktır.

Diğer seçenekler:

- Elde edilen örnekler mevcut çalışma dışında başka bir araştırma için kullanılmayacaktır.
- Çalışmamıza katılan gönüllü uygulama sırasında herhangi bir rahatsızlıkla ve riskle karşılaşmayacaktır.
- Bu araştırmada hastanın hiçbir şekilde hukuki ve mali sorumluluğu bulunmayacaktır ve tüm sorumluluk araştırıcıya aittir. Araştırmaya katılan gönüllü isteği üzerine mali ve hukuki yükümlülük olmaksızın çalışmadan ayrılabilir.
- Araştırmaya katılan gönüllünün kimlik bilgileri gizli tutulacaktır.
- Bu araştırma için gönüllüden herhangi bir ücret talep edilmeyecektir.
- Çalışmaya katılma zorunluluğunuz bulunmamaktadır.
- Çalışmaya katılmayı kabul etmemeniz halinde gelecekteki takip ve tedavinizin olumsuz şekilde etkilenmesi söz konusu değildir.
- Araştırmaya katılım süreci yalnızca kan örneği ve tükürük alınması ile sınırlıdır.
- Araştırmanın kabul kriterlerine uyum sağlamadığınız takdirde çalışmadan çıkarılabilirsiniz.
- Bu şekilde araştırma sonucunda elde edilen önemli bir bilgi olduğunda tarafınıza bildirilecektir. Ancak bu bilgiyi öğrenmemek sizin hakkınızdır. Araştırma neticesinde elde edilecek sonuçlarla uluslararası düzeyde yayınlanabilecek, hastalığınıza açıklık getirecek veriler elde edilebilir.

Hastalığınız hakkında elde edilen veriler, klinik bilgiler ve varsa çekilen fotoğraflar kimlik bilgileriniz saklı kalmak koşulu ile bilimsel dergilerde yayımlanabilir. Bu incelemelerin esas amacı sizde/aile bireyinizde saptanan anomalinin daha detaylı araştırmak ve genotip-fenotip ilişkisinin anlaşılmasına yardımcı olacak bilgilere ulaşmaktır.

İkinci Bölüm: Katılımcının Beyanı

"Doç. Dr. Gözde Yeşil Sayın tarafından İstanbul Üniversitesi Tıp Fakültesi, Tıbbi Genetik Anabilim Dalı'nda **"Nadir genetik sendromlu bir olgu ve ailesinde tüm ekzom ve/veya tüm genom analizleri ile genetik etiyolojinin belirlenmesi "** konuda bir araştırma yapılacağı belirtilerek bu araştırma ile ilgili yukarıdaki bilgiler bana aktarıldı. Bu bilgilendirmeden sonra böyle bir araştırmaya "gönüllü" olarak katılmak üzere davet edildim. Eğer bu araştırmaya katılırsam hekim ile aramda kalması gereken bana ait bilgilerin, gizlilik içinde, bu araştırma sırasında da büyük özen ve saygı ile korunacağı güvencesi verildi. Araştırma sonuçlarının eğitim ve bilimsel amaçlarla kullanımı sırasında da kişisel bilgilerimin özenle korunacağı konusunda bana yeterli güven verildi. Araştırmanın yürütülmesi sırasında herhangi bir sebep göstermeden araştırmadan çekilebilirim. Ayrıca tıbbi durumuma herhangi bir zarar verilmemesi koşuluyla araştırmacı tarafından araştırma dışı da tutulabilirim. Araştırma için yapılacak harcamalarla ilgili herhangi bir parasal sorumluluk altına girmiyorum. Tarafıma bir ücret ödenmeyeceğini biliyorum. İster doğrudan ister dolaylı olsun araştırma uygulamasından kaynaklanan nedenlerle olabilecek herhangi bir sağlık sorunumun ortaya çıkması halinde, her türlü tıbbi müdahalenin sağlanacağı konusunda gerekli güvence verildi. Bu tıbbi müdahalelerle ilgili olarak parasal bir yük altına da girmeyeceğim.

Araştırma sırasında bir sağlık sorunu ile karşılaştığımda; herhangi bir saatte, Arş. Gör. İsmail Musab Gülaçar'ı 05056519049 numaralı telefondan ve İstanbul Üniversitesi İstanbul Tıp Fakültesi Tıbbi Genetik Anabilim Dalı adresinden arayabileceğimi biliyorum. Bu araştırmaya katılmak zorunda değilim ve katılmayabilirim. Araştırmaya katılmam konusunda zorlayıcı bir davranışla karşılaşmış değilim. Eğer katılmayı reddedersem, bu durumun tıbbi bakımıma ve hekim ile olan ilişkime herhangi bir zarar getirmeyeceğini de biliyorum.

Bana yapılan tüm açıklamaları ayrıntılarıyla anlamış bulunmaktayım. Adı geçen bu araştırma projesinde **gönüllü** olarak yer alma kararını aldım. Bu konuda yapılan daveti gönüllülük içerisinde kabul ediyorum. İmzalı bilgilendirilmiş gönüllü olur formunun bir örneğinin bana verileceğini de biliyorum."

İmzalı bu formun bir kopyası bana verilecektir.

GÖNÜLLÜ ONAY FORMU
Yukarıda, gönüllüye araştırma öncesinde verilmesi gereken bilgileri okudum. Çalışma hakkında bana yazılı ve sözlü açıklamalar yapıldı. Bu koşullarla söz konusu klinik araştırmaya kendi rızamla hiçbir baskı olmaksızın katılmayı kabul ediyorum.

Çocuğunuzun (0-12 Yaş için)
Adı-soyadı:

Telefon no:
Adresi:

Veli/Vasinin (Gerekiyorsa);
Adı-soyadı:

İmzası:

Açıklamaları yapan araştırmacının
Adı-soyadı: Doç. Dr. Gözde Yeşil Sayın
İmzası:

Açıklamaları yapan araştırmacının
Adı-soyadı: Arş. Gör. İsmail Musab Gülaçar
İmzası:

İNTİHAL RAPORU İLK SAYFASI

TÜM GENOM DİZİLEME YÖNTEMİNİN GENETİK HASTALIKLARIN TANISINA KATKISI

ORİJİNALLİK RAPORU

%8	%7	%2	%2
BENZERLİK ENDEKSİ	İNTERNET KAYNAKLARI	YAYINLAR	ÖĞRENCİ ÖDEVLERİ

BİRİNCİL KAYNAKLAR

1	acikbilim.yok.gov.tr İnternet Kaynağı	%3
2	9lib.net İnternet Kaynağı	%1
3	en.mgitech.cn İnternet Kaynağı	%1
4	nek.istanbul.edu.tr:4444 İnternet Kaynağı	<%1
5	www.bioinforange.com İnternet Kaynağı	<%1
6	www.kefirtanesi.com İnternet Kaynağı	<%1
7	en.mgi-tech.com İnternet Kaynağı	<%1
8	ngs.med-gen.ru İnternet Kaynağı	<%1
9	Submitted to Istanbul University Öğrenci Ödevi	<%1