

ANKARA ÜNİVERSİTESİ
BİYOTEKNOLOJİ ENSTİTÜSÜ

TEMEL BİYOTEKNOLOJİ

DOKTORA TEZİ

NADİR MUTASYONLARIN TARAMASINA YÖNELİK DERİN DİZİLEME
ÇALIŞMALARI İÇİN ÇOK BOYUTLU HAVUZLAMA YAKLAŞIMININ
GELİŞTİRİLMESİ VE FAMİLYAL HİPERKOLESTEROLEMİ HASTA GRUBUNDA
UYGULANMASI

Haldun Doğan

Danışman Öğretim Üyesi

Prof. Dr. Hilâl Özdağ Sevgili

Aralık

2023

ETİK BEYAN

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

Haldun Doğan

İmzası

ONAY

Prof. Dr. Hilâl Özdağ Sevgili danışmanlığında Haldun Doğan tarafından hazırlanan bu çalışma 19/12/2023 tarihinde aşağıdaki jüri tarafından Biyoteknoloji Anabilim Dalı'nda doktora tezi olarak kabul edilmiştir.

Başkan: Prof. Dr. Lale ŞATIROĞLU TUFAN

İmza:

Üye: Prof. Dr. Hilâl ÖZDAĞ SEVGİLİ

İmza:

Üye: Doç. Dr. Bala GÜR DEDEOĞLU

İmza:

Üye: Prof. Dr. Serdar CEYLANER

İmza:

Üye: Prof. Dr. Hatice MERGEN

İmza:

Yukarıdaki sonucu onaylarım.

Prof. Dr. Demet CANSARAN DUMAN

Enstitü Müdürü

ÖZET

Doktora Tezi

Nadir Mutasyonların Taramasına Yönelik Derin Dizileme Çalışmaları İçin Çok Boyutlu Havuzlama Yaklaşımının Geliştirilmesi ve Familial Hiperkolesterolemi Hasta Grubunda Uygulanması

Haldun Doğan

Ankara Üniversitesi Biyoteknoloji Enstitüsü

Prof. Dr. Hilal Özdağ Sevgili

DNA dizi analizine dayanan testler düşük maliyetli olmamaları sebebiyle günümüzde genetik hastalıkların geniş kitlelerde taramasında sınırlı olarak kullanılabilmektedirler. Havuzlama yaklaşımı maliyet düşürücü etkisi sayesinde özellikle hastalık tarama programlarında tercih edilmektedir. Yeni Nesil Dizileme (YND/NGS) teknolojisinin havuzlama yaklaşımı ile birlikte kullanılması, genetik hastalıkların toplum çapında taranması önündeki en büyük engel olan maliyet problemine çözüm sağlayarak, kitlelerin genetik hastalıklarının DNA dizileme ile taranmasında kullanılabileceği öngörülmektedir.

Bu tez çalışmasında 2 ve 3 boyutlu farklı havuzlama yaklaşımları, basitlik ve kullanım kolaylığı ön planda tutularak test edilmiş, elde edilen sonuçlara göre seçilen 2 boyutlu 8x8'lik havuzlama ile 256 adet FH hastası ve 64 adet kontrol örneğinin 6 geni NGS teknolojisi kullanılarak nadir varyasyonlar açısından taranmıştır.

Çalışma kapsamında denenen havuzlama yaklaşımları arasından 2 boyutlu 8x8 ve 16x16'lık havuzlamalar başarılı, 24x24'lük havuzlama ise kısmen başarılı olarak değerlendirilmiştir. Daha yüksek havuzlamalar ise başarısız olarak değerlendirilmiştir. Taranan 320 adet örneğin 75'inde, 40 farklı mutasyondan toplam 89 adet genotip tespit edilmiştir.

2023, 176 sayfa

Anahtar kelimeler: Mutasyon tarama, NGS, yeni nesil DNA dizileme, derin dizileme, havuzlama, FH, familial hiperkolesterolemi

ABSTRACT

Ph.D. Thesis

Development of a Multi-Dimensional Pooling Approach for Scanning Rare Mutations in Deep Sequencing Experiments and Application on Familial Hypercholesterolemia Patient Group

Haldun Doğan

Ankara University Biotechnology Institute

Prof. Dr. Hilal Özdağ Sevgili

Tests based on DNA sequencing analysis have only limited application on screening of genetic diseases in masses, due to their high costs. The pooling approach especially used in disease screening programs, due to its cost reduction benefits. The combination of Next Generation Sequencing (NGS) technology and pooling approach can be used in screening of genetic diseases through the populations as it offers a solution to the high cost problem.

In this thesis, different 2 and 3 dimensional pooling approaches are tested without compromising the ease of use. Then, 256 FH patients and 64 control samples were screened through their 6 genes by using NGS technology and 2 dimensional 8x8 pooling approach, which is selected through testing different pooling approaches.

Among different pooling approaches tested, 2 dimensional 8x8 and 16x16 pooling strategies are evaluated as successful and 24x24 pooling is evaluated as partially successful. The higher pooling approaches are evaluated as unsuccessful. Among 320 samples screened, 89 different genotypes are detected from 40 different mutations in 75 of the samples.

2023, 176 pages

Keywords: Mutation screening, NGS, next-gen DNA sequencing, deep sequencing, pooling, FH, familial hypercholesterolemia

TEŞEKKÜR

Beni doktora yapmam konusunda ikna eden, öğrenciliğe kabul eden, yıllar geçmesine rağmen benden vazgeçmeyerek beni teşvik etmeye devam eden değerli hocam ve danışmanım Prof. Dr. Hilal Özdağ Sevgili'ye yürekten teşekkür ediyorum; kendisinin teşvikleri ve yol göstermesi olmasaydı bu çalışmayı bitirmem mümkün değildi.

Profesyonel ve özel hayatımda ne zaman ihtiyacım olsa bilgi, tecrübe ve desteklerini benden esirgemeyen ve tez çalışmamın maddi kaynağını sağlayan değerli patronlarım ve hocalarım Prof. Serdar Ceylaner ve Doç. Dr. Gülay Ceylaner'e yürekten teşekkür ediyorum.

Değerli vakitlerini ayırarak tez komitemde yer alan, fikirleri ile bana yol gösteren ve nezaketleri ile beni teşvik eden değerli hocalarım Prof. Dr. Davut Gül, Prof. Dr. Hatice Mergen ve Doç. Dr. Bala Gür Dedeoğlu'na yürekten teşekkür ediyorum.

Biyoinformatik alanındaki engin bilgi ve tecrübesi ile bana destek olan değerli arkadaşım Dr. Gökalg Çelik'e yürekten teşekkür ediyorum.

Tez çalışmam sırasında kullandığım materyallerin, 213S147 nolu projeye verdiği destek ile toplanmasını sağlayan TÜBİTAK'a yürekten teşekkür ediyorum.

Her zaman koşulsuz yanımda olan, benden maddi-manevi desteklerini esirgemeyen ve çalışmalarım için harcadığım vaktin bir kısmını kendilerine ayırmam gereken vakitten almama anlayış gösteren sevgili eşim Hatice ve canım çocuklarım Sinan, Nihan ve Rıdvan'a yürekten teşekkür ediyorum.

İyi ki varsınız...

Haldun Doğan

Ankara - 2023

İÇİNDEKİLER

ETİK BEYAN	i
ÖZET	iii
ABSTRACT	iv
TEŞEKKÜR.....	v
İÇİNDEKİLER.....	vi
ŞEKİLLER DİZİNİ.....	xii
ÇİZELGELER DİZİNİ.....	xv
SİMGELER DİZİNİ	xvi
1. GİRİŞ	1
2. KURAMSAL TEMELLER.....	2
2.1. DNA DİZİLEME	2
2.1.1. DNA DİZİLEMENİN TARİHÇESİ	3
2.1.1.1. MOORE YASASI VE DNA DİZİLEME	5
2.1.1.2. FREDERICK SANGER	7
2.1.2. DNA DİZİLEME TEKNOLOJİLERİ.....	9
2.1.2.1. SANGER DİZİLEME.....	9
2.1.2.2. YENİ NESİL DİZİLEME	10
2.1.2.2.1. DNA’NIN RASTGELE PARÇALANMASI (FRAGMENTASYON).....	11
2.1.2.2.2. KLONAL AMPLİFİKASYON	15
2.1.2.2.3. RASTGELE DİZİLEME	22

2.1.2.2.4. YÜKSEK HACİMLİ PARALEL DİZİLEME	30
2.1.2.2.5. ELDE EDİLEN KISA OKUMALARIN BİRLEŞTİRİLMESİ.....	32
2.1.2.2.6. YENİ NESİL DNA DİZİLEMENİN AVANTAJLARI	34
2.1.2.2.7. YENİ NESİL DNA DİZİLEMENİN DEZAVANTAJLARI	35
2.1.2.3. 3. NESİL DİZİLEME	36
2.1.2.3.1. SMRT SEKANS LAMA	37
2.1.2.3.2. NANOPOR SEKANS LAMA.....	39
2.1.3. FAZLI DİZİLEME YAKLAŞIMLARI	41
2.1.4. DNA DİZİLEME İLE NADİR VARYAN TLARIN TESPİTİ.....	43
2.1.4.1. DERİN DİZİLEME	43
2.2. HAVUZ LAMA YAKLAŞIMI.....	44
2.2.1. DNA DİZİLEME VE HAVUZ LAMA YAKLAŞIMI	44
2.2.2. ÇOK BOYUTLU HAVUZ LAMA YAKLAŞIMI	45
2.3. MUTASYON TARAMA.....	47
2.4. FAMILİYAL HİPERKOLESTEROLEMİ.....	48
2.4.1. BELİRTİLER.....	49
2.4.1.1. ARTEROSKLEROZ	50
2.4.2. TANI KRİTERLERİ.....	51
2.4.3. KALITIM.....	52
2.4.3.1. İLGİLİ GENLER.....	53
2.4.3.1.1. LDLR	53
2.4.3.1.2. APOB.....	53

2.4.3.1.3. PCSK9	54
2.4.3.1.4. ÇALIŞMA İÇİN SEÇİLEN GENLER	55
3. GEREKÇE VE AMAC.....	58
4. MATERYAL VE YÖNTEM.....	59
4.1. MATERYAL	59
4.1.1. HAVUZLAMA VALİDASYONU	59
4.1.2. FAMİLYAL HİPERKOLESTEROLEMİ HASTA GRUBU	59
4.2. YÖNTEM.....	60
4.2.1. HAVUZLAMA VALİDASYONU	60
4.2.1.1. DİZİNLERİN VE HAVUZLARIN OLUŞTURULMASI	60
4.2.1.2. MOLEKÜLER ÇALIŞMA.....	63
4.2.1.2.1. PCR.....	63
4.2.1.2.2. NGS	64
4.2.1.2.3. NGS VERİ ANALİZİ	64
4.2.2. FAMİLYAL HİPERKOLESTEROLEMİ HASTA GRUBU	65
4.2.2.1. DİZİNLERİN VE HAVUZLARIN OLUŞTURULMASI	65
4.2.2.2. MOLEKÜLER ÇALIŞMA.....	66
4.2.2.2.1. DNA İZOLASYONU	66
4.2.2.2.2. PCR.....	67
4.2.2.2.3. NGS	68
4.2.2.2.4. NGS VERİ ANALİZİ	68

4.2.2.3. POZİTİF ÖRNEKLERİN TESPİTİ.....	71
<u>5. ARAŞTIRMA BULGULARI.....</u>	72
5.1. HAVUZLAMA VALİDASYONU	72
5.1.1. JEL ELEKTROFOREZİ.....	72
5.1.2. NGS VERİ ANALİZİ	72
5.1.2.1. 2D08.....	73
5.1.2.2. 2D16.....	74
5.1.2.3. 2D24.....	75
5.1.2.4. 2D32.....	76
5.1.2.5. 3D08.....	77
5.1.2.6. TEKRARLANABİLİRLİK ÇALIŞMASI.....	77
5.2. FAMILİYAL HİPERKOLESTEROLEMİ HASTA GRUBU	77
5.2.1. JEL ELEKTROFOREZİ.....	77
5.2.2. NGS VERİ ANALİZİ	78
5.2.3. POZİTİF ÖRNEKLERİN TESPİTİ	78
<u>6. TARTIŞMA VE SONUÇ.....</u>	84
6.1. TARTIŞMA	84
6.1.1. DNA DİZİLEME	84
6.1.2. HAVUZLAMA VALİDASYONU ÇALIŞMASI	85
6.1.2.1. DAHA DENGELİ HAVUZLARIN OLUŞTURMASI.....	86
6.1.2.2. PCR.....	87
6.1.2.2.1. POLİMERAZ HATALARI	88

6.1.2.2.2. ALLEL DROP-OUT.....	89
6.1.2.2.3. KONTAMİNASYON.....	89
6.1.2.2.4. PCR ÖZET.....	91
6.1.2.3. UMI.....	92
6.1.2.4. DÜŞÜK HATA ORANINA SAHİP NGS VEYA 3. NESİL DİZİLEME	93
6.1.2.5. ANALİZ ALGORİTMALARI	94
6.1.3. FAMILİYAL HİPERKOLESTEROLEMİ HASTA GRUBU ÇALIŞMASI.....	95
6.1.3.1. HAVUZLAMA.....	95
6.1.3.2. TESPİT EDİLEN MUTASYONLAR	95
6.1.4. BU TEZ ÇALIŞMASININ YENİ LİTERATÜR İÇİNDEKİ YERİ	102
6.2. SONUÇ.....	106
<u>KAYNAKLAR.....</u>	<u>107</u>
<u>EKLER</u>	<u>128</u>
EK 1: Çalışmada kullanılan primerlerin dizileri ve elde edilen ampliconların boyları.....	128
EK 2: 2D08 analiz sonuçları dizin tablosu.	130
EK 3: 2D16 analiz sonuçları dizin tablosu	131
EK 4: 2D24 analiz sonuçları dizin tablosu	132
EK 5: 2D32 analiz sonuçları dizin tablosu	133
EK 6: Tekrarlanabilirlik çalışması sonuçları	134
EK 7: FH hasta grubu için tespit edilen 8 adet negatif havuz, 72 adet pozitif havuz, bulunan mutasyonlar ve frekansları.....	136

EK 8: Tespit edilen mutasyon çeşitleri.....	139
EK 9: Pozitif bulunan örnekler ve genotipleri.....	140
<u>ÖZGEÇMİŞ</u>	<u>143</u>
<u>TEZDEN ÇIKAN YAYINLAR.....</u>	<u>148</u>



ŞEKİLLER DİZİNİ

Şekil 2.1. Eksabayt	2
Şekil 2.2. DNA dizilemenin tarihçesi (23)	5
Şekil 2.3. Moore yasası (25).....	6
Şekil 2.4. Zaman içinde Mbp başına DNA dizileme verisi maliyeti	6
Şekil 2.5. NGS etkisi ile kırılmış DNA dizileme Moore yasası grafiği; bir insan genomu dizileme maliyeti	7
Şekil 2.6. Frederick Sanger (1918-2013)	8
Şekil 2.7. Sanger dizileme (32)	9
Şekil 2.8. Örnek fragmentasyon agaroz jel elektroforezi görüntüsü	12
Şekil 2.9. Tagmentasyon (46).....	14
Şekil 2.10. Adaptör (48)	16
Şekil 2.11. İndeksleme (50).....	17
Şekil 2.12. UMI (54)	18
Şekil 2.13. Emülsiyon PCR (57)	19
Şekil 2.14. Köprü amplifikasyonu (57)	20
Şekil 2.15. DNB (63).....	21
Şekil 2.16. Pirosekanslama (67)	23
Şekil 2.17. Sentez ile sekanslama (73)	25
Şekil 2.18. İyon sekanslama (83)	28
Şekil 2.19. Ligasyon ile sekanslama (86)	29

Şekil 2.20. Değişik akış hücreleri.....	31
Şekil 2.21. Örüntülü akış hücresi SEM görüntüsü (88).....	32
Şekil 2.22. IGV yazılımında bam dosyalarının görüntülenmesi	34
Şekil 2.23. SMRT sekanslama (107).....	38
Şekil 2.24. Nanopor sekanslama (112).....	40
Şekil 2.25. Fazlı dizileme (120).....	42
Şekil 2.26. 1, 2 ve 3 boyutlu havuzlama yaklaşımları (169)	46
Şekil 2.27. FH belirtiler: a) arkus kornea, b) ksantelezm, c) tendon ksantoma.....	49
Şekil 2.28. Hiperlipidemi sebebi ile görsel olarak değişime uğramış kan (178).....	50
Şekil 2.29. Ateroskleroz (181).....	50
Şekil 2.30. LDLR, APOB ve PCSK9 (198)	54
Şekil 4.1. Oluşturulan ana dizin, örnek numaraları ve kodları	61
Şekil 4.2. İlk dizin için oluşturulan havuzlar ve havuz kodları	62
Şekil 4.3. İlk 64 örnek için oluşturulan havuzlar ve havuz kodları	66
Şekil 4.4. Örnek IGV analizi	69
Şekil 4.5. L252 nolu örneğin doğrulanma ve zigozite belirlenmesi için seçimi	70
Şekil 4.6. L026, L028, L042 ve L044 nolu örneklerin doğrulanma ve zigozite belirlenmesi için seçimi.....	70
Şekil 5.1. Havuzlama çalışması örnek jel görüntüsü.....	72
Şekil 5.2. FH hasta gurubu çalışması örnek jel görüntüsü	77
Şekil 5.3. FH hasta gurubu 1. 8x8'lik dizin çalışması sonuçları	79

Şekil 5.4. FH hasta gurubu 2. 8x8'lik dizin çalışması sonuçları	80
Şekil 5.5. FH hasta gurubu 3. 8x8'lik dizin çalışması sonuçları	80
Şekil 5.6. FH hasta gurubu 4. 8x8'lik dizin çalışması sonuçları	81
Şekil 5.7. Kontrol gurubu 8x8'lik dizin çalışması sonuçları	81



ÇİZELGELER DİZİNİ

Çizelge 2.1. LDL-C seviyeleri ve klinik kategorizasyonu	51
Çizelge 2.2. Dutch kriterleri skrolama tablosu	52
Çizelge 2.3. Dutch kriterlerinde toplam skora göre tanı.....	52
Çizelge 4.1. Elde edilen dizin ve havuzlar	62
Çizelge 4.2. PCR Reaksiyon koşulları	63
Çizelge 4.3. Isıl döndürücü protokolü	63
Çizelge 4.4. PCR Reaksiyon koşulları	67
Çizelge 4.5. Isıl döndürücü protokolü	68
Çizelge 5.1. 2D08 analiz istatistikleri.....	73
Çizelge 5.2. 2D16 analiz istatistikleri.....	74
Çizelge 5.3. 2D24 analiz istatistikleri.....	75
Çizelge 5.4. 2D32 analiz istatistikleri.....	76
Çizelge 5.5. FH hasta grubu NGS veri analizi istatistikler.....	78
Çizelge 5.6. Pozitif örneklerin tespiti çalışması istatistikleri	79
Çizelge 5.7. Genotipler ve örnek sayıları	82
Çizelge 6.1. Değişik hasta grupları için ortalama LDL-C değerleri.....	98

SİMGELER DİZİNİ

µl	Mikrolitre
µm	Mikrometre
AT	Adenin-Timin
ATP	Adenozin Trifosfat
bp	Baz çifti
BWA	Burrows-Wheeler Aligner
CEO	Chief Executive Officer
dbSNP	Database of Single Nucleotide Polymorphisms
DGGE	Denaturing Gradient Gel Electrophoresis
dH ₂ O	Distile su
DHPLC	Denaturing High-Performance Liquid Chromatography
dk	Dakika
DNA	Deoksiribo nükleik asit
DNB	DNA nanotopları
dNTP	Deoksinükleotit trifosfat
dTTP	Deoksitimin trifosfat
dUTP	Deoksiurasil trifosfat
EDTA	Etilenediamintetraasetik asit
FH	Ailesel hiperkolesterolemi
fL	Femtolitre
GATK	Genome Analysis Toolkit
GC	Guanin-Sitozin
H ⁺	Hidrojen iyonu
HDL	Yüksek yoğunluklu lipoprotein
hg19	Human genome version 19

HPV	İnsan papilloma virüsü
HRM	High Resolution Melting
IGV	Integrative Genomics Viewer
İng.	İngilizcesi
Kbp	Kilo baz çifti
LDL	Düşük yoğunluklu lipoprotein
LDL-C	Düşük yoğunluklu lipoprotein – Kolesterol
MAF	Minör allel frekansı
MHz	Megahertz
NGS	Yeni nesil DNA dizileme
NIPT	Girişimsel olmayan prenatal test
nm	Nanometre
OMIM	Online Mendelian Inheritance in Man
PCR	Polimeraz zincir reaksiyonu
PGT-A	İmplantasyon öncesi genetik anöploidi testi
pH	Hidrojen potansiyeli, asitlik-bazlık ölçü birimi
Q	Phred kalite skoru
RNA	Ribo nükleik asit
SBS	Sentez ile sekanslama
SD	Standart sapma
SMA	Spinal müsküler atrofi
SMRT	Single Molecule Real Time
sn	Saniye
SSCP	Single-Strand Conformation Polymorphism
T	Timin
UDG	Urasil DNA glikozilaz

UMI	Unique Molecular Identifier
UNG	Urasil nükleotit glikozilaz
USD	Amerikan doları
UV	Ultraviyole
vcf	Variant Call File



1. GİRİŞ

DNA dizileme günümüzde en yaygın olarak kullanılan DNA analizi yöntemidir. İlk geliştirildiği 40 yıl öncesine göre müthiş bir hızla ilerleyerek, günümüzde dünya çapında yıllık yaklaşık 10 milyar dolarlık bir sektör haline gelmiş (1) ve hızla gelişmeye devam etmektedir.

Havuzlama yaklaşımı birçok biyolojik testte kullanılan bir yaklaşımdır. Özellikle maliyet düşürücü özelliği sebebi ile tarama testlerinde ve tarama programlarında kullanılabilmektedir. 2020 yılında başlayan SARS-CoV-2 pandemisi ile tekrar gündeme gelmiş ve bu alanda yapılan bilimsel çalışma ve yayınların sayısında ciddi bir artış olmuştur (2).

Hastalık tarama, birey ve toplum sağlığı için çok büyük bir öneme sahiptir. Hem hastalıkların önlenmesinde hem de hali hazırda oluşmuş hastalıkların erken teşhisi ile tedavisinin mümkün hale getirilmesinde ve/veya kolaylaştırılmasında etkin bir rol oynamaktadır. Bu sebeplerden dolayı ülkemiz de dahil olmak üzere birçok ülkede hastalık tarama programları uygulanmaktadır (3). Hastalık taramada DNA seviyesinde mutasyon taramanın uygulanması, günümüzde hala bir tarama testine göre yüksek olan maliyetleri sebebiyle son derece kısıtlıdır. Ancak gelişen DNA dizileme teknolojisi ve maliyet düşürücü havuzlama gibi yaklaşımların mevcut teknolojiler içerisine adapte edilmesi sayesinde, yakın gelecekte tüm genetik hastalıkların DNA bazında taranmaması için hiçbir sebep yoktur.

Familyal hiperkolesterolemi, yaklaşık 500 ila 200'de 1 olan frekansı ile tüm dünyada en yaygın görülen genetik ve kalıtsal hastalık olarak kabul edilmektedir (4). Hastalığın bireyde erken tespiti, hastalığın yönetilmesi, semptomlarının azaltılması veya geciktirilmesi açısından son derece faydalıdır. Bu özellikleri sebebi ile hastalık tarama programları ile taraması, bireysel ve toplumsal olarak çok büyük bir fayda sağlama potansiyeli olan bir hastalıktır.

2. KURAMSAL TEMELLER

2.1. DNA DİZİLEME

DNA dizileme (veya DNA sekanslama), bir DNA molekülündeki nükleotidlerin yönü ile beraber sırasının belirlenmesidir (5). DNA molekülü 4 adet nükleotidin (adenin, guanin, sitozin ve timin) oluşturduğu, dallanma yapmayan bir polimerdir. Polimer zincirinin bir yönü vardır; zincirin iki ucu aynı değildir. Nükleotidlerin yapısında bulunan deoksiriboz molekülünün karbon atomu numaralarından yola çıkarak, DNA molekülünün uçları 5' ve 3' olarak adlandırılır. DNA dizilimi -aksi belirtilmediği sürece- DNA polimerizasyon, transkripsiyon ve translasyonun da yönü olan 5' ucundan 3' ucuna olacak şekilde yazılır.

DNA dizileme günümüzde en yaygın olarak kullanılan DNA analiz yöntemidir. 2023 yılı itibarıyla dünya çapında pazar büyüklüğü yıllık yaklaşık 15 milyar dolar seviyesine çıkmıştır ve her yıl yaklaşık %15-20 büyümeye devam etmektedir (6). Yıllık üretilen DNA dizileme verisi ise yaklaşık 40 eksabayt seviyesindedir; ki bu miktar Youtube'a yüklenen yıllık verinin onlarca katına eşittir (7). 1 eksabayt yaklaşık olarak 10^{18} bayta eşittir (Şekil 2.1.).

yotta [Y]	10^{24}	= 1 000 000 000 000 000 000 000 000 000
zetta [Z]	10^{21}	= 1 000 000 000 000 000 000 000 000 000
exa [E]	10^{18}	= 1 000 000 000 000 000 000 000 000
peta [P]	10^{15}	= 1 000 000 000 000 000 000
tera [T]	10^{12}	= 1 000 000 000 000 000
giga [G]	10^9	= 1 000 000 000
mega [M]	10^6	= 1 000 000
kilo [k]	10^3	= 1 000
hecto [h]	10^2	= 100
deca [da]	10^1	= 10

Şekil 2.1. Eksabayt

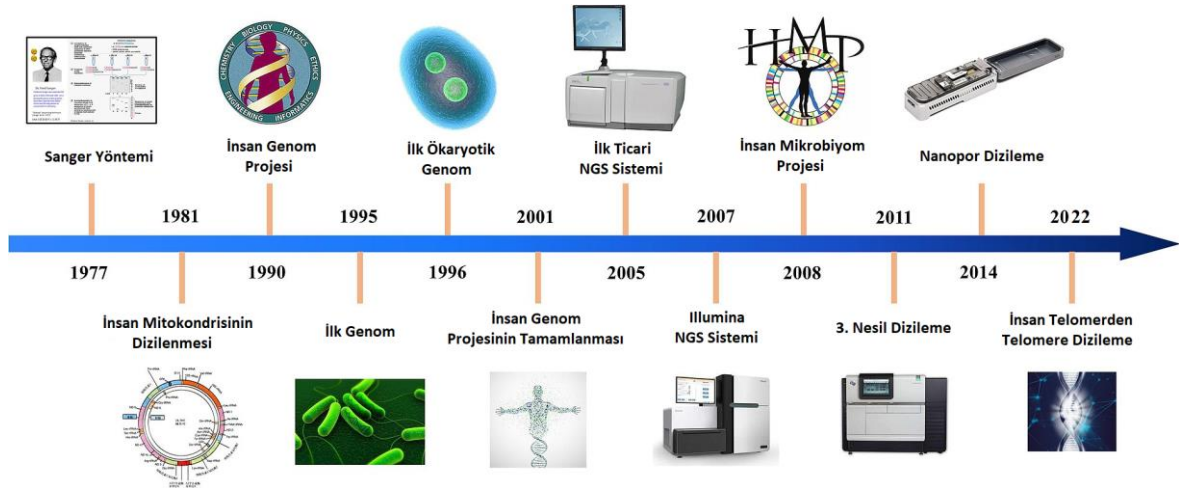
2.1.1. DNA DİZİLEMENİN TARİHÇESİ

DNA ilk olarak 1869 yılında Friedrich Miescher tarafından tanımlanmış ve izole edilmiştir (8). Uzun yıllar boyunca hücre içinde ne görev aldığı tam olarak anlaşılamamak ile birlikte, genel olarak kalıtsal bilginin nesiller arasında proteinler tarafından taşındığı düşüncesi hâkim olduğu için, DNA'nın kalıtsal molekül olabileceği düşünülmemiştir. DNA'nın kalıtsal molekül olabileceği fikri ilk olarak 1944 yılında Oswald Avery, Colin MacLeod ve Maclyn McCarty'nin bakteriler üzerinde yapmış ve yayınlamış oldukları çalışma ile gösterilmiş ve bu tarihten sonra kabul görmüştür (9). DNA molekülünün yapısının 1953 yılında, Rosalind Franklin ve Maurice Wilkins'in de yardımı ile, James Watson ve Francis Crick tarafından ortaya koyulması (10); ilk defa bir protein olan insülin molekülünün (daha sonra da 2000'li yılların ortalarına kadar kullanılan tek DNA dizileme yöntemini de geliştirecek olan Frederick Sanger tarafından) 1952 yılında dizilenmesinin tamamlanarak yapısının ortaya çıkartılması (11) ve 1958 yılında Francis Crick tarafından proteinlerin diziliminin, DNA'nın diziliminden geldiğinin öne sürülmesi (12) sonucunda, DNA molekülünün dizilenme gereksinimi had safhaya çıkmıştır. Buna rağmen DNA molekülünün ilk defa dizilenmesi 10 yıldan fazla sürmüştür; her ne kadar proteinler gibi polimer makromolekül olsa da DNA'nın dizilenmesi proteinlere göre çok daha zordur; proteinlerin dizilenmesinde kullanılan yaklaşım işe yaramamıştır. Bunun sebepleri; proteinlerin aksine DNA zincirinde sadece 4 farklı çeşit monomerin (nükleotid) olması (proteinler 20'den fazla çeşit amino asitten oluşurlar), nükleotidlerin moleküler ağırlıklarının birbirine çok yakın olması (en küçük nükleotid olan sitozin: 307 g/mol, en büyük nükleotid olan guanin: 347 g/mol, oran olarak 1:1,13; en küçük amino asit olan glisin: 75 g/mol, en büyük amino asit olan triptofan: 204 g/mol, oran olarak 1:2,72), tipik bir DNA zincirinin, tipik bir protein zincirine göre çok daha fazla monomerden oluşması, değişik DNA zincirlerinin kimyasal özelliklerinin proteinlerin aksine çok benzemesi sonucunda ayrılmasının zor olması ve dizi spesifik DNase enzimlerinin tam olarak keşfedilmemiş olmasını sayabiliriz (13). Bu sebeplerden dolayı DNA molekülünün ilk defa dizilenmesi ancak 1970 yılında Ray Wu tarafından sadece 8 nükleotid uzunluğunda gerçekleştirilebilmiştir (14). Geniş kitlelerce kullanılabilecek kadar pratik, yüksek kapasiteli ve gerçek anlamda DNA dizileme teknolojisi olarak kabul edilen 2 yöntem, 1977 yılında geliştirilmiştir (5); bunlar Walter Gilbert ve Allan Maxam'in geliştirdiği kimyasal degradasyon yöntemi (15) ve Frederick Sanger ve arkadaşlarının geliştirdiği dideoksi zincir sonlanma yöntemidir (16). Bu iki yöntemden

Frederick Sanger'in geliřtirdiđi zincir sonlanma yntemi, diđerine kıyasla ok daha gl bir yntem olduđu iin kısa sre iinde Maxam-Gilbert yntemi ortadan kalkmıř ve DNA dizileme tamamen Sanger yntemi zerine geliřmiřtir (5). Sanger yntemi ok gl bir yntemdir ve PCR gibi diđer ok gl bir teknolojinin 1985 yılında Kary Mullis tarafından geliřtirilip (17) Sanger yntemi ierisine dahil edilmesi ile teknoloji ok daha gl bir hale gelmiřtir. Aynı zamanda bařta tespit iin kullanılan radyoaktivite yerine floresan tespitinin geliřtirilmesi, yntem ierisinde daha verimli alıřacak DNA polimerazların geliřtirilmesi ve dizilemeyi gerekleřtirecek ok kapillerli otomatik kapiller elektroforez cihazlarının geliřtirilmesi ile Sanger dizileme ok daha glenmiřtir, kolaylařmıř, maliyeti dřmř, iřlem hacmi ykselmiřtir; yle ki 1980'lerin bařından 2000'li yılların ortalarına kadar (NGS geliřip pratikte kullanıma girene kadar) yaklaşık 25 yıl boyunca DNA dizileme sahasının tek oyuncusu olmuřtur (5, 18). İnsan Genom Projesi, btnyle Sanger dizileme yntemi ile yapılmıřtır (19, 20). Gnmzde bile, NGS kadar yaygın olmasa da hala kullanımına devam edilmektedir. Teknolojinin inanılmaz lde geliřtiđi bu 40 yıllık sre sonunda bile hala kullanılıyor olması ve daha uzun sre ortadan kalkacak gibi grnmemesi, tek bařına Sanger dizilemenin ne kadar gl bir yntem olduđunu kanıtlamaya yeter.

1990'lı yılların ortalarından itibaren, "ikinci nesil" olarak ta adlandırılan ancak daha yaygın olarak kullanılan ismi ile NGS teknolojileri geliřtirilmeye bařlanmıřtır. NGS tabanlı retilen ilk ticari cihaz, 454 Life Sciences tarafından 2000 yılında geliřtirilmiřtir. Teknolojiyi geliřtiren Jonathan Rothberg 2007 yılında bu cihazı Roche firmasına satarak btn dnya pazarına aılımlarını sađlamıřtır. Bu teknoloji 1990'ların sonlarına dođru geliřtirilen pirosekanslama yntemine dayanmaktadır (21). Bu tarihten itibaren Sanger yntemini geliřtirme zerine yapılan alıřmalar ve maddi yatırımlar hızla gerilemiř, bunu yerine teknoloji geliřtiren řirketler ok daha byk potansiyel barındıran NGS zerine yođunlařmıřtır. 2010'lu yıllardan itibaren ise bir sonraki nesil (3. nesil) DNA dizileme teknolojileri geliřtirilmeye bařlanmıřtır. Bununla beraber hangi teknolojinin NGS, hangi teknolojinin 3. nesil olarak kabul edilmesi gerektiđi ile ilgili tartıřmalar bařlamıřtır ve halen de devam etmektedir (5). 3. nesil DNA dizileme iin ok byk umutlar ile bařlanan alıřmalar beklendiđi kadar hızlı ilerleyememiřtir. Dnya zerinde ok fazla arařtırmacı ve firma 10 yıldan fazla bir sredir 3. nesil DNA dizileme teknolojileri zerinde alıřıyor olmasına rađmen, gnmzde ticari olarak arařtırmacıların kullanımına sunulan sadece iki teknoloji geliřtirilebilmiřtir; bunlardan biri olan PacBio firmasının teknolojisinin, modifiye

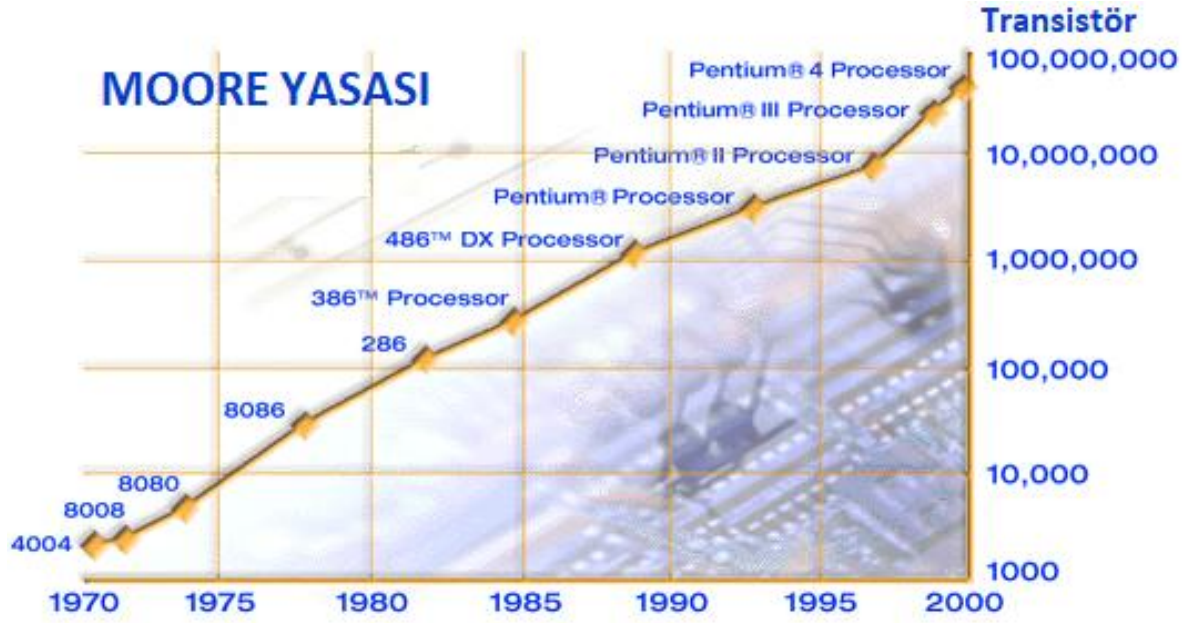
nükleotid ve enzimatik reaksiyon içermesi sonucunda NGS teknolojilerine benzemesi sebebi ile tam olarak 3. nesil olmayacağını düşünen araştırmacılar da vardır. Diğer teknoloji olan Oxford Nanopore Technologies firmasının teknolojisinin ise veri kalitesi hala beklenen seviyede değildir (22). Günümüzde 3. nesil DNA dizileme ve Sanger yöntemi aktif olarak kullanılmalarına rağmen oransal olarak o kadar küçüklerdir ki, tüm dünyada üretilen DNA dizileme verisinin tamamı NGS teknolojileri ile üretilmektedir denilirse çok ta abartmış olunmaz. DNA dizilemenin özet kronolojik tarihçesi Şekil 2.2.'de gösterilmiştir.



Şekil 2.2. DNA dizilemenin tarihçesi (23)

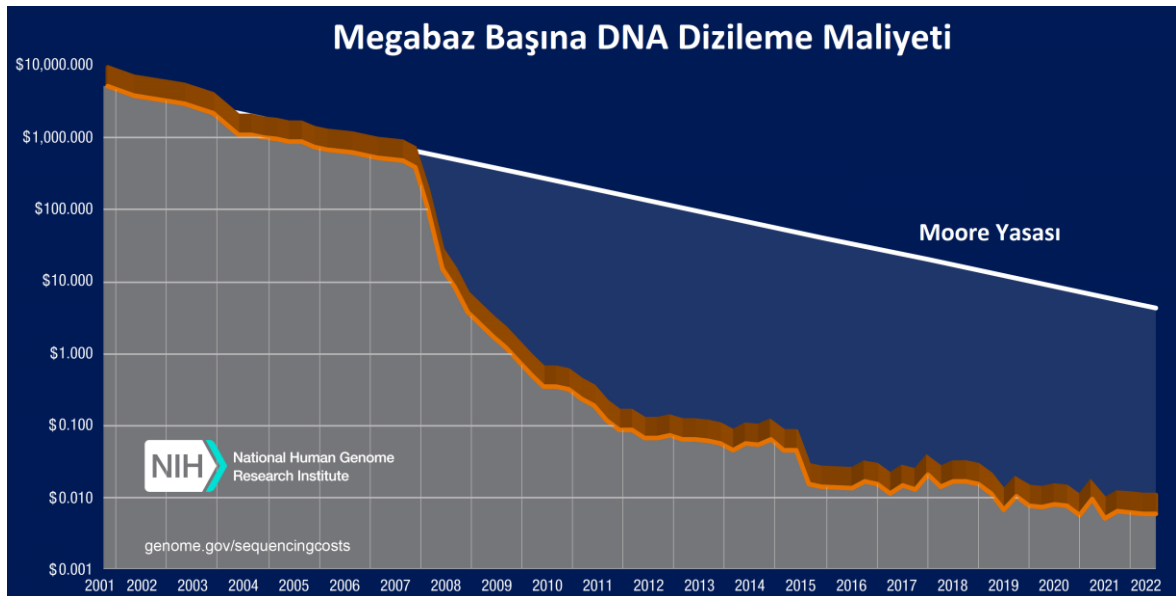
2.1.1.1. MOORE YASASI VE DNA DİZİLEME

Moore yasası olarak bilinen kavram, bilgisayar teknolojileri öncü firmalarından Intel firmasının CEO'su olan Gordon Moore'un, 1965 yılında bilgisayar teknolojilerinin ilerlemesi ile ilgili yapmış olduğu bir öngörüye dayanmaktadır. Moore, makalesinde özet olarak bilgisayar teknolojilerinin zaman içerisinde logaritmik olarak ilerleyeceğini, yani belirli bir zaman içerisinde teknolojinin kendisini ikiye katlayarak ilerlemesini sürdüreceğini, bunun sonucunda da teknolojinin çok hızlı ilerleyeceğini öngörmüştür (24). Zaman Gordon Moore'un öngörüsünde ne kadar haklı olduğunu göstermiş (Şekil 2.3.), bunun sonucunda da Moore yasası olarak ünlenen kavram, sadece bilgisayar teknolojileri değil, logaritmik olarak çok hızlı bir şekilde ilerleme kaydeden tüm gelişmeler için kullanılan genel bir jargon haline gelmiştir (25).



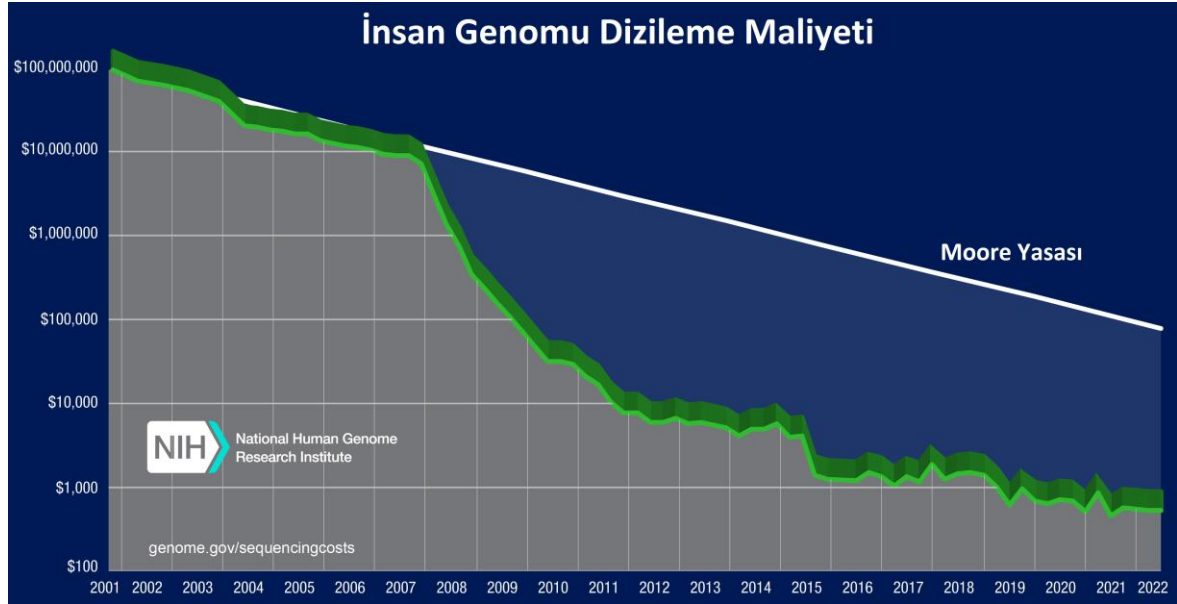
Şekil 2.3. Moore yasası (25)

DNA dizileme teknolojilerinin ilerlemesine bakıldığında zaman da, ilk geliştirilmeye başlandığı zamandan itibaren Moore yasasına uyduğu görülebilmektedir; DNA dizileme verisi birim maliyeti belirli bir zaman içinde yaklaşık olarak yarıya inerek ilerleme sağlanmıştır (26). Ancak çok ciddi bir farklılık göze çarpmaktadır; NGS teknolojilerinin geliştirilmesinden sonra grafikte bir kırılma yaşanmış ve DNA dizileme yine Moore yasasını takip etmekle birlikte bu sefer grafiğin farklı bir yerinden takip etmeye başlamıştır (Şekil 2.4.).



Şekil 2.4. Zaman içinde Mbp başına DNA dizileme verisi maliyeti

Günümüzde NGS teknolojileri geliştirilmemiş olsaydı, Moore yasasını takip eden Sanger teknolojisi ile yaklaşık 100.000 USD tutması beklenen bir insan genomunun dizilenme maliyeti, NGS teknolojileri sayesinde günümüzde birkaç yüz dolar seviyelerine kadar düşebilmiştir (Şekil 2.5.).



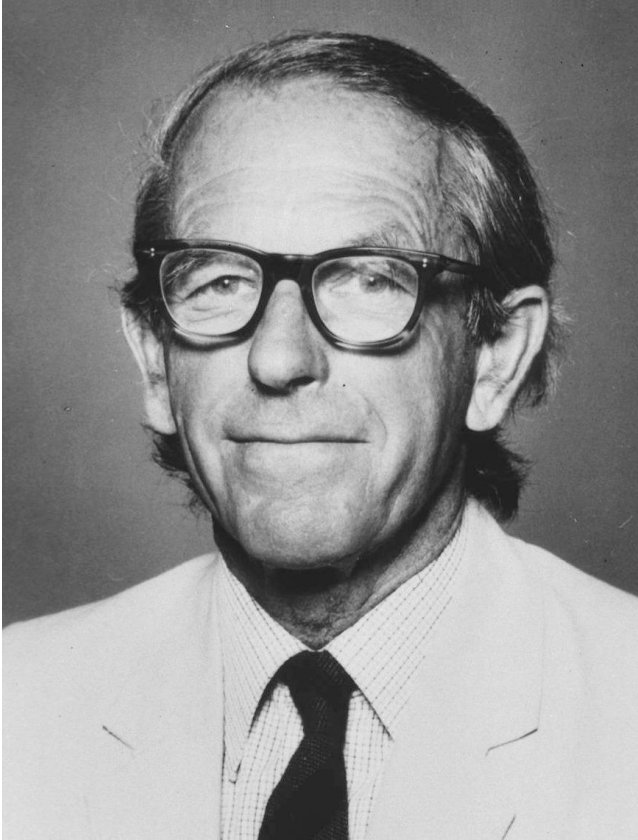
Şekil 2.5. NGS etkisi ile kırılmış DNA dizileme Moore yasası grafiği; bir insan genomu dizileme maliyeti

2.1.1.2. FREDERICK SANGER

Frederick Sanger, insanlığın son döneminde yaşamış en büyük bilim insanlarından bir olmasına rağmen, moleküler genetik alanı dışında hak ettiği kadar tanınmayan bir bilim insanıdır. Günümüzde genetik ve genomik alanlarındaki ilerlemede çok sayıda bilim insanının katkısı ile olmuş olsa da bunlardan en önemlilerinden bir -belki de en önemlisi- Frederick Sanger'dir; kendisi "genomiğin babası" olarak ta adlandırılmaktadır (27). Günümüzde bir insanın tüm genomunu dizilemeyi sağlayan NGS teknolojileri sayesinde çok sayıda hastalığın tanısı koyulabilmekte, kanser de dahil olmak üzere hastalıklar daha iyi anlaşılabilenmekte ve çok sayıda fazla hastalığa tedavi geliştirilebilmekteyse, bunun en önemli mihenk taşlarından bir tanesi Frederik Sanger'in Sanger yöntemini geliştirmesidir denebilir; İnsan Genom Projesi Sanger yöntemi ile yapılmıştır ve NGS teknolojileri Sanger yönteminin bir devamı olarak, ondan esinlenerek geliştirilmiştir.

1918 yılında İngiltere’de doğan ve 2013 yılında aramızdan ayrılan Frederick Sanger’in ne kadar büyük bir bilim insanı olduğunu anlamak için sadece şuna bakmak bile yeterlidir: 1901 yılından günümüze kadar Nobel ödülü ile ödüllendirilmiş bine yakın bilim insanı arasından 2 defa Nobel ödülü almış 5 bilim insanından biri, 2 defa aynı alanda Nobel ödülü almış 3 bilim insanından biri Frederick Sanger’dir. Kendisinin almış olduğu 2 Nobel ödülü dışında 2 doktora öğrencisi de kariyerlerinin ilerleyen yıllarında Nobel ödülü ile ödüllendirilmiştir (28).

Frederick Sanger’in hak ettiği kadar ünlü olmamasının belki de en önemli sebebi mütevazı bir kişiliği ve yaşam tarzı olmasıdır. Kendisi “Sir” olarak anılmak istemediği için şövalyelik ünvanını reddetmiştir. Aynı zamanda kendisini “sadece laboratuvarında takılan ve akademik olarak pek te parlak olmayan bir adam” olarak tanımlayan birisidir (28). Geçmişte yapmış olduğu çalışmalar ile günümüzde milyonlarca insanın hayatını direkt olarak etkileyen, kimisinin hayat kalitesini arttırırken kimisinin hayatını kurtaran, bir yandan da bu doktora çalışmasının yapılmasına vesile olmuş olan bu büyük ve mütevazı bilim insanını saygı ile anıyorum.



Şekil 2.6. Frederick Sanger (1918-2013)

İlk keşfedildiğinde genetik alanında bir devrim yaratan bu yöntem, kısa zamanda ciddi oranda geliştirilerek çok daha pratik ve otomatik bir hale getirildi; yaklaşık 10-15 yıl kadar sonra ilk geliştirildiği şekilde kullanımı neredeyse tamamen ortadan kalkmıştır. Yapılan geliştirmeler arasında PCR yönteminin geliştirilmesinden sonra Sanger yöntemi içerisine dahil edilmesi ile klonlama gereksinimin ortadan kalkması, radyoaktif tespit yerine floresan tespit yöntemi eklenmesi, değişik floresan özelliğe sahip moleküller kullanılarak çok kanallı tespit mümkün olması sonucu her nükleotid için ayrı reaksiyon yerine 4 nükleotidin de aynı anda dizilendiği tek reaksiyonun mümkün hale gelmesi, poliakrilamid jel elektroforezi yerine polimer bazlı kapiller elektroforezin geliştirilmesi, çok sayıda kapiller içeren otomatik kapiller elektroforez cihazları ile 384'e kadar örneğin tek seferde tam otomatik paralel dizilenmesi vb. geliştirmeleri sayabiliriz (29). İlk keşfedildiğinde birkaç yüz bazlık DNA dizileme bilgisinin elde edilmesi günler sürerken, tam otomatik çok kapillerli DNA dizileme cihazları ile bir cihazdan günde maksimum 5 Mbp DNA dizileme verisi alınabilecek hale gelmiştir (29). İnsan Genom Projesi ilk sürümü yayınlandığı 2000 yılındaki hali ile bu tip tam otomatik cihazlarda tamamlanmıştır (19). Bu ilk sürümünde eksik olan yaklaşık %8'lik kısım ise ancak 2022 yılında ve NGS ve uzun okuma yapabilen 3. nesil DNA dizileme cihazları sayesinde tamamlanabilmiştir (33).

2.1.2.2. YENİ NESİL DİZİLEME

Sanger dizileme, her ne kadar tek DNA dizileme yöntemi olarak kaldığı 25 yıllık dönem boyunca çok fazla geliştirilmiş olsa da temel prensipleri olarak çok standarttır ve tek bir yöntemdir denilebilir: zincir sonlandırma yöntemi (16). Ancak NGS böyle değildir; zaman içinde temel prensipleri farklı olan çok sayıda NGS teknolojisi geliştirilmiş, başarılı ve güçlü teknolojiler günümüze kadar gelebilmişmiş, tıpkı Maxam-Gilbert yönteminin Sanger yöntemi karşısında tutunamayıp ortadan kalkması gibi daha az başarılı olan ve daha az güçlü olan NGS teknolojileri ise zaman içinde kaybolmuştur (5, 34). NGS teknolojileri ilk geliştirilmeye başlandığı zaman hangi teknolojinin NGS olduğu konusunda bir tartışma söz konusu değil idi; çünkü mevcut tek teknoloji olan Sanger'den farklı olması NGS olması için yeterli idi. Ancak birbirinden farklı temele sahip teknolojilerin ortaya çıkmaya başlanması ile beraber, bir teknolojinin NGS mi yoksa bir sonraki nesil olan 3. nesil DNA dizileme teknolojisi mi olduğu ile ilgili tartışmalar da başladı. Temel prensipler olarak farklılıklar

içerse de tüm teknolojilerin, onları bir NGS teknoloji yapan genel olarak kabul görmüş özelliklerini şu şekilde sıralayabiliriz (34-37):

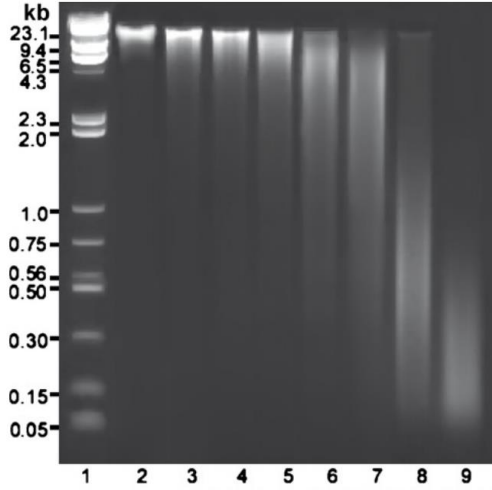
- DNA'nın rastgele parçalanması
- Klonal amplifikasyon
- Rastgele ve kısa dizileme
- Yüksek hacimli paralel dizileme
- Elde edilen kısa okumaların birleştirilmesi

2.1.2.2.1. DNA'NIN RASTGELE PARÇALANMASI (FRAGMENTASYON)

NGS'nin örnek hazırlama aşaması DNA'nın rastgele parçalanması ile başlar; bu aşama DNA fragmentasyonu olarak ta adlandırılır. Fragmentasyonun amacı dizilenmek istenen DNA'dan, küçük DNA parçaları (200-400 bp) oluşturmaktır. Fragmente edilmemiş yüksek moleküler ağırlıklı genomik DNA ve değişik seviyelerde fragmente edilmiş genomik DNA'nın örnek agaroz jel elektroforezi görüntüsü Şekil 2.8.'de gösterilmektedir (yüksek moleküler ağırlıklı genomik DNA (1) ve değişik seviyelerde fragmente edilmiş halleri (2-9)) (38). Parçalama aşaması nebulizasyon, sonikasyon vb. gibi fiziksel parçalama ile yapılabileceği gibi, DNase, transpozaz vb. gibi enzimatik yöntemler ile de yapılabilmektedir. DNA parçalandıktan sonra uçlarına dizileme aşamasında kullanılacak olan primer vb. gibi dizileri içeren özel DNA dizileri (adaptörler) eklenir. DNA'nın rastgele parçalanmasının istisnaları, uzun okuma yapabilen teknolojiler (genellikle 3. nesil olarak sınıflandırılırlar) ve dizilenmek istenen bölgenin küçük ve belirli bir bölge olduğu durumlardır. Uzun okuma yapabilen teknolojilerde onlarca Kbp kesintisiz okuma alınabilmektedir, bu sebeple DNA'nın küçük parçalara ayrılması yapılmaz. Şayet dizilenmek istenen bölge belirli küçük bir bölge ise (kullanılacak teknolojinin okuma uzunluğuna göre değişmekle beraber, yaklaşık 300 bp veya daha küçük ise), adaptör dizilerinin PCR ile eklenmesi sonucunda DNA'nın rastgele küçük parçalara ayrılmasına gerek kalmamaktadır.

Günümüzde DNA fragmentasyonu fiziksel ve enzimatik birkaç yöntem ile yapılmaktadır. Bu yöntemlerin çoğunda DNA parçalanması sonucunda yapışkan uçlar oluşmaktadır. Yapışkan uç, çift zincirli DNA'nın uç kısmında az sayıda nükleotidin tek zincir olması durumudur; yani 2 karşılıklı zincirin uç kısımlarının tam olarak birbirine karşılık

gelmemesidir. Yapışkan uçlar 5' fazlalığı olabileceği gibi 3' fazlalığı da olabilmektedir. Fragmentasyon sonucu oluşan bu yapışkan uçların bir sonraki aşama için düzeltilmesi gerekmektedir; bu amaç için de kısa kalan zinciri tamamlayan veya uzun kalan zinciri kesen enzimler kullanılmaktadır. Bu işlem uç tamiri olarak adlandırılmaktadır (39).



Şekil 2.8. Örnek fragmentasyon agaroz jel elektroforezi görüntüsü

FİZİKSEL FRAGMENTASYON

DNA hücre içinde korunaklı bir ortamdır ve oluşan göreceli az sayıda DNA hasarları tamir mekanizmaları tarafından tamir edilir. Ancak DNA hücre dışına alındığı zaman fiziksel etkenlere karşı çok hassas hale gelir ve uzun ince yapısı dolayısı ile DNA kırılması gerçekleşmesi kaçınılmazdır (40). DNA kırılmasına yol açan fiziksel etkenler arasında sıcaklık, basınç, titreşim, UV radyasyon vb. sayılabilir. Kontrollü fiziksel koşullar sağlanarak DNA'yı yaklaşık olarak belirli bir moleküler ağırlıkta (veya diğer bir tabir ile uzunlukta) olacak şekilde parçalamak mümkündür. DNA'nın fiziksel parçalanmasının avantajı, DNA solüsyonuna sonraki işlemlerde istenmeyen madde olma ihtimali olan değişik kimyasalların eklenmesine gerek olmaması ve son derece kontrollü olarak gerçekleştirilebilmesidir (41).

Sonikasyon

En yaygın olarak kullanılan fiziksel parçalama yöntemidir. Genel olarak 10 MHz'in altında frekansa sahip ultrason ses dalgaları kullanılarak DNA parçalanır (41). Standart sonikatörler daha geniş bir fragman boyu aralığı ile DNA fragmentasyonu yaptığı için, sadece DNA'yı daha verimli bir şekilde fragmente etme amaçlı, odaklı akustik parçalama prensibine dayalı

özel DNA sonikatörleri geliştirilmiştir (39). En ideal fragmentasyon yöntemi olarak nitelendirilse de, hem sarflarının pahalı olması, hem de DNA fragmentasyonundan başka bir işte kullanılmayan DNA sonikatörlerinin yüksek maliyetli olması sebebi ile özellikle küçük ve orta çaplı çoğu laboratuvar tarafından tercih edilmemektedir (42).

Nebulizasyon

Nebulizasyon, bir sıvının yüksek basınç ile küçük bir delikten geçmeye zorlanması sonucunda aerosol formunda çok küçük damlacıklar oluşturması sürecidir. İçinde DNA olan bir solüsyon nebulizasyon işlemine maruz bırakıldığında, işlem sırasında oluşan fiziksel zorlama, DNA'nın küçük parçalara ayrılmasına yol açar (42). DNA'nın nebulizasyon ile parçalanması aşamasında basınç kaynağı olarak nitrojen gazı kullanılır. Herhangi bir cihaz gerektirmemesi, sadece basınçlı nitrojen gazı tüpü ile basit bir plastik nebulizatör sarf malzemesi ile yapılması maliyet yönden ciddi bir avantaj sağlasa da, çok yüksek miktarda başlangıç DNA'sı gerektirmesi, yöntemin temeli olan aerosol oluşumunun kontaminasyon riskini arttırması ve düşük maliyetli enzimatik fragmentasyon alternatiflerinin geliştirilmesi sebebi ile günümüzde artık tercih edilmemektedir.

ENZİMATİK FRAGMENTASYON

DNA'yı modifiye eden enzimlerin bazıları, tek başlarına veya enzim kombinasyonları şeklinde DNA'nın fragmentasyonu için kullanılabilir. Fiziksel parçalamaya göre solüsyon içerisine bir takım kimyasallar eklenme gereksinimi, enzimlerin bozulabilmesi, reaksiyon içerisine karışmış olan bazı inhibitör moleküller sebebi ile reaksiyon verimliliğinin azalması veya tamamen kaybolması gibi dezavantajları olsa da; kullanım kolaylığı, moleküler genetik laboratuvarlarında standart olarak bulunan ekipman ile kolayca gerçekleştirilebilmesi ve kontaminasyon riskinin az olması gibi özellikleri sebebi ile günümüzde en yaygın olarak kullanılan DNA fragmentasyon yöntemi haline gelmişlerdir. Bu amaç için endonükleaz ve transpozaz enzimleri kullanılır.

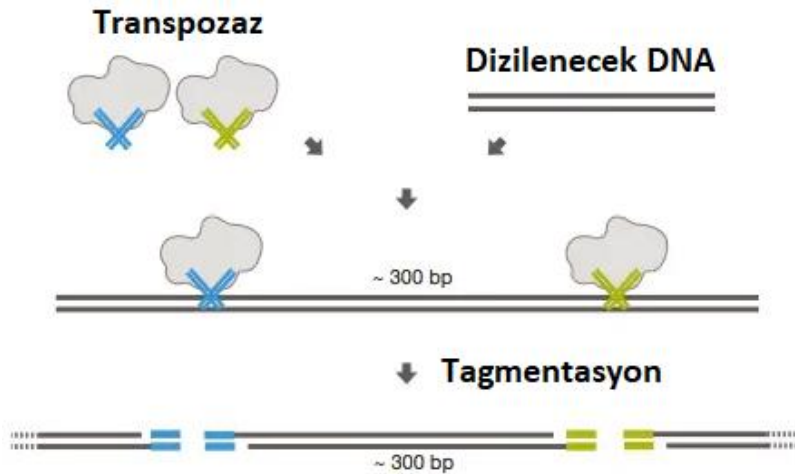
Endonükleazlar

Endonükleazlar çift zincir veya tek zincir DNA veya RNA moleküllerini, molekülün ortasından kesip molekülü daha küçük parçalara ayıran enzimlerdir. Endonükleazların bir kısmı diziye özgüdür; özel bir dizi veya dizi motifini tanıyarak dizi içerisinden veya dizi

dışındaki bir noktadan DNA'yı keserken (restriksiyon endonükleaz enzimleri gibi), bir kısmı da diziden bağımsız olarak rastgele kesim yaparlar (43). Bu rastgele kesim yapan enzimler NGS amaçlı DNA fragmentasyonunda kullanılabilir. Bazı endonükleazlar çift zincir DNA'nın tek zincirinde rasgele kesim gerçekleştirirler; bu enzimler tek zincir DNA kırıklarını tanıyarak diğer zinciri de kesen başka endonükleazlar ile birlikte kombine olarak kullanılarak DNA fragmentasyonunu gerçekleştirirler (42).

Tagmentasyon (Transpozaz)

Transpozaz enzimleri, transpozon olarak nitelendiren çift zincirli bir DNA parçasına bağlanıp, bu DNA parçasını başka bir çift zincirli DNA parçasının rastgele bir yerine ekleyen enzimlerdir. Tn5 transpozaz enzimi kullanılarak, NGS'in daha sonraki aşamalarında kullanılacak bazı diziler (adaptör dizilerinin bir kısmı), dizilenmek istenen DNA molekülleri içerisine rastgele eklenebilmektedir (44). Bu işlem "tagmentasyon" olarak adlandırılmaktadır. Tagmentasyon tam anlamıyla fragmentasyon olmasa da (DNA, çift zincir kırıkları ile parçalanmamaktadır), dizilenecek DNA molekülüne rastgele eklenen DNA parçaları, DNA'yı bir dizilenecek DNA, bir transpozaz tarafından eklenen DNA şeklinde parçalı bir hale getirmektedir (Şekil 2.9.). Tagmentasyon sonucu elde edilen DNA üzerinden, eklenen DNA bölgeleri primer bağlanma bölgeleri olacak şekilde kullanılarak yapılan PCR sonucunda ise, DNA üzerinden kısa ve rastgele parçalar elde edilmektedir ki bu şekilde DNA fragmentasyonu ile amaçlanan sonuca ulaşılmış olunmaktadır (44, 45).



Şekil 2.9. Tagmentasyon (46)

2.1.2.2.2. KLONAL AMPLİFİKASYON

Fragmentasyon sonrası küçük parçalara ayrılan DNA moleküllerinin çoğaltılması, daha yaygın kullanılan tabir ile “amplifiye edilmesi” gerekmektedir. Teorik olarak amplifikasyon yapılmadan, sadece tek molekül üzerinden de dizileme yapmak mümkündür. Ancak tek molekülden dizileme yapmanın büyük bir dezavantajı vardır; tek molekülden elde edilen sinyal miktarı çok düşük olduğu için dizileme yapmak çok zordur. Bu düşük miktardaki sinyal hem daha kirli ve daha hatalı bir dizilemeye yol açar, hem de bu düşük miktardaki sinyali daha verimli bir şekilde yakalamak için cihazlar çok daha kompleks ve pahalı olarak tasarlanmaları gerekir; ki bu da cihaz yatırım ve sürdürme maliyetlerini önemli ölçüde arttırır. Geçmişte teknolojisi Illumina teknolojisine çok benzeyen, ancak tek molekül dizileme teknolojisi kullanan, bir müddet piyasada tutunabilen, ancak daha sonra tek molekül dizilemenin dezavantajları sebebi ile iflas eden bir şirket te olmuştur (Helicos Inc.). Günümüzde tek molekül dizileme yapan şirketler vardır (Pacific Biosciences ve Oxford Nanopore), ancak bunlar hem tam olarak NGS olarak değerlendirilmemektedir (daha çok 3. nesil dizileme teknolojileri olarak kabul edilmektedir), hem de DNA dizileme sektöründeki payları çok düşüktür.

Tek molekül dizilemenin bu düşük sinyal dezavantajının ortadan kaldırılması DNA’nın klonal olarak amplifiye edilmesi ile gerçekleştirilir; yani her bir DNA molekülü ayrı ayrı çoğaltılır. Başlangıç kalıp DNA’sının çok sayıda molekülden oluştuğu standart PCR amplifikasyonu yaklaşımının aksine, klonal amplifikasyonda başlangıç kalıp DNA’sı tek bir DNA molekülüdür ve amplifikasyon sonunda tek bir DNA molekülünün klonları olan çok sayıda DNA molekülleri ayrı ayrı elde edilir. Amplifikasyon PCR olabileceği gibi zincir reaksiyon olmayan izotermal amplifikasyon da olabilmektedir. Çok farklı DNA dizilerine sahip DNA moleküllerinin tamamının aynı reaksiyon ortamında ve aynı koşullarda amplifiye edilebilmesinin günümüze kadar bulunan tek yolu, DNA dizilerinin uç kısımlarında bütün moleküller için ortak dizilerin olmasıdır (47). Bu ortak diziler, adaptör DNA moleküllerin dizilenmesi istenen DNA moleküllerinin uçlarına eklenmesi ile sağlanır. Milyarlarca farklı DNA molekülünün aynı ortamda PCR veya izotermal amplifikasyon ile amplifiye edilmesi bu şekilde mümkün olmaktadır.

Adaptör dizileri 3 farklı şekilde DNA dizilerine eklenebilir: DNA molekülleri uçlarına ligaz enzimi yardımı ile yapıştırılarak, transpozaz enzimi ile fragmentasyon esnasında veya adaptör dizileri içeren primerler kullanılarak yapılan PCR ile. Adaptörlerin eklenmesi ile DNA molekülleri klonal amplifikasyona hazırdır ve bu şekilde hazırlanan DNA'ya NGS kütüphanesi adı verilir.

ADAPTÖRLER

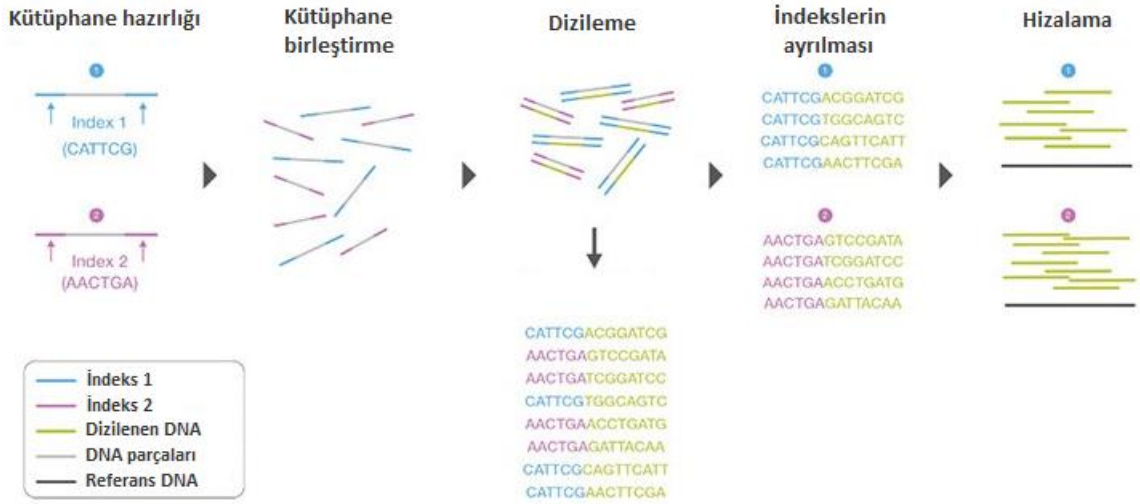
Adaptörler, dizilenmesi istenen DNA moleküllerinin uçlarına eklenen, dizilimi bilinen sentetik DNA molekülleridir. Adaptör dizileri içinde klonal amplifikasyonu mümkün kılan primer dizi bölgeleri ile birlikte, dizileme için kullanılacak primer bölgeleri ve farklı örneklerin fiziksel olarak aynı ortamda dizilenmesine rağmen verilerinin birbirinden ayrılmasını sağlayan indeks dizileri de bulunur (Şekil 2.10.).



Şekil 2.10. Adaptör (48)

İndeksleme

NGS sistemleri çok yüksek miktarda veriyi fiziksel olarak tek veya az sayıda kompartımanlarda elde ederler. Şayet büyük miktardaki veri tek bir örnekten elde edilmek istenmiyorsa, veri çok sayıda örnek arasında paylaştırılmak isteniyorsa bu amaç için bir çözüme ihtiyaç vardır: indeksleme. İndeksleme, fiziksel olarak aynı ortamda birden fazla örneğin dizilenmesine olanak sağlayan yaklaşımdır. Multipleksleme ve barkodlama olarak da adlandırılabilir. İndeksleme stratejisi; dizilimi bilinen ve birbirinden farklı yaklaşık 6-12 bp uzunluğa sahip DNA parçalarının, örnek hazırlama aşamasında adaptörler aracılığıyla her bir örneğe farklı olacak şekilde eklenmesi, dizileme sonunda ise bu indeks dizilerinin de ayrıca okunup hangi verinin hangi örnekten geldiğinin tespiti ile örneklere ait verilerin birbirinden ayrılması prensibine dayanır (49) (Şekil 2.11.).



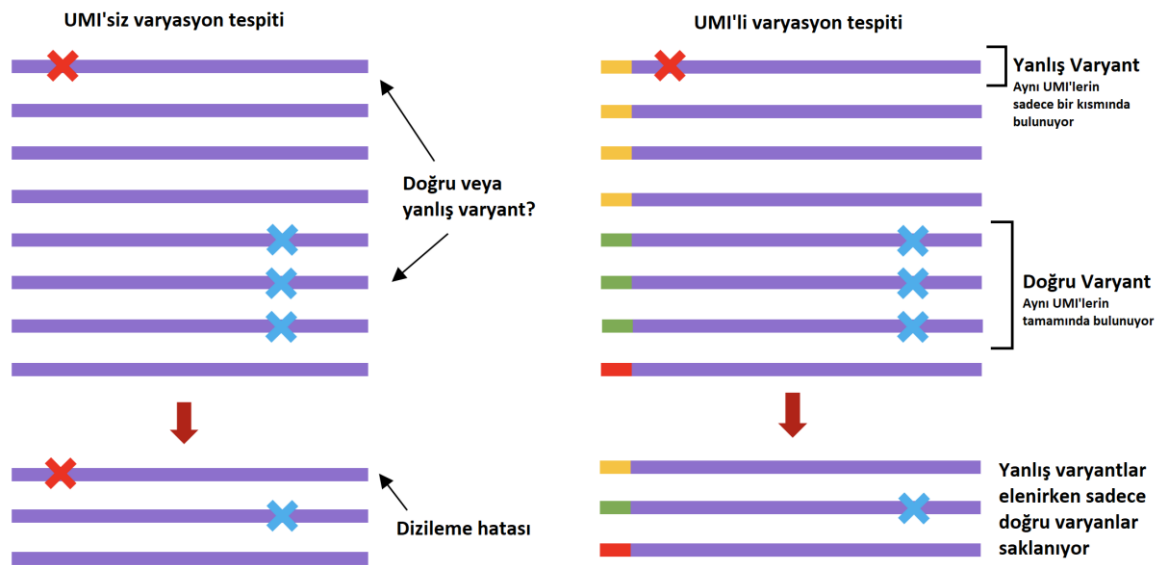
Şekil 2.11. İndeksleme (50)

İki farklı şekilde uygulanabilmektedir: tekli ve çiftli. Tekli indekslemede dizilenecek DNA moleküllerinin bir ucuna indeks diziler eklenirken, çiftli indekslemede moleküllerin iki ucuna farklı indeks dizileri eklenir. Çiftli indekslemenin avantajı, iki uca eklenecek indekslerin kombinasyonlarını kullanarak, az sayıda adaptör ile çok sayıda örneğin indekslenmesini sağlayabilmesidir; mesela moleküllerin bir ucuna eklenecek 8 farklı indeksi içeren 8 çeşit adaptör ve diğer ucuna eklenecek 12 farklı indeksi içeren 12 çeşit adaptör ile, $8 \times 12 = 96$ çeşit kombinasyon sayesinde, toplam 20 farklı adaptör ile 96 örnek multiplekslenebilmektedir (51). Tekli indekslemede kombinasyon olmadığı için bu mümkün değildir; ancak adaptör çeşidi kadar örnek multiplekslenebilir. Ancak çiftli indekslemenin de indeks zıplaması denilen bir dezavantajı vardır. İndeks zıplaması, kütüphane zenginleştirme aşamasında kimerik PCR ürünü oluşması ve bunun sonucunda indekslerin bir molekülden diğer bir moleküle “zıplaması”dır. Şayet çiftli indeksleme yaklaşımı kullanıldı ise, yeni oluşan kombinasyon başka bir örneğe ait olabileceği için örnekler arasında az da olsa birbirlerine veri sızması oluşturmaktadır. Bu sebeple özellikle somatik mutasyon çalışmaları gibi hassas çalışmalarda indeks kombinasyonları ile multipleksleme kapasitesi artırma yaklaşımı kullanılmamaktadır (52).

UMI

UMI (İng. Unique Molecular Identifier) veya Benzersiz Moleküler Tanımlayıcı, hassasiyetin artırılması için kullanılan bir adaptör modifikasyonu yöntemidir. UMI dizileri, adaptör moleküllerinin dizilenecek DNA ya bağlandığı noktaya eklenen genellikle 6-16 bp’lik

rastgele dizilerdir. Rastgele diziler oldukları için adaptör dizilerinin sentezi sırasında bu noktalarda dört nükleotid te rastgele birleşerek rastgele bir dizi oluşturur; böylece her bir molekül benzersiz denebilecek, uzunluğu ve molekül içindeki konumu bilinen ancak dizisi bilinmeyen bir DNA parçası ile işaretlenmiş olur. Dizileme sırasında UMI dizileri de ayrıca okunur. Aynı molekülden oluşmuş PCR kopyalarından elde edilen dizileme bilgileri ile UMI dizilerinin aynı olması gerekir; analiz sırasında şayet UMI dizisi aynı olan okumalar arasında dizide farklı noktalar tespit edilirse, bu farklı noktaların polimerizasyon veya dizileme hatası olduğu varsayılarak veri içerisinden çıkartılarak hatalardan daha fazla ayıklanmış bir veri elde edilir (Şekil 2.12.). Bu da hassasiyetin artmasına katkıda bulunur (53).



Şekil 2.12. UMI (54)

AMPLİFİKASYON

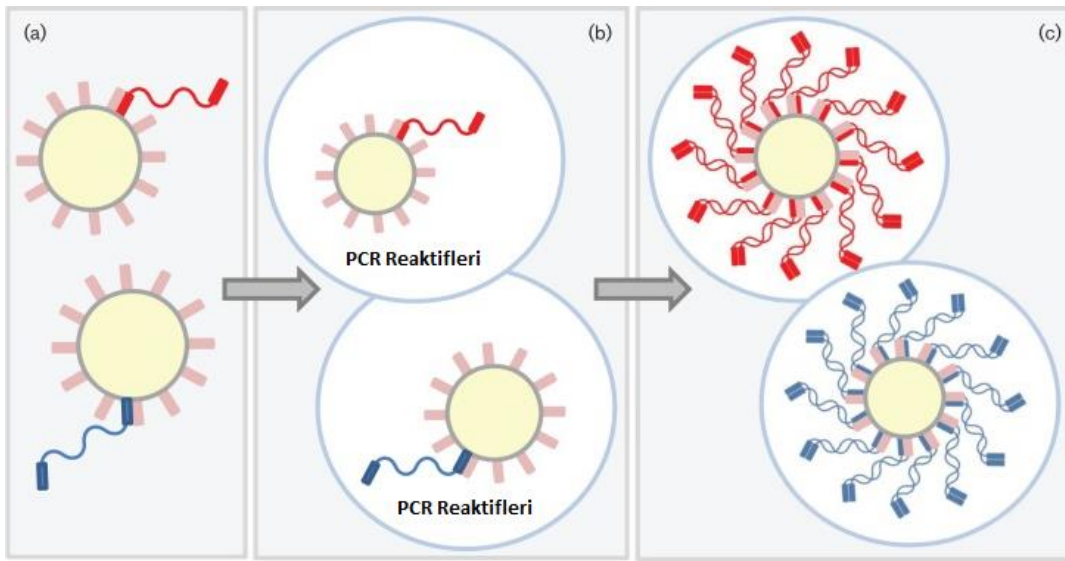
Klonal amplifikasyon için kullanılan mevcut 3 yaklaşım vardır: emülsiyon PCR, köprü amplifikasyon ve DNA nanotopları (DNB).

Emülsiyon PCR

Emülsiyon, birbiri içerisinde çözünmeyen iki sıvının karışımına verilen addır. Bu karışım homojen değildir; daha az miktarda olan bir sıvı (dağılan faz) dağılık/asılı kabarcıklar olarak daha fazla olan bir sıvının (kesintisiz faz) içinde dağılmış durumdadır. Emülsiyon PCR’da, hazırlanan PCR karışımı dağılan faz olacak şekilde, kesintisiz fazın ise yağ olduğu bir emülsiyon haline dönüştürülür; öyle ki su bazlı olan PCR karışımı, yağ ile çevrelenerek

birbirinden izole edilmiş, çapları yaklaşık 2-10 µm olan damlacıklar haline getirilirler. Bu damlacıkların her biri, içinde yeterli miktarda primer, dNTP, DNA polimeraz ve diğer PCR bileşenlerinin olduğu, ancak PCR karışımı içerisinde en az miktarda bulunan kalıp DNA'nın sadece bir kopya bulunduğu, birbirinden bağımsız minik reaksiyon kürecikleri haline gelmiş olurlar. Bu şekilde ısı döngüsüne tabi tutulan PCR karışımı sonunda, her bir reaksiyon küreciği içerisinde sadece tek bir DNA molekülünden çoğaltılmış PCR ürünleri oluşur (55, 56) (Şekil 2.13.).

Damlacıklar rastgele oluştuğu için her bir damla içerisine mutlaka bir adet kalıp DNA'nın gelmesi söz konusu değildir; bazı damlacıkların içerisine birden fazla kalıp DNA düşerken bazı damlacıkların içerisine hiç kalıp DNA denk gelmeyebilir. Bu istenmeyen damlacıklar sorun teşkil etmez, çünkü hacimleri yaklaşık olarak 100 fL olan bu küreciklerden yüz milyonlarca oluşturulduğu için, analiz için yeterli sayıda istenen şekilde damlacık bulunmaktadır; istenmeyen şekilde olan damlacıklar analiz esnasında göz ardı edilmektedir. NGS amacı ile yapılan emülsiyon PCR'da, fragmente edilmiş ve uçlarına adaptör takılmış DNA parçaları kalıp olarak kullanılır. Bu DNA parçaları sonradan kolayca geri toplanıp NGS'in diğer aşamalarına devam edilebilmesi amacı ile, PCR karışımı içerisindeki primerler mikron boyutta katı boncuklar üzerine tutturulmuş haldedir; bunun sonucu olarak reaksiyon sonunda elde edilen tüm amplikonlar (yani amplifikasyon aşamasında yeni oluşmuş DNA molekülleri) boncuk üzerinde sabitlenmiş şekilde oluşur. Oluşan bu yapılar "polony" adı verilmektedir.

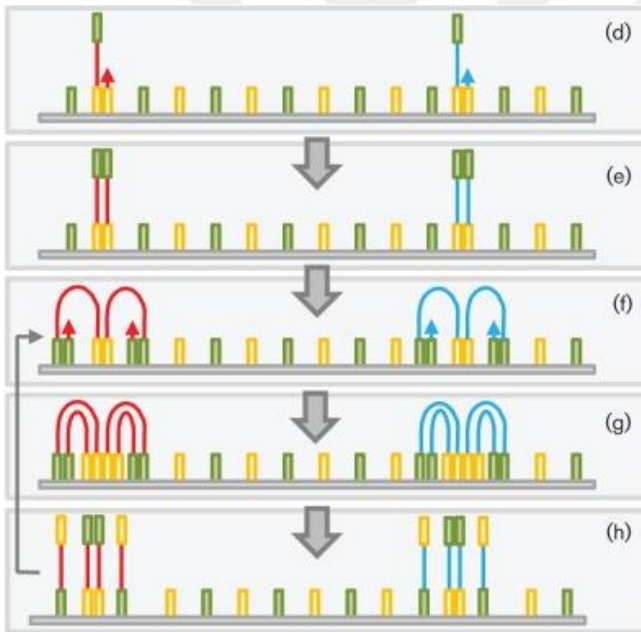


Şekil 2.13. Emülsiyon PCR (57)

Emülsiyon PCR, NGS uygulamaları dışında dijital PCR yönteminde de kullanılmaktadır. Emülsiyonu oluşturan damlacıkların çok daha standart olması ve farklı damlacıkların birleştirilmesi amacı ile, değişik uygulamalarda mikrofluidik sistemler de kullanılabilmektedir (55).

Köprü Amplifikasyonu

Illumina teknolojisinin geliştiricisi olan Solexa firmasının DNA dizileme denemeleri, akış hücresi (İng. flow cell) olarak adlandırılan bir reaksiyon odacığı içerisinde katı yüzey üzerine sabitlenmiş tek DNA moleküllerinin dizilenmeye çalışılması ile başlamıştır. Kısa süre içerisinde tek bir DNA molekülü yerine, tek bir molekülden çoğaltılmış DNA kümesi kullanılarak dizileme yapılmasının, sinyal miktarını ciddi oranda iyileştireceği ve dizilemeyi hem daha kolay hem de daha kaliteli bir hale getireceği fark edilmiştir. Bunun üzerine köprü amplifikasyonu geliştirilmiştir.



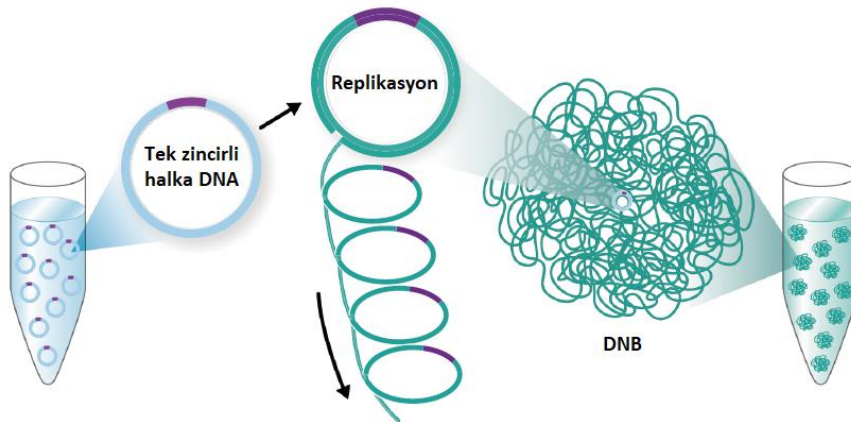
Şekil 2.14. Köprü amplifikasyonu (57)

Köprü amplifikasyonunda adaptör dizilerine karşılık gelen PCR primerleri akış hücresi içerisindeki katı yüzey üzerine sabitlenmiştir. İki ucunda adaptör dizileri olan DNA parçaları akış hücresi içerisine gönderilir ve yüzey üzerine rastgele bağlanır. Bağlandıkları yerde PCR gerçekleştirilir. PCR esnasında, her bir döngüde, tek zincir DNA molekülleri bir uçlarından katı yüzeye sabitlenmiş olduklarından, diğer uçları ile yakında bulunan yüzeye sabit başka bir primer molekülüne bağlanır. Polimerizasyon sonucunda katı yüzeye sabit tek bir tek

zincir DNA molekülünden, birbirine karşılık gelen iki adet tek zincir DNA molekülü oluşturulmuş olur. DNA moleküllerinin boyu yaklaşık birkaç yüz bp olduğu için diğer uçtan yakalanan primerin fiziksel olarak çok yakında olması gerekir; DNA molekülünün boyu yaklaşık 200 nm dir. PCR döngülerinin tamamlanmasından sonra, yüzey üzerine sabit, rastgele yerlerde oluşmuş, tek bir DNA molekülünden klonal olarak çoğaltılmış, yaklaşık 1000 molekül içeren ve yaklaşık 1 µm çapında DNA kümecikleri oluşur. Bu kümecikleri oluşturma işlemine köprü amplifikasyonu, kümeciklerin kendilerine de küme (İng. cluster) adı verilmiştir (58, 59) (Şekil 2.14.).

DNA Nanotopları

DNA nanotopları, tek zincir DNA'nın sıvı içerisinde, zincir içindeki nükleotidlerin zayıf etkileşimi sonucu gevşek bir katlanma motifi oluşturarak, küresel bir şekil alması sonucu oluşan yapıdır. Bu yapı, bir yün ipinin toparlanarak oluşturulan gevşek bir yün topuna benzetilebilir. NGS için kullanılan DNA nanotopları oluşturulurken adaptör dizilerinden faydalanılır. DNA ilk önce enzimatik degradasyon ile tek zincir haline getirilir. Adaptör dizileri kullanılarak tek zincirin iki ucu birleştirilerek halka yapıda DNA oluşturulur. Daha sonra yine bu adaptör dizileri primer bölgeleri olarak kullanılarak zincir değişim (İng. strand displacement) özelliği bulunan bir DNA polimeraz enzimi ile izotermal amplifikasyon yapılır (60). Amplifikasyon sonucunda tek bir DNA molekülünden oluşmuş, molekülün yüzlerce kopyasını arka arkaya eklenmiş bir şekilde içeren, tek zincirli, yaklaşık 300 nm çapında DNA nanotopları oluşturulmuş olunur (61, 62) (Şekil 2.15.). DNB'nin diğer yöntemlere göre bir avantajı da zincirleme reaksiyon olmadığı için Sayfa 87'da tartışılan PCR yönteminin dezavantajlarından büyük ölçüde etkilenmemektedir.



Şekil 2.15. DNB (63)

2.1.2.2.3. RASTGELE DİZİLEME

NGS cihazlarında dizileme rastgele yapılır. Buna İngilizce olarak “shotgun” (av tüfeği) dizileme denilmiştir (av tüfeğinden çıkan çok sayıda saçmanın hedefe doğru, ancak rastgele yönelmesi, saçmaların bir kısmının hedefi ıskalmasına rağmen hedefi vuranların işi görmeye yetmesi benzetmesinden yola çıkılarak) (47, 64). Sanger dizilemenin yapıldığı kapiller elektroforez sistemlerinde nerede (hangi kapillerde) hangi DNA parçasının dizlendiği bilinmektedir, ancak NGS sistemlerinde dizileme sırasında nerede hangi DNA parçasının dizlendiği bilinmez, rastgele olarak dizileme yapılır. Cihazın herhangi bir noktasında dizilen DNA parçasının nereye ait olduğu ancak dizileme tamamlanıp elde edilen dizinin analiz edilmesi ile tespit edilebilir.

NGS sistemlerinde, DNA dizileme kısmında değişik yöntemler kullanılmaktadır. Bunların başlıcaları, pirosekanslama, sentez ile sekanslama (SBS olarak da bilinir), ligasyon ile sekanslama ve iyon yarıiletken sekanslamadır; günümüzdeki NGS sistemlerinin tamamı bu yöntemlerden bizzat bir tanesini veya modifiye edilmiş bir çeşidini kullanmaktadırlar.

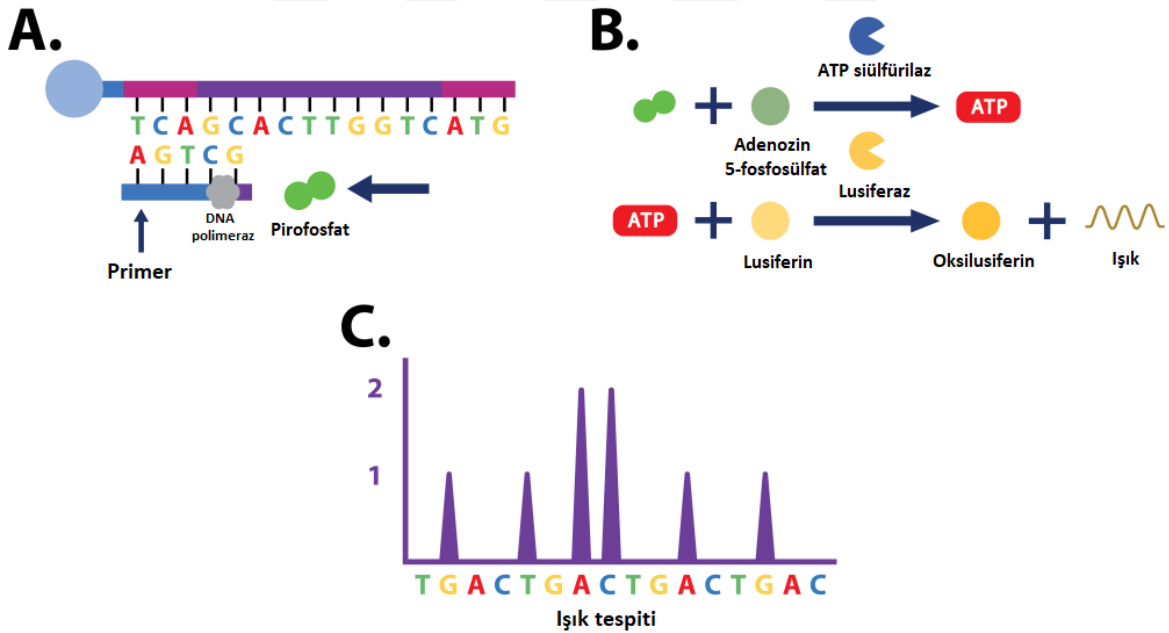
PIROSEKANSLAMA

Pirosekanslama, DNA sentezi sırasında DNA polimeraz enziminin nükleotidleri uzayan zincirin 3’ ucuna eklemesi ile oluşan değişikliklerin, birtakım enzimler yardımı ile ışık sinyaline dönüştürülmesi prensibine dayanan DNA dizileme yöntemidir (65). Pål Nyren ve arkadaşları tarafından 1993 yılında geliştirilmiştir (66). Bu dizileme yöntemi ilk olarak minidizileme, yani kısa DNA dizilerini okuma amaçlı kullanılsa da, 1998 yılında yine aynı ekip tarafından geliştirilerek yeni nesil dizileme teknolojisine uyarlanabilecek hale getirilmiştir (21). Ticari olarak 2005 yılında piyasaya sürülmüş olan ilk NGS cihazı pirosekanslama temeline dayanmaktadır (65, 66).

Pirosekanslama, doğada ateşböceği gibi bazı canlıların gerçekleştirdiği biyoluminesans ile ışık üretme olayından faydalanarak geliştirilmiştir. Pirosekanslamda, DNA polimeraz enzimi dışında iki adet enzim daha kullanılmaktadır; ATP-sülfürilaz ve lusiferaz enzimleri. DNA polimeraz enziminin nükleotid-trifosfat moleküllerini uzayan DNA zincirinin 3’ ucuna eklemesi ile ortaya di-fosfat (pirofosfat) molekülü ile hidrojen iyonu ortaya çıkar. Ortamda bulunan ATP-sülfürilaz enzimi, reaksiyon içindeki adenosin 5-fosfosülfat molekülünü

pirofosfat ile birleştirerek ATP molekülünü oluşturur. Oluşan ATP molekülü de lüsiferaz enziminin katalizörlüğü ile ortamda bulunan lüsiferin molekülü ile tepkimeye girerek bir ışık parlaması oluşturur (65, 66).

Pirosekanslamada katı bir yüzeye bağlanan tek zincirli DNA parçaları üzerine her bir döngüde her bir nükleotid sıra ile gönderilerek, DNA polimerizasyonunun gerçekleştirilmesi ve reaksiyon ortamında bulunan diğer enzim ve reaktifler sayesinde bir ışık sinyali alınması sağlanır. Şayet dizilemede sıradaki nükleotid, ortama gönderilen nükleotid değilse polimerizasyon gerçekleşmeyecek ve ışık sinyali alınmayacaktır. Şayet sırada gönderilen nükleotidden ardışık olarak birden fazla bulunuyorsa da birden fazla nükleotid aynı anda uzayan DNA dizisine eklenecek ve aynı oranda fazla reaksiyon ve dolayısıyla fazla ışık oluşacaktır. Sonuçta her bir döngüde ve her bir nükleotidde ışık sinyali alınıp alınmadığının tespiti ve alındı ise ne miktarda alındığının ölçülmesi ile dizileme gerçekleştirilir (65) (Şekil 2.16.).



Şekil 2.16. Pirosekanslama (67)

DNA dizilimi üzerinde aynı bazın iki veya daha fazla ardışık olarak tekrar ettiği durumlarda bu bazların aynı döngü içerisinde aynı zamanda dizilenmesi bir avantaj gibi görünse de (her bir döngüde sadece bir nükleotid dizilenmesine göre yaklaşık %50'si kadar fazla DNA

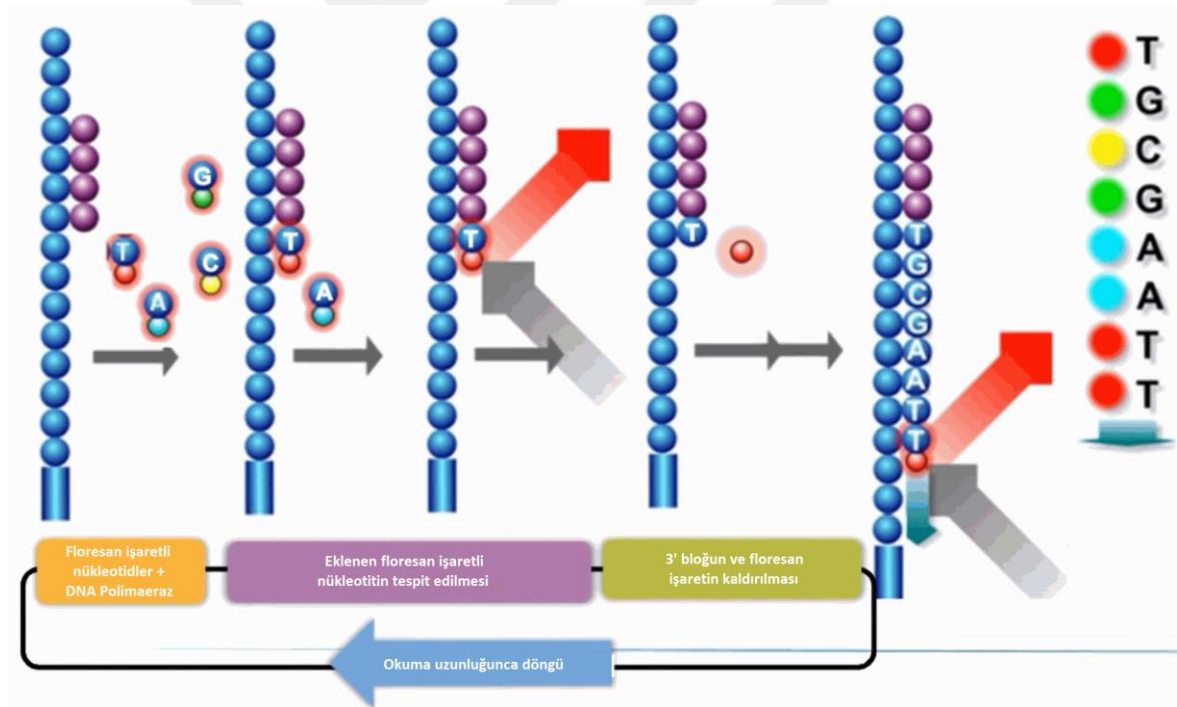
dizilimi elde edilmektedir), bu özellik pirosekanslamanın (ve aynı prensiple çalışan iyon sekanslamanın) en büyük dezavantajı olmuştur: homopolimer problemi.

Homopolimer, tek bir yapı taşından oluşmuş polimere verilen addır. DNA’da homopolimer denildiği zaman da DNA dizisi içinde aynı nükleotidin ardışık olarak çok sayıda tekrar ettiği kısımları anlaşılır. Homopolimer bölgelerini aynı döngü içerisinde beraber dizileyen teknolojilerde, dizilenirken elde edilen toplam sinyalin tam olarak kaç bazdan geldiğinin belirlenmesi gerekmektedir. Homopolimeri oluşturan baz sayısı ne kadar fazla ise, dizinin doğru belirlenmesi o kadar zordur; çünkü her baz için olan sapma miktarı toplanarak homopolimer bölgesi için daha büyük bir sapma miktarı oluşturacaktır. Sonunda oluşan toplam sapma baz sayısının hatalı olarak bir fazla veya bir az olarak belirlenmesine yol açabilmektedir. Yani mesela 7 adet T’nin tekrar ettiği bir DNA homopolimeri, dizilenme sonunda hatalı olarak 6 T veya 8 T olarak dizilenebilmektedir. Hata oranları homopolimer uzunluğu dışında kütüphane hazırlık aşaması, kullanılan polimerazın özellikleri, homopolimerin okumanın neresinde olduğu, okuma derinliği, analiz algoritmaları vb. çok sayıda etkene de bağlıdır; bu sebeple literatürde yüzde 1-2 ile yüzde 80-90 gibi çok geniş bir aralıkta hata oranları bildirilmiştir. Ancak kesin olan 4 ten fazla baz içeren homopolimerlerde hatalı dizileme ihtimali belirgin bir şekilde artmakta ve baz sayısı arttıkça da dramatik olarak daha da artmaktadır (68-71). Geliştirilen özel DNA polimerazlar ve yazılım algoritmaları ile homopolimer problemi azaltılmasına rağmen, özellikle düşük derinlikli ve büyük hedefli insan klinik genetiği çalışmaları (tüm ekzom dizileme ve tüm genom dizileme gibi) yapan araştırmacıların büyük çoğunluğu sırf bu homopolimer probleminden dolayı bu teknolojileri tercih etmemektedirler.

SENTEZ İLE SEKANSLAMA

NGS sistemlerinde kullanılan sentez ile sekanslama yönteminin temeli, zincir sonlandıran ve floresan işaretli, ancak hem zincir sonlandırma özelliği hem de floresan özelliği kimyasal olarak geri döndürülebilir olan (İngilizce olarak “Reversible Dye Terminator” olarak adlandırılır) nükleotid analoglarının kullanılmasıdır. Sanger dizilemede de floresan işaretli ve zincir sonlandıran nükleotid analogları kullanılır, ancak NGS’de kullanılan moleküllerin farkı geri döndürülebilir olmasıdır; yani bu nükleotidler değişik kimyasallara maruz bırakılarak hem floresan özellikleri hem de zincir sonlandırma özellikleri ortadan kaldırılabilmektedir (72).

Sentez ile sekanslama sırasında, her bir döngüde ortama DNA polimeraz ve 4 nükleotid çeşidi için floresan işaretli zincir sonlandırıcı nükleotidler içeren reaksiyon karışımı gönderilir. DNA polimeraz her bir zincire bu nükleotidlerden sadece bir tane ekleyebilir. Reaksiyona girmeyen reaksiyon karışımı ortamdan uzaklaştırılır ve floresan tespit yapılır. Her bir nükleotid çeşidi ayrı özellikte bir floresan molekül içerdiği için, 4 nükleotidin de tespiti aynı işlemle gerçekleştirilir. Daha sonra dizilenecek DNA moleküllerinin bir sonraki döngüye hazır hale getirilmesi için kimyasal olarak nükleotidlerin floresan işaretleri ve zincir sonlandırma özellikleri ortadan kaldırılır. Reaksiyon karışımının ortama gönderilmesi ile bir sonraki döngü başlar. Bu şekilde her döngüde dizilenecek milyonlarca molekülün (Illumina sistemlerindeki “cluster” yapılarının) sadece bir nükleotidi dizilenecek, toplam döngü sayısı tamamlandığında milyonlarca farklı DNA molekülünden gelen belirli uzunlukta bir DNA dizileme bilgisi elde edilmiş olunur (59) (Şekil 2.17.). Her döngüde sadece bir nükleotid dizilendiği için bu yöntemin homopolimer problemi yoktur.



Şekil 2.17. Sentez ile sekanslama (73)

Günümüzde üretilen toplam NGS verisinin çok büyük bir kısmı (hatta tamamına yakını) bu yöntem veya bu yöntemin modifiye edilmiş bir hali kullanılarak elde edilmektedir. Modifiye edilmiş yöntemde 4 yerine 2 farklı floresan molekül kullanılmakta ve 4 nükleotid bu moleküllerin kombinasyonu ile birbirinden ayrıştırılmaktadır; mesela T ve C farklı 2 floresan

molekül taşıırken, A her iki molekülü de taşımakta, G ise iki molekülü de taşımamaktadır (74). 4 yerine 2 farklı floresan kullanmanın en büyük avantajı, çok daha basit optik yola ve tespit sistemlerine sahip cihazların üretilebilmesine olanak sağlamasıdır; bu sayede günümüzde “benchtop” diye isimlendirilen, Real-Time PCR cihazlarında bile daha düşük maliyetli NGS cihazları üretilebilmiştir.

Geri döndürülebilir özellikli nükleotid analogları (ve bu doğal olmayan nükleotidlerin polimerizasyonunu gerçekleştirebilen tasarlanmış DNA polimerazlar) bu yöntemi mümkün kılmak ile beraber, aynı zamanda teknolojiyi sınırlandıran etken de olmuşlardır. Doğal DNA polimerizasyonunda bulunmayan bu kimyasal süreçlerin verimliliklerinin göreceli olarak düşük olması sebebi ile oluşan hatalar (her döngüde mutlaka olması gereken polimerizasyonun ve floresan özellik ve blokların kaldırılmasının gerçekleşmemesi veya fazladan bir baz daha eklenmesi) fazlama oluşmasına sebep olmaktadır. Fazlama, dizilenen klonal moleküllerin bir kısmının önden, bir kısmının da arkadan gelmesidir; yani moleküllerin bir kısmının olması gereken doğru bazdan sinyal verirken, bir kısmının bir önceki bazdan, bir kısmının da bir sonraki bazdan sinyal vermesidir. Ve hatalar döngüler boyunca birikerek arttığı için belirli bir döngü sonrasında fazlama o kadar yüksek hale gelmektedir ki, hatalı baz okumaları artmakta, dizileme kalitesi düşmekte ve daha da fazla dizilemenin bir anlamı kalmamaktadır (75, 76). Bu sebep ile bu yöntemden elde edilen okuma uzunlukları kısadır; günümüzde en uzun okuma uzunlukları yaklaşık 300 bp civarındadır. Aslında 300 bp bile biraz zorlanmış bir uzunluktur; 200-250 bazdan sonra okuma kalitesi çok düşmektedir. Bu düşüş, okuma hatalarının daha kolay tahammül edilebildiği mikrobiyal uygulamalar gibi bazı uygulamalar için büyük problem teşkil etmemek ile birlikte; özellikle düşük derinlikli ve büyük hedefli insan klinik genetiği çalışmalarında (tüm ekzom dizileme ve tüm genom dizileme gibi) sıkıntı oluşturmaktadır. Bu sebeple günümüzde NGS'nin en yaygın yapılma amacı olan insan klinik çalışmalarında kullanılan, en yaygın dizileme yönteminin, en yaygın okuma uzunluğu sadece 100-150 bp civarındadır. Sentez ile sekanslamanın ilk geliştirildiği 2000'li yıllarda okuma uzunluğu çok daha kısaydı; en uzun okumalar 30-35 bp civarındaydı. Zaman içinde geliştirilen kimyasal süreçler ile daha verimli reaksiyonlar elde edilerek hatalar azaltıldı ve geliştirilen biyoinformatik analiz algoritmaları ile fazlama miktarları hesaplanarak daha verimli sinyal düzeltmeleri yapıldı; bu sayede 30-35 bp uzunluktan günümüzdeki haline gelebildi. Ancak uzun süredir sentez ile sekanslamanın okuma uzunlukları artmamaktadır; görünüme göre bu

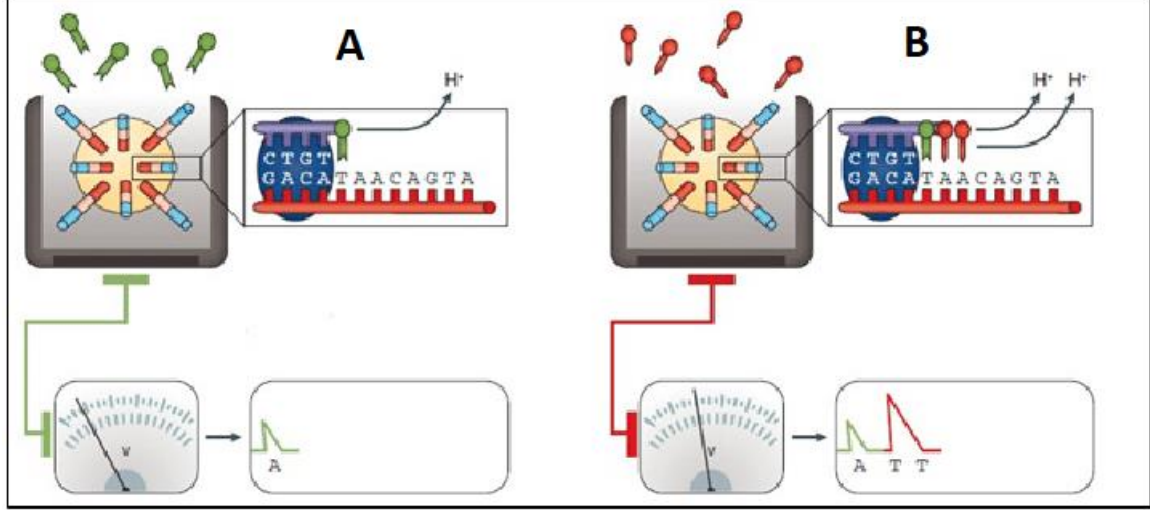
teknoloji okuma uzunluđu konusunda limitine ulaşmıştır ve bu yöntem ile daha uzun okumlar elde edilemeyecek gibi görünmektedir. Bu dizileme yöntemini kullanan yeni üretilen cihazların tamamı okuma uzunlukları benzer şekilde kısa olan ancak daha fazla DNA parçası okumak sureti ile daha fazla dizileme verisini daha ucuza üretmeyi vadeden cihazlardır; uzun okumalar 3. nesil teknolojilerden gelmektedir (77, 78).

İYON SEKANSLAMA

454 Life Sciences ile ilk NGS sistemini geliştirilen Jonathan Rothberg daha sonra iyon sekanslama yöntemini geliştirmiştir. İyon yarı iletken sekanslama olarak da bilinen iyon sekanslama, DNA sentezi sırasında DNA polimeraz enziminin nükleotidi, uzayan zincirin 3' ucuna eklemesi ile ortaya çıkan hidrojen iyonu (H^+) ve pirofosfat molekülünden, hidrojen iyonunun tespiti sayesinde gerçekleştirilir (Şekil 2.18.). Ortaya çıkan H^+ iyonunun -pH'ta değişikliğe yol açtığı için- bir pH metre ile gerçek zamanlı tespit edilebileceği 1992'den beri bilinmektedir (79). Ancak bu prensibi kullanarak, yarıiletken mikroçip sensörler üzerinde yüksek hacimli DNA dizileme yöntemi -biraz da gecikmeli olarak- yaklaşık 20 yıl sonra geliştirilebilmiştir (80).

İyon sekanslama ışık tespiti dışında başka bir tespit yöntemi kullanan ilk DNA dizileme yöntemidir (Sanger ve Maxam-Gilbert dizileme yöntemleri de dahil). Işık tespiti (optik tespit), üretilen NGS cihazlarını daha karmaşık, hassas, pahalı, yavaş ve büyük boyutlu yapmaktadır; optik tespitin ortadan kaldırılması bir anda bu teknolojiyi rakiplerine göre daha avantajlı hale getirmiştir. Elektronik alanındaki mikroçip teknolojisinin zaman içindeki logaritmik gelişimi ve yaygınlığı da düşünülecek olursa, mikroçip bazlı bir DNA dizileme yönteminin avantajı bir kat daha artmaktadır. Ancak bu yöntem ile üretilen teknolojinin -aynı pirosekanslamada olduğu gibi- homopolimer dezavantajı vardır; iyon sekanslamadan daha önce geliştirilmiş olan pirosekanslamanın homopolimer problemi, bütün olarak iyon sekanslama için de geçerlidir. Bu sebeple sentez ile sekanslama yöntemi ile rekabetinde geri kalmıştır; her ne kadar günümüzde sentez ile sekanslamadan sonra en yaygın kullanılan ikinci DNA dizileme yaklaşımı olsa da sentez ile sekanslama ile karşılaştırıldığında çok daha az yaygınlaşabilmiştir. Aynı pirosekanslamada olduğu gibi özellikle düşük derinlikli ve büyük hedefli insan klinik genetiği çalışmaları (tüm ekzom dizileme ve tüm genom dizileme gibi) yapan araştırmacıların büyük çoğunluğu sırf bu homopolimer probleminden dolayı bu teknolojiyi tercih etmemektedirler. En yaygın kullanıldığı alanlar, dizileme hatalarının daha

iyi tolere edildiği girişimsel olmayan prenatal test (İng. Noninvasive Prenatal Testing, NIPT) ve implantasyon öncesi genetik anöploidi testi (İng. Pre-Implantation Genetic Testing for Aneuploidy, PGT-A) gibi etiket bazlı çalışmalardır (81, 82).



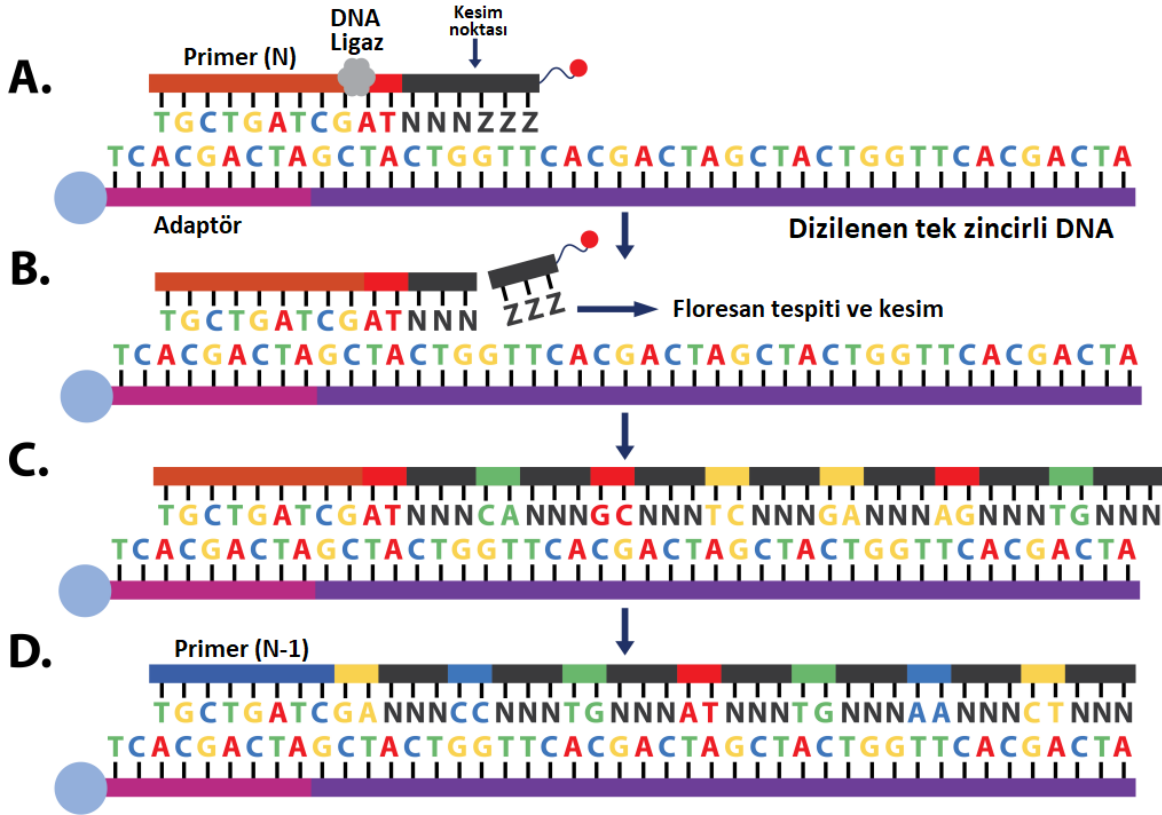
Şekil 2.18. İyon sekanslama (83)

LİGASYON İLE SEKANSLAMA

DNA ligaz enziminin floresan işaretli kısa DNA parçalarını DNA molekülüne eklemesi ile yapılan DNA dizileme yöntemidir. DNA ligaz enzimleri özel koşullarda iki DNA molekülünü birbirine bağlayan enzimlerdir. Hücre içinde genellikle DNA rekombinasyonu ve DNA tamiri işlemlerinde görev alırlar (84). Ligasyon ile sekanslamada her döngüde bir veya 2 bazı kodlanmış ve farklı floresan molekül içeren kısa tek zincir DNA problemleri ve DNA ligaz enzimi içeren reaksiyon karışımı ortama gönderilir (Şekil 2.19.). Hibridizasyon ve ligasyon sonrasında reaksiyona girmeyen reaksiyon karışımı yıkanarak uzaklaştırılır. Floresan sinyal tespiti ile kodlanmış olan bir veya iki bazın okuması gerçekleştirilir. Bir sonraki döngü için ya floresan sinyal ve ligasyon bloğu kaldırılır, veya hibridizasyon primeri tamamen uzaklaştırılarak bir veya iki baz daha uzun primer gönderilerek bir veya iki baz kayma ile okuma döngülerine devam edilir (85).

Dezavantajları olan kısa okuma uzunluğu (sadece 35 bp) ve çok uzun olan dizileme sürelerine rağmen başlarda maliyet açısından avantajlı olan ligasyon ile sekanslama, piyasada bir müddet tutunabilmiş, ancak hızla gelişen diğer teknolojiler karşısında

dezavantajlı durumu gitgide artınca, piyasada ticari olarak bulunan tek teknoloji olan SOLiD sekanslama da rekabeti kaybedip 2016’da tüm ticari faaliyetlerini sonlandırmıştır (5).



Şekil 2.19. Ligasyon ile sekanslama (86)

PHRED KALİTE SKORLAMA

Mevcut NGS sistemleri (ve 3. nesil sistemler) yukarıda bahsedilenler gibi farklı dizileme teknolojileri kullandıkları için hata oranları da birbirinden farklı olabilmektedir. Dizileme yapılırken aynı zamanda herhangi bir nokta için okunulan bazın güvenilirliğini belirten bir kalite skoruna ihtiyaç vardır; ki herhangi bir noktada tespit edilen bir değişikliğin ne kadar güvenilir bir sonuç olduğu değerlendirilebilsin. Bu amaç için Phred skorlama metodu geliştirilmiştir (87); Phred skorlamada elde edilen sinyal kalitesi ile bir muhtemel hata ihtimali hesaplanır (P) ve bu ihtimal aşağıdaki formül ile Phred kalite skoruna (Q) dönüştürülür:

$$Q = -10 \log_{10} P$$

Formüle göre Phred skorunun 40 olması, tahmin edilen hata ihtimalinin 10.000’de 1, 30 olması 1.000’de 1, 20 olması 100’de 1 ve 10 olması 10’da 1 olması anlamına gelir. Okuma yapılan her bir baz ile birlikte o baza ait Phred skoru (Q skoru da denir) oluşturulan fastq dosyasına işlenir.

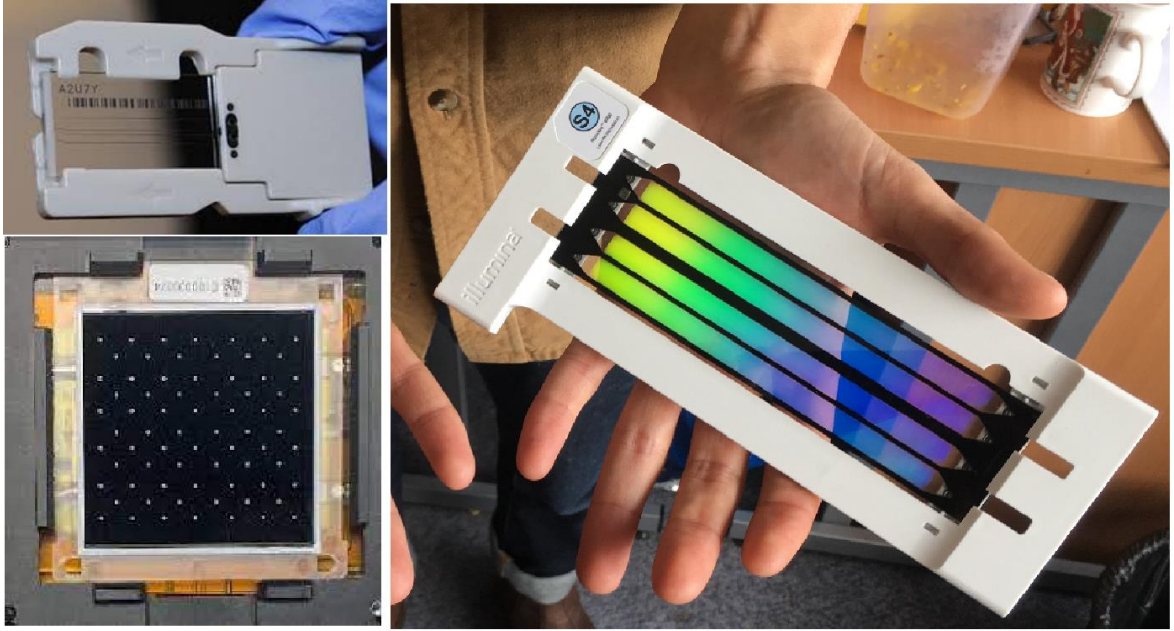
Her ne kadar Q skorları 40’ın üzerine çıkabilse de günümüzde yaygın olarak kullanılan NGS sistemlerinin hata ihtimallerini belirtmek için 40’ın üzerindeki skorlar (yani 10.000’de 1’den daha az hata ihtimalleri) biraz fazla kalmaktadır; çünkü genelde hata ihtimalleri genelde 10.000’de 1’den çok daha yüksektir. Bu sebeple genel kullanılan Q skorlamada maksimum skor 40 olarak skorlama yapılmaktadır; nadir olarak 40’ın üzerinde olarak hesaplanan skorlar 40’a çekilmektedir. Genel olarak Q skorunun 30-40 arasında olması çok yüksek kalitede okuma, 20-30 arasında olması yüksek kalitede okuma anlamına gelir denilebilir. Q skorunun 10-20 arasında olması kalitenin düşük olmasına rağmen yine de kullanılabilir olduğu anlamına gelirken, 10’dan düşük olması durumu ise okumaya kesinlikle güvenilememesi anlamına gelir.

2.1.2.2.4. YÜKSEK HACİMLİ PARALEL DİZİLEME

NGS’nin en büyük avantajı, göreceli olarak kısa sürede, göreceli olarak çok miktarda veri üretebilmesidir. NGS’ye asıl gücünü veren özelliği budur da denebilir. Bu sebeple NGS aynı zamanda “Yüksek Hacimli Paralel Dizileme” (İng. Massively Parallel Sequencing) olarak ta isimlendirilebilmektedir. Paralel dizileme, birden fazla dizilemenin aynı cihazda ve aynı anda beraber olarak yapılmasıdır. Paralel dizileme NGS’ye özgü bir özellik değildir; Sanger dizilemenin yapıldığı kapiller elektroforez cihazlarında da paralel dizileme gerçekleştirilebilmektedir. Ancak bu yüksek hacimli değildir; cihazın kapiller sayısına göre 4-48 adet örnek paralel dizilenebilmektedir (teknolojinin ulaştığı son noktada 384 kapillere sahip cihazlar geliştirilmiştir, ancak kullanım zorluğu ve NGS’nin geliştirilmesi sebepleri ile yaygınlaşamamıştır). NGS cihazlarında ise yüksek hacim söz konusudur, mevcut cihazlardan en yüksek hacimli tek bir seferde 10 milyarlarca DNA parçasını paralel olarak dizileyebilmektedir (en düşük kapasiteli NGS cihazlarında bile bu sayı milyonlar seviyesindedir).

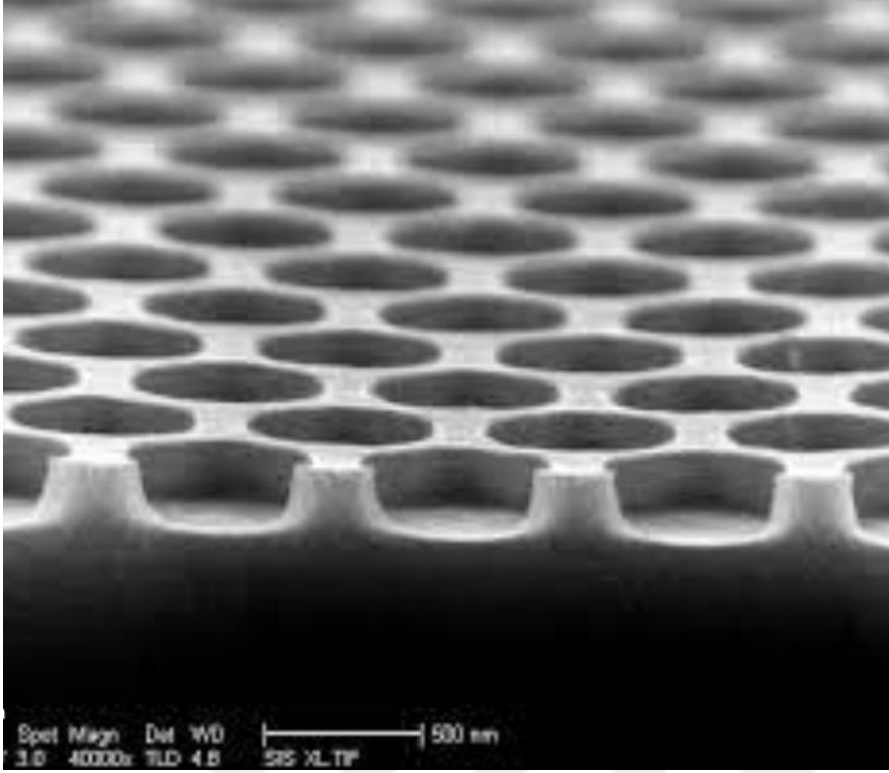
AKIŞ HÜCRESİ

NGS sistemlerine yüksek hacimli paralel dizileme özelliğini kazandıran, dizilemenin gerçekleştirildiği akış hücresi (İng. flow cell) yapılarıdır; akış hücreleri bütün NGS sistemlerinde bulunurlar. Akış hücreleri, klonal olarak amplifiye edilmiş DNA parçalarından küçük (maksimum birkaç cm^2) ama çok yoğun (yaklaşık birkaç milyon adet / mm^2) bir dizin oluşturan ve iki ucundaki küçük açıklıktan faydalanarak oluşturulabilen tek yönlü akış sayesinde içerisine istenilen sıvı karışımının gönderilmesine izin veren kapalı bir odacık veya odacıkları içeren, yaklaşık 5-15 cm uzunluğunda yapılardır. İki ucunda delik olan içi boş bir cam gibi basit bir yapıda da olabilirken, iyon sekanslamada gereken mikroçip tespiti için gerekli bazı elektronik birleşenleri de içeren daha karmaşık yapılar da olabilmektedirler. Şekil 2.20.'de değişik NGS cihazları için üretilmiş, cihazlara özel örnek akış hücreleri görülebilir.



Şekil 2.20. Değişik akış hücreleri

Akış hücrelerinin çok büyük bir kısmı, oluşturulan klonal amplifiye edilmiş DNA parçalarının yerleştirileceği veya oluşturulacağı mikrometre boyutundaki boşlukların düzenli bir şekilde önceden oluşturulduğu örüntülü (İng. patterned) akış hücreleridir (Şekil 2.21.); ancak bazı sistemlere özgü örüntüsüz akış hücreleri de vardır. Örüntüsüz akış hücrelerinde klonal amplifiye edilmiş DNA, akış hücresi içerisine rastgele dağılır; tespit esnasında akış hücresi üzerindeki koordinatları belirlenerek analiz edilebilir.



Şekil 2.21. Örüntülü akış hücresi SEM görüntüsü (88)

2.1.2.2.5. ELDE EDİLEN KISA OKUMALARIN BİRLEŞTİRİLMESİ

NGS sistemlerinde genelde birkaç yüz bp'lik kesintisiz okumalar elde edilir (en yaygın olarak kullanılan cihazlarda yaklaşık 100-200 bp arasında). Bu okumalar Sanger dizilemeye göre biraz kısadır; Sanger dizileme ile 1 Kbp'den biraz daha fazla kesintisiz okuma elde etmek mümkündür. Elde edilen dizilemeler hem küçük hem de aynı zamanda rastgele olduğu için bir yazılım tarafından ortak yerlerinden birleştirilerek uzun dizilerin oluşturulması gerekmektedir.

NGS ile ilgili şu şekilde bir analogi kurulabilir:

SORU: Elimizde çok büyük bir fotoğraf var, mesela bir kenarı 2 m olan bir kare şeklinde. Elimizde bulunan küçük bir tarayıcı ile bu fotoğrafı nasıl tarayabiliriz?

ÇÖZÜM: Fotoğrafı tek seferde tarayıcıya sığdırıp tarayamayacağımıza göre, fotoğrafı tarayıcıya sığacak boyutlarda, çok sayıda küçük parçaya böleriz (bu örnek için birkaç yüz parça), her parçayı ayrı ayrı tararız, sonra bir resim düzenleme yazılımı yardımı ile elde

ettiğimiz küçük parçaları bir “puzzle”ın parçalarını birleştiriyor gibi birleştirerek resmin tamamını elde ederiz.

NGS’nin mantığı bu örneğe çok benzemektedir; kesintisiz olarak dizileyemediğimiz uzun bir DNA parçasını dizileyebileceğimiz boyutlarda küçük parçalara bölüp, parçaları ayrı ayrı dizileyip, daha sonra bir yazılım yardımı ile dizilenen parçaları birleştirmek.

Elde edilen kısa okumaların birleştirilme aşamasında iki farklı yaklaşım söz konusudur: şayet dizilenen bölge için herhangi bir kalıp bilgisi mevcut ise, diziler bu kalıp dizi üzerine hizalanabilir (İng. alignment); şayet dizilenen bölge için herhangi bir kalıp bilgisi mevcut değil ise veya kullanılmak istenmiyorsa, diziler kalıp bilgisi olmadan da birleştirilebilir (İng. assembly veya denovo assembly). Bu amaçlar için geliştirilmiş, farklı özelliklere sahip yazılımlar ve biyoinformatik algoritmalar mevcuttur.

Kalıp bilgisi olmayan DNA’nın dizilenmesi için yeni dizileme (İng. denovo sequencing), kalıp bilgisi olan DNA’nın dizilenmesi için ise yeniden dizileme (İng. resequencing) terimleri kullanılır.

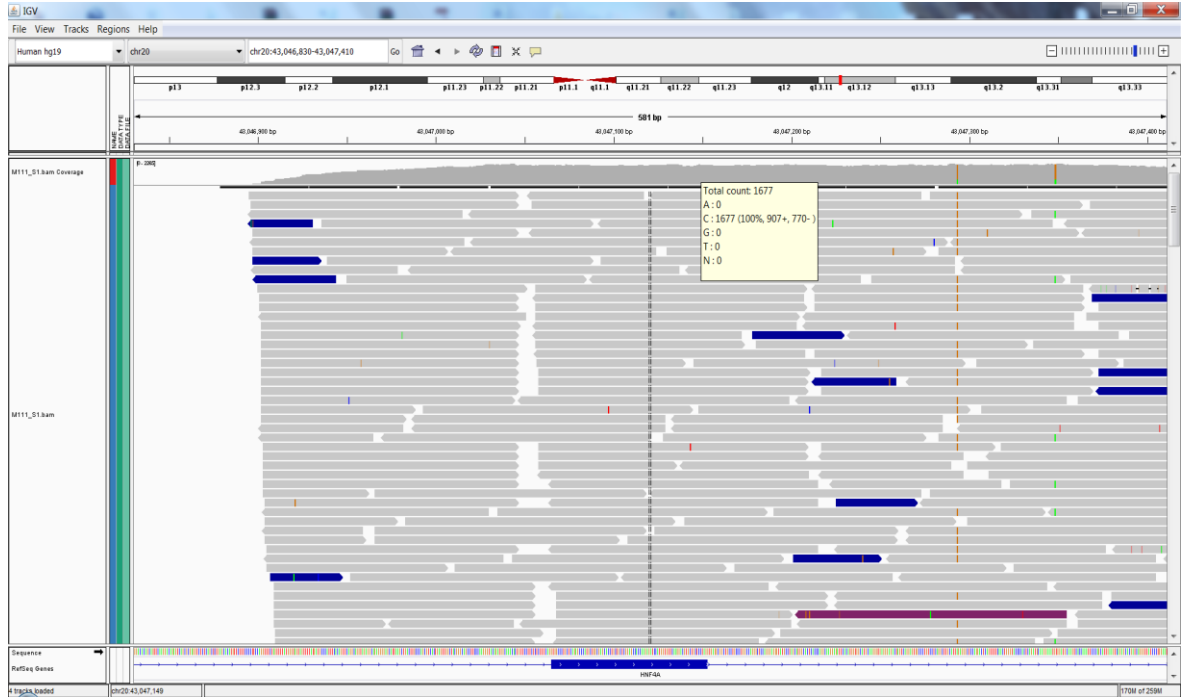
HİZALAMA

İngilizce “alignment” olarak adlandırılır. Elde edilen dizilerin, kalıp bir dizi üzerine en uygun olarak yerleştiği yerin bulunarak o noktalara hizalanmasıdır. Değişik amaçlar için değişik algoritmalar ve algoritmalara ait değişik ayarlar kullanılabilir. En yaygın olarak kullanılan hizalama araçları BWA (Burrows-Wheeler Aligner) ve Bowtie 2 araçlarıdır (89, 90).

IGV

IGV, veya Integrative Genomics Viewer, Broad Institute tarafından, büyük boyutlu genomik veri dosyalarının standart masaüstü bilgisayarlarda da görüntülenmesinin sağlanması amacıyla geliştirilmiş bir yazılımdır (91). Tamamen ücretsiz ve açık kaynak bir yazılım olduğu için zaman içinde hem araştırmacılar arasında çok popüler bir yazılım haline gelmiş, hem de açık kaynak olması sebebi ile dünyanın birçok yerinden araştırmacıların katkısı ile iyileştirilmiş ve halen iyileştirilmeye devam etmektedir (92). Hizalama sonrası oluşturulan hizalama dosyaları (sam, bam veya cram dosyaları) ile birlikte çok sayıda başka dosya

formatı, IGV sayesinde çok kolay ve ücretsiz bir şekilde görüntülenebilmektedir (Şekil 2.22.).



Şekil 2.22. IGV yazılımında bam dosyalarının görüntülenmesi

BİRLEŞTİRME

İngilizce “assembly” olarak adlandırılır. Elde edilen dizilerin, herhangi bir kalıp dizi bilgisi bulunmadan, sadece ortak dizileri yardımı ile birleştirilerek daha büyük kesintisiz dizilelerin oluşturulmasıdır. Değişik veri setlerinde, değişik amaçlar için kullanılmak üzere çok sayıda açık kaynak ve ticari birleştirme aracı bulunmaktadır (93).

Birleştirme sonucunda oluşturulan kesintisiz dizilere İngilizce bitişik anlamına gelen “contiguous” kelimesinden türetilmiş kontig (İng. contig) adı verilir. Kontiglerin kesintili olarak birleştirilmiş (iki kontig arasındaki boşluğun uzunluğunun yaklaşık olarak bilindiği) hallerine de İngilizce “scaffold” adı verilmektedir.

2.1.2.2.6. YENİ NESİL DNA DİZİLEMENİN AVANTAJLARI

NGS’nin en büyük avantajı, yüksek hacimli paralel dizileme özelliği sebebi ile yüksek miktardaki dizileme bilgisini, hızlı ve düşük maliyet ile üretebilmesidir (94). Bu da özellikle yüksek veri gereksinimi duyan uygulama alanları için çok büyük bir avantajdır; Sanger

dizileme ile sadece bir hayal olan uygulamaları günümüzde mümkün hale getirmiştir. 2000’li yılların başında tamamlanan ve tamamı Sanger dizileme ile yapılan İnsan Genom Projesi, 10 yıldan fazla sürmüş ve yaklaşık olarak 2,7 milyar dolara mal olmuştu (95). İnsan Genom Projesi’nde yapılan şey için sadece bir bireyin genomunun dizilenmesi diyebiliriz. Üzerinden geçen yaklaşık 20 yıldan sonra, günümüzde bir insanın genomunun dizilenmesi 10 yıldan fazla süreden birkaç güne, milyarlarca dolardan da birkaç yüz dolara düşmüştür. Bu muazzam düşüşü NGS teknolojilerinin geliştirilmesine borçluyuz. Kısacası NGS’nin düşük maliyetli dizileme verisi üretebilmesi genetik alanında yeni bir çağın başlangıcı olmuştur denebilir.

NGS aynı zamanda klonal amplifikasyon özelliği sayesinde Sanger dizilemeye göre çok daha hassastır (96). Sanger dizileme ile ayrıca klonlama yaparak gerçekleştirilebilecek çalışmalar, NGS ile klonlama gereksinimi olmadan yapılabilmektedir. Bu da yüksek hassasiyet gerektiren çalışmalar için fazladan kolaylık ve maliyet düşüşünü beraberinde getirmektedir.

2.1.2.2.7. YENİ NESİL DNA DİZİLEMENİN DEZAVANTAJLARI

NGS’nin yüksek miktarda veri üretme kapasitesi, bazı durumlarda dezavantaj haline gelebilmektedir. NGS yüksek miktardaki dizileme bilgisini, hızlı ve düşük maliyet ile üretebilmektedir ancak çalışma başına maliyet çok daha yüksektir ve her bir çalışma daha uzun sürede tamamlanmaktadır. Çalışmada elde edilecek verinin -paralel dizileme mantığı ile- tek seferde tamamı elde edilmek zorundadır, parça parça elde edilemez. Bu sebeplerden dolayı küçük projelerde ve hızlı sonuç alınması gerektiği durumlarda Sanger dizileme daha avantajlı hale gelebilmektedir.

Son birkaç seneye kadar NGS cihazlarının maliyetlerinin, Sanger cihazlarına göre daha pahalı olması NGS için bir dezavantaj teşkil ediyordu; ancak son yıllarda üretilen daha düşük kapasiteli “benchtop” NGS cihazları sayesinde günümüzde küçük laboratuvarlar da Sanger cihazlarına göre çok daha uygun fiyatlara NGS cihazlarını alabilmektedirler. Böylece cihaz yatırım maliyeti, NGS için bir dezavantaj olmaktan çıkmıştır.

NGS’nin Sanger dizilemeye olan bir dezavantajı da enformatiğe olan ihtiyacının daha yüksek olmasıdır. Özellikle daha yüksek kapasiteli NGS cihazlarına sahip olan

laboratuvarların, yüksek veri çıktısı sebebi ile hem işlem ve depolama gücü daha yüksek bilgisayarlara, hem de veriyi bu bilgisayarlarda işleyebilecek deneyimli biyoinformatik uzmanlarına ihtiyaçları vardır.

NGS'nin bir diğer dezavantajı da standart olmamasıdır. Sanger dizileme çok standart bir yöntemdir, kullanılan kit ve cihazlarda ufak farklılıklar olsa da temel yöntem aynıdır; dünyanın diğer ucunda bir araştırmacının sadece yaptığı çalışmada Sanger dizilemeyi kullandığını söylemesi bile, dünyanın diğer ucundaki araştırmacının yapılan çalışmayı -ufak detaylar hariç- anlaması için yeterlidir. Ancak NGS'de böyle değildir; çok sayıda farklı prensiplere ve avantaj/dezavantajlara sahip olan cihazlar, kitler, analiz algoritmaları, yazılımlar vb. vardır (34). NGS'nin Sanger dizilemeye göre çok büyük avantajlarının olmasına rağmen, çoğu araştırmacının Sanger dizilemeyi halen DNA dizilemenin “altın standardı” olarak görmesinin sebebi de Sanger in “altın” olma özelliğinden ziyade “standart” olma özelliğinden gelmektedir ve bu konuda haksız da değildir. Yine de “altın standart” olarak nitelendirilme özelliği, yapılan birtakım çalışmaların ışığında zaman içerisinde yavaş yavaş Sanger dizilemeden NGS'ye doğru kaymış ve halen kaymaya devam etmektedir (97-99).

2.1.2.3. 3. NESİL DİZİLEME

3. nesil dizileme teknolojileri, NGS teknolojilerinden daha sonra geliştirilen ve bazı özellikleri ile NGS'den ayrılan DNA dizileme teknolojileridir. Bu teknolojilerde bulunması gereken/beklenen özellikler şu şekilde özetlenebilir (100-104):

- Amplifikasyon olmadan tek molekülden dizileme yapılması
- Nükleotid analoglarının kullanılmaması
- Polimeraz, ligaz vb. DNA'yı modifiye eden enzimlerin kullanılmaması
- Gerçek zamanlı dizileme ve analiz yapılması (döngü bazlı olmaması)
- Uzun okuma yapabilmesi
- Düşük hata oranına sahip olması
- Hızlı olması
- Epigenetik modifikasyonların tespit edilebilmesi
- AT-GC sapmasının olmaması
- Yüksek miktarda dizileme verisinin düşük maliyetli üretilmesi

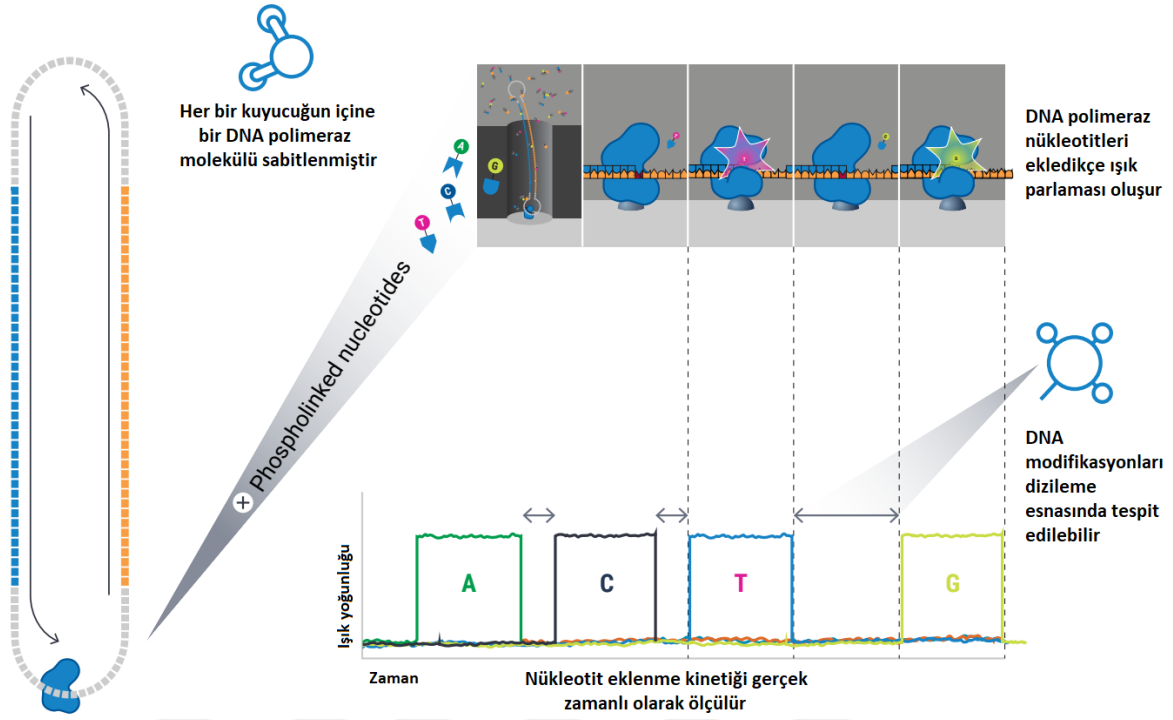
Günümüzde bu özelliklerin tamamına sahip olan bir teknoloji bulunmamaktadır; 3. nesil olduğu öne sürülen mevcut teknolojilerin tamamı bunların sadece bir kısmına sahiptir. “Bu özelliklerden hangilerine sahip olan 3. nesil olarak kabul edilebilir?” sorusunun cevabı hakkında akademik çevrelerde bir uzlaşma da yoktur; bazı sistemler bazı bilim insanları tarafından 3. nesil olarak nitelendirilirken aynı sistemler bazı bilim insanlarınca NGS olarak nitelendirilmektedir. Hatta bu özelliklerden bazılarına sahip sistemler için 3. nesil, daha fazlasına sahip olanlar için 4. nesil tanımlaması yapanlar da vardır. 3. nesil algısı zaman içinde de değişmiş gibi görünmektedir; şu an herkes tarafından NGS olarak kabul edilen bazı teknolojiler ilk çıktıkları zaman bazı bilim insanları tarafından 3. nesil olarak ta nitelendirilebilmiştir. Kısacası 3. nesil DNA dizileme teknolojileri hakkında tam bir uzlaşma yoktur; kişisel ve dönemsel yorumlara göre değişik göstermektedir.

3. nesil DNA dizileme teknolojilerinin bir avantajı da epigenetik modifikasyonların en önemlisi olan DNA metilasyonunu, bisülfid çevrimine ihtiyaç duymadan direkt dizileme esnasında tespit edebilmesidir (105).

Günümüzde genel olarak 3. nesil olarak kabul gören iki teknoloji vardır: Pacific Bioscience firmasının SMRT (Single Molecule Real Time) sekanslama teknolojisi ve Oxford Nanopore firmasının nanopor sekanslama teknolojisi.

2.1.2.3.1. SMRT SEKANSLAMA

SMRT, İngilizce “Gerçek Zamanlı Tek Molekül” anlamına gelen “Single Molecule Real Time” ifadesinin baş harflerinden oluşur. Bu teknolojide adaptörler takılan, tek zincire indirgenen ve halka yapısı kazandırılan tek DNA molekülleri üzerinden, yaklaşık 70 nm çapında ve 100 nm derinliğindeki kuyucuklar içerisinde ve kuyular içine sabitlenmiş olan tek bir DNA polimeraz enzimi molekülü sayesinde DNA polimerizasyonu gerçekleştirilir. Kullanılan özel tasarlanmış ve her bir nükleotid için farklı bir floresan molekül içeren nükleotid analogları, uzayan zincire eklendikçe floresan moleküllerini bırakırlar ve floresan ışımalara sebep olurlar. Bu ışımların özel bir kamera sistemi ile gerçek zamanlı takibi ile de dizileme gerçekleştirilmiş olunur (106) (Şekil 2.23.).



Şekil 2.23. SMRT sekanslama (107)

SMRT sekanslamanın en önemli özelliklerinden biri klasik bir NGS sistemine göre çok daha uzun okumalar yapabilmesidir; ortalama okuma uzunluğu 10-25 Kbp, maksimum okuma uzunluğu ise 175 Kbp'ye kadar çıkabilmektedir. Bunu, tek molekülden ve döngü bazlı olmadan dizileme yapmasına borçludur (fazlama problemi yok). Bu uzunluklar ortalama 150 bp civarında okuma uzunluğu olan NGS sistemlerine göre inanılmaz bir uzunluktur; genom boyutunda kısa okumalar ile çözülmesi mümkün olmayan birçok sorun bu şekilde çözülebilmektedir (insan genomunu telomerden telomere kesintisiz dizileyebilmek gibi). Dezavantajı ise hata oranının yüksek olması (yaklaşık %10) ve baz başına dizileme maliyetinin yüksek olmasıdır. Yüksek hata oranı problemi, halkasal olan DNA'nın döndürülerek tek seferde birden fazla dizlenmesi ile ortadan kaldırılabilir; bu şekilde hata oranı %1'in altına düşerek rakip NGS sistemleri ile benzer düzeye inmektedir (bu durumda okuma uzunlukları kısalmakta ve 10-20 Kbp'ye düşmektedir, ama bu büyük bir dezavantaj sayılmaz, çünkü 10-20 Kbp'de gayet uzun bir okumadır). Bu sebeple bu teknolojinin en büyük dezavantajı dizileme maliyetidir; baz başına dizileme maliyeti rakip NGS sistemlerine göre 3-5 kat daha fazladır.

SMRT sekanslama yukarıda belirtilen 3. nesil dizileme teknolojilerinin özelliklerinden

- Amplifikasyon olmadan tek molekülden dizileme yapılması
- Gerçek zamanlı dizileme ve analiz yapılması (döngü bazlı olmaması)
- Uzun okuma yapılabilmesi
- Hızlı olması
- Epigenetik modifikasyonların tespit edilebilmesi

özelliklerine uyarken,

- Nükleotid analoglarının kullanılmaması
- Polimeraz, ligaz vb. DNA'yı modifiye eden enzimlerin kullanılmaması
- Yüksek miktarda dizileme verisinin düşük maliyetli üretilebilmesi
- AT-GC sapmasının olmaması

özelliklerine uymamaktadır. Düşük hata oranına sahip olması özelliğine de kısmi olarak uyduğu söylenebilir.

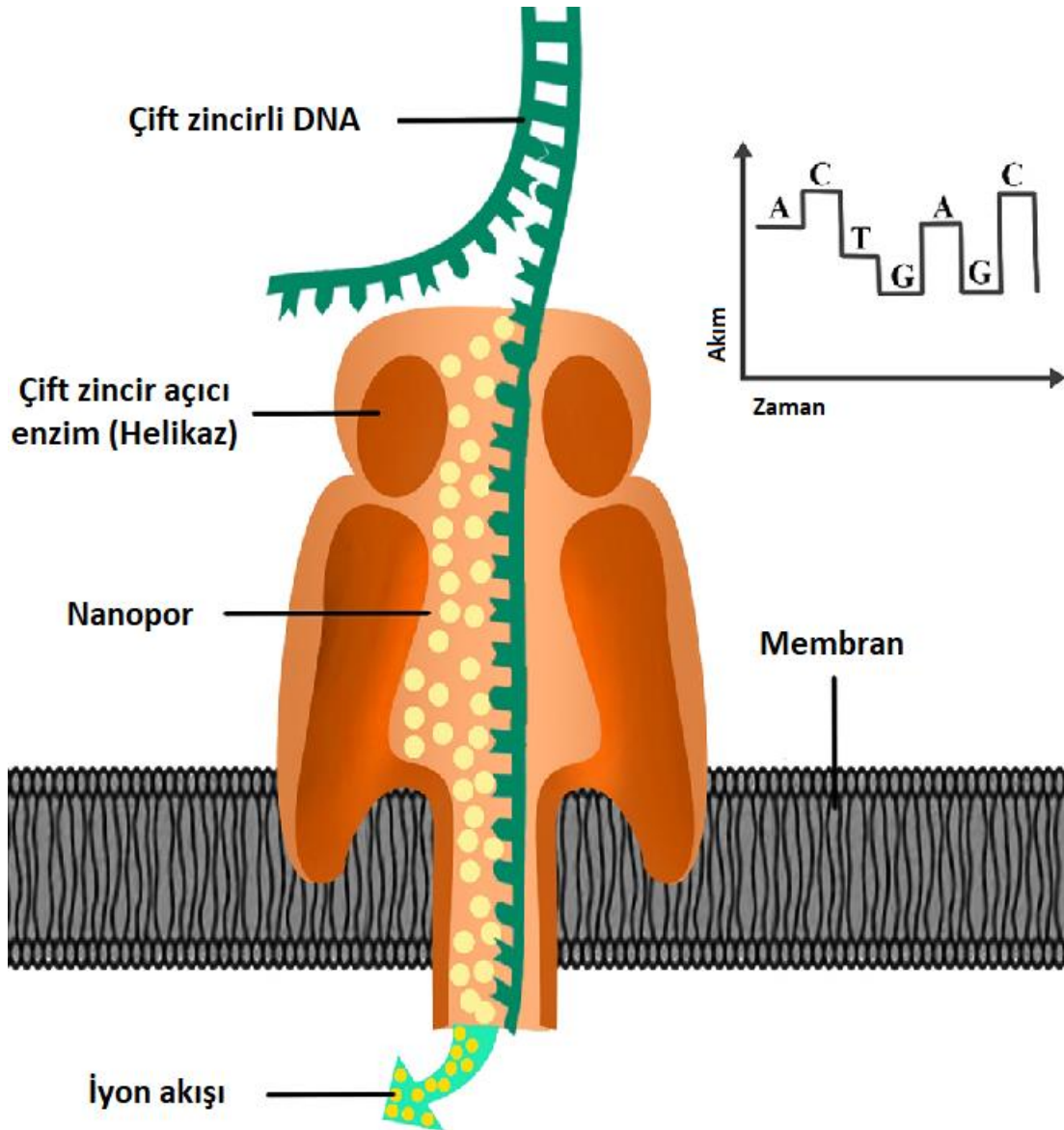
2.1.2.3.2. NANOPOR SEKANSLAMA

Nanopor sekanslama, DNA'nın bir membran üzerine sabitlenmiş biyolojik veya kimyasal bir nanopor içerisinden geçirilerek, geçirilme esnasında oluşan fiziksel değişikliklerin gerçek zamanlı tespiti ile gerçekleştirilen DNA dizilemedir. Teorik temelleri 1990'lı yılların başında atılmış, ilk deneysel çalışmaları 1990'lı yılların ortalarında başlamış, 2000'li yıllarda iyice geliştirilerek 2010'lu yılların başında bir firma tarafından ticari bir ürüne dönüşebilmiştir (108-110).

Nanopor sekanslamanın ticari bir ürün olarak piyasaya çıkması ve çok sayıda başka firmanın nanopor bazlı başka teknolojiler üzerine çalışmaya ve yol almaya başlaması, 3. nesil dizilmenin geleceğinin nanopor üzerine olduğu ve kısa sürede piyasaya çıkacak ve geliştirilecek başka nanopor teknolojileri ile birlikte kısa sürede nanopor sekanslamanın piyasaya hükmedeceği, NGS sistemlerinin yerini alacağı (aynen NGS'in Sanger dizilemeye yaptığı gibi) şeklinde bir beklenti oluşmuştur. Ancak üzerinden geçen yaklaşık 10 yıla rağmen ne ticari ürüne dönüşebilen başka bir nanopor teknolojisini/firması çıkabilmiş, ne de mevcut teknoloji gerektiği kadar geliştirilebilmiştir (NGS'nin ilk çıktığı 2005 yılından 2015

yılına kadar 10 yıl içinde ne kadar büyük bir ilerleme kaydettiği düşünülecek olursa, nanopor teknolojisi için 10 yılda neredeyse hiç gelişmemiş bile denilebilir).

Mevcut nanopor sekanslama teknolojisinde nanoporlardan geçen tek zincir DNA, nanopor içindeki iyon akışını yavaşlatır. Bazıların 3 boyutlu yapıları birbirinden farklı olduğu için, farklı bazlar iyon akışını farklı bir şekilde etkiler. Her bir nanopordaki iyon akışının gerçek zamanlı olarak görüntülenmesi sonucunda herhangi bir nanopordan geçen tek DNA molekülünün dizilimi gerçek zamanlı olarak ortaya çıkartılmış olunur (111) (Şekil 2.24.).



Şekil 2.24. Nanopor sekanslama (112)

Bu teknolojinin günümüzdeki en büyük dezavantajı çok yüksek olan hata oranıdır; hata oranı %10 civarındadır (113). Bu da klinik insan genetiği amaçlı kullanımını neredeyse imkansız hale getirmektedir. Mevcut NGS sistemlerinde bu oran %1'in altındadır (114).

Nanopor sekanslama yukarıda belirtilen 3. nesil dizileme teknolojilerinin özelliklerinden

- Amplifikasyon olmadan tek molekülden dizileme yapılması
- Gerçek zamanlı dizileme ve analiz yapılması (döngü bazlı olmaması)
- Uzun okuma yapılabilmesi
- Hızlı olması
- Epigenetik modifikasyonların tespit edilebilmesi
- Nükleotid analoglarının kullanılmaması
- Polimeraz, ligaz vb. DNA'yı modifiye eden enzimlerin kullanılmaması
- AT-GC sapmasının olmaması

özelliklerine uyarken,

- Yüksek miktarda dizileme verisinin düşük maliyetli üretilebilmesi
- Düşük hata oranına sahip olması

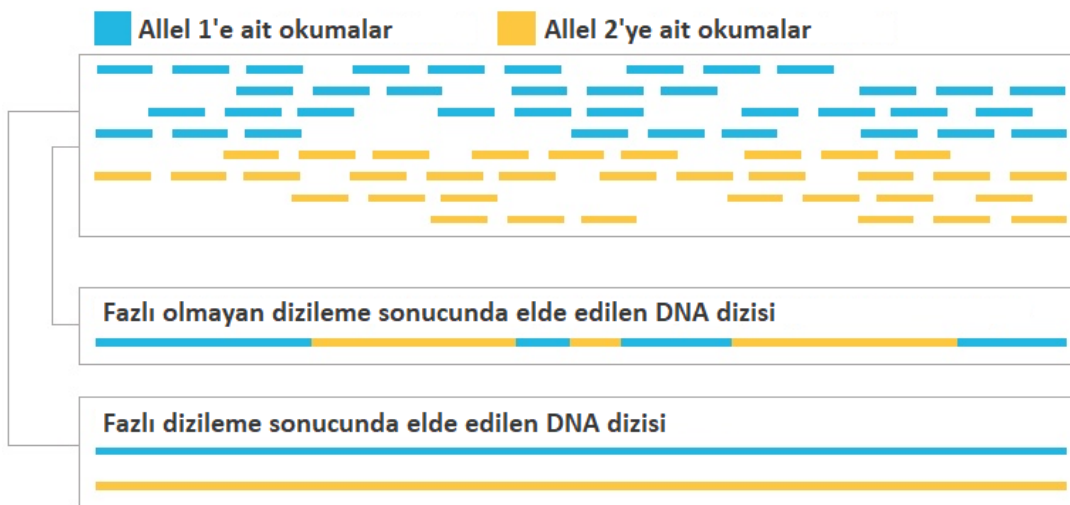
özelliklerine uymamaktadır. Görüleceği üzere nanopor sekanslama, SMRT sekanslamaya göre daha fazla 3. nesil dizileme özelliğine sahiptir. Bu fark sebebi ile bazı araştırmacılar tarafından günümüzdeki tek 3. nesil dizileme teknolojisi, bazı araştırmacılar tarafından da 4. nesil dizileme teknolojisi olarak tanımlanmaktadır.

2.1.3. FAZLI DİZİLEME YAKLAŞIMLARI

Fazlı dizileme, poliploid genoma sahip organizmalardan elde edilen DNA dizileme verisinden elde edilen birden fazla varyasyonun aynı DNA zinciri üzerinde olup olmadığının tespit edildiği DNA dizileme yaklaşımıdır (115). İnsan gibi diploid bir organizmada genler çok büyük oranda çiftler halinde bulunur; her bir ebeveynden alınan homolog kromozomlar üzerinde aynı lokusta bulunan aynı genin birer kopyası şeklinde. DNA dizileme sırasında bu iki kopya gen (veya homolog kromozomların aynı lokusları) beraber dizilenir. Dizilenen bölgede şayet birden fazla heterozigot varyasyon tespit edilirse, bu varyasyonların aynı DNA zinciri üzerinde mi, farklı DNA zincirleri üzerinde mi olduğu; diğer bir tabir ile aynı

ebeveyninden mi, farklı ebeveynlerden mi geldikleri bilgisi çoğu durumda çok önemli bir bilgidir. Bu bilgiyi de içeren DNA dizilemeye fazlı dizileme denir (115). Kısacası fazlı dizilemede homolog kromozomlar üzerindeki diziler beraber değil, ayrı ayrı okunur (Şekil 2.25.).

Sanger dizileme yönteminde, tek molekül dizileme gerçekleştirilmediği için dizileme verisi içinden fazlama yapmak mümkün değildir; fazlama ancak bireyin genetik materyalini aldığı ebeveynleri de dizilenerek (116) veya dizilenecek bölgeye ait alleller klonlama ile birbirinden ayrılarak yapılabilir (115). NGS’de ise, tek molekül dizilendiği için fazlı dizileme yapmak mümkündür ancak okuma uzunluğu ile sınırlıdır; okuma uzunluğu kısa olan teknolojiler için bu kısıtlayıcı bir durumdur (117). Bu kısıtlamayı aşmak için değişik yaklaşımlar geliştirilmiştir. Bunlardan birisi mate-pair kütüphaneler kullanmaktır. Mate-pair kütüphanelerde örnek hazırlama aşamasında birkaç kilobazdan oluşan DNA parçalarının iki ucu birleştirilerek, birbirinde belirli bir uzaklıktaki iki bölgeden okuma elde edilir; böylece fazlama için daha uzak bölgelerdeki varyasyonlar da kullanılabilir hale getirilmiş olunur (118). Bu şekilde hazırlanan kütüphaneler aynı zamanda birleştirme ile analiz edilecek dizileme verisinde, dizilenen bölgedeki tekrar bölgeleri gibi problemli bölgelerin daha iyi çözülmesinde de yardımcı olur (118). Bir diğer yaklaşım ise dizilenecek DNA molekülüne dizisi bilinen DNA parçalarının (etiketlerin) rastgele eklenmesidir. Bu şekilde hazırlanan kütüphanelerde etiketler yazılım tarafından tanınarak hem analiz sırasında dizi içerisinde çıkarılır, hem de elde edilen okumanın hangi DNA klonundan geldiği belirlenerek fazlama yapılabilir (119).



Şekil 2.25. Fazlı dizileme (120)

2.1.4. DNA DİZİLEME İLE NADİR VARYANTLARIN TESPİTİ

Derin dizileme, veya daha doğru tabir ile derin DNA dizileme, NGS yönteminin klonal amplifikasyon özelliği sayesinde, belirli bir DNA parçası üzerinden tek seferde çok sayıda klonal DNA dizileme verisi alınarak, o bölgenin çok daha hassas dizilendiği NGS yaklaşımıdır (49, 121, 122).

2.1.4.1. DERİN DİZİLEME

Okuma derinliği (İng. depth of sequence), NGS teknolojilerinin klonal amplifikasyon özelliğinden kaynaklanan, hassasiyeti arttıran ve Sanger yönteminde uygulaması olmayan bir özelliğidir. Okuma derinliği, verinin hizalama veya birleştirme sonrasında, herhangi bir noktası için kaç farklı DNA parçasında ayrı ayrı okunduğudur. Derinlik kavramının dışında bir de kapsam (İng. coverage) kavramı vardır. Derinlik ve kapsam birbirine yakın kavramlar olsalar ve birçok durumda birbirlerinin yerine kullanılabilir olsalar da tamamen aynı değildir; kapsam tek bir noktada ziyade bir bölge için kullanılır ve o bölgedeki ortalama derinliği ifade eder (49, 123). İnsan gibi diploid bir organizmanın dizilenmesinde herhangi bir nokta için 20 veya daha büyük bir okuma derinliği, kalıtsal varyasyonların ve zigoitelerinin tespiti için çoğu durumda yeterlidir (123). Hatta düşük döngü sayılı PCR içeren örnek hazırlama prosedürleri ile yapılan çalışmalar için, 10-12 civarında okuma derinlikleri bile duruma göre kabul edilebilir seviyelerdedir (124). Herhangi bir nokta için okuma derinliği ne kadar fazla ise, elde edilen sonuç o nokta için o kadar fazla güvenilirdir denilebilir. Ancak belirli bir derinlikten sonra okuma derinliğini daha da arttırmak güvenilirliği daha da arttırmaz; çünkü sonuç zaten gereğinden çok daha fazla güvenilirdir. Bu tip durumlarda okuma derinliğini arttırmak ise hassasiyeti artırarak, somatik varyasyonlar gibi dizilenen örnek içerisinde daha nadir bulunabilecek varyasyonların tespitinde kullanılabilir (125). Bu şekilde derinliğin normal ihtiyacın üzerinde arttırılarak, nadir varyasyonların tespiti için daha hassas bir dizileme yapılmasına derin dizileme denir. Aynı zamanda okuma derinliğinin bunun da üzerine çıkarılmasına ultra-derin dizileme de denilmektedir. Her ne kadar kesin bir tanımı olmasa da kabaca okuma derinliğinin 100'ün üzerinde olduğu dizilemeler için derin dizileme, okuma derinliğinin 1000'in üzerinde olduğu dizilemeler için de ultra-derin dizileme denilebilir (49).

2.2. HAVUZLAMA YAKLAŞIMI

Havuzlama, örnek havuzlama veya havuz testi, birden fazla örneğin birleştirilerek tek bir örnekmiş gibi test edilmesi yaklaşımına verilen genel bir isimdir. Özellikle mikrobiyal testlerde, patojen tespit alanında sıklıkla kullanılmaktadır. Doğru kullanıldığı zaman hem test maliyetlerini hem de test sürelerini belirgin bir şekilde düşürebilmektedir. Tarihsel olarak ilk defa 1943 yılında 2. Dünya Savaşı'na katılan Amerikan askerlerinde frengi taramasında kullanılmıştır (126).

Havuzlama yaklaşımının başarılı bir şekilde kullanılabilmesi için iki önemli durumun mevcut olması gerekmektedir; birincisi kullanılan testin, örnekler birleştirildikten sonra dahi verimli bir şekilde tespit yapabilmesi için hassas bir test olması gerekmektedir; ikincisi ise aranılan parametrenin aranılan grup içinde nadir bulunuyor olması gerekmektedir (127). Havuzlama yaklaşımında örneklerin testin hassasiyeti ve beklenen pozitiflik oranı gözetilerek, belirli sayılarda birleştirilmesi ile örnek havuzları oluşturulur, havuzlar aranan parametre açısından test edilir, negatif çıkan havuzların içindeki tüm örneklerin negatif olduğu kabul edilir, pozitif çıkan havuzların içindeki örnekler ise teker teker test edilerek pozitif olan örnek ya da örnekler tespit edilir (126). Havuzlar ne kadar çok örnekten oluşursa, havuzlamanın getirdiği avantajlar o kadar fazla olur. Ancak havuzları oluşturacak örnek sayısının bir sınırı vardır, birleştirilen örnek sayısı testin hassasiyetini etkileyecek ve yanlış negatif sonuç verebilecek kadar çok olmaması gerekmektedir. Kullanılan test çok hassas olsa ve hassasiyet kaynaklı pratikte bir sınır olmasa bile, havuzları oluşturacak örnek sayısının beklenen pozitiflik oranına göre ayarlanması gerekir; şayet havuzların içerisindeki örnek sayısı çok fazla tutulursa bu sefer de havuzların tamamına veya büyük kısmına pozitif örnek düşeceği için, havuzlama avantajı değil tam tersine dezavantajı hale gelmiş olacaktır (128). Kısacası havuzlama, çok sayıda örneğin toplu halde negatif olarak değerlendirilmesi sonucunda yapılan toplam test sayısının, bunun sonucunda da toplam maliyet ve test sürelerinin azaltılması prensibine dayanmaktadır (126).

2.2.1. DNA DİZİLEME VE HAVUZLAMA YAKLAŞIMI

Havuzlama tekniğini kullanarak yapılan DNA analizi, 1990'lı yıllarda nakil için toplanan kanlardaki enfeksiyon ajanlarının kolay ve uygun maliyetler ile taranması ile rutin olarak kullanılmaya başlanmıştır (129, 130). Havuzlama yaklaşımının DNA dizilemede ilk

kullanımı ise 2000 yılında yayınlanan bir mini-dizileme çalışmasıdır; bu çalışma ile ATP7B geninde tanımlanmış 10 patojenik mutasyonun toplumdaki frekansları, 2640 örneğin havuzlama ile analiz edilmesi ile belirlenmeye çalışılmıştır (131). NGS teknolojilerinin yaygınlaşması ile birlikte 2009 yılından sonra konu ile ilgili yayınlar çıkmaya başlamıştır. Günümüze kadar 30'dan fazla yayın çıkmıştır. Bunların bir kısmı havuzlama yaklaşımında havuzların oluşturulması veya elde edilen verini analizi ile ilgili matematiksel veya bilgisayar modelleme üzerine yapılmış biyoinformatik çalışmalardır (132-143). Aynı zamanda gerçek veri üzerinden yapılmış biyoinformatik odaklı çalışmalar da vardır (144-146). Literatür de bir tane de havuzlanmış örneklerden varyant tespiti için geliştirilmiş bir yazılıma ait yayın da bulunmaktadır (147). Havuzlanmış örneklerden üretilmiş gerçek NGS verisi üzerinden varyasyon tespiti amaçlı çalışmaların yaklaşık yarısı ise varyasyon keşfi amaçlı olup, varyasyonların bireyde verimli bir şekilde tespiti amaçlı genotipleme çalışmaları değildir (148-158). Literatürde karşılaşılan genotipleme amaçlı çalışmalar da vardır, ancak bunların tamamı 5-20'li havuzlama gibi göreceli olarak güvenli havuzlar üzerinde ve bilinmeyen örneklerden yapılmış genotipleme çalışmalarıdır (131, 159-167). Bu çalışmalar içinde 2 tanesi 2 boyutlu havuzlama tekniğini kullanan çalışmalardır (161, 162). Daha önce genotiplenmiş örnekler ile yapılmış validasyon çalışmasına ise literatürde sadece 1 tane rastlanmıştır, o da tek boyutlu ve tek bir havuz çeşidi içeren ve sadece 20 örnekten oluşan bir çalışmadır (168).

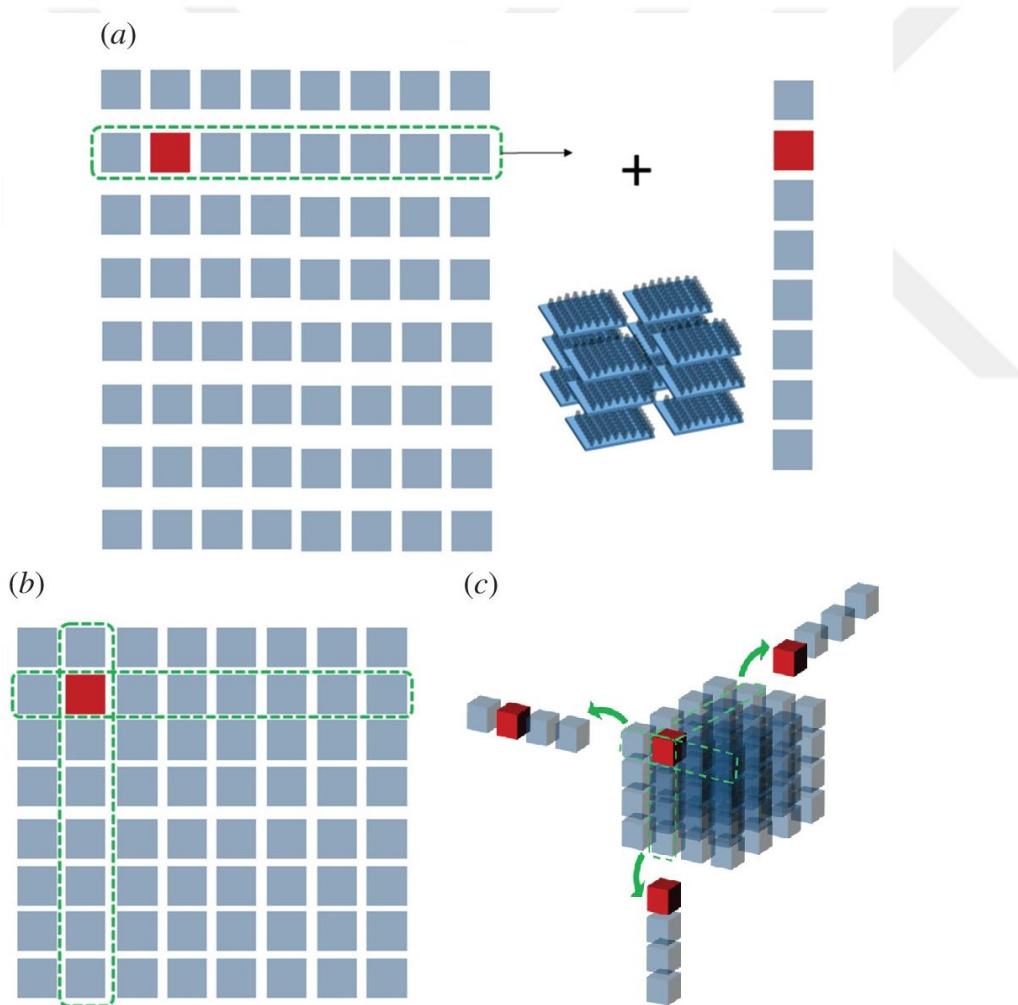
Tezin validasyon kısmı, literatürde 2 ve 3 boyutlu havuzlama yaklaşımının genotipi bilinen çok sayıda örnek üzerinden ve farklı sayılarda örnekten oluşan havuzlardan aynı çalışma içinde denendiği ilk çalışma olacaktır.

2.2.2. ÇOK BOYUTLU HAVUZLAMA YAKLAŞIMI

Kullanılan havuzlama yaklaşımları çoğunlukla tek boyutludur, yani örnekler gruplara ayrılarak karıştırılıp havuzları oluştururlar. Bu durumda pozitif örneklerin belirlenebilmesi için pozitif çıkan havuzların tamamındaki örneklerin teker teker çalışmaya zorunluluğu vardır (Şekil 2.26.). Ancak çok boyutlu havuzlama yapmak ta mümkündür. Mesela 2 boyutlu havuzlamada örnekler 2 boyutlu bir dizin oluşturacak şekilde dizilir ve X ve Y olmak üzere 2 boyutta havuzlama yapılır. Bunun sonucunda her örnek tek boyutta olduğu gibi bir havuz içerisinde değil, biri X ve biri Y boyutlarında olmak üzere iki adet havuz içerisinde temsil

edilir. Veya aynı şekilde 3 boyutlu havuzlama yaklaşımında örnekler X, Y ve Z boyutlarında olmak üzere üç havuz içinde temsil edilirler (169). Çok boyutlu havuzlamada pozitif çıkan havuzlar pozitif örneklerin “koordinatlarını” belirteceği için, sadece kesişim noktalarındaki örneklerin doğrulanma amaçlı çalışılması yeterlidir (Şekil 2.26.).

Çok boyutlu havuzlama yaklaşımının tek boyutlu havuzlamaya göre avantajları ve dezavantajları bulunmaktadır. Doğru kullanıldığı zaman toplam test sayısını ve maliyeti bir miktar daha düşürebilir. Aynı zamanda olası her pozitif örnek boyut sayısına göre birden fazla havuz içerisinde test edileceği için, yanlış negatiflik ihtimali de düşmektedir. Dezavantajı ise çok boyutlu havuzlamada havuzları oluşturmak, tek boyutluya göre daha zor ve karmaşıktır ve boyut sayısı arttıkça çok daha zor ve karmaşık hale gelmektedir.



Şekil 2.26. 1, 2 ve 3 boyutlu havuzlama yaklaşımları (169)

2.3. MUTASYON TARAMA

Genetik hastalıkların kesin tanısı için hastalığa sebep olan mutasyon veya mutasyonların hasta bireyde tespit edilmesi ve zigozitesinin belirlenmesi gerekmektedir. İnsanın haploid 3,2 milyar bazdan oluşan devasa bir genomu sahip olduğunu ve çoğu genetik hastalık durumunda hastalığa sebep olan mutasyonların sadece bir veya birkaç noktada olduğunu düşünecek olursak, bu mutasyonların tespit edilmesi adeta samanlıkta iğne aramaya benzetilebilir.

Mutasyonların tespiti, şu ana kadar tanımlanmış 6 bin civarında kalıtsal hastalık içerisinde göreceli olarak yaygın görülen ve yaygın mutasyonları olan çok az sayıda hastalığı, büyük delesyon ve duplikasyon gibi DNA dizileme ile tespit edilemeyen mutasyonları ve tekrar dizisi hastalıklarındaki tekrar mutasyonlarını istisna olarak kabul edersek, ancak DNA dizileme yöntemi ile mümkündür. DNA dizileme, NGS teknolojilerinin geliştirilmesi ve yaygın olarak kullanılmaya başlanması ile düşük maliyetli ve pratik bir yöntemle dönüşmüştür. Ancak geçmişte bu böyle değildi; özellikle büyük bölgelerin DNA dizileme çalışmaları hem maliyetli hem de zahmetli çalışmalar idi. Bu sebep ile, DNA dizilemesinden önce kullanılarak mutasyonların bulunabileceği bölgelerin tespiti amaçlı, DNA dizilemeye göre daha düşük maliyetli ve daha pratik yöntemler geliştirilmiştir. Çoğu, mutasyon içeren DNA zinciri ile mutasyon içermeyen DNA zincirinin bir araya gelerek oluşturduğu heterodupleks DNA parçalarının tespiti esası üzerine kurulu bu yöntemler arasında HRM, DGGE, DHPLC ve SSCP'yi sayabiliriz. Dizilenmek istenen bölge ilk önce bu yöntemler ile taranıyor, tarama sonunda pozitif çıkan bölgeler dizilenerek varyasyon tespit ediliyor, bu şekilde yapılan dizileme miktarı ile birlikte maliyet de azaltılıyordu. Ancak zaman içinde DNA dizilemenin maliyetinin dramatik bir şekilde düşmesi ile birlikte artık bu yöntemleri mutasyon tarama amaçlı kullanmak anlamsız hale gelmiştir; günümüzde tarama maliyetinin çok daha altına dizilemek mümkündür. Yine de hasta bireylerin dışında, taşıyıcıların veya hastalık belirtilerinin daha ortaya çıkmadığı ama erken tanının tedavide çok büyük öneme sahip olduğu genetik hastalıklara sahip bireylerin taranması amaçlı olarak DNA dizilemeyi kullanmak, günümüzde hala çok maliyetli bir çalışma olarak karşımıza çıkmaktadır. Bu sebeple günümüzde bu amaç için yapılan tarama çalışmalarında, hassasiyeti ve özgünlüğü DNA dizilemeye göre daha düşük olan, ancak çok daha uygun maliyetli yöntemler kullanılmaktadır.

Ülkemizde taranan kalıtsal hastalıklar SMA, beta talasemi, fenilketonüri, konjenital hipotiroidi, biyotidinaz, kistik fibrozis ve konjenital adrenal hiperplazi hastalıklarıdır (3). Bu hastalıklardan SMA ve beta talasemi evlilik öncesi yetişkinlerden ve taşıyıcıların tespiti amacı ile yapılırken, diğer hastalıklar yenidoğanlardan ve hasta bireylerin erken teşhis alması amacı ile yapılmaktadır. Bu hastalıkların hepsi genetik hastalıklardır, ancak SMA dışındaki hastalıkların tamamının taraması moleküler genetik yöntemler dışındaki yöntemler ile yapılmaktadır. Moleküler genetik yöntemlerin hassasiyetinin çok daha yüksek olmasına ve kesin tanı için mutlaka gerekmelerine rağmen, pahalı oldukları için tarama programlarında kullanılamamaktadırlar.

Şayet düşük maliyetli ve DNA dizileme temelli, yüksek hassasiyet ve özgüllüğe sahip bir tarama yöntemi geliştirilebilirse bu hastalıklar dahil olmak üzere çok sayıda başka genetik hastalık ta toplumun değişik seviyelerinde verimli bir şekilde taranarak, hem doğan yeni hasta bireylerin sayısı azaltılabilir, hem de erken tanı ile mevcut hastaların tedavi süreçleri çok daha başarılı bir şekilde yönetilebilir.

2.4. FAMILİYAL HİPERKOLESTEROLEMİ

Familiyal hiperkolesterolemi (FH) (OMIM 143890), lipitlerin kanda anormal seviyede bulunması durumu olan dislipidemi durumunun en yaygın formu olan hiperkolesteroleminin bir türüdür. Hiperkolesterolemiye yol açan sebepler primer ve sekonder olmak üzere 2 gruba ayrılabilir. Primer sebepler (ailesel sebepler de denir), kolesterol veya trigliseritlerin kandaki seviyelerini ayarlayan genlerde oluşan mutasyonlar sonucu oluşan ve kalıtsal olan sebeplerdir. Sekonder sebepler ise genetik olmayan, çevresel sebepleri içerir. Familiyal hiperkolesterolemi, primer sebepler ile oluşan kalıtsal bir dislipidemi türüdür (4, 170, 171).

FH’de, mutasyonlar sonucu kolesterol ve lipit metabolizması bozulmakta ve kişide “hiperkolesterolemi, ciltte ksantomlar, arkus kornea, poliartrit ve tendosinovit ve koroner hastalığa yatkınlık” gözlenmektedir (4). Bu hastalık, tüm hiperkolesterolemili hastaların yaklaşık %5’ini oluşturmaktadır. Hastalık, toplumda en sık görülen iki genetik hastalıktan biri olarak kabul edilmektedir (172). Tüm toplumlarda görülen bu hastalığın prevalansı 1:200 ile 1:500 arasında olduğu düşünülmektedir (173). Bazı topluluklarda is “kurucu etki” nedeniyle daha yüksek oranda görülmektedir. Dünyada ise 10 milyon kişinin etkilenmiş olduğu sanılmaktadır (174).

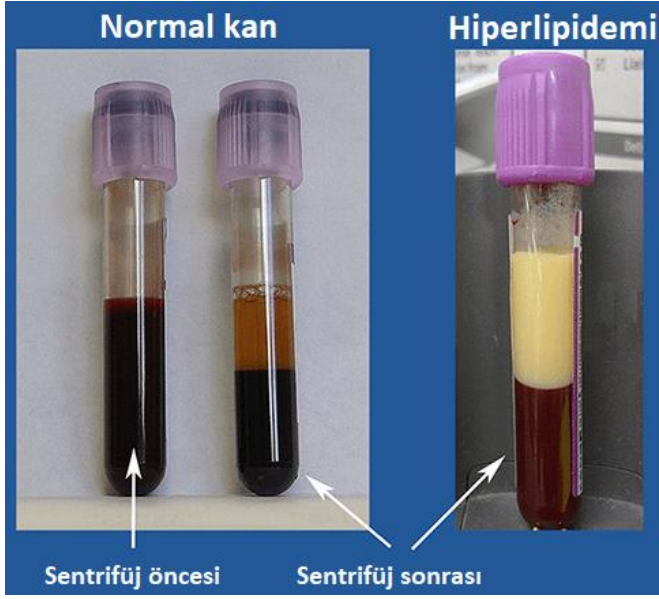
Çok önemli bir hastalık olmakla birlikte ne yazık ki olguların %90 kadarı eksik tanı almakta veya hiç tanı alamamaktadır (175). Bazı gelişmiş ülkelerde dahi hastaların küçük bir kısmı tanı alabilmektedir; Dünya Sağlık Örgütü (WHO) etkilenmiş Kanada'lıların ancak %15 kadarına tanı konulabildiğini bildirmiştir (176). Ancak hastalarda semptomsuz dönemi uzatmak, uygun tedaviyi sağlamak, hastalığın takibi ve komplikasyonları önlemek için erken tanıya gereksinim vardır. Hastaya erken tanı konulması halinde yaşam biçimi, diyet, egzersiz ve lipid düşürücü ilaçlar kullanılarak ateroskleroz gelişiminin engellenmesi ve yaşam kalitesinin artması ile beraber beklenen yaşam süresinin 25-30 yıl kadar uzaması mümkündür (177).

2.4.1. BELİRTİLER

FH'nin en önemli ve tipik belirtisi hiperkolesterolemidir; kandaki LDL kolesterol (LDL-C) seviyesi ne kadar yüksek ise, kişinin FH olma ihtimali o kadar yüksektir; belirli bir seviyenin üzerinde ise kişi kesinlikle FH olarak değerlendirilmektedir (170). Bunun dışında sekonder belirti olarak arkus kornea, tendon ksantoma ve ksantelezma da görülebilir (Şekil 2.27.). FH sebebi ile dolaşımda olan yüksek miktardaki LDL partüküllerinin hızlı bir şekilde ateroskleroz gelişimini tetiklemesi ve erken yaşlarda ortaya çıkan kardiyovasküler hastalıklara sebep olmasıdır (170, 171). FH'nin bazı uç nokta vakalarında kandaki lipid miktarı o kadar artabilmektedir ki, kanın görsel olarak rengi bile değişebilmektedir (Şekil 2.28.).



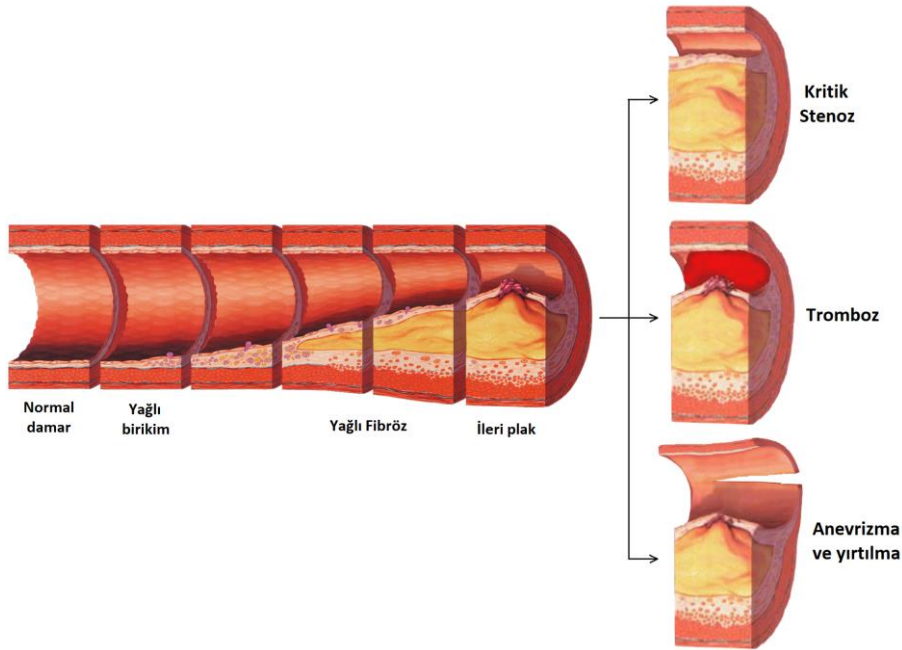
Şekil 2.27. FH belirtiler: a) arkus kornea, b) ksantelezma, c) tendon ksantoma



Şekil 2.28. Hiperlipidemi sebebi ile görsel olarak değişime uğramış kan (178)

2.4.1.1. ARTEROSKLEROZ

Kardiyovasküler hastalıklar gelişmiş ülkelerde en önce gelen ölüm sebeplerinden biridir. Hatta Afrika ülkeleri dışındaki en önemli ölüm sebebidir (179). Kardiyovasküler hastalıkların büyük çoğunluğu ise ateroskleroz sebebi ile oluşan ve damar sertliği olarak adlandırılan arteriyoskleroz sebebi ile oluşur (180). Ateroskleroz damar içerisinde lipitler, kan hücreleri ve başka maddelerin oluşturduğu plakların birikmesi durumudur (Şekil 2.29.).



Şekil 2.29. Ateroskleroz (181)

Arteriosklerozun sebebi tam olarak bilinmemek ile birlikte birçok sebebin etkili olduđu multifaktoriyel bir durum olduđu düşünölmektedir. Risk faktörleri içinde hiperkolesterolemi, yüksek seviyedeki enflamasyon belirteçleri, yüksek kan basıncı, diyabet, sigara tüketimi, obezite, aile öyküsü ve sağlıklısız beslenme sayılabilir (182, 183).

2.4.2. TANI KRİTERLERİ

Tanıda en önemli kriter, en önemli ve tipik belirti olan hiperkolesterolemidir. Ancak tek başına yeterli değildir. Bunun sebebi hiperkolesteroleminin seviyeleri olması ve hiperkolesterolemiye yol açabilecek sekonder sebeplerdir. Hangi seviye LDL-C'nin hiperkolesterolemi sayılacağı net değildir; Çizelge 2.1.'de LDL-C seviyeleri ve anlamları verilmiştir (184).

Çizelge 2.1. LDL-C seviyeleri ve klinik kategorizasyonu

LDL-C seviyesi (mg/dL)	Kategori
< 100	Optimal
100 - 129	Optimal Üzeri
130 - 159	Yüksek Sınırında
160 - 189	Yüksek
> 190	Çok Yüksek

Her ne kadar LDL-C seviyeleri 190'dan büyük olunca çok yüksek kabul edilse de FH'nin otozomal resesif formunu taşıyan bazı hastalar da bu seviye 500'ün üzerine bile çıkabilmektedir (185). Görüldüğü üzere yüksek kolesterol, siyah/beyaz şeklinden ziyade “grinin tonları” şeklindedir, bunun üzerine yüksek kolesterole yol açabilecek sekonder sebepleri de katacak olursak FH'nin tanısındaki güçlüğü gözler önüne sermiş oluruz. FH'den şüphelenilen bir durumda yapılması gereken ilk şey, ikincil sebep olabilecek sekonder klinik sebeplerin (hipotroidi, diyabet, nefrotik sendrom, obstüraktif karaciğer hastalığı vb.) ve çevresel faktörlerin (ilaç kullanımı, yüksek miktarda alkol kullanımı, sağlıklısız beslenme, hareketsiz yaşam tarzı vb.) elimine edilmesidir (170). Bu aşamadan sonra bireyin FH açısından değerlendirilmesi için kullanılan skollama bazlı tanı kriterlerinden birine başvurulması önerilmektedir (170, 171, 186). Bu kriterlerin en önemlilerinden biri Dutch kriterleri olarak ta bilinen “The Dutch Lipid Clinic Network Diagnostic Criteria”dır (187). Dutch kriterlerinde Çizelge 2.2.'deki skor değerlerine göre skollama yapılır.

Çizelge 2.2. Dutch kriterleri skorelama tablosu

Aile Öyküsü	Skor
Birinci derece akrabalarda prematür koroner ve vasküler hastalığın olması (Erkeklerde 55 yaş, kadınlarda 60 yaşın altında olması) ya da Birinci derece akrabalarda LDL kolesterol düzeyinin 95 persentilin üstünde olması (yaş ve cinsiyete göre)	1
Birinci derece akrabalarda tendon ksantoma ve/veya arkus kornea bulunması ya da 18 yaş altı çocuklarda LDL kolesterol düzeyinin 95 persentilin üstünde olması (yaş ve cinsiyete göre)	2
Klinik Öykü	Skor
Premature rofile arter hastalık olması (Erkeklerde 55 yaş, kadınlarda 60 yaşın altında olması)	2
Prematüre serebral ya da vasküler hastalık olması (yaş ve cinsiyete göre)	1
Fizik Muayene	Skor
Tendon ksantoma	6
Arkus kornea (<45 yaş)	4
LDL-kolesterol (mmol/L)	Skor
LDL-C $\geq 8,5$ (>330 mg/dL)	8
LDL-C 6,5-8,4 (250-329 mg/dL)	5
LDL-C 5,0-6,4 (190-249 mg/dL)	3
LDL-C 4,0-4,9 (155-189 mg/dL)	1
Genetik Analiz	Skor
DNA analizi – LDLR, ApoB, PCSK9 geninde patojenik mutasyon tespiti	8

Elde edilen toplam skora göre çıkarımı yapılan tanı da aşağıdaki Çizelge 2.3.'te verilmiştir.

Çizelge 2.3. Dutch kriterlerinde toplam skora göre tanı

Toplam Skor	Tanı / Yorum
> 8	Kesin FH
6 – 8	Muhtemel FH
3 - 5	Düşük Olasılıkla FH
< 3	Olası Olmayan FH

Görüleceği üzere Dutch kriterlerine göre LDL-C seviyesi 8,5 mmol/l veya 330 mg/dL'nin üzerinde olan bireyler kesin olarak FH hastası kabul edilirler.

2.4.3. KALITIM

Familiyal hiperkolesterolemi, otozomal dominant geçiş gösteren bir tek gen hastalığıdır. Otozomal resesif formları olsa da çok nadirdir (186, 188). Penetransı %90'ın üzerindedir; yani mutasyonu taşıyan bireyde çok büyük oranda semptomlar ortaya çıkar (177, 186).

2.4.3.1. İLGİLİ GENLER

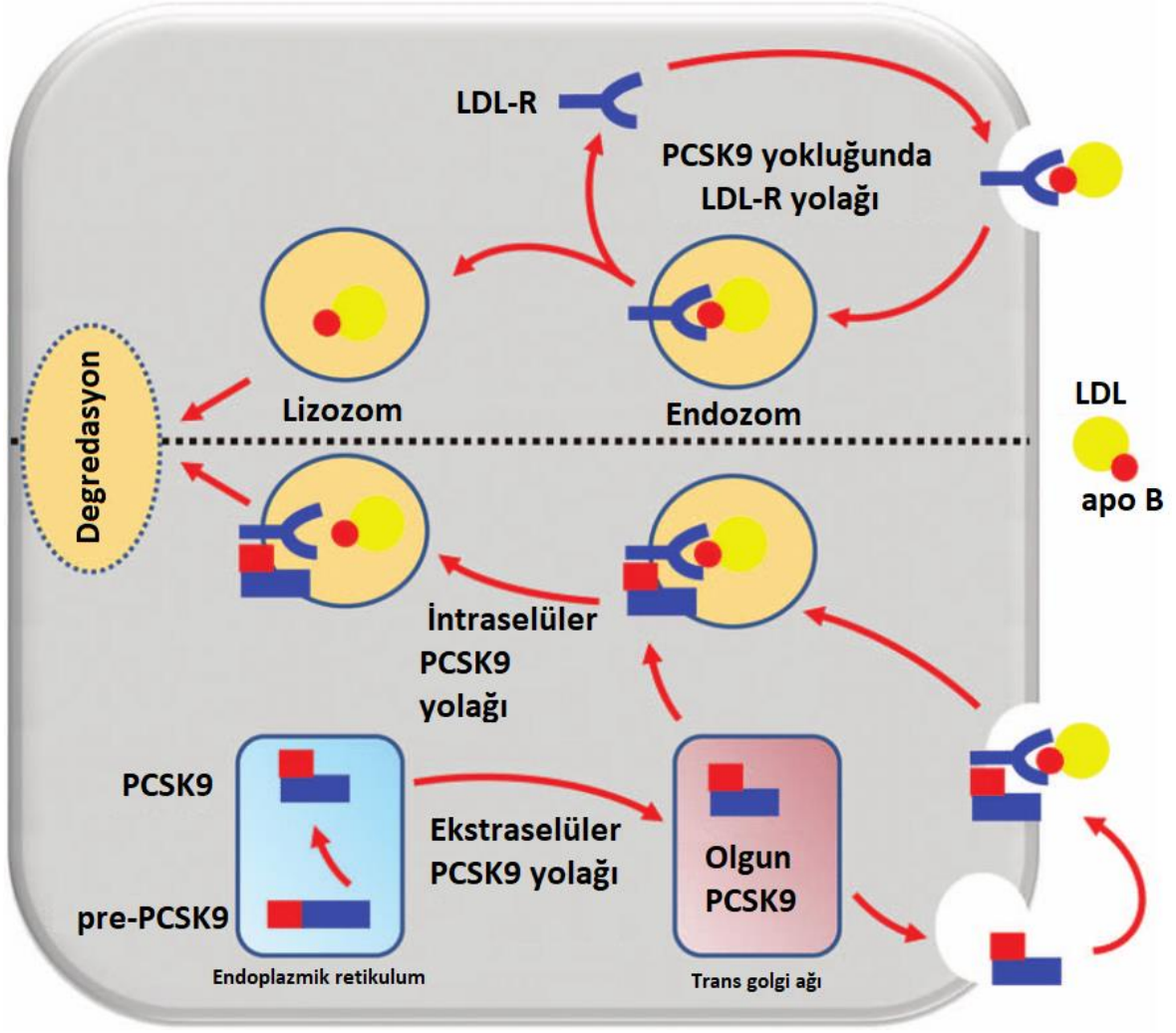
FH ile ilgili günümüze kadar 8 adet gen tanımlanmıştır. Bu genler LDLR (OMIM: 606945), APOB (OMIM: 107730), PCSK9 (OMIM: 607786), APOA2 (OMIM: 107670), GHR (OMIM: 600946), GSBS (OMIM: 604088), EPHX2 (OMIM: 132811) ve LDLRAP1 (OMIM: 605747)'dir (4, 189-191). Bu genlerin dışında olup, hiperkolesterolemi ile seyredip klinik olarak FH ye benzeyen ve iyi bir klinik değerlendirme yapılmadığı takdirde FH ile karışma ihtimali olan kolesterol ester depo hastalığına yol açan LIPA (OMIM: 613497) geni vardır. Bu genlerden en önemlisi kesin FH olarak değerlendirilen vakaların %70'ini, mutasyon tespit edilebilen vakaların ise %90'ından fazlasından sorumlu olan LDLR genidir (186). Diğer iki önemli gen ise vakaların yaklaşık %5'inden sorumlu olan APOB geni ve vakaların yaklaşık %1'inden sorumlu olan PCSK9 genidir (192).

2.4.3.1.1. LDLR

Kandaki kolesterol dengesinde çok önemli bir rolü oynayan bir hücre yüzeyi reseptörüdür. 19. kromozomun p koluna lokalizedir. Protein ürünü 860 amino asitten oluşur. 18 kodlayan ekzona sahip, yaklaşık 44 Kbp boyunda orta boy bir gendir (193). Kanda lipidlerin taşındığı lipoprotein yapılarından biri olan LDL lipoproteininin hücre, tarafından tanınarak hücre içerisine alınmasında rol oynar (194) (Şekil 2.30.). Mutasyona uğraması sonucunda LDL lipoproteinleri verimli bir şekilde hücre içerisine alınamaz, bunun sonucunda da kandaki LDL miktarı artarak hiperkolesterolemi durumuna yol açar (195).

2.4.3.1.2. APOB

LDL parçacıklarını oluşturan ana apolipoproteindir. 2. kromozomun p koluna lokalizedir. Her ne kadar yaklaşık 44 Kbp boyu ile genom üzerindeki yerleşimi açısından orta boy bir gen olsa da, 4563 amino asit kodlayan 26 ekzonlu büyük bir gendir (196). Görevi kandaki lipitlere bağlanarak LDL parçacıklarını oluşturmaktır (Şekil 2.30.). Mutasyona uğraması durumunda LDL parçacıklarının LDL reseptörlerine bağlanmasını zayıflatıp, LDL moleküllerinin hücre içine alınmasını azaltabilmektedir. Bunun sonucunda da hiperkolesterolemi durumu ortaya çıkabilmektedir (197).



Şekil 2.30. LDLR, APOB ve PCSK9 (198)

2.4.3.1.3. PCSK9

LDL reseptörlerinin parçalanarak miktarlarının düşmesinde rol oynayan bir serin proteaz enzimidir. 1. kromozomun p koluna lokalizedir. Protein ürünü 692 amino asitten oluşur. 12 kodlayan ekzona sahip, yaklaşık 25 Kbp boyunda orta boy bir genidir (199). LDLR ve APOB genlerinin aksine, PCSK9 geni LDL reseptörlerinin yıkımında rol aldığı için (Şekil 2.30.), gendeki aktive edici mutasyonlar hiperkolesterolemiye yol açar (200). Bu gendeki fonksiyon kaybı mutasyonları ise, LDL reseptörlerinin yıkımını azalttığı ve miktarlarını arttırdığı için tam tersi olarak hipokolesterolemiye yol açar (201). PCSK9, fonksiyonu azaldığı zaman kolesterol seviyelerini düşürücü etki yaptığından, hiperkolesterolemi hastalarında kullanılmak üzere iyi bir ilaç hedefidir; PCSK9 inhibitörleri 2015 yılından beri kullanılmaktadır (202).

2.4.3.1.4. ÇALIŞMA İÇİN SEÇİLEN GENLER

Çalışma için kullanılacak materyaller, LDLR, APOB ve PCSK9 genleri dizilenmiş ve mutasyon bulunamamış olası FH hastaları olduğu için, FH ile ilgili olduğu bildirilmiş diğer genlerden 6 tanesi çalışmaya alınmıştır. Çalışma için seçilen genler:

LDLRAP1

LDL reseptörünün sitoplazmik kısmı ile etkileşime girip, LDL parçacıklarının hücre içerisine alınmasında rol oynayan proteini kodlayan genidir. 1. kromozomun p koluna lokalizedir. Protein ürünü 308 amino asitten oluşur. 9 adet kodlayan ekzona sahip, yaklaşık 25 Kbp boyunda orta boy bir genidir (203). İlgili diğer genlerin aksine LDLRAP1 geninin sadece otozomal resesif kalıtım formu vardır; heterozigot bireylerde klinik gözlemlenmez. Birey homozigot mutant veya birleşik heterozigot olarak fonksiyon kaybettirici bir mutasyon veya mutasyonlar taşıyorsa, LDL parçacıkları hücre içerisine verimli bir şekilde alınamaz ve hiperkolesterolemi durumu ortaya çıkar (204, 205).

APOA2

APOA2 geninin ürünü olan Apolipoprotein A-II, bir lipoprotein olan HDL parçacıklarının yapısında olan ana proteinlerden biridir. APOA2 geni 1. kromozomun q koluna lokalizedir. Protein ürünü 100 amino asitten oluşan, 3 adet kodlayan ekzona sahip, yaklaşık 1,5 Kbp boyunda küçük bir genidir (206). Mekanizması belli olmamak ile birlikte, yapılan bir çalışmada bu gendeki mutasyonların lipid profilini etkileyebileceği gösterilmiştir (207).

ITIH4

Bu genin ürünü olan peptid, plazma serin inhibitörlerinden biri olan inter-alfa-tripsin inhibitörünü oluşturan peptid zincirlerinden biridir. 3. kromozomun p koluna lokalizedir. Protein ürünü 930 amino asitten oluşan, 24 adet kodlayan ekzona sahip, yaklaşık 18 Kbp boyunda orta boy bir genidir (208). Mekanizması belli olmamak ile birlikte, yapılan bir çalışmada bu gendeki mutasyonların lipid profilini etkileyebileceği gösterilmiştir (209).

GHR

Bu genin ürünü olan büyüme hormonu reseptörüne bağlanan büyüme hormonu, hücreler arası bir sinyal iletim yolağını aktive ederek hedef hücrede mitojenik ve anabolik tepkimeyi tetikleyerek büyümeye yol açar. 5. kromozomun p koluna lokalizedir. Her ne kadar yaklaşık 300 Kbp boyu ile genom üzerindeki yerleşimi açısından büyük bir gen olsa da, 638 amino asit kodlayan 10 ekzonlu orta boy bir gendir (210). Gendeki mutasyonlar, otozomal dominant ve otozomal resesif olarak kısa boyluluk ve cücelik gibi fenotipler ile ilişkilendirilmekle beraber, yayınlanan bir çalışma ile gendeki mutasyonların kandaki lipit seviyesini etkileyebileceği gösterilmiştir (211-213).

PPP1R17

Şu ana kadar üzerinde çok az çalışma yapılmış ve çok az bilgiye sahip olduğumuz bu genin ürünü olan protein, bir protein fosfataz 1 inhibitörüdür. Nöronlarda hücre içerisine besin alımı ile ilgili işlevi olduğu bildirilmiştir (214). 7. kromozomun p koluna lokalizedir. Protein ürünü 155 amino asitten oluşan, 4 adet kodlayan ekzona sahip, yaklaşık 20 Kbp boyunda küçük bir gendir (215). Mekanizması belli olmamak ile birlikte, yapılan bir çalışmada bu gendeki mutasyonların lipit profilini etkileyebileceği gösterilmiştir (216).

LIPA

Lizozomal asit lipaz enzimini kodlayan gendir. 5. kromozomun p koluna lokalizedir. Protein ürünü 399 amino asitten oluşan, 10 adet kodlayan ekzona sahip, yaklaşık 38 Kbp boyunda orta boy bir gendir (217). Lizozomal asit lipaz enzimi hücre içerisinde yağ asitleri ve kolesterolün lizozomlardan salınmasında rol oynayan bir enzimdir. Enzimin yeterli miktarda olmadığı durumlarda lizozomlar içerisinde lipid birikimi olur. Sitoplazmaya yeterli miktarda lipid salınımı olmadığı için hücre lipid seviyelerinde düşüş olduğunu düşünüp lipit sentezi mekanizmalarını aktive eder. Bunun sonucunda lipidler lizozomlar içerisinde birikirken yeni sentezlenen lipidler kana salınarak kandaki lipid seviyelerini yükseltir (218). İki tür hastalığa yol açar, kolesterol ester depo hastalığı ve Wolman hastalığı.

Wolman hastalığı enzim aktivitesini tamamen ortadan kaldıran mutasyonların her 2 allelde de bulunduğu, ağır klinik tablo ile seyreden ve -tıbbi müdahale ile bile- doğumdan sonra birkaç ay içerisinde hastanın ölmesi ile sonuçlanan çok ağır bir otozomal resesif hastalıktır.

Kolesterol ester depo hastalığı ise Wolman hastalığının daha hafif formudur; bu hastalık otozomal resesif kalıtım göstermekle beraber bulunan mutasyonların etkisine göre lizozomal asit lipaz enziminin aktivitesi tamamen ortadan kalkmamıştır. Sonuçta bulunan mutasyonlara ve kalan enzim aktivitesine göre değişik klinik tablo ortaya çıkabilmektedir; bir kısım hasta küçük yaşta belirti vererek tanı alabilmekte, ancak klinik tablosu daha hafif hastalar ise ailesel hiperkolesterolemi yanlış tanısı alabilmekte veya hiç tanı almadan yaşamlarını sürdürebilmektedir (218, 219). Aynı zamanda LIPA geninde heterozigot formda bulunan mutasyonların serum lipit seviyelerini arttırdığı ile ilgili bulgular da vardır (220). Kısacası LIPA geninde bulunan mutasyonlar, benzer klinik verebildiği için ailesel hiperkolesterolemi tanısı düşünülen hastalar arasında tespit edilebilmektedir.



3. GEREKÇE VE AMAÇ

Hastalık tarama programları, hastalığın yayılmasının azaltılması ve/veya erken tanı ile tedavi verimliliğinin artırılması amacıyla, hasta ve/veya taşıyıcı bireylerin, mevcut durumda semptom göstermeyen sağlıklı bireyler arasından düşük maliyet ile belirlendiği sağlık uygulamalarıdır. Çoğu durumda hem bireyler hem toplum hem de devletlerin sağlık sistemleri için son derece faydaları olan uygulamalardır. Ülkemizde ve dünyada hali hazırda, önemli bir kısmı enfeksiyon hastalıkları ile ilgi çok sayıda tarama yapılmaktadır. Ülkemizde Sağlık Bakanlığı'nın yürütmekte olduğu HPV taraması ve son dönemde yaşadığımız pandemi sebebi ile özel şirketlerin bile kendi imkanları ile çalışanlarına yaptırdıkları SARS-CoV-2 virüsü taramaları örnek olarak gösterilebilir. Enfeksiyon hastalıkları dışında kalıtsal hastalıkların taranması da hem ülkemizde hem de dünyanın çeşitli ülkelerinde yapılabilmektedir. Ülkemizde yetişkinlerden yapılan hemoglobinopati ve SMA tarama programları ve yeni doğanlardan yapılan kalıtsal hastalık tarama programları da buna örnek olarak verilebilir. Kalıtsal hastalıkların taraması, SMA taraması hariç, DNA seviyesinde yapılmamaktadır, çünkü DNA seviyesinde yapılan testler her ne kadar hassasiyeti ve özgüllüğü yüksek testler olsa da yüksek maliyetli oldukları için tarama mantığına uygun değildir. Genetik hastalıkların sebeplerinin en büyük çoğunluğunu oluşturan dizi varyasyonu mutasyonların taranması için geliştirilecek maliyet düşürücü yaklaşımların, hızla ilerleyen ve maliyeti düşen, yüksek hassasiyet ve özgüllüğe sahip olan DNA dizileme yönteminin, tarama programlarında yaygın olarak kullanılmasını kolaylaştırması kaçınılmazdır.

Bu tez çalışması ile tarama programlarında sıklıkla başvuru alan havuzlama yaklaşımının, çok boyutlu olarak NGS DNA dizileme yöntemine uygulanmasını geliştirilmesi, sınırlarının araştırılması ve pilot ölçekte denenmesi ile uygulanabilirliğinin gösterilmesi amaçlanmıştır.

4. MATERYAL VE YÖNTEM

4.1. MATERYAL

4.1.1. HAVUZLAMA VALİDASYONU

Çalışmaya dahil edilen örnekler 213S147 nolu “Ülkemizde familyal hiperkolesterolemiye yol açan sık görülen tek gen hastalıklarının dağılımı, sık ve nadir görülen mutasyonların belirlenmesi, tanı ve toplum taramasına yönelik genetik test kiti geliştirilmesi.” adlı Tübitak 1003 projesi kapsamında çalışılarak tüm LDLR geni dizilenen örnekler arasından seçilmiştir. Proje, Gülhane Askeri Tıp Akademisi Komutanlığı, 16 Aralık 2013 tarihli 29 nolu oturumu ile amaç, yöntem ve yaklaşım bakımından etik ilkelere uygun olarak değerlendirilmiştir. Pozitif örnekler LDLR geninin 4. ekzonunda mutasyon taşıyan örneklerden seçilmiştir; bunun sebebi proje sonunda LDLR geninin en fazla çeşit mutasyon içeren ekzon olduğunun görülmesidir (yayınlanmamış veri).

Toplamda 506 örnek çalışmaya dahil edilmiştir. 506 örnekten 48 tanesi (%9,5) LDLR geni 4. ekzonunda 1 adet patojenik mutasyonu heterozigot veya homozigot olarak taşıyan pozitif örnektir, 458 (%90,5) tanesi ise negatif örnektir (LDLR geni 4. ekzonunda patojenik mutasyon bulundurmamaktadır). 48 pozitif örnek toplamda 26 çeşit mutasyon içermektedir, bunlardan 20 tanesi heterozigot iken 6 tanesi homozigot, 2 tanesi de hem homozigot hem de heterozigot olarak bulunmaktadır; toplamda 28 farklı genotip bulunmaktadır. Pozitif örneklerden 2 farklı mutasyonu heterozigot olarak bulunduran 2 tanesi tekrarlanabilirlik çalışması için 4 defa kullanılmıştır; böylece kullanılan toplam örnek sayısı 512, pozitif örnek sayısı ise 54 (%10,5) olmuştur. Örnek olarak 213S147 nolu Tübitak projesi kapsamında EDTA’lı tüpe alınan periferik kan örnekleri kullanılmıştır.

4.1.2. FAMILİYAL HİPERKOLESTEROLEMİ HASTA GRUBU

Çalışmaya dahil edilen 320 örnek, 213S147 nolu “Ülkemizde familyal hiperkolesterolemiye yol açan sık görülen tek gen hastalıklarının dağılımı, sık ve nadir görülen mutasyonların belirlenmesi, tanı ve toplum taramasına yönelik genetik test kiti geliştirilmesi.” adlı Tübitak 1003 projesi kapsamında toplanan örnekler arasından seçilmiştir. Bu örnekler LDLR, APOB ve PCSK9 geni dizilenen ve patojenik mutasyon tespit edilmeyen örneklerdir. 256 adet örnek

FH belirtisi olan örnek, 64 adet örnek ise FH belirtisi olmayan kontrol örneğidir. Materyal olarak EDTA'lı tüpe alınan periferik kan örnekleri kullanılmıştır.

4.2. YÖNTEM

Bu çalışma 2 ana kısımdan oluşmaktadır: havuzlama validasyonu ve FH hasta ve kontrol grubu için 6 gende mutasyon taranması.

4.2.1. HAVUZLAMA VALIDASYONU

4.2.1.1. DİZİNLERİN VE HAVUZLARIN OLUŞTURULMASI

8x8'lik 8 adet dizin oluşturulmuş ve dizinin her bir pozisyonu 1'den 512'ye kadar numaralandırılmıştır. Örnekler dizin içerisine rastgele dağıtılmış; tam bir rastgeleliğin sağlanması için kura çekilerek dağıtım yapılmıştır. Sadece tekrarlanabilirlik çalışması için ayrılan 2 örneğe ait 8 pozisyon rastgele dağıtılmamış, belirli yerlere yerleştirilmiştir. Yerleşim sonrası dizin üzerindeki numaralar ve örnek kodları Şekil 4.1.'de gösterilmiştir; sarı ve turuncu olarak işaretli pozisyonlar, tekrarlanabilirlik çalışması için rastgele olmadan yerleştirilmiş 2 örneğe aittir.

Örnekler hazırlanan dizine göre EDTA'lı tüplerden 8'li strip tüplerine 200 µl alınmıştır; bu şekilde her 8x8'lik dizin için 8 adet 0,2 ml hacimli 8'li strip tüp elde edilmiştir. Bunun yapılmasının sebebi havuzların pratik bir şekilde oluşturulabilmesidir.

2 boyutlu 8'li havuzlama (2D08) için her bir 8x8'lik dizinin içerisindeki örnekler X ve Y düzleminde, her bir örnekten 8 kanallı pipet yardımı ile 20 şer µl alınmak sureti ile 0,2 ml hacimli 8'li strip tüplerde birleştirilerek, 8 örnek içeren örnek havuzları oluşturulmuştur. Vorteks yardımı ile iyice karıştırılmış ve her bir havuza bir kod verilmiştir. Bu şekilde her bir 8x8'lik dizin için 16 adet havuz, toplam 8 dizin için de 128 adet havuz oluşturulmuştur. İlk dizin için oluşturulan havuzlar ve havuz kodları Şekil 4.2.'de gösterilmiştir.

2 boyutlu 16'lı havuzlama (2D16) için 8 dizin, 4 dizin içeren 2 gruba ayrılmıştır. Her gruptaki 8x8'lik 4 dizin birleştirilmiş ve bu şekilde 2 adet 16x16'lı dizin elde edilmiştir. Dizinleri X ve Y düzlemindeki bileşenleri, 2 boyutlu 8'li havuzlama için oluşturulan 8 örneklik havuzlardan 2'sinin birleşiminden oluştuğu için, örnekler baştan tek tek birleştirilmemiştir; uygun olan 8 örneklik havuzlardan 10'ar µl birleştirilerek 16 örneklik

havuzlar meydana getirilmiştir. Bu şekilde iki adet 16x16'lı dizinden 64 adet havuz meydana getirilmiştir.

2 boyutlu 24'lü havuzlama için (2D24) 9 adet 8x8'lik dizine ihtiyaç olduğu için bir havuz iki defa kullanılmıştır. 9 adet 8x8'lik dizinden 24x24'lü bir adet dizin elde edilmiştir. Havuzlar oluşturulurken uygun olan 8'li havuzlardan 10'ar ıl alınarak birleştirilmiştir. Toplamda 24 örnek içeren 48 adet havuz oluşturulmuştur.

1	2	3	4	5	6	7	8	257	258	259	260	261	262	263	264
FH0403	FH06014	FH01501	FH032046	FH029012	FH014012	FH014011	FH0209	FH032018	FH07015	FH015016	FH0605	FH01106	FH04011	FH042025	FH030035
9	10	11	12	13	14	15	16	265	266	267	268	269	270	271	272
FH019010	FH03102	FH032056	FH029014	FH015034	FH0205	FH04902	FH034017	FH015027	FH04602	FH01704	FH030018	FH01002	FH0702	FH032065	FH03104
17	18	19	20	21	22	23	24	273	274	275	276	277	278	279	280
FH01405	FH033015	FH02804	FH032012	FH0609	FH029022	FH03307	FH033022	FH0806	FH042010	FH029017	FH07017	FH03303	FH029020	FH04601	FH32010
25	26	27	28	29	30	31	32	281	282	283	284	285	286	287	288
FH01703	FH0808	FH07010	FH015033	FH02909	FH032021	FH032026	FH011043	FH043010	FH029010	FH03208	FH01504	FH033010	FH033025	FH061052	FH033018
33	34	35	36	37	38	39	40	289	290	291	292	293	294	295	296
FH032043	FH0207	FH042044	FH0908	FH01402	FH03202	FH015035	FH01507	FH015040	FH032032	FH0701	FH0404	FH06019	FH02307	FH0606	FH033041
41	42	43	44	45	46	47	48	297	298	299	300	301	302	303	304
FH02807	FH06010	FH030017	FH01902	FH011023	FH015031	FH04908	FH0603	FH04309	FH0809	FH01602	FH061039	FH030015	FH011026	FH042036	FH061030
49	50	51	52	53	54	55	56	305	306	307	308	309	310	311	312
FH061055	FH03302	FH011045	FH01909	FH02302	FH03204	FH030028	FH0902	FH011027	FH02803	FH09017	FH030024	FH01107	FH015043	FH0703	FH03006
57	58	59	60	61	62	63	64	313	314	315	316	317	318	319	320
FH042027	FH032028	FH02808	FH032040	FH033021	FH01607	FH01401	FH01608	FH032049	FH01108	FH019011	FH08039	FH015048	FH03301	FH011021	FH08025

65	66	67	68	69	70	71	72	321	322	323	324	325	326	327	328
FH09014	FH030029	FH033032	FH029019	FH08032	FH019013	FH0405	FH033028	FH011039	FH033040	FH033023	FH032027	FH032038	FH015045	FH08023	FH09012
73	74	75	76	77	78	79	80	329	330	331	332	333	334	335	336
FH017011	FH0403	FH061050	FH03101	FH033015	FH032036	FH01409	FH030021	FH0907	FH01903	FH03304	FH02904	FH01702	FH01905	FH061034	FH030032
81	82	83	84	85	86	87	88	337	338	339	340	341	342	343	344
FH08033	FH033017	FH029021	FH015014	FH08031	FH02801	FH04204	FH011031	FH011024	FH01109	FH061049	FH029013	FH01907	FH032048	FH02303	FH061020
89	90	91	92	93	94	95	96	345	346	347	348	349	350	351	352
FH07020	FH015030	FH032013	FH03307	FH0406	FH043012	FH061013	FH011047	FH061046	FH032031	FH08019	FH061016	FH0901	FH0708	FH06017	FH033024
97	98	99	100	101	102	103	104	353	354	355	356	357	358	359	360
FH061056	FH032033	FH033026	FH011040	FH0206	FH0401	FH01604	FH032029	FH011011	FH08018	FH06016	FH032016	FH02020	FH011038	FH032025	FH030025
105	106	107	108	109	110	111	112	361	362	363	364	365	366	367	368
FH01101	FH09010	FH028013	FH0402	FH08039	FH042049	FH032025	FH03308	FH02802	FH015025	FH07022	FH02907	FH011048	FH06019	FH02021	FH061044
113	114	115	116	117	118	119	120	369	370	371	372	373	374	375	376
FH032015	FH032057	FH014017	FH01801	FH03007	FH08015	FH07023	FH038022	FH08035	FH09013	FH02902	FH0707	FH061021	FH04906	FH014015	FH01001
121	122	123	124	125	126	127	128	377	378	379	380	381	382	383	384
FH028011	FH014014	FH032011	FH07019	FH032053	FH01406	FH04013	FH011025	FH03309	FH011037	FH061051	FH011012	FH07026	FH015038	FH011013	FH015019

129	130	131	132	133	134	135	136	385	386	387	388	389	390	391	392
FH030022	FH04904	FH02809	FH033012	FH061038	FH032058	FH011044	FH032047	FH08020	FH032066	FH011030	FH06013	FH030019	FH033030	FH02012	FH015011
137	138	139	140	141	142	143	144	393	394	395	396	397	398	399	400
FH032019	FH029011	FH04605	FH042050	FH032050	FH01102	FH028012	FH08034	FH08021	FH030013	FH03004	FH011014	FH03207	FH02906	FH07012	FH08011
145	146	147	148	149	150	151	152	401	402	403	404	405	406	407	408
FH042023	FH014013	FH0403	FH061057	FH061059	FH01103	FH0601	FH042021	FH033029	FH01601	FH03005	FH08038	FH011015	FH01701	FH034021	FH06012
153	154	155	156	157	158	159	160	409	410	411	412	413	414	415	416
FH015029	FH038027	FH011032	FH033037	FH03305	FH011028	FH014016	FH042032	FH0904	FH03805	FH030011	FH028010	FH0408	FH07011	FH0807	FH032059
161	162	163	164	165	166	167	168	417	418	419	420	421	422	423	424
FH032052	FH030030	FH09011	FH01605	FH032051	FH03009	FH03203	FH01104	FH015023	FH042039	FH08012	FH01709	FH08016	FH061014	FH015026	FH015044
169	170	171	172	173	174	175	176	425	426	427	428	429	430	431	432
FH07018	FH03306	FH015032	FH030026	FH08014	FH032014	FH04008	FH04308	FH015042	FH030016	FH042011	FH04407	FH07016	FH0705	FH0802	FH032030
177	178	179	180	181	182	183	184	433	434	435	436	437	438	439	440
FH032055	FH03809	FH030034	FH08036	FH0602	FH032054	FH01908	FH03803	FH08022	FH08024	FH01706	FH033039	FH02901	FH01904	FH06019	FH032062
185	186	187	188	189	190	191	192	441	442	443	444	445	446	447	448
FH07014	FH0903	FH02306	FH0906	FH0607	FH01105	FH02806	FH015017	FH03008	FH032063	FH08031	FH07024	FH03103	FH030020	FH01407	FH011022

193	194	195	196	197	198	199	200	449	450	451	452	453	454	455	456
FH09016	FH0805	FH032039	FH0801	FH01705	FH015010	FH033011	FH02015	FH02301	FH033014	FH061011	FH011016	FH033031	FH02309	FH0804	FH011041
201	202	203	204	205	206	207	208	457	458	459	460	461	462	463	464
FH033020	FH011033	FH05004	FH06018	FH032060	FH08025	FH06107	FH04010	FH01707	FH04907	FH03002	FH02017	FH04603	FH08028	FH08013	FH03209
209	210	211	212	213	214	215	216	465	466	467	468	469	470	471	472
FH015028	FH08017	FH043023	FH08037	FH033020	FH038021	FH015039	FH03201	FH032064	FH011036	FH032037	FH017010	FH023010	FH015041	FH043019	FH032035
217	218	219	220	221	222	223	224	473	474	475	476	477	478	479	480
FH033042	FH01606	FH019012	FH0403	FH030010	FH01906	FH04208	FH08017	FH034026	FH011017	FH032023	FH06108	FH0706	FH032022	FH029015	FH030023
225	226	227	228	229	230	231	232	481	482	483	484	485	486	487	488
FH07027	FH015047	FH03908	FH033038	FH032061	FH02701	FH02908	FH07025	FH01404	FH061043	FH015022	FH0704	FH01503	FH061041	FH032045	FH029016
233	234	235	236	237	238	239	240	489	490	491	492	493	494	495	496
FH015037	FH0803	FH01901	FH030027	FH030037	FH011034	FH06109	FH03001	FH011018	FH014018	FH015036	FH09015	FH011020	FH015024	FH033019	FH06015
241	242	243	244	245	246	247	248	497	498	499	500	501	502	503	504
FH033035	FH08010	FH015012	FH032041	FH01508	FH01708	FH07021	FH061058	FH032044	FH04012	FH0409	FH011029	FH05902	FH015021	FH032020	FH061025
249	250	251	252	253	254	255	256	505	506	507	508	509	510	511	512
FH0909	FH015015	FH033013	FH023011	FH01403	FH0608	FH042014	FH02022	FH014010	FH042017	FH033027	FH033016	FH032017	FH0709	FH03003	FH06019

Şekil 4.1. Oluşturulan ana dizin, örnek numaraları ve kodları

1	2	3	4	5	6	7	8	2D08x01
FH0403	FH06014	FH01501	FH032046	FH029012	FH014012	FH014011	FH0209	
9	10	11	12	13	14	15	16	2D08x02
FH019010	FH03102	FH032056	FH029014	FH015034	FH0205	FH04902	FH034017	
17	18	19	20	21	22	23	24	2D08x03
FH01405	FH033015	FH02804	FH032012	FH0609	FH029022	FH03307	FH033022	
25	26	27	28	29	30	31	32	2D08x04
FH01703	FH0808	FH07010	FH015033	FH02909	FH032021	FH032026	FH011043	
33	34	35	36	37	38	39	40	2D08x05
FH032043	FH0207	FH042044	FH0908	FH01402	FH03202	FH015035	FH01507	
41	42	43	44	45	46	47	48	2D08x06
FH02807	FH06010	FH030017	FH01902	FH011023	FH015031	FH04908	FH0603	
49	50	51	52	53	54	55	56	2D08x07
FH061055	FH03302	FH011045	FH01909	FH02302	FH03204	FH030028	FH0902	
57	58	59	60	61	62	63	64	2D08x08
FH042027	FH032028	FH02808	FH032040	FH033021	FH01607	FH01401	FH01608	
2D08y01	2D08y02	2D08y03	2D08y04	2D08y05	2D08y06	2D08y07	2D08y08	

Şekil 4.2. İlk dizin için oluşturulan havuzlar ve havuz kodları

2 boyutlu 32’li havuzlama (2D32) için 16 adet 8x8’lik dizine ihtiyaç olduğu için 8x8’lik 8 adet dizin ikişer defa kullanılarak 16 adet dizin elde edilmiştir. Aynı dizinlerin X ve Y düzleminde aynı pozisyonlara düşmemesi ve tamamen aynı havuzların oluşmamasına özen gösterilmiştir. Bu şekilde 32x32’li bir adet dizin oluşturulmuştur. Havuzların oluşturulması için yine 8’li havuzlar kullanılmıştır; uygun 8’li havuzlar 10’ar ıl birleştirilerek 32’li havuzlar oluşturulmuştur. 32’li toplam 64 adet havuz oluşturulmuştur.

3 boyutlu 8’li havuzların (3D08) oluşturulması için, 8 adet 8x8’lik dizin, üst üste yerleştirilerek bir küp biçiminde bir dizine dönüştürülmüştür. X, Y ve Z boyutlarında 64 örneğin birleşmesi ile oluşacak havuzlar, diğer havuzlamalarda olduğu gibi uygun 8’li havuzun 10’ar ıl birleştirilmesi ile oluşturulmuştur.

Sonuç olarak elde edilen havuz sayısı ve havuz içindeki örnek sayısı, havuzlama biçimi açısından Çizelge 4.1.’de gösterilmiştir.

Çizelge 4.1. Elde edilen dizin ve havuzlar

Boyut	Dizinin bir boyutu	Toplam dizin sayısı	Toplam havuz sayısı (328)	Havuz içindeki örnek sayısı	Taranan toplam örnek sayısı
2	8	8	128	8	512
2	16	2	64	16	512
2	24	1	48	24	576
2	32	1	64	32	1024
3	8	1	24	64	512

4.2.1.2. MOLEKÜLER ÇALIŞMA

Havuzlama validasyonu moleküler çalışması için elde edilen havuzlardan sırası ile PCR, NGS ve NGS veri analizi aşamaları gerçekleştirilmiştir.

4.2.1.2.1. PCR

LDLR geninin 4. ekzonunun PCR amplifikasyonu için kandan direkt PCR yaklaşımı seçilmiştir. Bunun sebebi sadece tek PCR yapılacak olması ve toplam havuz sayısının yüksek olmasıdır (toplam 328 havuz); bu şekilde maliyet ve iş gücünden tasarruf edilmiştir. Polimeraz olarak kandan direkt PCR yapmaya uygun, inhibitörlere dirençli Phire II Hot Start (Thermo Fisher Inc.) DNA polimeraz enzimi kullanılmıştır. Reaksiyon koşulları Çizelge 4.2.'de gösterilmiştir.

Çizelge 4.2. PCR Reaksiyon koşulları

Reaksiyon bileşeni	Miktar (µl)
dH ₂ O	16
5x Reaksiyon tamponu (PhireII)	5
dNTP karışımı (her birinden 10mM)	0,5
İleri primer (5 µM)	1
Geri primer (5 µM)	1
DNA Polimeraz PhireII (0,5u/µl)	0,5
Kan	1
Toplam	25

5x Reaksiyon tamponu içerisinde yeteri miktarda MgCl₂ (15 mM) olduğu için bu bileşenden fazladan eklenmemiştir.

Primer olarak bölüm 4.2.2.1.2.'de tarif edilen primerler kullanılmıştır.

Isıl döndürücü cihaz olarak T100 ısı döndürücü (Biorad Inc.) kullanılmıştır. Kullanılan protokol Çizelge 4.3.'te gösterilmiştir.

Çizelge 4.3. Isıl döndürücü protokolü

Sıcaklık	Süre (dk:sn)	Döngü sayısı
95°C	02:00	1
95°C	00:20	40
60°C	00:20	
72°C	00:40	
72°C	02:00	1
4°C	∞	1

PCR sonrası reaksiyonlar %2'lik agaroz jel elektroforezi ile kontrol edilmiş, başarısız reaksiyonlar tekrar edilmiştir.

4.2.1.2.2. NGS

NGS çalışması için Miseq (Illumina Inc.) platformu kullanılmıştır. Kütüphane hazırlık aşaması için transpozaz bazlı NexteraXT kiti (Illumina Inc.) kullanılmıştır. Başlangıç DNA olarak pürifiye edilmiş ve konsantrasyonları 0,5 ng/μl olarak normalize edilmiş PCR ürünleri kullanılmıştır. PCR ürünlerinin pürifikasyonu için filtrasyon bazlı NucleoFast 96 PCR pürifikasyon kiti (MACHEREY-NAGEL GmbH) kullanılmıştır. Pürifikasyon sonrası ürünlerin spektrofotometrik konsantrasyon tayini için Nanodrop ND1000 (Thermo Fisher Inc.) mikrohacim spektrofotometresi kullanılmıştır. Pürifiye edilmiş ve konsantrasyonları belirlenmiş PCR ürünleri dH₂O ile 0,5 ng/μl ye seyreltilerek kütüphane hazırlık aşamasına hazır hale getirilmiştir. Daha sonra NexteraXT kiti kullanılarak kütüphane hazırlığı yapılmıştır. Dizileme Miseq platformunda 300V2 dizileme kartuşları kullanılarak gerçekleştirilmiştir.

4.2.1.2.3. NGS VERİ ANALİZİ

Düşük örnek sayılı havuzlar (2 boyutlu 8'li, 16'lı ve 24'lü) için ayrı, yüksek örnek sayılı havuzlar (2 boyutlu 32'li ve 3 boyutlu 8'li) için ayrı analiz yaklaşımları kullanılmıştır. Düşük örnek sayılı havuzlar için kullanılan analiz yöntemi basit, yüksek miktarda biyoinformatik yetkinliğine ihtiyaç duymayan ve herkesin kolayca uygulayabileceği bir yaklaşımdır. Yüksek örnek sayılı havuzların analizi için ise -basit yaklaşım yeterli olmadığından- daha kompleks bir yaklaşım seçilmiştir.

DÜŞÜK ÖRNEK SAYILI HAVUZLAR

Düşük örnek sayılı havuzlar (2 boyutlu 8'li ve 16'lı) için Miseq platformunda, Miseq Reporter yazılımı ile hg19 referans insan genomuna hizalanmış bam dosyaları, IGV yazılımında açılarak gözle analiz yapılmıştır. Havuzlanmış örneklerin analizi için gerekli olan yüksek hassasiyet, allel frekans eşik değeri 0,02'ye ayarlanarak sağlanmıştır. IGV'de MAF (minör allel frekansı) 0,02'den yüksek varyasyon olan noktalar pozitif olarak kabul edilmiştir. Pozitif sonuçlar beklenen genotipler tablosu ile karşılaştırılarak havuzlamanın başarısı incelenmiştir.

YÜKSEK ÖRNEK SAYILI HAVUZLAR

Yüksek örnek sayılı havuzlarda (2 boyutlu 24'li, 2 boyutlu 32'li ve 3 boyutlu 8'li) 0,02 allel frekans eşik değeri yeterli olmadığı için eşik değerini 0,02'nin altına düşürmek gerekmektedir. Ancak 0,02'nin altındaki eşik değerlerinde çok fazla okuma hatası kaynaklı yalancı pozitif sonuç çıkmaktadır. Bu sebeple bu havuzların analizinde bu amaç için geliştirilmiş hazır algoritmalar kullanılmıştır.

Veri 2 farklı algoritma ile analiz edilmiştir: GATK HaplotypeCaller ve GATK Mutect2 (221). GATK HaplotypeCaller ile yapılan varyant çağırma süreçlerinde havuz içerisindeki toplam kromozom sayısı dikkate alınarak –ploidy parametresi buna göre şekillendirilmiştir. GATK Mutect2 algoritmasında ise veri 10, 20 ve 30 Phred skorlarına göre filtrelenerek varyantların çağırılması gerçekleştirilmiştir; bu şekilde 3 farklı eşik değeri kullanılarak 3 adet sonuç elde edilmiştir.

GATK HaplotypeCaller analizi ile beraber elde edilen toplamda 4 adet sonuç değerlendirilmiştir. GATK HaplotypeCaller analizinin çok fazla sayıda yalancı pozitif ürettiği görülmüştür. GATK Mutect2 algoritmasında ise yalancı pozitif oranının beklendiği üzere Phred skoru eşik değeri yükseldikçe düştüğü, ancak yine yükselen Phred skoru eşik değeri ile birlikte hassasiyetin de azaldığı görülmüştür. En ideal sonucun GATK Mutect2 ve Phred skoru 20 olan analiz olduğu görülerek analizlerin tamamı bu sonuçlar ile yapılmıştır; algoritma sonucunda üretilen vcf dosyaları IGV yazılımında açılarak analiz gerçekleştirilmiştir.

4.2.2. FAMILİYAL HİPERKOLESTEROLEMİ HASTA GRUBU

Familiyal hiperkolesterolemi hasta grubu mutasyon taraması çalışması için 2 boyutlu 8'li havuz yaklaşımı seçilmiştir.

4.2.2.1. DİZİNLERİN VE HAVUZLARIN OLUŞTURULMASI

Seçilen 256 adet FH belirtisi olan örnek (örnek kodları L001-L256 arası), 8x8'lik 4 adet dizin içerisine dağıtılmıştır. 64 adet kontrol örneğinden (örnek kodları L257-L320 arası) ise 1 adet 8x8'lik dizin oluşturularak toplam 5 adet 8x8'lik dizin elde edilmiştir. Dizin içerisindeki örnekler X ve Y düzleminde, her bir örnekten 8 kanallı pipet yardımı ile 25'er

µl alınmak sureti ile 0,2 ml hacimli 8’li strip tüplerde birleştirilerek, 8 örnek içeren toplam 200 µl hacimli örnek havuzları oluşturulmuştur. Vorteks yardımı ile iyice karıştırılmıştır ve her bir havuza bir kod verilmiştir. Bu şekilde her bir 8x8’lik dizin için 16 adet havuz, toplam 5 dizin için de 80 adet havuz oluşturulmuştur. İlk 64 örneğin oluşturduğu dizin için oluşturulan havuzlar ve havuz kodları Şekil 4.3.’te gösterilmiştir.

L001	L002	L003	L004	L005	L006	L007	L008	LX1-1
L009	L010	L011	L012	L013	L014	L015	L016	LX1-2
L017	L018	L019	L020	L021	L022	L023	L024	LX1-3
L025	L026	L027	L028	L029	L030	L031	L032	LX1-4
L033	L034	L035	L036	L037	L038	L039	L040	LX1-5
L041	L042	L043	L044	L045	L046	L047	L048	LX1-6
L049	L050	L051	L052	L053	L054	L055	L056	LX1-7
L057	L058	L059	L060	L061	L062	L063	L064	LX1-8

LY1-1 LY1-2 LY1-3 LY1-4 LY1-5 LY1-6 LY1-7 LY1-8

Şekil 4.3. İlk 64 örnek için oluşturulan havuzlar ve havuz kodları

4.2.2.2. MOLEKÜLER ÇALIŞMA

Familiyal hiperkolesterolemi hasta grubu mutasyon taraması moleküler çalışması için elde edilen havuzlardan sırası ile DNA izolasyonu, PCR, NGS ve NGS veri analizi aşamaları gerçekleştirilmiştir.

4.2.2.2.1. DNA İZOLASYONU

Örnek sayısının az, örnek başına kurulacak PCR sayısının fazla olması sebebiyle validasyon çalışmasında olduğu gibi kandan direkt PCR yaklaşımı yerine, kanlardan DNA izolasyonu gerçekleştirilip, sonraki basamaklara izole edilmiş kanlar ile devam edilmesi yaklaşımı tercih edilmiştir.

Kandan DNA izolasyonları, havuzlardan elde edilen 200 µl kan ve QIAamp DNA Mini Kit (Qiagen GmbH) kullanılarak gerçekleştirilmiştir. Elde edilen izolatlar -20°C’de muhafaza edilmiştir.

4.2.2.2.2. PCR

6 genin toplamda 59 kodlayan ekzonu bulunmaktadır. Tüm kodlayan bölge ile birlikte ekzon-intron kesişim bölgelerinden en az 20’şer bazı kapsayacak şekilde toplamda 36 adet PCR reaksiyonu için 36 çift primer tasarlanmıştır (LIPA için 9, LDLRAP1 için 6, APOA2 için 1, ITIH4 için 8, GHR için 8 ve PPP1R17 için 4 çift). Tasarım için hg38 referans insan genomu ve Primer Designer V2.0 (Scientific & Educational Softwarre) yazılımı kullanılmıştır. Amplikon boyları 234 bp ile 2.100 bp arasında değişmektedir. Primerlerin dizileri ve elde edilen amplikonların boyları EK 1’de verilmiştir.

Polimeraz olarak MyTaq Hot Start (Bioline Inc.) polimeraz kullanılmıştır. Reaksiyon koşulları Çizelge 4.4.’te gösterilmiştir.

Çizelge 4.4. PCR Reaksiyon koşulları

Reaksiyon bileşeni	Miktar (µl)
dH ₂ O	15,8
5x Reaksiyon tamponu (MyTaq)	5
İleri primer (5 µM)	1
Geri primer (5 µM)	1
DNA Polimeraz PhireII (0,2u/µl)	0,2
DNA	2
Toplam	25

5x Reaksiyon tamponu içerisinde yeteri miktarda MgCl₂ (15 mM) ve dNTP karışımı (her birinden 5 mM) olduğu için bu bileşenlerden fazladan eklenmemiştir.

Isıl döndürücü cihaz olarak T100 ısı döndürücü (Biorad Inc.) kullanılmıştır. Kullanılan protokol Çizelge 4.5.’te gösterilmiştir.

PCR sonrası reaksiyonlar %2’lik agaroz jel elektroforezi ile kontrol edilmiş, başarısız reaksiyonlar tekrar edilmiştir. Her havuza ait başarılı reaksiyonlar birleştirilerek, havuza ait tüm amplikonların tek bir tüpte toplanması sağlanmıştır. Birleştirme sırasında her amplikonun karışım içinde kabaca eşit miktarda temsil edilmesi için reaksiyon verimliliği ve amplikon boyları göz önünde bulundurulmuştur; zayıf amplifikasyon gösteren ve boyu

daha büyük olan PCR ürünlerinden daha fazla alınarak karışım oluşturulmuştur. Bu şekilde toplam 80 havuzun her biri için, içinde 6 gene ait kodlayan bölge ve ekzon-intron kesişim noktalarının PCR ampikonlarını içeren 80 adet PCR havuzu oluşturulmuştur.

Çizelge 4.5. Isıl döndürücü protokolü

Sıcaklık	Süre (dk:sn)	Döngü sayısı
95 ⁰ C	10:00	1
95 ⁰ C	00:45	45
60 ⁰ C	00:45	
72 ⁰ C	00:45	
72 ⁰ C	10:00	1
4 ⁰ C	∞	1

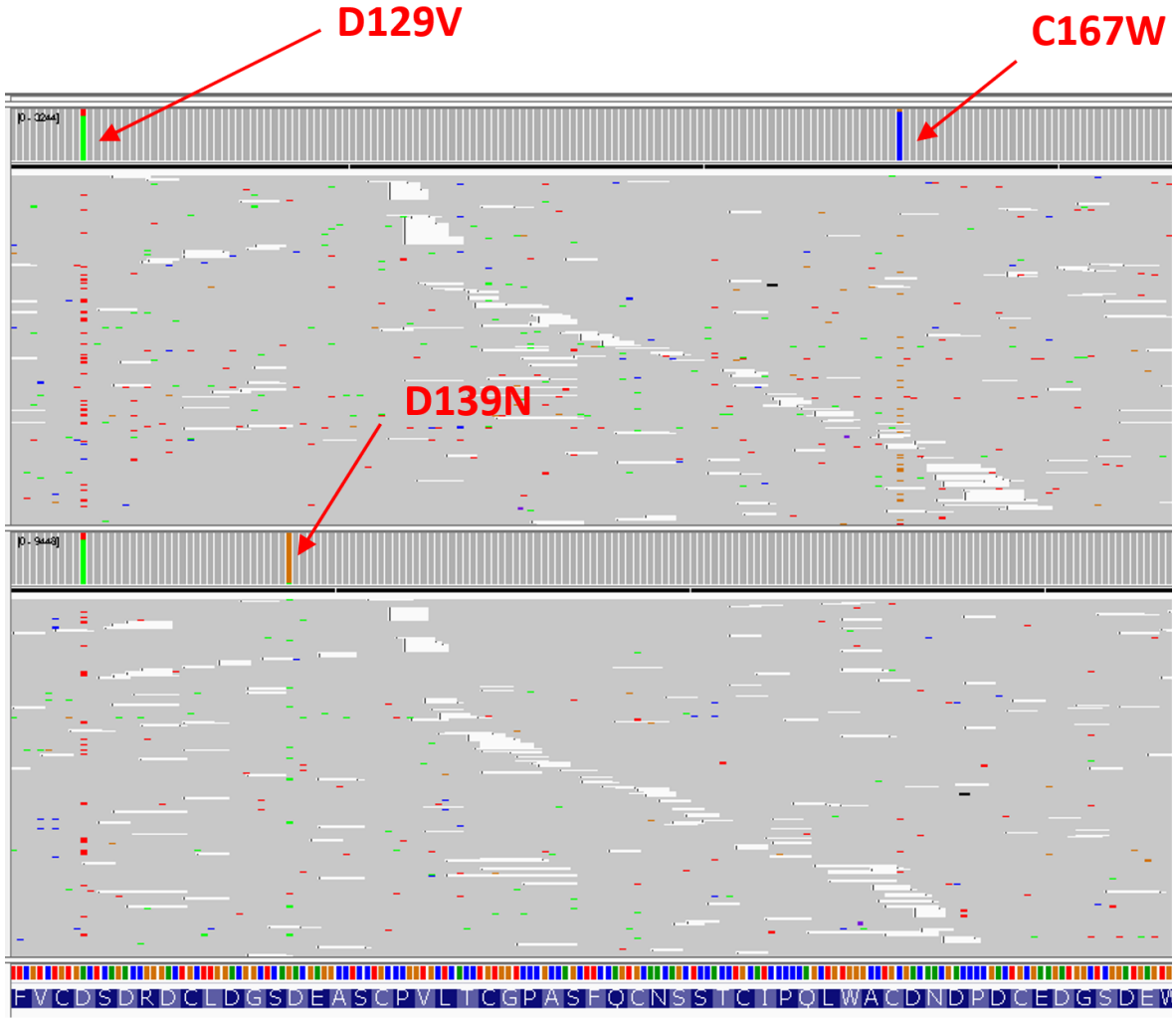
4.2.2.2.3. NGS

NGS çalışması için bölüm 4.2.1.2.2. tarif edilen yöntemin aynısı kullanılmıştır.

4.2.2.2.4. NGS VERİ ANALİZİ

NGS analizinde, bölüm 4.2.1.2.3.'te tarif edilen, düşük örnek sayılı havuzlar için yapılan analizi yaklaşımı kullanılmıştır; Miseq platformunda, Miseq Reporter yazılımı ile hg19 referans insan genomuna hizalanmış bam dosyaları, IGV yazılımında açılarak gözle analiz yapılmıştır. Havuzlanmış örneklerin analizi için gerekli olan yüksek hassasiyet, allel frekans eşik değeri 0,02'ye ayarlanarak sağlanmıştır (Şekil 4.4.).

IGV de 0,02 ve daha yüksek MAF'lı olarak belirlenen noktalar pozitif olarak kabul edilmiştir. Pozitif noktalardaki bulunan değişikliğin popülasyon allel frekansı, 1000 Genome Projesi (222) faz3 verilerine göre şayet 0,02'nin üzerindeyse, değişiklik yaygın bir polimorfizm olarak değerlendirilerek çalışma dışı bırakılmıştır. Popülasyon allel frekansı 0,02 ve altında olan değişiklikler ise mutasyon olarak değerlendirilmiş ve pozitif olarak değerlendirilen havuzlar dizin tablosuna işlenmiştir.



Şekil 4.4. Örnek IGV analizi

POZİTİF ÖRNEKLERİN TESPİTİ İÇİN ÖRNEK SEÇİMİ

X ve Y düzlemlerinden alınan pozitif sonuçlara göre pozitifliklerin kesiştiği nokta veya noktalarda bulunan örnekler, hem doğrulanma hem de zigozitenin belirlenmesi için seçilmiştir. Örnek olarak LX4-8 ve LY4-4 havuzlarında, LIPA geninin 7. ekzonunda I252L mutasyonu tespit edildiği için, L252 nolu örnek tek başına LIPA geni 7. ekzonunun dizilenmesi için seçilmiştir (Şekil 4.5.).

Birden fazla aynı mutasyon için pozitif örnek barındıran havuzlarda, hangi örneklerin pozitif olduğunu tespit etmek için birden fazla kesişim noktasının dizilenmesi gerekmektedir. Mesela LDLRAP1 geninde R238W mutasyonunun LX1-4, LX1-6, LY1-2 ve LY1-4 havuzlarında tespit edilmesi sonucunda, her ne kadar gerçekte bu durum 2 örneğin pozitif olmasından kaynaklansa da (L028 ve L042) bunun ortaya çıkartılabilmesi için kesişim

noktalarında bulunan 4 örneğin de (L026, L028, L042 ve L044) dizilenmesi gerekmektedir (Şekil 4.6.).

L193	L194	L195	L196	L197	L198	L199	L200	LX4-1
L201	L202	L203	L204	L205	L206	L207	L208	LX4-2
L209	L210	L211	L212	L213	L214	L215	L216	LX4-3
L217	L218	L219	L220	L221	L222	L223	L224	LX4-4
L225	L226	L227	L228	L229	L230	L231	L232	LX4-5
L233	L234	L235	L236	L237	L238	L239	L240	LX4-6
L241	L242	L243	L244	L245	L246	L247	L248	LX4-7
L249	L250	L251	L252	L253	L254	L255	L256	LX4-8

LY4-1 LY4-2 LY4-3 LY4-4 LY4-5 LY4-6 LY4-7 LY4-8

Şekil 4.5. L252 nolu örneğin doğrulanma ve zigozite belirlenmesi için seçimi

L001	L002	L003	L004	L005	L006	L007	L008	LX1-1
L009	L010	L011	L012	L013	L014	L015	L016	LX1-2
L017	L018	L019	L020	L021	L022	L023	L024	LX1-3
L025	L026	L027	L028	L029	L030	L031	L032	LX1-4
L033	L034	L035	L036	L037	L038	L039	L040	LX1-5
L041	L042	L043	L044	L045	L046	L047	L048	LX1-6
L049	L050	L051	L052	L053	L054	L055	L056	LX1-7
L057	L058	L059	L060	L061	L062	L063	L064	LX1-8

LY1-1 LY1-2 LY1-3 LY1-4 LY1-5 LY1-6 LY1-7 LY1-8

Şekil 4.6. L026, L028, L042 ve L044 nolu örneklerin doğrulanma ve zigozite belirlenmesi için seçimi

Kısacası aynı mutasyonun tespit edildiği bütün havuzların dizin içindeki bütün kesişim noktalarındaki örnekler, bir sonraki aşama için seçilmiştir.

4.2.2.3. POZİTİF ÖRNEKLERİN TESPİTİ

Havuzların çalışılması sonucunda belirlenen pozitif havuzlara göre seçilen örneklerin sadece mutasyonu taşıyan ampliconları teker teker dizilenerek pozitif örnekler tespit edilmiş, doğrulanmış ve zigoziteri belirlenmiştir.

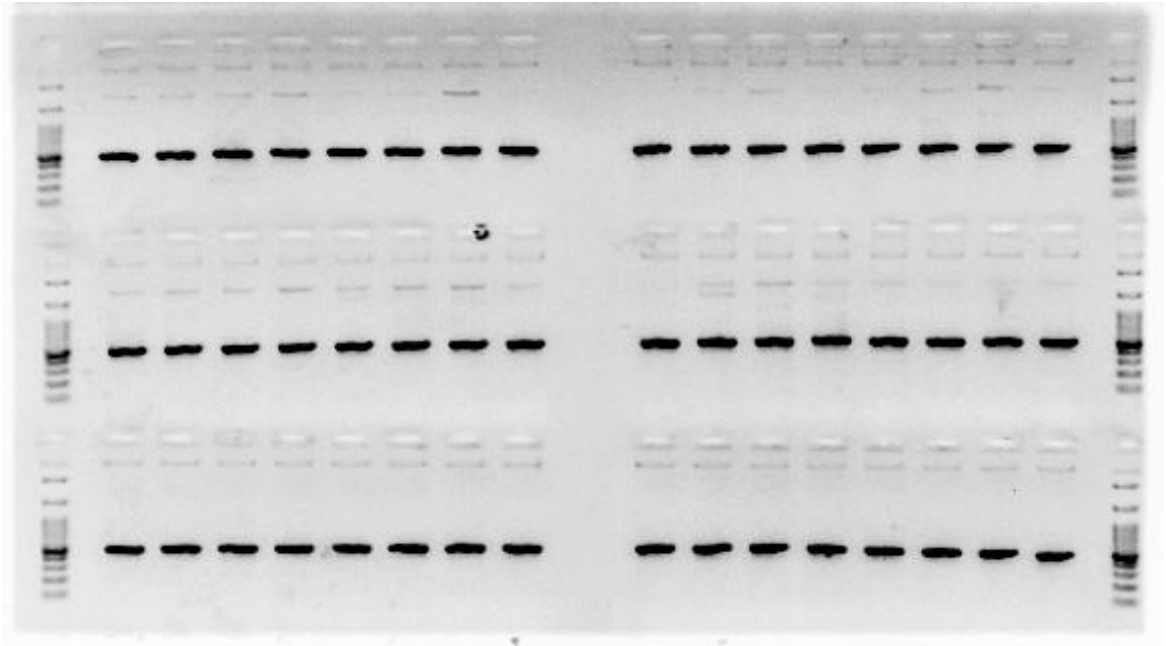
Yöntem olarak -üç noktada farklılık ile birlikte- havuzlama validasyonu çalışması için 4.2.1.'de tarif edilen yöntem kullanılmıştır. Birinci farklılık dizin oluşturma; havuzlama validasyonunda çalışılan materyal örnek havuzları iken pozitif örneklerin tespitinde örnekler teker teker çalışılmıştır. İkinci farklılık çalışılan bölge; havuzlama validasyonunda sadece LDLR geninin 4. ekzonu dizilenirken pozitif örneklerin tespitinde her bir örnek için hangi genin hangi ekzonunda/ekzonlarında pozitiflik tespit edilmişse o amplicon dizilenmiştir, bu sebeple PCR için bölüm 4.2.2.1.2.'de tarif edilen yöntem kullanılmıştır. Üçüncü farklılık ise analiz; havuzlama validasyonunda havuzlama sonucunda gereken hassasiyet artırımı için 0,02 ye düşürülen allel frekans eşik değeri pozitif örneklerin tespitinde sadece heterozigot/homozigot ayrımı için yeterli olan 0,2 de tutulmuştur.

5. ARAŞTIRMA BULGULARI

5.1. HAVUZLAMA VALİDASYONU

5.1.1. JEL ELEKTROFOREZİ

LDLR geni 4. ekzon için yapılan PCR'ların agaroz jel elektroforezi sonunda elde edilen örnek jel görüntüsü Şekil 5.1.'de gösterilmiştir; örnek jelde 48 adet PCR ürünü 100 bp'lik DNA merdiveni ile yürütülmüştür.



Şekil 5.1. Havuzlama çalışması örnek jel görüntüsü

5.1.2. NGS VERİ ANALİZİ

Analiz sonrası tespit edilen pozitif havuzların ve bilinen örnek genotiplerinin dizin tablosuna işlenmiştir. Dizin tablosunda genotipi bilinen pozitif örnekler kırmızı çerçeveye alınmıştır; üst kısma bilinen genotip, alt kısma da X ve Y boyutlarında tespit edilen mutasyon yazılmıştır. Aynı dizin içine birden fazla anı mutasyona sahip örnek düşmesi durumunda oluşan kesişim noktaları birden fazla olduğundan, sadece 2D08 ve 2D16 grupları için alt hücrelerin hepsine mutasyon isimleri yazılmamış, bunun yerine mutasyonlar renk kodları ile belirtilmiştir. Diğer gruplar için -kesişim noktalarının sayısı çok fazla olduğundan- böyle bir renk kodu yapılmamıştır. Tekrarlanabilirlik çalışması için rastgele dağıtılmayıp belirli yerler

yerleştirilen örnekler ise tabloda sarı ve turuncu olarak işaretlenmiştir. Her 2 boyutun birinde dahi tespit edilemeyen örnekler kırmızı olarak işaretlenmiştir

5.1.2.1. 2D08

2D08 çalışması için hazırlanan dizin tablosu EK 2’de gösterilmiştir.

Çalışma ile ilgili istatistikler Çizelge 5.1.’de verilmiştir.

Çizelge 5.1. 2D08 analiz istatistikleri

Toplam örnek	512
Toplam havuz	128
Pozitif havuz	74
Pozitif örnek	54
Tespit edilen pozitif örnek	54 (%100)
Yapılması gereken dizileme	512
Yapılan dizileme	196 (%38)
Maksimum kapsam	493
Minimum kapsam	37.830
Ortalama kapsam	8.839
Maksimum MAF (homozigot)	0,28
Maksimum MAF (heterozigot)	0,20
Minimum MAF	0,02
Ortalama MAF	0,10

2D08 çalışması sonucunda, havuzlama yaklaşımı kullanılmadan yapılması gereken dizileme sayısının %38’i yapılarak pozitif örneklerin %100’ü tespit edilebilmiştir.

5.1.2.2. 2D16

2D16 çalışması için hazırlanan dizin tablosu EK 3'te gösterilmiştir.

Çalışma ile ilgili istatistikler Çizelge 5.2.'de verilmiştir.

Çizelge 5.2. 2D16 analiz istatistikleri

Toplam örnek	512
Toplam havuz	64
Pozitif havuz	55
Pozitif örnek	54
Tespit edilen pozitif örnek (2 boyuttan en az birinde)	53 (%98)
Tespit edilen pozitif örnek (2 boyutta da)	51 (%94)
Yapılması gereken dizileme	512
Yapılan dizileme	176 (%34)
Maksimum kapsam	43152
Minimum kapsam	1274
Ortalama kapsam	10291
Maksimum MAF (homozigot)	0,15
Maksimum MAF (heterozigot)	0,10
Minimum MAF	0,02
Ortalama MAF	0,06

2D16 çalışması sonucunda, havuzlama yaklaşımı kullanılmadan yapılması gereken dizileme sayısının %34'ü yapılarak pozitif örneklerin %94'ü tespit edilebilmiştir.

5.1.2.3. 2D24

2D24 çalışması için hazırlanan dizin tablosu EK 4'te gösterilmiştir.

Çalışma ile ilgili istatistikler Çizelge 5.3.'te verilmiştir.

Çizelge 5.3. 2D24 analiz istatistikleri

Toplam örnek	576
Toplam havuz	48
Pozitif havuz	47
Pozitif örnek	62
Tespit edilen pozitif örnek (2 boyuttan en az birinde)	59 (%95)
Tespit edilen pozitif örnek (2 boyutta da)	47 (%76)
Maksimum kapsam (filtre sonrası)	5168
Minimum kapsam (filtre sonrası)	624
Ortalama kapsam (filtre sonrası)	2456
Maksimum MAF (homozigot)	0,087
Maksimum MAF (heterozigot)	0,067
Minimum MAF	0,004
Ortalama MAF	0,033
Tespit edilen yalancı pozitif, X eksen	11
Tespit edilen yalancı pozitif, Y eksen	15
Tespit edilen yalancı pozitif, X-Y eksenleri ortak	0

2D24 çalışması sonucunda, havuzlama yaklaşımı kullanılarak yapılması gereken dizileme sayısı, hassasiyetin düşük ve dizin içindeki pozitif örnekler -tarama testinde olmaması gerekecek kadar- fazla olması sebebi ile hesaplanmamıştır. Pozitif örneklerin %76'sı tespit edilebilmiştir.

5.1.2.4. 2D32

2D32 çalışması için hazırlanan dizin tablosu EK 5'te gösterilmiştir

Çalışma ile ilgili istatistikler Çizelge 5.4.'te verilmiştir.

Çizelge 5.4. 2D32 analiz istatistikleri

Toplam örnek	1024
Toplam havuz	64
Pozitif havuz	64
Pozitif örnek	108
Tespit edilen pozitif örnek (2 boyuttan en az birinde)	90 (%83)
Tespit edilen pozitif örnek (2 boyutta da)	30 (%28)
Toplam heterozigot örnek	94
Toplam homozigot örnek	14
Toplam tespit edilen heterozigot örnek	20 (%21)
Toplam tespit edilen homozigot örnek	10 (%71)
Maksimum kapsam (filtre sonrası)	6609
Minimum kapsam (filtre sonrası)	1467
Ortalama kapsam (filtre sonrası)	3688
Maksimum MAF (homozigot)	0,058
Maksimum MAF (heterozigot)	0,041
Minimum MAF	0,003
Ortalama MAF	0,028
Tespit edilen yalancı pozitif, X eksen	23
Tespit edilen yalancı pozitif, Y eksen	19
Tespit edilen yalancı pozitif, X-Y eksenleri ortak	0

2D32 çalışması sonucunda, havuzlama yaklaşımı kullanılarak yapılması gereken dizileme sayısı, hassasiyetin düşük ve dizin içindeki pozitif örnekler -tarama testinde olmaması gerekecek kadar- fazla olması sebebi ile hesaplanmamıştır. Pozitif örneklerin %21'i tespit edilebilmiştir.

5.1.2.5. 3D08

3D08 çalışması için hazırlanan havuzlar 64 örnekten ve 128 allelden oluştuğundan dolayı elde edilen sonuçlarda bu seviyede bir analiz yapabilmek için çok az gerçek pozitif ve çok fazla yalancı pozitif olduğu görülmüş ve analiz yapılmamıştır.

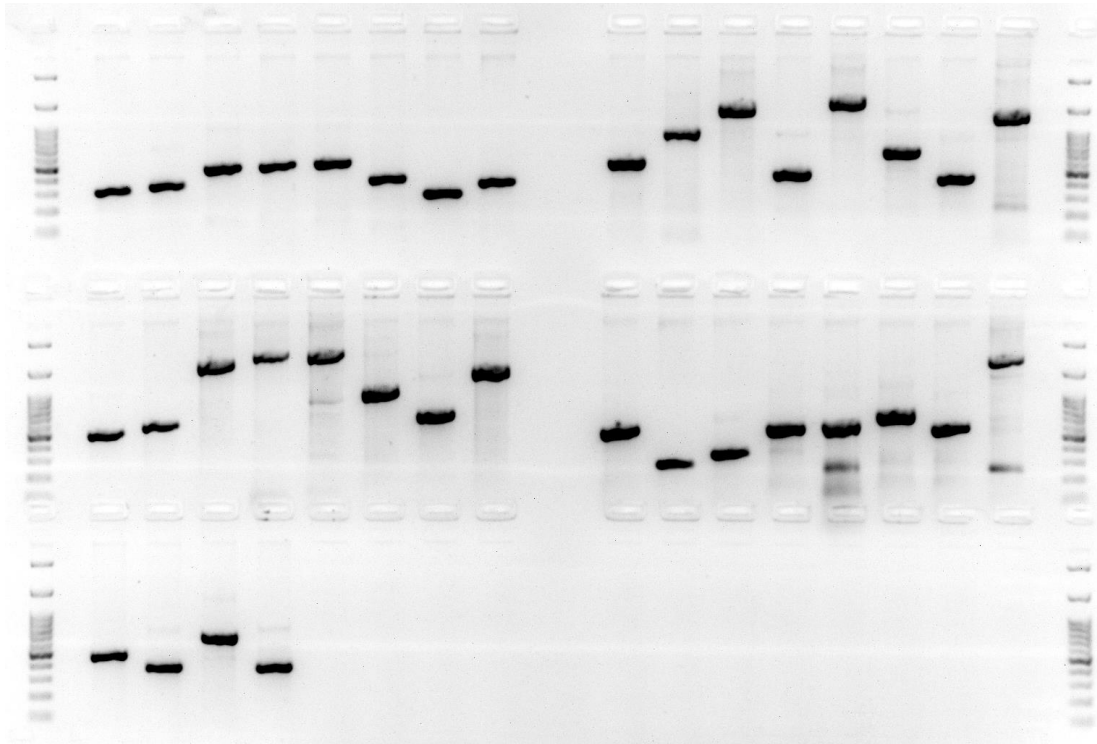
5.1.2.6. TEKRARLANABİLİRLİK ÇALIŞMASI

Çalışmada tekrarlanabilirlik çalışması için 2 örnek rastgele dağıtılmamış, 4'er belirli pozisyona olacak şekilde toplamda 8 pozisyona yerleştirilmiştir. 2D32 çalışmasında çok fazla tespit edilemeyen mutasyon olduğu için (13/32) tekrarlanabilirlik çalışmasından çıkartılmıştır. Elde edilen dizileme verisindeki allel frekansları EK 6'da verilmiştir.

5.2. FAMILİYAL HİPERKOLESTEROLEMİ HASTA GRUBU

5.2.1. JEL ELEKTROFOREZİ

FH hasta grubu için yapılan PCR'ların agaroz jel elektroforezi sonunda elde edilen örnek jel görüntüsü Şekil 5.2.'de gösterilmiştir; tek bir havuz için yapılan 36 adet PCR ürünü 100 bp'lik DNA merdiveni ile yürütülmüştür.



Şekil 5.2. FH hasta gurubu çalışması örnek jel görüntüsü

5.2.2. NGS VERİ ANALİZİ

Analizde tespit edilen 8 adet negatif havuz, 72 adet pozitif havuz, bulunan mutasyonlar ve frekansları EK 7’de verilmiştir. Analize ait istatistikler Çizelge 5.5.’te gösterilmiştir.

Çizelge 5.5. FH hasta grubu NGS veri analizi istatistikler

Toplam hasta sayısı	256
Toplam kontrol sayısı	64
Toplam örnek sayısı	320
Gen sayısı	6
Havuzlama	2 Boyut, 8x8’lik
Toplam havuz	80
Pozitif havuz	72
Toplam mutasyon	150
Toplam mutasyon, LIPA	10 (%7)
Toplam mutasyon, LDLRAP1	68 (%45)
Toplam mutasyon, APOA2	0 (%0)
Toplam mutasyon, ITIH4	44 (%29)
Toplam mutasyon, GHR	24 (%16)
Toplam mutasyon, PPP1R17	4 (%3)
Maksimum MAF	0,66
Minimum MAF	0,03
Ortalama MAF	0,12
Toplam mutasyon çeşidi	40
Toplam mutasyon çeşidi, LIPA	2 (%5)
Toplam mutasyon çeşidi, LDLRAP1	10 (%25)
Toplam mutasyon çeşidi, APOA2	0 (%0)
Toplam mutasyon çeşidi, ITIH4	18 (%45)
Toplam mutasyon çeşidi, GHR	9 (%23)
Toplam mutasyon çeşidi, PPP1R17	1 (%2)

Tespit edilen toplamda 40 çeşit mutasyon, genomik pozisyonları, dbSNP veritabanı kodları, 1000 Genom Projesi faz 3 verilerine göre popülasyon frekansları ve ClinVar veri tabanına göre klinik önemleri EK 8’de gösterilmiştir.

5.2.3. POZİTİF ÖRNEKLERİN TESPİTİ

NGS analizi sonucunda seçilen 111 örnekten yapılan pozitiflik tespiti, doğrulama ve zigozite belirleme çalışmasına ait istatistikler Çizelge 5.6.’da verilmiştir.

Çizelge 5.6. Pozitif örneklerin tespiti çalışması istatistikleri

Seçilen örnek sayısı	111
Çalışılan toplam bölge	134
Pozitif bulunan örnek	75
Toplam mutasyon	89
Heterozigot mutasyon	63
Homozigot mutasyon	26

FH hasta grubuna ait toplam 4 adet 8x8'lik dizinin sonuçları Şekil 5.3. ile Şekil 5.6. arasında gösterilmiştir. Kontrol grubuna ait 1 adet 8x8'lik dizinin sonuçları ise Şekil 5.7.'de gösterilmiştir. Pozitif örnekler ve taşıdıkları mutasyonlara ait hücreler renkli işaretlenerek belirtilmiştir. İşaretsiz olmayan ancak sağ üst kısımlarında kırmızı belirteç taşıyan örnekler çalışma sonunda mutasyonu taşımadığı tespit edilen örneklerdir. Birden fazla renk ile işaretlenmiş örnekler farklı genlerinde birden fazla mutasyon tespit edilmiş örneklerdir.

L001	L002	L003	L004	L005	L006	L007	L008	LIPA	I252L
								LIPA	Q298Q
L009	L010	L011	L012	L013	L014	L015	L016	LDLRAP1	I74M
								LDLRAP1	IVS3+1G>T
L017	L018	L019	L020	L021	L022	L023	L024	LDLRAP1	Q136X
								LDLRAP1	IVS4-1G>A
L025	L026	L027	L028	L029	L030	L031	L032	LDLRAP1	R238W
L033	L034	L035	L036	L037	L038	L039	L040	ITIH4	R56X
								ITIH4	E392E
L041	L042	L043	L044	L045	L046	L047	L048	ITIH4	IVS12+5G>A
								GHR	L7L
L049	L050	L051	L052	L053	L054	L055	L056	GHR	R229H
L057	L058	L059	L060	L061	L062	L063	L064	GHR	A613A
								GHR	L622L

Şekil 5.3. FH hasta gurubu 1. 8x8'lik dizin çalışması sonuçları

L065	L066	L067	L068	L069	L070	L071	L072	LDLRAP1	I74M
L073	L074	L075	L076	L077	L078	L079	L080	LDLRAP1	Q136X
L081	L082	L083	L084	L085	L086	L087	L088	LDLRAP1	S202H
L089	L090	L091	L092	L093	L094	L095	L096	LDLRAP1	E217X
L097	L098	L099	L100	L101	L102	L103	L104	LDLRAP1	R238W
L105	L106	L107	L108	L109	L110	L111	L112	LDLRAP1	V290V
L113	L114	L115	L116	L117	L118	L119	L120	ITI4	E332K
L121	L122	L123	L124	L125	L126	L127	L128	GHR	M530T
								PPP1R17	V143M

Şekil 5.4. FH hasta gurubu 2. 8x8'lik dizin çalışması sonuçları

L129	L130	L131	L132	L133	L134	L135	L136	LIPA	I252L
L137	L138	L139	L140	L141	L142	L143	L144	LDLRAP1	S202H
L145	L146	L147	L148	L149	L150	L151	L152	LDLRAP1	R238W
L153	L154	L155	L156	L157	L158	L159	L160	ITI4	D87D
L161	L162	L163	L164	L165	L166	L167	L168	ITI4	G249R
L169	L170	L171	L172	L173	L174	L175	L176	ITI4	V327L
L177	L178	L179	L180	L181	L182	L183	L184	ITI4	N333N
L185	L186	L187	L188	L189	L190	L191	L192	GHR	R229H

Şekil 5.5. FH hasta gurubu 3. 8x8'lik dizin çalışması sonuçları

L193	L194	L195	L196	L197	L198	L199	L200	LIPA	I252L
L201	L202	L203	L204	L205	L206	L207	L208	LDLRAP1	c.246delG
L209	L210	L211	L212	L213	L214	L215	L216	LDLRAP1	Q86X
L217	L218	L219	L220	L221	L222	L223	L224	LDLRAP1	Q136X
L225	L226	L227	L228	L229	L230	L231	L232	LDLRAP1	R238W
L233	L234	L235	L236	L237	L238	L239	L240	ITIH4	E332K
L241	L242	L243	L244	L245	L246	L247	L248	ITIH4	T832M
L249	L250	L251	L252	L253	L254	L255	L256	ITIH4	T832M
								GHR	R179C
								GHR	V499M
								GHR	D564E

Şekil 5.6. FH hasta gurubu 4. 8x8'lik dizin çalışması sonuçları

L257	L258	L259	L260	L261	L262	L263	L264	LDLRAP1	E217X
L265	L266	L267	L268	L269	L270	L271	L272	LDLRAP1	R238W
L273	L274	L275	L276	L277	L278	L279	L280	ITIH4	L150L
L281	L282	L283	L284	L285	L286	L287	L288	ITIH4	G249R
L289	L290	L291	L292	L293	L294	L295	L296	ITIH4	E332K
L297	L298	L299	L300	L301	L302	L303	L304	ITIH4	M358I
L305	L306	L307	L308	L309	L310	L311	L312	ITIH4	G418S
L313	L314	L315	L316	L317	L318	L319	L320	ITIH4	R689C
								ITIH4	W850C
								ITIH4	T896T
								ITIH4	T586T
								ITIH4	P919P
								GHR	E62K

Şekil 5.7. Kontrol gurubu 8x8'lik dizin çalışması sonuçları

Tespit edilen 89 adet genotip ve örnek sayıları Çizelge 5.7.'de verilmiştir.

Çizelge 5.7. Genotipler ve örnek sayıları

Gen	Mutasyon	Zigotite	Adet
LDLRAP1	Q136X	homozigot	19
LDLRAP1	R238W	heterozigot	14
LIPA	I252L	heterozigot	5
GHR	R179C	heterozigot	3
ITIH4	E332K	heterozigot	3
LDLRAP1	S202H	heterozigot	3
GHR	R229H	heterozigot	2
GHR	V499M	heterozigot	2
ITIH4	G249R	heterozigot	2
ITIH4	P919P	heterozigot	2
LDLRAP1	E217X	heterozigot	2
LDLRAP1	I74M	homozigot	2
LDLRAP1	R238W	homozigot	2
GHR	A613A	heterozigot	1
GHR	D564E	heterozigot	1
GHR	E62K	heterozigot	1
GHR	L622L	heterozigot	1
GHR	L7L	heterozigot	1
GHR	M530T	heterozigot	1
ITIH4	D87D	heterozigot	1
ITIH4	D892D	heterozigot	1
ITIH4	E392E	heterozigot	1
ITIH4	G418S	heterozigot	1
ITIH4	IVS12+5G>A	heterozigot	1
ITIH4	L150L	heterozigot	1
ITIH4	M358I	heterozigot	1
ITIH4	N333N	heterozigot	1
ITIH4	R56X	heterozigot	1
ITIH4	R689C	heterozigot	1
ITIH4	T586T	heterozigot	1
ITIH4	T832M	heterozigot	1
ITIH4	T896T	heterozigot	1
ITIH4	V327L	heterozigot	1
ITIH4	W850C	heterozigot	1
LDLRAP1	c.246delG	homozigot	1
LDLRAP1	IVS3+1G>T	homozigot	1
LDLRAP1	IVS4-1G>A	homozigot	1
LDLRAP1	Q86X	heterozigot	1
LDLRAP1	V290V	heterozigot	1
LIPA	Q298Q	heterozigot	1
PPP1R17	V143M	heterozigot	1

Pozitif bulunan örnekler ve genotipleri EK 9’da verilmiştir.

Çalışma sonunda 320 örneğin 6 geni taranmış ve 75 örnekte toplamda 89 adet mutasyon tespit edilmiştir. Çalışma havuzlama yaklaşımı kullanılmadan yapılmış olsaydı 320 örnek için toplamda 11.520 adet amplikonun dizilenmesi gerekirken, havuzlama yaklaşımı sayesinde 80 havuzun taranması için 2.880 ve havuzlar içindeki pozitif örneklerin tespiti için 134 olacak şekilde toplamda 3.014 amplikon dizileme ile, toplamda dizilenmesi gereken amplikonun sadece %26’sı yapılarak çalışma tamamlanabilmiştir.



6. TARTIŞMA VE SONUÇ

6.1. TARTIŞMA

6.1.1. DNA DİZİLEME

DNA dizileme, teknolojinin çok hızlı geliştiği son 50 yıl içerisinde ilk önce keşfedilmiş, daha sonra teknolojinin bu hızına ayak uydurarak son derece hızlı gelişmiş ve günümüzdeki halini almıştır. İlk keşfedildiği yıllarda “insan genomu mutlaka bir gün dizilenecek, ancak bu yüz yıl, belki de bin yıl sürecektir” şeklinde düşünen bazı araştırmacılara inat, yaklaşık 25 yıl sonra İnsan Genom Projesi’nin ilk ayağı tamamlanmış, yaklaşık 45 yıl sonra ise de telomerden-telomere tüm insan genomun kesintisiz dizilimi tamamlanabilmiştir (223). Bu hızlı gelişimin sebebi, 1. nesil olarak kabul edilen Sanger dizileme ve 2. nesil olan NGS dizilemenin kendi içlerinde logaritmik olarak gelişmesi ve bir teknolojiden diğer teknolojiye geçilirken gerçekleşen sıçrama ile logaritmik gelişmenin de ötesine geçilmesi ile mümkün olmuştur. 2. nesil hala logaritmik olarak ilerlemeye devam etmektedir, 3. nesil ise gelişmeye devam etse de diğer ikisine göre daha yavaş gelişmektedir ve görünüşe göre etkisi 1. ve 2. nesil kadar dramatik olmayacaktır.

NGS dizileme, özellikle insan genetik hastalıklarının tanımlanması, anlaşılması, önlenmesi, tedavi olanaklarının geliştirilmesi vb. alanlarda geçmişte hayal dahi edilemeyecek bir etkide bulunarak insanlığa çok büyük bir katkı sağlamıştır. Kısa bir süre öncesine kadar pahalı ve zor olan genetik çalışmalara ait kaynakların insan için sadece hastalar ve hastalıklar için kullanılmasının gerektiği, bu kaynakların sağlıklı insanlar üzerinde çalışma yapılmamasını gerektirecek kadar kıymetli olduğunu düşünen bir çok araştırmacının fikri, NGS dizilemenin işleri çok daha ucuz, kolay, hızlı ve ulaşılabilir hale getirmesi sebebi ile hızla değişmiştir ve değişmektedir. Bu bağlamda son yıllarda artan nutrigenetik, sporcu genetiği vb. hastalık odağında olmayan çalışmalar dışında, her ne kadar hastalık odaklı olsa da hasta odaklı olmayan kalıtsal hastalık taşıyıcı tespiti çalışmaları da hem yaygınlaşmış hem de insanların hizmetine sunulmaya başlanmıştır; birçok merkez özellikle yeni evlenen veya evlenmeyi düşünen sağlıklı çiftler için geniş kapsamlı taşıyıcılık tarama testlerini insanların hizmetine sunmaya başlamışlardır. Her ne kadar bu tip testler çoğu birey tarafından karşılanabilecek maliyetlerde olsa da geniş kapsamlı popülasyon taşıyıcılık tarama programları için hala çok pahalı testlerdir. Bu noktada havuzlama yaklaşımı ile birleştirilen NGS dizileme

teknolojisinin yakın gelecekte bu maliyet problemini çözerek, geniş kapsamlı toplum taramalarının devletler tarafından geniş kitleler üzerinde uygulanmasına olanak sağlayarak taşıyıcı bireylerin tespit edilmesine ve hastalıkların oluşmadan önce engellenmesine olanak sağlayacağı öngörülebilir.

6.1.2. HAVUZLAMA VALİDASYONU ÇALIŞMASI

Bu çalışma ile havuzlama yaklaşımı ve NGS dizileme teknolojisinin beraber kullanılabilirliği gösterilerek sınırlarının belirlenmesi amaçlanmıştır. Bunda kısmen başarılı olunmuştur; her ne kadar daha önce genotiplendirilmiş çok sayıda örnek kullanılarak yapılan çalışma sonucunda güvenli tarama çalışmaları yapılabileceği gösterilmiş olsa da sınırları konusunda ancak kaba bir fikir elde edilebilmiştir. Çalışma sonunda 8'li havuzlar oluşturmanın hassasiyeti çok yüksek olarak değerlendirilmiş (%100), 16'lı havuzlar oluşturmanın hassasiyeti çok yüksek olarak bulunmasa da (%94 - %98) yine de bir tarama çalışması için kabul edilebilir seviyede olduğu sonucuna varılmış, 24'lü havuzlar oluşturmanın düşük hassasiyet (%95 - %76) ve yüksek yalancı pozitif sonuç vermesi sebebi ile fayda ve zarar değerlendirilmesi çok iyi yapılmak sureti ile ancak kullanılabileceği sonucuna varılmış, 32'li ve daha yüksek havuzlar oluşturmanın ise, hassasiyetin tarama çalışması için bile olsa kabul edilebilir bir seviyede olmadığı (<%71) için tamamen mantıksız olduğu sonucuna varılmıştır.

Kısacası bu sonuca göre 8-16'lı havuzlar üzerine çıkmamak gerekmektedir. Ancak bu çalışma yüksek hassasiyet oluşturacak koşullar sağlanarak yapılan bir çalışma olmak yerine; NGS teknolojisine sahip herhangi bir laboratuvarın gerçekleştirebileceği şekilde basit, hızlı, düşük çalışma ve tasarım maliyetli olacak şekilde tasarlanmış bir çalışmadır. Bu sebeple teknolojinin sınırını 8-16'lı havuzlama şeklinde düşünmek çok yanlış olacaktır; bir takım geliştirme ve çalışmaya eklenen bazı yaklaşımlar ile elde edilebilecek çok daha hassas bir çalışma sayesinde, çok daha yüksek sayıda havuzlar oluşturularak güvenli tarama yapmak mümkündür. Bu çalışmanın hassasiyetini düşürdüğü düşünülen sebepler, yapılabilecek geliştirmeler ve sürece eklenebilecek yaklaşımlar bu çalışmada tercih edilmeme sebepleri ile birlikte aşağıda tartışılmıştır.

6.1.2.1. DAHA DENGELİ HAVUZLARIN OLUŞTURMASI

Havuzlamanın başarısının artırılması için ön koşul dengeli havuzların oluşturulması, yani havuzlar içindeki tüm örneklerin eşit miktarda temsil edilmesidir. Şayet bu sağlanamazsa bundan sonraki aşamalar hassasiyeti ne kadar arttıracak önlemler içerse de başarı düşük olacaktır; havuzlar içinde az miktarda temsil edilen pozitif örneklerin tespit edilme şansı düşük olacaktır.

Genetik çalışmalar için iki seviyede havuz oluşturma yapılabilir; kan veya DNA izolatu. En dengeli havuzlar DNA'ları aynı yöntem ile izole edilip, nitelik ve nicelik olarak değerlendirilen örneklerin karıştırılması ile oluşturulabilir; nitelikleri benzer olan örnekler birbiri ile karıştırılacak şekilde ve nicelikleri göz önünde bulundurularak karıştırılırsa çok dengeli havuzlar oluşturmak mümkündür. Ancak bu hem maliyeti hem de iş gücünü arttırıcı bir faktördür; bu durumda her örnekten DNA izolasyonu yapıp, bir de DNA izolatlarının hem nicelik hem de nitelik olarak değerlendirilmesi gerekmektedir. Bu da bir tarama testinde olması gereken pratiklik ve düşük maliyet özelliğine aykırıdır. Bu sebeple -her ne kadar bir miktar denge problemi oluştursa da- tarama amaçlı genetik bir havuzlama çalışmasında örneklerin kan seviyesinde karıştırılarak havuzların oluşturulmasının daha uygun olacağı düşünülebilir (DNA seviyesinde karıştırmak yerine kan seviyesinde karıştırıp, oluşacak dengesizliğin telafisi için daha düşük örnek içeren havuzlama yaklaşımı kullanmak çoğu zaman çok daha verimli olacaktır).

Kanların karıştırılması noktasında dikkat edilebilecek ve/veya yapılabilecek şeyler vardır. Mesela farklı koşullarda saklanmış/muamele edilmiş kanların karıştırılmaması (eski kanlar ile yeni kanların veya dondurulmuş kanlar ile dondurulmamış kanların karıştırılmaması gibi) veya taze kanların beyaz küre sayımı yapılarak ve örneklerin havuz içindeki miktarlarının, kandaki DNA'nın nerdeyse tamamını oluşturan beyaz küre sayımlarına göre ayarlandığı bir havuz oluşturma metodu kullanılabilir. Beyaz küre sayımı da maliyeti arttırıcı bir etken olmakla beraber, DNA izolasyonu ile birlikte DNA nicelik ve nitelik tayinine göre daha ucuz ve pratik bir yöntemdir.

Çalışmada şayet sadece DNA'nın izole edilmesi ancak nitelik ve nicelik açısından değerlendirilmemesi planlanıyorsa, bunun yerine kan seviyesinde karıştırma yapmak daha uygundur; çünkü kan örnekleri (özellikle aynı şekilde saklanmış/muamele edilmiş ise) DNA

izolatlarına göre çok daha standarttır. Sağlıklı bir bireyde kandaki beyaz küre miktarı mikrolitre başına $4-10 \times 10^4$ seviyesindedir ve çoğu durumda iki sağlıklı bireyin beyaz küre miktarları arasında çok az fark vardır (yapılan bir çalışmada 34 binden fazla sağlıklı bireyin beyaz küre ortalaması $5,6 \times 10^4$ ve standart sapma $1,4 \times 10^4$ olarak bulunmuştur, yani sadece %25'lik bir standart sapma vardır (224)). Ancak DNA izole edildiği zaman örnekler arasındaki nicelik ve nitelik farkları daha fazla açılabilir; bu sebeple şayet DNA izolatları kullanılacaksa mutlaka nicelik ve nitelik tayini yapılmalı, şayet ikisi birden yapılamıyorsa nicelik tayini mutlaka yapıp karıştırılacak örnekler ve karıştırma miktarları ona göre belirlenmelidir.

Validasyon çalışmasında (FH grubu hasta çalışmasında da aynı şekilde) pratiklik ve düşük maliyete daha fazla önem verildiği için DNA izolatları kullanılması hiç düşünülmemiştir. Ancak bu tez çalışmasında kullanılan örneklerin tamamı yaklaşık 2 yıllık bir süre içerisinde toplanmış ve havuzlar örnekler toplandıktan yaklaşık 3 yıl sonra oluşturulmuş olduğu için örnekler arasında ciddi bir fark oluşmuştur. Bu süre zarfında kan örneklerinin oda sıcaklığında saklanmış olması da bu farkı daha fazla açan bir etken olmuş olabilir. Şayet bunun yerine çok daha benzer koşullarda saklanmış kan örnekleri (mümkünse taze kan örnekleri) kullanılmış olsaydı daha dengeli havuzların oluşturulabileceği ve bulunan hassasiyetin daha yüksek olabileceği düşünülmektedir.

6.1.2.2. PCR

PCR yönteminin 1985 yılında geliştirilmesi, moleküler genetik alanındaki en önemli dönüm noktalarından biridir. Genomik DNA'yı -çok büyük bir molekül olduğu için- yüksek kopya sayılarında elde etmek mümkün değildir; 1 mol insan haploid genomu yaklaşık 1 milyon tondur (300 m'lik 4-5 adet yük gemisinin toplam ağırlığı). Bu da Genomik DNA'dan yapılan çalışmaların çok az sayıda molekülden yapılma sorununu beraberinde getirmektedir. Bu soruna PCR yöntemi çok ideal bir çözüm getirmiştir; PCR genomik DNA'nın sadece ilgilenilen bölge veya bölgelerinin kopya sayısını milyonlarca/milyarlarca arttırarak yapılacak tespiti çok daha kolay hale getirmektedir. Keşfedildiği ilk yıllardan itibaren hızla yaygınlaşmıştır ve günümüzde içerisinde PCR veya PCR benzeri bir aşama olmayan moleküler genetik bir çalışma neredeyse hiç kalmamıştır.

PCR'in avantajları çok fazla olmasına rağmen dezavantajlar da vardır. Bu dezavantajlar aşağıda tartışılmıştır.

6.1.2.2.1. POLİMERAZ HATALARI

DNA polimeraz enzimleri az da olsa hata yapabilmektedir; yani DNA replikasyonu sırasında olmaması gereken bir nükleotidi olmaması gereken bir yere ekleyebilmektedir. Bu hata oranları her ne kadar polimerazdan polimeraza ve reaksiyon koşullarına göre değişiklik gösterse de genel olarak 10^{-6} ile 10^{-4} arasında bir değerdedir (225); yani DNA polimeraz enzimleri yaklaşık olarak 10 bin ile 1 milyonda 1 nükleotidi yanlış ekleyebilmektedirler. Bu oran çok az olmasına rağmen genomik DNA'nın içermiş olduğu nükleotid sayısı göz önüne alındığı zaman aslında çok yüksektir. Ancak hücre içinde bulunan DNA tamir mekanizmaları bu tip DNA polimeraz hatalarının tamamına yakınıni tespit edip düzeltir; bu şekilde hata oranı 10^{-9} seviyelerine kadar düşürülmüş olur (226).

Hücre dışı reaksiyon ortamlarında DNA tamir mekanizmaları olmadığı için DNA polimeraz kaynaklı hatalar direkt olarak sonuca yansıyabilmektedir. Her ne kadar 10 binde 1 hata hala çok düşük bir oran gibi görünse de (%0,01) aslında böyle değildir; çünkü bir öndeki döngüde yeni oluşturulan DNA zincirinin bir sonraki döngüde kalıp olarak kullanıldığı zincirleme reaksiyon mantığına göre yapılan hatalar da döngüler boyunca birikerek artar. Bu şekilde biriken hatalar sonucunda kullanılan polimeraz enziminin özelliğine ve reaksiyon koşullarına göre hata oranı 10^{-2} veya daha yüksek seviyelere bile çıkabilmektedir. Sadece PCR kaynaklı hata oranı bu kadar yüksek olan bir çalışmada ise havuzlama yaklaşımını kullanmak mümkün değildir.

PCR sırasında oluşan polimeraz hatalarını minimuma indirmek için yapılacak iki temel şey vardır; hata oranı düşük, hata düzetme özelliğine sahip (İng. proofreading) DNA polimeraz enzimi kullanmak ve hataların birikmesini engelleme amaçlı olarak düşük döngü sayısına sahip ısı döndürücü protokolleri kullanmak.

Validasyon çalışmasında pratiklik ve maliyet göz önünde bulundurulduğu için DNA izolasyonu olmadan kandan direkt PCR yapılmıştır. Ancak kandan direkt PCR yapabilmek için bu amaç için özel olarak geliştirilmiş, tam kan içerisinde bulunan ve PCR'ı engelleyen inhibitörlere dayanıklı bir DNA polimeraz enzimi kullanılması gerekmektedir; ki bu

enzimlerden de çok fazla yoktur, seçenek azdır. Çalışmada kullanılan ve kandan direkt PCR yapmaya uygun bir DNA polimeraz olan Phire II DNA polimeraz enziminin de hata düzeltme özelliği yoktur. Şayet kandan direkt PCR yapılmasaydı hata düzeltme özelliğine sahip bir DNA polimeraz enzimi seçilerek hassasiyet seviyeleri arttırılabilirdi.

6.1.2.2.2. ALLEL DROP-OUT

PCR'in bir dezavantajı da allel drop-out olabilmesi, yani PCR reaksiyon karışımı içerisine eklenen kalıp DNA'daki bir allelin, PCR sonunda elde edilen amplikonlar içerisinde hiç bulunmaması veya oransal olarak daha az bir oranda temsil edilmesi durumudur (227). PCR bir zincirleme reaksiyon olduğu için başlangıç durumundaki küçük farklılıkların reaksiyon sonucunda büyük farklara dönüşmesi potansiyelini barındırmaktadır. Genomik DNA'nın PCR ile amplifiye edilecek bölgesi farklı allellerde birbirine çok benzese de farklı allelerin en az bir noktasında farklılık bulunur (yoksa zaten farklı allel olmazlardı). Bu farklılık her ne kadar küçük olsa da bazı allellerin verimli bir şekilde amplifiye olmasını engelleyerek reaksiyon sonucunda tespit limitlerinin altında kalacak şekilde azalmasına sebep olabilir. Allel drop-out riskini minimize etmek için bazı önlemler alınabilse de (iyi bir primer tasarımı, ilerleyiciliği (İng. processivity) yüksek bir DNA polimeraz kullanmak, daha düşük döngü sayısına sahip ısı döndürücü protokolü kullanmak vb.) bu riski tamamen ortadan kaldırmak mümkün değildir.

Validasyon çalışmasında primer tasarımı kısmında özen gösterilmiş ve ilerleyiciliği yüksek bir DNA polimeraz enzimi kullanılmıştır. Ancak amplikon dizileme yaklaşımı seçildiği için kullanılan toplam PCR döngü sayısı yüksektir. Allel drop-out etkisini incelemek başlı başına ayrı bir çalışma olduğu için bu çalışma içerisine dahil edilmemiştir, bu sebeple etkisinin olup olmadığını söylemek zordur. Ancak özellikle yüksek örnek içeren havuzlardaki hassasiyet düşüşünün bir kısmının sebebi bu olabilir.

6.1.2.2.3. KONTAMİNASYON

PCR'in bir diğer dezavantajı da yüksek kontaminasyon riski içermesidir. Bu da aynı allel drop-out dezavantajında olduğu gibi PCR'in zincirleme reaksiyon olmasından kaynaklanmaktadır; başlangıçta veya reaksiyon sırasında reaksiyon içerisine giren bir kontaminasyon kaynağı, zincirleme reaksiyon mantığı ile reaksiyon sırasında yüksek miktarda amplifiye edileceği için, reaksiyon sonunda oluşan amplikonların çalışmayı

etkileyecek bir kısmının veya tamamının kontaminasyon kaynaklı olma ihtimali vardır (228).

PCR kontaminasyonunun -her ne kadar tek sebebi olmasa da- en büyük sebebi daha önceki PCR'lar sonucu üretilmiş olan ampliconlardır; ampliconlar PCR sırasında milyonlarca/milyarlarca kopya üretildikleri ve aynı zamanda PCR reaksiyonunda kalıp olarak kullanıldıkları için, daha önce tamamlanmış bir PCR reaksiyonunun çok çok küçük bir kısmı dahi tamamlanmamış başka bir reaksiyonun içerisine bir şekilde girse, reaksiyonu tamamen kontamine etmesi işten bile değildir.

Yoğunluklu olarak PCR yapılan bir ortam kontamine olarak kabul edilir. Oluşan ampliconlar buharlaşamayacak kadar büyük moleküllerdir (buharlaşacak enerjiye ulaşana kadar çoktan parçalanırlar) ancak DNA tozayabilir; herhangi bir yere sıçrayan ve kuruyan PCR ürünündeki ampliconlar zaman içerisinde oluşan hava akımları ile birlikte koparak havaya karışarak bir nevi amplicon tozu oluştururlar. PCR reaksiyonlarının mikropipet ile bir yerden bir yere aktarılması esnasında oluşabilen mikrovessiküller de aynı şekilde havada amplicon tozu oluştururlar. Şayet dikkat edilmezse ve yeterli önlemler alınmazsa bu amplicon tozları hava yolu ile veya kişiler/malzemeler yolu ile reaksiyon kurulan ortama girip sonra kurulacak olan PCR reaksiyonlarını kontamine etmesi işten bile değildir.

PCR kontaminasyonu riskini minimize etmenin en önemli yolu moleküler genetik çalışmaya uygun olarak tasarlanmış bir laboratuvar, özenli, titiz ve kurallara uyarak çalışmaktan geçer. Düşük döngü sayılı ısıtma/dondurma protokoller kullanılması da kontaminasyon riskinin azaltılmasında faydası vardır. Aynı zamanda mümkün olan durumlarda mutlaka PCR reaksiyonu sırasında negatif (kalıp DNA olmayan) kontrol reaksiyonu kurularak genel bir kontaminasyon problemi olup olmadığı rutin olarak görüntülenmelidir.

PCR kontaminasyonu riskini azaltma amaçlı geliştirilmiş bir yaklaşım vardır; bu yaklaşımda PCR sırasında nükleotidlerden timin nükleotidinin (dTTP) urasil nükleotidi (dUTP) ile değiştirilmesi sonucu ampliconlar içerisinde tamamen veya kısmen urasil bulunması ve yapılacak yeni reaksiyonların urasil DNA glikozilaz (UDG) enzimi (veya urasil nükleotid glikozilaz enzimi, UNG) ile muamele edilmesi prensibi kullanılmaktadır (229). UDG enzimi urasil bazını DNA zincirinden kopartarak reaksiyon içerisinde bulunması muhtemel amplicon kontaminasyon kaynaklarının kalıp olarak kullanılmasını engellerken, kalıp

DNA üzerinde urasil olmadığı için kalıp DNA'nın amplifiye olmasını engellemez. Faydalarına rağmen UDG kullanımının da dezavantajları vardır. Öncelikle her DNA polimeraz dUTP'yi uzayan DNA zincirine verimli olarak ekleyememekte, bu sebeple UDG yaklaşımı kullanılacak çalışmalarda DNA polimeraz enzimi seçenekleri azalmaktadır. Aynı zamanda UDG kullanımı maliyeti arttırmaktadır. Bazı çalışmalarda ise PCR sonrasında ampliconlar içerisinde uracil bulunması, çalışmanın sonraki basamaklarını zorlaştırmaktadır. Bu sebepler ile her çalışmada UDG yaklaşımı tercih edilmemektedir; en çok tercih edildiği alan, çalışılan örnekler arasında çok fazla fark bulunan, dolayısıyla çok ufak bir kontaminasyonun bile sonucu çok ciddi bir şekilde etkileyebileceği ve dinamik aralığın çok yüksek olmasının gerektiği çalışmalardan olan patojen tespit ve kopya sayısı belirleme çalışmalarıdır.

Her türlü önlem alınmış olsa dahi PCR'da kontaminasyon riski her zaman için vardır. Amaç PCR kontaminasyon riskini imkansız olan sifıra indirmek değil, bu riski en ideal bir şekilde yönetebilmektir.

Validasyon çalışması uygun bir laboratuvar da ve titiz olarak yapılarak kontaminasyon riski azaltılmıştır. Aynı zamanda kurulan negatif kontrol PCR reaksiyonlarında herhangi bir amplifikasyon olmadığı için PCR kontaminasyonunun bu çalışmada negatif bir etkisi olmadığı düşünülmektedir. Ancak amplicon dizileme yaklaşımının seçilmiş olması sebebi ile mevcut risk bir miktar daha yüksektir; hassasiyetin düşük çıktığı yüksek örnek içeren havuzlama kısmında kontaminasyonun etkisi olmuş olabilir.

6.1.2.2.4. PCR ÖZET

PCR, yöntem olarak genetik çalışmaların çok büyük bir problemine çözüm olmuş çok güçlü bir yöntemdir ve günümüzde moleküler genetiğin vazgeçilmez araçlarından biridir. Ancak beraberinde getirdiği yukarıda bahsedilen dezavantajlar sebebi ile dikkatli kullanılmalı, bu dikkat miktarı da yapılacak çalışmanın gerektirdiği hassasiyet ve özgüllüğün miktarı ile birlikte orantılı olarak arttırılmalıdır.

PCR'sız NGS kütüphane hazırlama kitleri geliştirilmiş olsa da bu kitlerin bir kısmı 2-3 döngü gibi az miktarda olsa da PCR aşaması içermektedir. Hiç PCR aşaması içermeyen kitlerin bir kısmı da kütüphane hazırlığı sonrası klonal amplifikasyonu aşamasında PCR'a

tabi tutulmaktadır. Tamamen PCR'sız bir çalışma mümkün olmak ile birlikte her çalışmayı bu şekilde yapmak mümkün değildir; tamamen PCR'sız çalışmalarda hem yüksek nicelik ve nitelikte DNA gereksinimi olmakta hem de günümüzde küçük hedefli çalışmaları bu yaklaşım ile yapmak mümkün olmamaktadır. Kısacası PCR'dan tamamen kaçmak çok zordur, birçok durumda mümkün değildir. Ancak yine de çalışmanın başından sonuna kadar ne kadar az sayıda toplam döngü PCR yapılıyorsa PCR kaynaklı hatalar o derecede az olacaktır.

Validasyon çalışmasında (ve FH hasta grubu çalışmasında) dizilenecek hedef küçük olduğu için amplikon dizileme yaklaşımı seçilmiştir, bu sebep ile toplamda düşük döngü sayılı bir çalışma elde etmek mümkün olmamaktadır. Döngü sayısı olarak 40 döngülü bir ısıt döndürücü protokolü kullanılmıştır, örnek hazırlama ve klonal amplifikasyon aşamalarında ise toplamda yaklaşık 25 döngü daha olması sebebiyle toplamda yaklaşık 65 döngü PCR yapılmıştır. Örnek hazırlama ve klonal amplifikasyon aşamaları için yapılabilecek bir şey olmamak ile beraber iyi bir PCR optimizasyon ile 40 döngülü kandan direkt yapılan PCR 35, belki 30 döngü civarlarına düşürülebilirdi. Ancak bu kadar az miktardaki bir düşüşün en sonunda elde edilen hassasiyeti çok ta değiştirmeyeceği düşünülmektedir.

Havuzlama yaklaşımının genetik çalışmalarda kullanılması doğal olarak testin hassasiyet ve özgüllüğünün yüksek olmasını gerektirdiği için PCR'dan tamamen kaçılmasa da çalışmadaki toplam döngü sayısı mümkün olduğunca azaltılarak daha yüksek hassasiyet seviyesi elde etmek ve daha yüksek örnek içeren havuzlar kullanılan çalışmalar yapmak mümkündür.

6.1.2.3. UMI

UMI (İng. Unique Molecular Identifier) veya Benzersiz Moleküler Tanımlayıcı, hassasiyetin artırılması için kullanılan bir adaptör modifikasyonu yöntemidir. UMI dizileri, adaptör moleküllerinin dizilenecek DNA ya bağlandığı noktaya eklenen genellikle 6-16 bp'lik rastgele dizilerdir. Kuramsal temeller kısmı, sayfa 17'de işlenmiştir. UMI yaklaşımını kullanmak, DNA polimeraz kaynaklı okuma hatalarını büyük ölçüde elimine ettiği için hassasiyeti arttırmaktadır. UMI yaklaşımı sayesinde hata oranı yaklaşık 100 kat düşürülebilmekte ve hassasiyet te bu şekilde 100 kat artarak %0,01 seviyelerine kadar arttırılabilmektedir (230, 231).

Validasyon çalışmasında UMI yaklaşımı kullanılmamıştır. Bunun sebebi UMI yaklaşımının maliyet olarak yüksek olması ve analiz olarak ta daha kompleks olmasıdır; bunun yerine basit ve düşük maliyetli bir yaklaşım seçilerek bu yaklaşım ile neler yapılabileceği üzerine odaklanılmıştır. Aynı zamanda çalışma yaklaşımı olarak ampikon dizileme seçildiği için UMI yaklaşımı kullanılmış olsaydı dahi ilk PCR aşamasında oluşan DNA polimeraz hatalarını elimine etmek mümkün olmayacaktı, UMI yaklaşımı kısmi bir iyileşmeye sebep olacaktı.

Ampikon dizileme olmayan bir yöntem ile birlikte kullanılan UMI yaklaşımı sayesinde elde edilebilen %0,01 (10.000’de 1) hassasiyet, kolayca birkaç yüz örnekten oluşan havuzların oluşturulabilmesine olanak verebilecektir; teorik olarak birkaç bin örnekli havuzların bile oluşturulabilmesi mümkün olacaktır.

6.1.2.4. DÜŞÜK HATA ORANINA SAHİP NGS VEYA 3. NESİL DİZİLEME

Dizileme sonucunda elde edilen okumalar içerisinde hatalı olan okumalar, kütüphane hazırlık aşaması ve öncesinde meydana gelebileceği gibi, cihazda dizilenirken de meydana gelebilir. Mevcut NGS sistemleri (ve 3. nesil sistemler) farklı teknolojiler kullandıkları için hata oranları da birbirinden farklı olabilmektedir. Bazı sistemlerin hata oranı belirgin bir şekilde düşük olsa da, en yaygın olarak kullanılan sistemlerin hata oranları birbirine benzemektedir ve Phred kalite skorumaya yöntemine göre 40 veya daha az skor ile gösterilirler. Phred skoru 30-40 arasında olan okumalar çok kaliteli olarak değerlendirilir; okuma sonucu yanlış baz olma ihtimali 1.000’de 1 ‘den daha küçüktür. Ancak dizileme içerisinde daha düşük Q skorlarına sahip okumalar da olabilir. Bu düşük kalitede okumalar hassasiyet açısından sorun oluştururlar; hata oranı ne kadar yüksekse o noktadaki hassasiyet te o kadar azalmış olur. Rastgele oluşan düşük kalitedeki okumalar, bu çalışmada da olduğu gibi Q skoru filtresi kullanılarak elimine edilebilirler. Ancak sistematik olan düşük kaliteli okumaların (örneğin GC oranı yüksekliği gibi diziye özgü bazı durumlar, bazı noktalarda kalitenin düşmesine yol açabilmektedir) olduğu durumlarda Q skoru filtresi de sadece bir miktar işe yarayabilmektedir. Kısacası kaliteli okuma elde etmek her zaman için bir avantajdır.

Mevcut yaygın kullanılan sistemlerin her ne kadar kalite skorları benzer olsa da, yeni geliştirilen ve kalite açısından değerlendirildiği zaman maksimum Q skoru olarak 40’ın

yeterli olmadığı, okumaların çoğunluğunun kalitesinin Q40'tan fazla olduğu teknolojiler de geliştirilmeye başlanmıştır (232, 233). Yakın gelecekte hata oranı çok daha düşük teknolojilerin geliştirilmesi de muhtemeldir.

Validasyon çalışmasında Illumina firmasının Miseq sistemi kullanılmıştır ve Miseq sisteminin Phred skorlaması maksimum 40'tır. Havuzlama yaklaşımının kullanılacağı DNA dizileme temelli bir tarama çalışmasında daha yüksek Phred skorlarına çıkabilen bir platformun tercih edilmesi, çalışmanın hassasiyetini de arttıracaktır.

6.1.2.5. ANALİZ ALGORİTMALARI

Çalışmada iki farklı analiz yaklaşımı kullanılmıştır; düşük örnek sayılı havuzlar için basit ve manüel bir analiz yaklaşımı, yüksek örnek sayılı havuzlar için somatik mutasyon tespit için geliştirilmiş hazır bir algoritma. Yüksek hassasiyet oranlarına çıkabilmek için iki yöntem de yeterli değildir. Her ne kadar somatik mutasyon tespiti için geliştirilen yazılımın amacı düşük oranlı mutasyonların keşfi olsa ve bu sebeple bu çalışma için uygun gibi görünse de; daha kapsamlı ve sadece bu amaç için üretilmiş, analizi yaparken sadece o havuzu değil, aynı boyuttaki ve diğer boyuttaki diğer havuzları da beraber değerlendirip pozitif örnekleri direkt tespit edebilen, UMI yaklaşımı ile uyumlu, yapay zeka ve makine öğrenmesi ile desteklenmiş bir yazılım ile çok daha hassas analizlerin yapılması ve havuz başına çok daha fazla örneğin karıştırılması mümkün olabilir. Yaklaşık 20 yıllık NGS serüveninde elde edilen geliştirmelerin bir kısmı işin kimya/reaksiyon aşamalarındaki geliştirmelerden kaynaklanırken, yazılım teknolojisindeki geliştirmelerin katkısı göz ardı edilemez; aynı verinin daha gelişmiş algoritmalar ile analiz edilmesi sonucunda çok daha uzun ve düşük hata oranına sahip DNA dizileme elde etmek mümkün olabilmektedir. Aynı şekilde çok boyutlu havuzlar içinde nadir mutasyonları taşıyan bireylerin tespiti için özel olarak geliştirilecek algoritmaların da süreç içindeki pozitif etkisini mutlaka olacağı düşünülmektedir.

6.1.3. FAMILİYAL HİPERKOLESTEROLEMİ HASTA GRUBU ÇALIŞMASI

Çalışmada toplam 256 hasta ve 64 kontrol örneğinde 6 genin taraması gerçekleştirilmiştir.

6.1.3.1. HAVUZLAMA

Havuzlama yaklaşımı olarak validasyon çalışmasında %100 başarılı olarak tespit edilen 2 boyutlu 8x8'lik dizinler kullanılmıştır.

Havuzlar oluşturulurken 213S147 nolu Tübitak 1003 projesi kapsamında toplanan örneklerden uygun olanlar, sırası ile oluşturulan dizinlere eklenmiştir. Sıra ile eklenmesi bir uygunsuzluk oluşturmuştur; örnekler Türkiye çapında merkezlerden genelde gruplar halinde geldiği için Tübitak projesindeki sıraları rastgele değildi, örnekler dizinlere Tübitak projesindeki sıra ile eklenince de örnekler dizinlere rastgele olarak dağılmamış oldu. Bunun sonucunda da özellikle ilk iki 8x8'lik dizin içerisine Hatay-Adana yöresinden gelen ve bu bölgede yaygın olarak görüldüğü açığa çıkartılmış olan Q136X patojenik mutasyonuna sahip hastalardan olması gerekenden fazla düşmüş oldu (ilk gruba 14 adet, ikinci gruba 4 adet, diğer 2 hasta grubuna ise toplamda 1 adet). Özellikle ilk gruba düşen 14 adet pozitif örnek çok daha büyük problem çıkartabilirdi, ancak pozitif örneklerin 2 grup halinde arka arkaya nerdeyse mükemmel bir şekilde sıralanması sonucunda X ekseninde pozitif çıkan havuz sayısı sadece 3 olmuş ve 14 örneğin tespiti için sadece 21 örneğin teker teker dizilenmesi yetmiştir. Şayet örnekler bu şekilde sıralanmamış olsalardı, 2 boyutta tespit edilen pozitif havuzların kombinasyonlarından dolayı dizin içindeki örneklerin tamamının teker teker dizilenmesi bile gerekebilirdi. Yapılması gereken şey bu çalışma planlanırken Tübitak projesindeki örneklerin rastgele olmadığı ve havuz oluşturmada problem çıkarabileceği öngörülüp, aynı validasyon çalışmasında olduğu gibi örneklerin dizinler içerisine kura ile rastgele dağıtılmasının sağlanmasıydı. Buradan çıkartılabilecek sonuç: rastgelelik havuzlama yaklaşımı için çok önemlidir.

6.1.3.2. TESPİT EDİLEN MUTASYONLAR

Çalışılan 6 genden sadece bir tanesi FH ile direkt bağlantılıdır: otozomal resesif kalıtım gösteren LDLRAP1 geni. Bu sebep ile çalışma sonunda hasta grubu içerisinde bu gende homozigot patojenik mutasyon taşıyan bireylerden bulunması bekleniyordu ve öyle de oldu. Ancak bulunan homozigot örnek sayısı beklenenden fazla çıktı (toplam 24 homozigot örnek,

toplam hasta grubu örnek sayısının %9'u). Bunun sebebi de yine Hatay-Adana bölgesinden gelen ve Q136X mutasyonu homozigot taşıyan hastaların grup halinde çalışma içerisine girmiş olmasıdır (toplam 19 Q136X homozigot örnek, toplam hasta grubu örnek sayısının %7,4'ü). Bu 24 örneğin beklendiği gibi LDL-C değerleri çok yüksektir (ilk tanı ortalaması 494 mg/dl, n=14; örnek alım ortalaması 395 mg/dl, n=23; tüm hasta grubu ilk tanı ortalaması 274 mg/dl, n=103; tüm hasta grubu örnek alım ortalaması 240 mg/dl, n=211). Bu 24 örneğin 14'ü için (%59) akraba evliliği var bilgisi varken, 8 tanesi için (%33) akraba evliliği yok bilgisi vardır, 2'si için (%8) ise herhangi bir bilgi yoktur. Bu örneklerin tamamı için olmasa da tamamına yakınında akraba evliliği öyküsü bulunması beklenirken %33'ünde akraba evliliği yok bilgisinin olması beklenmeyen bir sonuç olmuştur. Bu noktada verinin hatalı olduğu düşünülmektedir; form doldurulurken yanlış doldurulmuş veya aslında akraba evliliği öyküsü olmasına rağmen bunun hasta tarafından bilinmemesi veya sadece yakın akrabalık ilişkisinin kastedildiğinin düşünülmesi sonucunda yanlış beyanda bulunulmuş olabilir.

LDLRAP1 geninde heterozigot mutasyon tespit edilen 15 örneğin LDL-C seviyelerinde ise bir arış gözlemlenmemiştir (ilk tanı ortalaması 250 mg/dl, n=5; örnek alım ortalaması 231 mg/dl, n=13). Ancak bu 15 örnek içinde dikkat çeken iki örnek vardır; LDL-C seviyeleri örnek alım sırasında 344 mg/dl ve 390 mg/dl olan ve stop kodonuna sahip mutasyonları (Q86X ve E217X) taşıyan örnekler. Bu örneklerin LDL-C seviyeleri normal olamayacak kadar yüksektir ve homozigot mutasyon taşıyan örnekler civarındadır. Mutasyonları stop kodonu olduğu için patojenitelerinde soru işareti de yoktur. Akraba evliliği durumu bulunmadığı bilgisi olan bu örneklerde muhtemel olarak tespit edilemeyen bir mutasyon daha vardır; amplikon dizileme ile tespit edilemeyecek olan heterozigot büyük insertsiyon/delesyon veya derin intronik mutasyon bulunma ihtimali yüksektir. Taramadan kaçmış bir mutasyon olma ihtimali de vardır ancak bu ihtimal daha düşüktür, çünkü heterozigot bulunan mutasyonların oranının 0,08-0,11 arasında olması havuzlama aşaması ile ilgili bir problem olmadığını düşündürmektedir. Yine de kontaminasyon veya allel drop-out sebebi ile bu iki örnekte bulunan heterozigot mutasyonların taramada tespit edilememiş olma ihtimali az da olsa vardır.

LDLRAP1 geni dışında homozigot mutasyon tespit edilen bir gen olmamıştır.

LIPA geninde homozigot mutasyon bulunan bireylerin lipid profili FH ile karışabileceği için hasta grubu içerisinde LIPA geninde homozigot mutasyon bulunduran bireylerin de tespit edilmesi beklenmekteydi, ancak böyle olmadı. Bunun sebebi LIPA geni mutasyonlarının daha nadir olması ve/veya LIPA genindeki homozigot mutasyonların önemli bir kısmının FH kliniği ile karışmayacak kadar farklı olan Wolman hastalığına yol açması olabilir. LIPA genindeki heterozigot mutasyonların da FH'ye benzer bir klinik oluşturabileceğine yönelik mevcut bir yayına (220) binaen hasta grubu içinde anlamlı miktarda LIPA geni heterozigot mutasyon taşıyan hasta olabileceği düşünülmüştü, ancak bu şekilde bir sonuç ta elde edilemedi. LIPA geninde heterozigot mutasyon tespit edilen toplamda 6 hastanın 3'ünde LDLRAP1 geninde homozigot mutasyon tespit edildiği için LIPA geninden elde edilen sonuç anlamsızlaşmış oldu. Geri kalan 3 hastanın 1'inin LDL-C miktarı LIPA genindeki heterozigot bir mutasyon ile açıklanamayacak kadar yüksek (örnek alım değeri 388 mg/dl) olması sebebi ile, bu hastanın klinik durumuna sebep olan başka bir mutasyonun olması daha muhtemeldir (bu mutasyon LIPA geninde de olabilir). Geriye kalan 2 hastanın 1'inin lipid profili ile ilgili bir bilgi mevcut değildir. Kalan son hastanın LDL-C seviyesi ise (241 mg/dl) her ne kadar yüksek olsa da heterozigot bir mutasyon ile açıklanabilecek seviyededir. Ancak bu hastada heterozigot olarak tespit edilen I252L mutasyonunun klinik sınıflaması -her ne kadar bazı çalışmalarda belirsiz olarak sınıflandırılmış olsa da- çoğunluk olarak iyimser (İng. benign) veya iyimsere yakın (İng. likely benign) olarak değerlendirildiği ve simülasyon analizlerinde de iyimser olarak görüldüğü için patojenik olma ihtimali çok düşüktür; bu hastanın da mevcut kliniğinin başka bir mutasyon sebebi ile olmuş olması ihtimali çok daha yüksektir. Özet olarak bu çalışmada LIPA geni heterozigot mutasyonlarının FH benzeri bir klinik oluşturabileceğine dair bir bulgu elde edilememiştir.

LDLRAP1 dışındaki diğer genlerde bulunan heterozigot mutasyonların FH kliniği ile ilgili net bir ilişkisi yoktur, bildirilen ilişki sadece yatkınlık oluşturabilecekleri ve/veya modifiye etkileri olabilecekleri yönündedir. Hatta çalışmaya başlandığı tarihte FH'ye yatkınlık oluşturabileceği yönünde OMIM'de fenotip ilişkisi görünen ITIH4 geninin bu ilişkisi çalışmanın tamamlandığı tarihlerde OMIM veri tabanından kaldırılmış durumdadır. Şayet hasta grubunda bu genlerden herhangi biri için homozigot bir mutasyon tespit edilmiş olsaydı o zaman bu homozigot genotip ile FH'nin bir ilişkisi olabileceği düşünülebilirdi, ancak herhangi bir homozigot mutasyona rastlanmadı. LDLRAP1 dışındaki genler için hasta grubunda toplamda 256 örnek içinde 30 adet (%12) heterozigotluk tespit edilirken kontrol

grubundaki 64 örnekte 11 adet (%17) heterozigotluk tespit edilmiştir. Bunun sonucunda da LDLRAP1 geni dışındaki diğer genlerde bulunan heterozigot değişikliklerin FH kliniği oluşturacağına dair bir bulgu elde edilemedi sonucuna varılabilir.

Tübitak projesi kapsamında hastalardan toplanan klinik bilgiler içinde LDL-C parametresi için iki farklı değer vardır: ilk tanı ve örnek alım seviyeleri. İlk tanı, hastaya ilk tanısı koyulduğu ve tedavisi başlamadan önceki değerdir. Örnek alım değeri ise proje için örnek alınması sırasında ölçülen LDL-C değeridir. İlk tanı seviyesi daha önemlidir, çünkü ilk tanıdan sonra tedavi süreci başladığı için LDL-C değerleri de düşmektedir. Ancak örnek alım LDL-C seviyesi değeri daha fazla hasta için toplanabildiği için onun da avantajı vardır. Değişik hasta grupları için ortalama LDL-C değerleri Çizelge 6.1.'de özetlenmiştir.

Çizelge 6.1. Değişik hasta grupları için ortalama LDL-C değerleri

	İlk tanı LDL-C değeri (mg/dl)	Örnek alım LDL-C değeri (mg/dl)
Tüm hasta grubu n=256	274 n=103	240 n=211
LDLRAP1 geni homozigot mutasyon tespit edilen n=24	494 n=14	395 n=23
LDLRAP1 geni heterozigot mutasyon tespit edilen n=15	250 n=5	231 n=13
Herhangi bir geninde sadece heterozigot mutasyon tespit edilen n=35	216 n=13	226 n=29
Herhangi bir mutasyon tespit edilmeyen n=197	244 n=76	220 n=159

Çalışma grubu içerisinde örnek alım tarihindeki LDL-C seviyesi bilgisi olan 211 hastanın 38'inin LDL-C seviyesi 300 mg/dl ve üzerindedir. Bunlardan 18'inde (%47) homozigot, 5'inde (%13) ise heterozigot mutasyon tespit edilmiştir. Geri kalan 15'inde (%40) ise herhangi bir mutasyon tespit edilememiştir. 5 adet heterozigot mutasyon tespit edilen hastanın 2'sinin mutasyonu LDLRAP1 genindedir ve klinik durumları gende tespit edilememiş muhtemel başka bir mutasyon ile açıklanabilir (Q86X ve E217X mutasyonlarını heterozigot olarak taşıyan hastalar). Geri kalan genlerinde heterozigotluk tespit edilen 3 hasta ile beraber 15 adet mutasyon tespit edilmeyen, toplamda 18 adet hastanın (%47) LDL-

C seviyeleri çok yüksek olmalarına rağmen kliniklerini açıklayacak herhangi bir genotip tespit edilememiştir. Bu örneklerin LDL-C seviyeleri çok yüksek olduğu için mutlaka bir mutasyona sahip oldukları ve bu mutasyonun da homozigot olma ihtimalinin yüksek olduğu düşünülebilir. Bu hastalarda havuzlama ile tarama yaklaşımı ile tespit edilemeyecek olan büyük insertsiyon/delesyon mutasyonları olabilir. Veya bu hastaların bir kısmında Tübitak projesi kapsamında çalışılan LDLR, APOB ve PCSK9 genlerinde bir mutasyonu olup, çalışma esnasında yanlış genotiplendirilmiş olabilirler. Bir başka ihtimal yüksek LDL-C seviyelerine yol açan ve bu çalışma ve Tübitak projesi kapsamında çalışılan toplamda 9 genden başka bir genlerinde mutasyon olmasıdır. Son olarak örnekler arasında klinik bilgileri yanlış kaydedilmiş, aslında LDL-C seviyeleri belirtilen kadar yüksek olmayan örnekler olmuş olabilir.

İnsan hastalıklarının moleküler genetik çalışmalarında çok yaygın olarak kullanılan iki terim, mutasyon ve polimorfizm terimleridir. Bu terimlerin kelime anlamlarına bakacak olursak polimorfizm çoklu form/yapı/morfoloji anlamına gelirken mutasyon ise DNA üzerinde bir hata sonucu oluşmuş herhangi bir değişiklik anlamına gelmektedir. Ancak bu terimler insan genetik hastalıkları alanında kullanılırken ilginç bir şekilde tam tersi olmasa da tersine yakın anlamlarda kullanılmaktadır: mutasyon DNA üzerindeki herhangi bir değişiklik değil, nadir görülen ve çoğunlukla morfolojiyi -olumsuz yönde olmak koşulu ile-etkileyerek çoklu morfoloji oluşturabilen anlamında kullanılırken; polimorfizm ise morfolojiyi değiştirip çoklu morfolojiye sebep olan değil, yaygın görülüp çoğu durumda herhangi bir morfolojik değişikliğe yol açmayan anlamında kullanılmaktadır. Bu çalışma bir insan genetik hastalıkları moleküler çalışması olduğu için mutasyon ve polimorfizm terimleri kelime anlamlarında değil, yaygın kullanılan anlamlarında kullanılmıştır. Patojeniteye sahip olan veya sahip olma ihtimali olan ve nadir görülen mutasyonlar ile patojeniteye sahip olmayan ve yaygın görülen polimorfizmlerin arasındaki çizgi de net değildir. Bazı değişiklikler için net bir şekilde mutasyon veya polimorfizm olarak sınıflandırılabilirken, aradaki gri alanda kalıp net bir sınıflandırma yapılamayan değişiklikler de vardır. Bu çalışmada mutasyon/polimorfizm ayrımı için popülasyon allel frekansı bilgisi kullanılmıştır; tespit edilen herhangi bir değişikliğin tüm popülasyonlardaki allel frekansı 1000 Genome Projesi faz3 verilerine göre şayet 0,02'nin üzerindeyse değişiklik polimorfizm olarak değerlendirilmiş, şayet 0,02 veya altında bir değerse mutasyon olarak değerlendirilmiştir. Genelde benzer çalışmalar için 0,01 allel frekans eşik değeri

kullanılırken bu çalışma için 0,02'nin seçilmesinin sebebi çalışılan hastalığın yaygın bir hastalık olmasıdır. Çalışma sonucunda tespit edilen ve mutasyon olarak değerlendirilen 40 değişikliğin 39'unun allel frekansı 0,01 ve altında bir değerdir, ancak 1 tanesinin, LDLRAP1 geni R238W mutasyonunun frekansı 0,02 olarak belirtilmiştir. Bu değişikliği incelediğimiz zaman tüm popülasyonlardaki frekansının 0,02 olması karşın bazı popülasyonlarda 0,05'e kadar bildirildiği görülmektedir (1.006 allel içinde 53). Bu da patojenik etki beklenmeyecek kadar yüksek bir frekanstır. Aynı zamanda veri tabanlarının neredeyse tamamında klinik sınıflaması iyimser veya iyimsere yakın olarak değerlendirilmiştir. Bu sebeple aslında bu değişikliği mutasyon yerine polimorfizm olarak sınıflamak daha yerinde olacaktır. Diğer tüm değişikliklerin frekansı 0,01 veya daha düşük olduğu için çalışmanın en başında eşik değeri olarak 0,02 yerine 0,01 kullanılsaydı daha iyi olurdu denilebilir. Çalışmada R238W değişikliği toplam 14 örnekte heterozigot, 2 örnekte ise homozigot olarak tespit edilmiştir. Homozigot olarak tespit edilen örneklerde LDLRAP1 geninde başka bir homozigot mutasyon da tespit edildiği için bu mutasyonun çalışmadan çıkartılması çalışmadaki homozigot bulunan mutasyonlardan yapılan çıkarımları etkilemeyecektir. Çalışmada heterozigot bulunan mutasyonların klinik etkisi ile ilgili önemli bir çıkarım yapılamadığı için, R238W mutasyonunu çalışmadan çıkartmak her ne kadar tespit edilen mutasyonların istatistiğini değiştirse de sonuçta yapılan çıkarımı değiştirmeyecektir. Özet olarak, şayet eşik değeri 0,02 yerine 0,01 olarak tercih edilseydi ve R238W mutasyonu polimorfizm olarak kabul edilip çalışma dışı bırakılsaydı daha uygun olacaktı ancak sonuçta elde edilen çıkarım açısından değerlendirildiğinde R238W'nin çalışma içerisinde olup olmaması arasında bir fark olmadığı için çok ta önemli değildir denilebilir.

Çalışmada tespit edilen 40 çeşit mutasyondan 37 tanesi daha önce bildirilmiş değişiklikler olurken, 3 tanesi daha önce hiç bildirilmemiş değişikliklerdir.

1000 Genom Projesi faz 3 verilerine göre popülasyon frekansları; 1 adet mutasyon için 0,02, 1 adet mutasyon için 0,01, 31 adet mutasyon için ise 0,01 den küçük olarak bildirilmiştir. 7 adet mutasyon için ise hiçbir veri yoktur; hiçbir veri olmaması bu mutasyonların ya daha önce hiç tespit edilmediklerini veya daha önce tespit edildikleri halde frekans verisi çıkarılamayacak kadar nadir olduklarını göstermektedir.

Bulunan 40 mutasyonun klinik anlamlarına ClinVar veri tabanından bakılacak olursa; 5 adedinin patojenik, 4 adedinin belirsiz öneme sahip, 4 adedinin iyimsere yakın, 2 adedinin iyimser olarak sınıflandırıldığı görülmekle beraber, büyük çoğunluğunu oluşturan 25 adedi için herhangi bir veri olmadığı görülmektedir. Bu 25 taneden 17 tanesi ITIH4 genine aittir. ITIH4 geni her ne kadar 1995 yılında ilk tanımlanmış (234) ve bu çalışmada çıktığı üzere çok sayıda varyasyon içeren bir gen olmasına rağmen, aradan geçen yaklaşık 30 yılın ardından sadece yüksek kolesterol fenotipi ile zayıf bir şekilde ilişkilendirilmiş olması bu genin yaşam için önemli olmayan bir gen olduğunu göstermektedir.

Çalışma içerisindeki en küçük gen olan APOA2 geninde hem hasta hem de kontrol grubunda hiç mutasyon tespit edilmemiştir.

Tez çalışmasında 320 örneğin 6 geni taranmıştır. Çalışma havuzlama yaklaşımı kullanılmadan yapılmış olsaydı 320 örnek için toplamda 11.520 adet amplikonun dizilenmesi gerekirken, havuzlama yaklaşımı sayesinde 80 havuzun taranması için 2.880 ve havuzlar içindeki pozitif örneklerin tespiti için 134 olacak şekilde toplamda 3.014 amplikon dizileme ile, toplamda dizilenmesi gereken amplikonun sadece %26'sı yapılarak çalışma tamamlanabilmiştir. Maliyet olarak bakıldığında zaman ise havuzlamanın etkisi daha azdır; havuzlama olmadan yapılmış olsaydı sarf malzeme maliyeti örnek başına 63 USD civarlarında olacak iken, havuzlama sayesinde örnek başı sarf malzeme maliyeti 28 USD civarlarına düşmüştür. Yani dizilenen amplikon olarak bakıldığında %74 olan düşüş, maliyet olarak hesaplandığında %56 civarındadır. Bunun sebebi havuzlama yaklaşımının çok daha derin dizilemeye ihtiyaç duyması ve bunun sonucunda da çalışmanın dizi verisi elde etme kısmında maliyet açısından ciddi bir düşüş olmamasıdır. Aynı zamanda kullanılan platformun (Illumina - Miseq) göreceli olarak düşük kapasiteli yüksek maliyetli bir platform olmasının da bu açıdan katkısı olmuştur; dizi verisi üretme maliyeti diğer maliyetler yanında yüksek kalmaktadır. Dizileme verisi maliyeti çok daha düşük olan yüksek kapasiteli cihazlarda, çok daha kapsamlı kütüphane hazırlama kitleri kullanılarak yapılacak çalışmalarda sarf malzeme maliyet düşüşü çok daha dramatik olacaktır. Mesela toplamda 5.000 genin dizilenmesinde kullanılan bir klinik ekzom kitinin sarf malzeme maliyeti, yüksek kapasiteli cihazlarda 110 USD civarında gelirken, bu maliyetin 2 boyutlu 16'lı havuzlama yaklaşımı ile 25 USD, 2 boyutlu 24'lü havuzlama yaklaşımı ile ise 16 USD civarlarına düşürmek mümkündür. Bu şekilde sırası ile %77 ve %85 sarf malzeme maliyeti

düşüşü sağlanabilir. Görüldüğü gibi bu tez çalışmasında örnek başına 28 USD sarf malzeme maliyeti ile 6 gen taranmışken, daha büyük ölçekli çalışmalar tasarlanarak çok daha az bir sarf malzeme maliyeti ile çok daha fazla gen taramak mümkündür (sadece 16 USD'ye 5.000 gen gibi). Ancak büyük kapsamlı çalışmalarda cihaz yatırım maliyetleri, işletme maliyetleri, analiz maliyetleri vb. daha fazla olacağı için örnek başı maliyetleri hesaplarken bu kalemlerin de mutlaka göz önünde bulundurulması gerekecektir.

NGS teknolojileri hızla ilerlemektedir ve elde edilen dizileme verisi maliyetleri çok hızlı (logaritmik olarak) düşmektedir; buna karşılık tüm çalışma içerisindeki dizileme verisi maliyetleri dışındaki maliyetler ise -her ne kadar onlar da zaman içerisinde bir miktar düşseler de- dizileme verisi üretme maliyeti kadar hızlı düşmemektedirler. Örnek olarak son 5 yıl içerisinde 10 kat düşen dizi verisi üretme maliyetlerine karşılık, diğer maliyetler toplamı en fazla yarisına kadar inebilmiştir. Bu nedenle ilerleyen zaman içerisinde düşen dizileme maliyetlerinin, havuzlama stratejisi için ek bir avantaj getireceğini söylemek yanlış olmaz.

6.1.4. BU TEZ ÇALIŞMASININ YENİ LİTERATÜR İÇİNDEKİ YERİ

Havuzlama yaklaşımı iyi bilinen ve yaygın kullanılan bir tekniktir. NGS teknolojisi geliştirilmeden önce DNA analizinde de kullanılan bir yaklaşım olmakla beraber, kullanılan yöntemlerin hassasiyetleri çok yüksek olmadığından sınırlı olarak kullanılmaktaydı. 2005 yılından sonra ticari olarak ulaşılabilir hale gelmesi ve 2008 yılından sonra yaygınlaşmaya başlamasıyla, NGS'in havuzlama yaklaşımı ile beraber kullanılabilirlik potansiyeli hemen dikkat çekmiş ve 2009 yılından itibaren arka arkaya yayınlar çıkmaya başlamıştır. Dikkat çeken bir şey de ilk yıllarda enformatik alanında çalışan araştırmacıların dikkatini çekmiş olmasıdır; çok sayıda havuzlama ile NGS'in beraber kullanılmasının enformatik boyutunu inceleyen yayın çıkmıştır. Bu çalışmalar genellikle havuz hazırlama stratejileri, varyant tespiti ve yalancı pozitiflerin ayıklanması ile ilgilidir (133, 144).

Enformatik alanı dışında gerçek örneklerde mutasyon/polimorfizm çalışmaları da 2009 yılından itibaren çıkmaya başlamıştır. Özet olarak zaten bilinen havuzlama yaklaşımı, NGS kullanan araştırmacılar tarafından hemen keşfedilmiş ve çalışılmaya başlanmış bir yaklaşımdır. Günümüze gelinceye kadar 40'tan fazla yayın çıkmıştır. Bu yayınları kabaca şu şekilde gruplayabiliriz:

- Bir kısmı çok sayıda örneğin havuzlandığı, patojenik nadir mutasyonların tespiti yerine bağlantı veya varyasyon keşfi amaçlı yapılan polimorfizm keşfi / frekans belirleme çalışmalarıdır (örnek (148)).
- Bir kısmı çok sayıda örneğin, havuz başına az sayıda örnek içerecek şekilde havuzlandığı patojenik mutasyon tespit çalışmalarıdır. Bu çalışmaların hedef bölgeleri genelde küçüktür (örnek (159)).
- Bir kısmı az sayıda örneğin, havuz başına az sayıda örnek içerecek şekilde havuzlandığı patojenik mutasyon tespit çalışmalarıdır. Bu çalışmaların hedef bölgeleri genelde büyüktür (örnek (164)).
- Bir kısmı çok sayıda örneğin, havuz başına çok sayıda örnek içerecek şekilde havuzlandığı patojenik mutasyon tespit çalışmalarıdır. Bu çalışmaların hedef bölgeleri büyük te olabilmektedir, küçük te (örnek (157)).

Bunun dışında yapılan çalışmalar ile ilgili dikkat çeken bazı noktalar da tespit edilmiştir:

- Neredeyse tamamı DNA aşamasında karıştırılan havuzlardan oluşmaktadır.
- Neredeyse tamamı tek boyutlu havuzlama stratejisini kullanmaktadır.
- Birden fazla havuzlama stratejisini deneyip karşılaştıran kapsamlı bir çalışmaya rastlanmamıştır.
- Genotipi bilinen örnekleri dahil eden az sayıda çalışma vardır.

Kan aşamasında karıştırılarak havuzların oluşturulduğu bir çalışmaya hiç rastlanmamıştır. Bu dikkat çekici durumun sebeplerinden birinin DNA aşamasında -DNA'ların nicelik ve nitelik olarak değerlendirilmiş olma ön koşulu ile- karışımın yapıldığı durumun en dengeli havuzları oluşturması olduğu düşünülmektedir. Kan aşamasında karıştırmak daha dengesiz havuzlara yol açabilir; özellikle çok fazla örneğe ihtiyaç duyulan ve hepsinin aynı anda toplanmasının çok zor olduğu araştırma projeleri için. Aynı zamanda -şayet çalışma toplum taraması amaçlı yapılmıyorsa- çalışmalarda DNA izolasyonunun maddi payı küçük olduğu için araştırmacılar düşük maliyet yerine yüksek hassasiyeti tercih ediliyor olabilir. Hatta az da olsa örneklerin DNA seviyesinde değil, daha sonraki aşama olan PCR aşamasından sonra karıştırılarak havuzların oluşturulduğu çalışmalara bile rastlanmıştır (151, 154). PCR aşamasından sonra karıştırılan örnekler ile hazırlanan havuzlar ile yapılan çalışma, DNA aşamasında karıştırılanlara göre çok daha dengeli olma potansiyelini barındırmaktadır. Bu

alıřmalardan birinde hata dzeltici DNA polimeraz ile beraber PCR sonrası karıřtırılan rneklerde tespit limitinin %0,1'e kadar inebileceęi belirtilmiř ve hata dzeltici enzimin nemine vurgu yapılmıřtır (151). Bu sebepler ile arařtırmacıların gznde daha dengeli havuzlar ile birlikte daha yksek hassasiyetin aęır bastıęı dřnlebilir. Ancak bu kadar yksek hassasiyette de olsa PCR sonrası karıřtırılan rneklerden oluřan bir yntemin - pratiklik ve maliyet aısından- rutinde toplum taramasında kullanılması mmkn deęildir.

Dięer dikkat eken nokta yapılan alıřmaların nerdeyse tamamının tek boyutlu havuzlama stratejisini kullanmıř olmalarıdır. İki boyutlu iki adet yayına rastlanmıřtır (161, 162); ki bunlardan birinin havuzlama stratejisi tam anlařılamamıřtır (161), bir boyutlu da olabilir.  boyutluya ise hi rastlanmamıřtır. Ancak her ne kadar ok boyutlu havuzlamaya ok fazla rastlanmasa da, bařka havuzlama stratejilerine rastlanmıřtır; bunlar akıřan havuzlama (138) (İng. overlapping pooling) ve aęırlıklı havuzlama (136) (İng. weighted pooling). akıřan havuzlama aslında 2 boyutlu havuzlamaya benzemektedir; farklı rnek kombinasyonlarını ieren havuzlar oluřturup, rneklerin birden fazla havuzda alıřılmasını saęlayıp, sonra pozitif ve negatif ıkan havuzlardan yola ıkarak direkt pozitif rneęi bulma veya ihtimalleri azaltma zerine kurulmuř bir stratejisi vardır. Aęırlıklı havuzlama ise tek boyutludur ve her bir rneęin bir defa sefer dizilenir, ancak bu stratejide de her bir rnekten farklı miktarlarda alınarak oluřturulan havuzlar sz konusudur. Sonuta ıkan pozitiflik oranına gre pozitiflięin hangi rnekten geldięinin tahmin edilebileceęi hipotezi zerine kurulmuř bir stratejisi vardır. Byle bir stratejinin mutasyon taraması iin ok uygun olmadıęını dřnmemize raęmen, bazı zel durumlarda uygulama alanı olabilir.

Literatrde NGS ile birlikte birden fazla havuzlama stratejisini aynı alıřmada karřılařtırarak deneyen bir alıřmaya da rastlanmamıřtır.

Genotipi bilinen rneklerin dahil edildięi alıřma da azdır. Genelde hasta grubu havuzlara ayrılarak alıřılmıř, bulunan deęiřiklikler ile devam edilmiř, bulunamayan muhtemel mutasyonlar ile pek ilgilenilmemiřtir.

Literatrde tez alıřmasının validasyon kısmına en ok benzeyen yayın olarak deęerlendirilen yayında, tek boyutlu 20'li tek bir havuz ierisine genotipi bilinen 20 rnek DNA ařamasında karıřtırılarak alıřma yapılmıř ve tespit edilen mutasyonlarda %100 bařarı saęlanmıřtır (168). alıřma rnek sayısının az olması, pozitiflik sayısının az olması, farklı

stratejileri barındırmaması ile tez çalışmasına göre daha zayıf olarak değerlendirilmiştir. Çalışmanın daha güçlü olduğu tek kısım kapsamıdır; çalışma 5 geni 116 PCR ile amplifiye ederek ampikon dizilemesini gerçekleştiren bir çalışmadır.

Literatürde tez çalışmasının FH grubu kısmına en çok benzeyen yayın olarak değerlendirilen yayında, 576 örneğin genomunun 250 Kbp lik bir bölümü, 2 boyutlu 12x12'lik havuzlarla taranmıştır (162). Bu çalışma DNA aşamasında karıştırılan havuzlardan oluşması, 12'li havuzlar kullanması, kapsamının 250 Kbp olması ve herhangi bir hasta grubu üzerinde değil, sadece varyasyon tespiti amaçlı rastgele örneklerden yapılmış olması dışında, tez çalışmasının FH grubu kısmına çok benzemektedir. Çalışma kandan olduğu için toplum taraması için tasarlandığı söylenemez. Aynı zamanda hasta grubu çalışılmadığı için örnek bazlı klinik bir çıkarım da yapılamamıştır.

Özet olarak literatürde DNA havuzlama ve NGS yöntemlerinin beraber kullanıldığı çok sayıda çalışma vardır, ancak bu çalışmaların çok azı genetik hastalıkların toplum taraması odaklı sayılabilecek çalışmalardır. Bunun yansıması pratikte de görülmektedir; devletler binlerce bireyin genomunun dizilenmesi için genom merkezlerine büyük miktarlarda yatırım yaparken, benzer maliyetlere milyonlarca taşıyıcı bireyi veya FH gibi yaygın görülen hastalıklar için hasta bireyleri tespit etme yolunda yeterince ilerleme kaydedilememiş ve günümüze kadar uygulamaya geçen bu şekilde bir tarama programı olmamıştır.

6.2. SONUÇ

Hastalık tarama, birey ve toplum sağığı için çok önemli bir kavramdır. Hastalık tarama sayesinde hastalıklar etkilerini göstermeden çok daha önce tespit edilebilmektedir. Erken tanı bazı hastalıklar için yaşam ve ölüm arasındaki sınırı belirleyebilecek etkiye bile sahip olabilmektedir.

DNA dizi analizi birçok hastalığın tanısında kullanılmaktadır. Günümüzde çok gelişmiş ve hala gelişmekte olan DNA dizi analizi yöntemi ve bir tarama yaklaşımı olan havuzlamanın güçleri birleştirilerek, toplumun değişik kesimlerinde DNA dizileme bazlı hastalık taraması yapılabilir ve birey ve toplum sağığına dramatik fayda sağlanabilir.

Bu tez çalışmasında havuzlama yaklaşımı ile birlikte kullanılan DNA dizi analiz yönteminin, genetik hastalıkların taranması açısından basit ve etkili bir şekilde nasıl kullanılabileceğı gösterilmiştir.

KAYNAKLAR

1. Latest Research On "DNA Sequencing Market" Analysis | 2023-2029. Available from: <https://www.marketwatch.com/press-release/latest-research-on-dna-sequencing-market-analysis-2023-2029-2023-05-26>.
2. Lagopati N, Tsioli P, Mourkioti I, Polyzou A, Papaspyropoulos A, Zafiropoulos A, et al. Sample pooling strategies for SARS-CoV-2 detection. J Virol Methods. 2021;289:114044. Epub 20201211. doi: 10.1016/j.jviromet.2020.114044. PubMed PMID: 33316285; PubMed Central PMCID: PMC7834440.
3. T.C. Sağlık Bakanlığı Halk Sağlığı Genel Müdürlüğü Tarama Programları. Available from: <https://hsqm.saglik.gov.tr/tr/programlar/tarama-programlari.html>.
4. HYPERCHOLESTEROLEMIA, FAMILIAL, 1; FHCL1. Available from: <https://www.omim.org/entry/143890>.
5. Heather JM, Chain B. The sequence of sequencers: The history of sequencing DNA. Genomics. 2016;107(1):1-8. Epub 20151110. doi: 10.1016/j.ygeno.2015.11.003. PubMed PMID: 26554401; PubMed Central PMCID: PMC4727787.
6. DNA Sequencing Market Size. Available from: <https://www.novaoneadvisor.com/report/dna-sequencing-market>.
7. Big Data among Big Data: Genome Data. Available from: <https://www.3billion.io/blog/big-data-among-big-data-genome-data>.
8. Dahm R. Discovering DNA: Friedrich Miescher and the early years of nucleic acid research. Hum Genet. 2008;122(6):565-81. Epub 20070928. doi: 10.1007/s00439-007-0433-0. PubMed PMID: 17901982.
9. Avery OT, Macleod CM, McCarty M. Studies on the Chemical Nature of the Substance Inducing Transformation of Pneumococcal Types : Induction of Transformation by a Desoxyribonucleic Acid Fraction Isolated from Pneumococcus Type Iii. J Exp Med. 1944;79(2):137-58. doi: 10.1084/jem.79.2.137. PubMed PMID: 19871359; PubMed Central PMCID: PMC2135445.
10. Watson JD, Crick FH. Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. Nature. 1953;171(4356):737-8. doi: 10.1038/171737a0. PubMed PMID: 13054692.
11. Stretton AO. The first sequence. Fred Sanger and insulin. Genetics. 2002;162(2):527-32. doi: 10.1093/genetics/162.2.527. PubMed PMID: 12399368; PubMed Central PMCID: PMC1462286.
12. Crick FH. On protein synthesis. Symp Soc Exp Biol. 1958;12:138-63. PubMed PMID: 13580867.

13. Hutchison CA, 3rd. DNA sequencing: bench to bedside and beyond. *Nucleic Acids Res.* 2007;35(18):6227-37. Epub 20070912. doi: 10.1093/nar/gkm688. PubMed PMID: 17855400; PubMed Central PMCID: PMC2094077.
14. Wu R. Nucleotide sequence analysis of DNA. I. Partial sequence of the cohesive ends of bacteriophage lambda and 186 DNA. *J Mol Biol.* 1970;51(3):501-21. doi: 10.1016/0022-2836(70)90004-5. PubMed PMID: 4321727.
15. Maxam AM, Gilbert W. A new method for sequencing DNA. *Proc Natl Acad Sci U S A.* 1977;74(2):560-4. doi: 10.1073/pnas.74.2.560. PubMed PMID: 265521; PubMed Central PMCID: PMC392330.
16. Sanger F, Nicklen S, Coulson AR. DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci U S A.* 1977;74(12):5463-7. doi: 10.1073/pnas.74.12.5463. PubMed PMID: 271968; PubMed Central PMCID: PMC431765.
17. Saiki RK, Scharf S, Faloona F, Mullis KB, Horn GT, Erlich HA, et al. Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. *Science.* 1985;230(4732):1350-4. doi: 10.1126/science.2999980. PubMed PMID: 2999980.
18. Chen CY. DNA polymerases drive DNA sequencing-by-synthesis technologies: both past and present. *Front Microbiol.* 2014;5:305. Epub 20140624. doi: 10.3389/fmicb.2014.00305. PubMed PMID: 25009536; PubMed Central PMCID: PMC4068291.
19. Venter JC, Adams MD, Myers EW, Li PW, Mural RJ, Sutton GG, et al. The sequence of the human genome. *Science.* 2001;291(5507):1304-51. doi: 10.1126/science.1058040. PubMed PMID: 11181995.
20. Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, et al. Initial sequencing and analysis of the human genome. *Nature.* 2001;409(6822):860-921. doi: 10.1038/35057062. PubMed PMID: 11237011.
21. Ronaghi M, Uhlen M, Nyren P. A sequencing method based on real-time pyrophosphate. *Science.* 1998;281(5375):363, 5. doi: 10.1126/science.281.5375.363. PubMed PMID: 9705713.
22. Ranasinghe D, Jayadas TTP, Jayathilaka D, Jeewandara C, Dissanayake O, Guruge D, et al. Comparison of different sequencing techniques for identification of SARS-CoV-2 variants of concern with multiplex real-time PCR. *PLoS One.* 2022;17(4):e0265220. Epub 20220404. doi: 10.1371/journal.pone.0265220. PubMed PMID: 35377884; PubMed Central PMCID: PMC8979425.
23. Wikipedia - DNA sequencing. Available from: https://en.wikipedia.org/wiki/DNA_sequencing.
24. Gordon Moore - Cramming more components onto integrated circuits. Available from: <http://cva.stanford.edu/classes/cs99s/papers/moore-crammingmorecomponents.pdf>.

25. Moore's law. Available from: https://en.wikipedia.org/wiki/Moore%27s_law.
26. DNA Sequencing Forges Ahead. Available from: <https://pubsapp.acs.org/cen/coverstory/87/8750cover2.html>.
27. Frederick Sanger (1918–2013). Available from: <https://www.nature.com/articles/505027a>.
28. Frederick Sanger. Available from: https://en.wikipedia.org/wiki/Frederick_Sanger.
29. Chan EY. Advances in sequencing technology. *Mutat Res.* 2005;573(1-2):13-40. doi: 10.1016/j.mrfmmm.2005.01.004. PubMed PMID: 15829235.
30. Franca LT, Carrilho E, Kist TB. A review of DNA sequencing techniques. *Q Rev Biophys.* 2002;35(2):169-200. doi: 10.1017/s0033583502003797. PubMed PMID: 12197303.
31. Sanger F. Determination of nucleotide sequences in DNA. *Biosci Rep.* 1981;1(1):3-18. doi: 10.1007/BF01115145. PubMed PMID: 6456784.
32. Apollo Institute - Sanger Sequencing. Available from: <https://apollo-institute.org/sanger-sequencing/>.
33. Nurk S, Koren S, Rhie A, Rautiainen M, Bizikadze AV, Mikheenko A, et al. The complete sequence of a human genome. *Science.* 2022;376(6588):44-53. Epub 20220331. doi: 10.1126/science.abj6987. PubMed PMID: 35357919; PubMed Central PMCID: PMC9186530.
34. Kumar KR, Cowley MJ, Davis RL. Next-Generation Sequencing and Emerging Technologies. *Semin Thromb Hemost.* 2019;45(7):661-73. Epub 20190516. doi: 10.1055/s-0039-1688446. PubMed PMID: 31096307.
35. Ambardar S, Gupta R, Trakroo D, Lal R, Vakhlu J. High Throughput Sequencing: An Overview of Sequencing Chemistry. *Indian J Microbiol.* 2016;56(4):394-404. Epub 20160709. doi: 10.1007/s12088-016-0606-4. PubMed PMID: 27784934; PubMed Central PMCID: PMC5061697.
36. Pickrell WO, Rees MI, Chung SK. Next generation sequencing methodologies--an overview. *Adv Protein Chem Struct Biol.* 2012;89:1-26. doi: 10.1016/B978-0-12-394287-6.00001-X. PubMed PMID: 23046880.
37. Niedringhaus TP, Milanova D, Kerby MB, Snyder MP, Barron AE. Landscape of next-generation sequencing technologies. *Anal Chem.* 2011;83(12):4327-41. Epub 20110525. doi: 10.1021/ac2010857. PubMed PMID: 21612267; PubMed Central PMCID: PMC3437308.
38. Georgiou CD, Papapostolou I. Assay for the quantification of intact/fragmented genomic DNA. *Anal Biochem.* 2006;358(2):247-56. Epub 20060811. doi: 10.1016/j.ab.2006.07.035. PubMed PMID: 16942746.

39. Bronner IF, Quail MA. Best Practices for Illumina Library Preparation. *Curr Protoc Hum Genet.* 2019;102(1):e86. doi: 10.1002/cphg.86. PubMed PMID: 31216112.
40. Zimmermann J, Hajibabaei M, Blackburn DC, Hanken J, Cantin E, Posfai J, et al. DNA damage in preserved specimens and tissue samples: a molecular assessment. *Front Zool.* 2008;5:18. Epub 20081023. doi: 10.1186/1742-9994-5-18. PubMed PMID: 18947416; PubMed Central PMCID: PMC2579423.
41. Elsner HI, Lindblad EB. Ultrasonic degradation of DNA. *DNA.* 1989;8(10):697-701. doi: 10.1089/dna.1989.8.697. PubMed PMID: 2693020.
42. Knierim E, Lucke B, Schwarz JM, Schuelke M, Seelow D. Systematic comparison of three methods for fragmentation of long-range PCR products for next generation sequencing. *PLoS One.* 2011;6(11):e28240. Epub 20111130. doi: 10.1371/journal.pone.0028240. PubMed PMID: 22140562; PubMed Central PMCID: PMC3227650.
43. Miyazaki K. Random DNA fragmentation with endonuclease V: application to DNA shuffling. *Nucleic Acids Res.* 2002;30(24):e139. doi: 10.1093/nar/gnf139. PubMed PMID: 12490730; PubMed Central PMCID: PMC140088.
44. Adey A, Morrison HG, Asan, Xun X, Kitzman JO, Turner EH, et al. Rapid, low-input, low-bias construction of shotgun fragment libraries by high-density in vitro transposition. *Genome Biol.* 2010;11(12):R119. Epub 20101208. doi: 10.1186/gb-2010-11-12-r119. PubMed PMID: 21143862; PubMed Central PMCID: PMC3046479.
45. Picelli S, Bjorklund AK, Reinius B, Sagasser S, Winberg G, Sandberg R. Tn5 transposase and tagmentation procedures for massively scaled sequencing projects. *Genome Res.* 2014;24(12):2033-40. Epub 20140730. doi: 10.1101/gr.177881.114. PubMed PMID: 25079858; PubMed Central PMCID: PMC4248319.
46. ABP Biosciences - Tn5 Transposase. Available from: <https://www.abpbio.com/product/tn5-transposase/>.
47. Messing J. The universal primers and the shotgun DNA sequencing method. *Methods Mol Biol.* 2001;167:13-31. doi: 10.1385/1-59259-113-2:013. PubMed PMID: 11265313.
48. Lexogen - Next-Generation RNA-Seq. Available from: <https://www.lexogen.com/rna-lexicon-next-generation-sequencing/>.
49. Sims D, Sudbery I, Ilott NE, Heger A, Ponting CP. Sequencing depth and coverage: key considerations in genomic analyses. *Nat Rev Genet.* 2014;15(2):121-32. doi: 10.1038/nrg3642. PubMed PMID: 24434847.
50. Illumina Inc. - Glossary. Available from: <https://www.illumina.com/science/technology/next-generation-sequencing/beginners/glossary.html>.
51. Illumina multiplexing. Available from: <https://www.illumina.com/techniques/sequencing/ngs-library-prep/multiplexing.html>.

52. Illumina Dual Index. Available from: <https://www.illumina.com/techniques/sequencing/ngs-library-prep/multiplexing/unique-dual-indexes.html>.
53. Girardot C, Scholtalbers J, Sauer S, Su SY, Furlong EE. Je, a versatile suite to handle multiplexed NGS libraries with unique molecular identifiers. *BMC Bioinformatics*. 2016;17(1):419. Epub 20161008. doi: 10.1186/s12859-016-1284-2. PubMed PMID: 27717304; PubMed Central PMCID: PMC5055726.
54. Geneious Biologics - Understanding Single Cell technologies: Barcodes and UMIs. Available from: <https://help.geneiousbiologics.com/hc/en-us/articles/4781289585300-Understanding-Single-Cell-technologies-Barcodes-and-UMIs>.
55. Zhu Z, Jenkins G, Zhang W, Zhang M, Guan Z, Yang CJ. Single-molecule emulsion PCR in microfluidic droplets. *Anal Bioanal Chem*. 2012;403(8):2127-43. Epub 20120327. doi: 10.1007/s00216-012-5914-x. PubMed PMID: 22451171.
56. Kanagal-Shamanna R. Emulsion PCR: Techniques and Applications. *Methods Mol Biol*. 2016;1392:33-42. doi: 10.1007/978-1-4939-3360-0_4. PubMed PMID: 26843044.
57. Radford AD, Chapman D, Dixon L, Chantrey J, Darby AC, Hall N. Application of next-generation sequencing technologies in virology. *J Gen Virol*. 2012;93(Pt 9):1853-68. Epub 20120530. doi: 10.1099/vir.0.043182-0. PubMed PMID: 22647373; PubMed Central PMCID: PMC3709572.
58. Balasubramanian S. Solexa sequencing: decoding genomes on a population scale. *Clin Chem*. 2015;61(1):21-4. Epub 20141020. doi: 10.1373/clinchem.2014.221747. PubMed PMID: 25332311.
59. Bentley DR, Balasubramanian S, Swerdlow HP, Smith GP, Milton J, Brown CG, et al. Accurate whole human genome sequencing using reversible terminator chemistry. *Nature*. 2008;456(7218):53-9. doi: 10.1038/nature07517. PubMed PMID: 18987734; PubMed Central PMCID: PMC2581791.
60. Blanco L, Bernad A, Lazaro JM, Martin G, Garmendia C, Salas M. Highly efficient DNA synthesis by the phage phi 29 DNA polymerase. Symmetrical mode of DNA replication. *J Biol Chem*. 1989;264(15):8935-40. PubMed PMID: 2498321.
61. Drmanac R, Sparks AB, Callow MJ, Halpern AL, Burns NL, Kermani BG, et al. Human genome sequencing using unchained base reads on self-assembling DNA nanoarrays. *Science*. 2010;327(5961):78-81. Epub 20091105. doi: 10.1126/science.1181498. PubMed PMID: 19892942.
62. Porreca GJ. Genome sequencing on nanoballs. *Nat Biotechnol*. 2010;28(1):43-4. doi: 10.1038/nbt0110-43. PubMed PMID: 20062041.
63. MGI DNBSEQ™ Sequencing Technology. Available from: <https://en.mgitech.cn/news/18/>.

64. Messing J, Crea R, Seeburg PH. A system for shotgun DNA sequencing. *Nucleic Acids Res.* 1981;9(2):309-21. doi: 10.1093/nar/9.2.309. PubMed PMID: 6259625; PubMed Central PMCID: PMC326694.
65. Harrington CT, Lin EI, Olson MT, Eshleman JR. Fundamentals of pyrosequencing. *Arch Pathol Lab Med.* 2013;137(9):1296-303. doi: 10.5858/arpa.2012-0463-RA. PubMed PMID: 23991743.
66. Nyren P, Pettersson B, Uhlen M. Solid phase DNA minisequencing by an enzymatic luminometric inorganic pyrophosphate detection assay. *Anal Biochem.* 1993;208(1):171-5. doi: 10.1006/abio.1993.1024. PubMed PMID: 8382019.
67. Apollo Institute - Pyrosequencing. Available from: <https://apollo-institute.org/pyrosequencing/>.
68. Fujita S, Masago K, Okuda C, Hata A, Kaji R, Katakami N, et al. Single nucleotide variant sequencing errors in whole exome sequencing using the Ion Proton System. *Biomed Rep.* 2017;7(1):17-20. Epub 20170517. doi: 10.3892/br.2017.911. PubMed PMID: 28685054; PubMed Central PMCID: PMC5492560.
69. da Fonseca PG, Paiva JA, Almeida LG, Vasconcelos AT, Freitas AT. Empirical assessment of sequencing errors for high throughput pyrosequencing data. *BMC Res Notes.* 2013;6:25. Epub 20130122. doi: 10.1186/1756-0500-6-25. PubMed PMID: 23339526; PubMed Central PMCID: PMC3852801.
70. Ivady G, Madar L, Dzsudzsak E, Koczok K, Kappelmayer J, Krulisova V, et al. Analytical parameters and validation of homopolymer detection in a pyrosequencing-based next generation sequencing system. *BMC Genomics.* 2018;19(1):158. Epub 20180221. doi: 10.1186/s12864-018-4544-x. PubMed PMID: 29466940; PubMed Central PMCID: PMC5822529.
71. Gilles A, Meglecz E, Pech N, Ferreira S, Malausa T, Martin JF. Accuracy and quality assessment of 454 GS-FLX Titanium pyrosequencing. *BMC Genomics.* 2011;12:245. Epub 20110519. doi: 10.1186/1471-2164-12-245. PubMed PMID: 21592414; PubMed Central PMCID: PMC3116506.
72. Mardis ER. Next-generation sequencing platforms. *Annu Rev Anal Chem (Palo Alto Calif).* 2013;6:287-303. doi: 10.1146/annurev-anchem-062012-092628. PubMed PMID: 23560931.
73. DNA Tech - Illumina High Throughput Sequencing. Available from: <https://dnatech.genomecenter.ucdavis.edu/illumina-high-throughput-sequencing/>.
74. Illumina 2-Channel SBS Technology. Available from: <https://www.illumina.com/science/technology/next-generation-sequencing/sequencing-technology/2-channel-sbs.html>.
75. Fuller CW, Middendorf LR, Benner SA, Church GM, Harris T, Huang X, et al. The challenges of sequencing by synthesis. *Nat Biotechnol.* 2009;27(11):1013-23. Epub 20091106. doi: 10.1038/nbt.1585. PubMed PMID: 19898456.

76. Massingham T, Goldman N. All Your Base: a fast and accurate probabilistic approach to base calling. *Genome Biol.* 2012;13(2):R13. Epub 20120229. doi: 10.1186/gb-2012-13-2-r13. PubMed PMID: 22377270; PubMed Central PMCID: PMC4053729.
77. Warburton PE, Sebra RP. Long-Read DNA Sequencing: Recent Advances and Remaining Challenges. *Annu Rev Genomics Hum Genet.* 2023. Epub 20230419. doi: 10.1146/annurev-genom-101722-103045. PubMed PMID: 37075062.
78. Method of the Year 2022: long-read sequencing. *Nat Methods.* 2023;20(1):1. doi: 10.1038/s41592-022-01759-x. PubMed PMID: 36635552.
79. Sakurai T, Husimi Y. Real-time monitoring of DNA polymerase reactions by a micro ISFET pH sensor. *Anal Chem.* 1992;64(17):1996-7. doi: 10.1021/ac00041a040. PubMed PMID: 1416046.
80. Rothberg JM, Hinz W, Rearick TM, Schultz J, Mileski W, Davey M, et al. An integrated semiconductor device enabling non-optical genome sequencing. *Nature.* 2011;475(7356):348-52. Epub 20110720. doi: 10.1038/nature10242. PubMed PMID: 21776081.
81. Xie MJ, Liang ZK, He D, Xu WW, Wu YS, Yang XX, et al. Noninvasive Prenatal Testing of Rare Autosomal Aneuploidies by Semiconductor Sequencing. *DNA Cell Biol.* 2018;37(3):174-81. Epub 20180130. doi: 10.1089/dna.2017.4075. PubMed PMID: 29381401.
82. Gui B, Zhang Y, Liang B, Kwok YKY, Lui WT, Yeung QSY, et al. Semiconductor Sequencing for Preimplantation Genetic Testing for Aneuploidy. *J Vis Exp.* 2019(150). Epub 20190825. doi: 10.3791/59273. PubMed PMID: 31498328.
83. Semantic Scholar - Generations of Sequencing Technologies: From First to Next Generation. Available from: <https://www.semanticscholar.org/paper/Generations-of-Sequencing-Technologies%3A-From-First-Kchouk-Gibrat/a016143b0cddd6a47722f062e719de3f1062cbec>.
84. Lehman IR. DNA ligase: structure, mechanism, and function. *Science.* 1974;186(4166):790-7. doi: 10.1126/science.186.4166.790. PubMed PMID: 4377758.
85. Metzker ML. Sequencing technologies - the next generation. *Nat Rev Genet.* 2010;11(1):31-46. Epub 20091208. doi: 10.1038/nrg2626. PubMed PMID: 19997069.
86. Apollo Institute - Solid Sequencing. Available from: <https://apollo-institute.org/solid-sequencing/>.
87. Ewing B, Green P. Base-calling of automated sequencer traces using phred. II. Error probabilities. *Genome Res.* 1998;8(3):186-94. PubMed PMID: 9521922.
88. Illumina Inc. - HiSeq X Ten Datasheet. Available from: <https://www.illumina.com/content/dam/illumina-marketing/documents/products/datasheets/datasheet-hiseq-x-ten.pdf>.

89. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*. 2009;25(14):1754-60. Epub 20090518. doi: 10.1093/bioinformatics/btp324. PubMed PMID: 19451168; PubMed Central PMCID: PMC2705234.
90. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods*. 2012;9(4):357-9. Epub 20120304. doi: 10.1038/nmeth.1923. PubMed PMID: 22388286; PubMed Central PMCID: PMC3322381.
91. Robinson JT, Thorvaldsdottir H, Winckler W, Guttman M, Lander ES, Getz G, et al. Integrative genomics viewer. *Nat Biotechnol*. 2011;29(1):24-6. doi: 10.1038/nbt.1754. PubMed PMID: 21221095; PubMed Central PMCID: PMC3346182.
92. IGV. Available from: <https://igv.org/>.
93. Jung H, Winefield C, Bombarely A, Prentis P, Waterhouse P. Tools and Strategies for Long-Read Sequencing and De Novo Assembly of Plant Genomes. *Trends Plant Sci*. 2019;24(8):700-24. Epub 20190614. doi: 10.1016/j.tplants.2019.05.003. PubMed PMID: 31208890.
94. Hert DG, Fredlake CP, Barron AE. Advantages and limitations of next-generation sequencing technologies: a comparison of electrophoresis and non-electrophoresis methods. *Electrophoresis*. 2008;29(23):4618-26. doi: 10.1002/elps.200800456. PubMed PMID: 19053153.
95. The Cost of Sequencing a Human Genome. Available from: <https://www.genome.gov/about-genomics/fact-sheets/Sequencing-Human-Genome-cost>.
96. Serrati S, De Summa S, Pilato B, Petriella D, Lacalamita R, Tommasi S, et al. Next-generation sequencing: advances and applications in cancer diagnosis. *Onco Targets Ther*. 2016;9:7355-65. Epub 20161202. doi: 10.2147/OTT.S99807. PubMed PMID: 27980425; PubMed Central PMCID: PMC5144906.
97. Artech-Lopez A, Avila-Fernandez A, Romero R, Riveiro-Alvarez R, Lopez-Martinez MA, Gimenez-Pardo A, et al. Sanger sequencing is no longer always necessary based on a single-center validation of 1109 NGS variants in 825 clinical exomes. *Sci Rep*. 2021;11(1):5697. Epub 20210311. doi: 10.1038/s41598-021-85182-w. PubMed PMID: 33707547; PubMed Central PMCID: PMC7952542.
98. McCourt CM, McArt DG, Mills K, Catherwood MA, Maxwell P, Waugh DJ, et al. Validation of next generation sequencing technologies in comparison to current diagnostic gold standards for BRAF, EGFR and KRAS mutational analysis. *PLoS One*. 2013;8(7):e69604. Epub 20130726. doi: 10.1371/journal.pone.0069604. PubMed PMID: 23922754; PubMed Central PMCID: PMC3724913.
99. Beck TF, Mullikin JC, Program NCS, Biesecker LG. Systematic Evaluation of Sanger Validation of Next-Generation Sequencing Variants. *Clin Chem*. 2016;62(4):647-54. Epub 20160204. doi: 10.1373/clinchem.2015.249623. PubMed PMID: 26847218; PubMed Central PMCID: PMC4878677.

100. Hu T, Chitnis N, Monos D, Dinh A. Next-generation sequencing technologies: An overview. *Hum Immunol.* 2021;82(11):801-11. Epub 20210319. doi: 10.1016/j.humimm.2021.02.012. PubMed PMID: 33745759.
101. Ozsolak F. Third-generation sequencing techniques and applications to drug discovery. *Expert Opin Drug Discov.* 2012;7(3):231-43. Epub 20120202. doi: 10.1517/17460441.2012.660145. PubMed PMID: 22468954; PubMed Central PMCID: PMC3319653.
102. Xiao T, Zhou W. The third generation sequencing: the advanced approach to genetic diseases. *Transl Pediatr.* 2020;9(2):163-73. doi: 10.21037/tp.2020.03.06. PubMed PMID: 32477917; PubMed Central PMCID: PMC7237973.
103. Schadt EE, Turner S, Kasarskis A. A window into third-generation sequencing. *Hum Mol Genet.* 2010;19(R2):R227-40. Epub 20100921. doi: 10.1093/hmg/ddq416. PubMed PMID: 20858600.
104. Athanasopoulou K, Boti MA, Adamopoulos PG, Skourou PC, Scorilas A. Third-Generation Sequencing: The Spearhead towards the Radical Transformation of Modern Genomics. *Life (Basel).* 2021;12(1). Epub 20211226. doi: 10.3390/life12010030. PubMed PMID: 35054423; PubMed Central PMCID: PMC8780579.
105. Hardy A, Duharcourt S, Defrance M. DNA Modification Patterns Filtering and Analysis Using DNAModAnnot. *Methods Mol Biol.* 2023;2624:87-114. doi: 10.1007/978-1-0716-2962-8_7. PubMed PMID: 36723811.
106. Eid J, Fehr A, Gray J, Luong K, Lyle J, Otto G, et al. Real-time DNA sequencing from single polymerase molecules. *Science.* 2009;323(5910):133-8. Epub 20081120. doi: 10.1126/science.1162986. PubMed PMID: 19023044.
107. PacBio - Epigenetics. Available from: <https://www.pacb.com/wp-content/uploads/Epigenetics.png>.
108. Kasianowicz JJ, Brandin E, Branton D, Deamer DW. Characterization of individual polynucleotide molecules using a membrane channel. *Proc Natl Acad Sci U S A.* 1996;93(24):13770-3. doi: 10.1073/pnas.93.24.13770. PubMed PMID: 8943010; PubMed Central PMCID: PMC19421.
109. Deamer DW, Akeson M. Nanopores and nucleic acids: prospects for ultrarapid sequencing. *Trends Biotechnol.* 2000;18(4):147-51. doi: 10.1016/s0167-7799(00)01426-8. PubMed PMID: 10740260.
110. Mikheyev AS, Tin MM. A first look at the Oxford Nanopore MinION sequencer. *Mol Ecol Resour.* 2014;14(6):1097-102. Epub 20140924. doi: 10.1111/1755-0998.12324. PubMed PMID: 25187008.
111. Lin B, Hui J, Mao H. Nanopore Technology and Its Applications in Gene Sequencing. *Biosensors (Basel).* 2021;11(7). Epub 20210630. doi: 10.3390/bios11070214. PubMed PMID: 34208844; PubMed Central PMCID: PMC8301755.

112. Shivashakarappa K RV, Tupakula VK, Farnian A, Vuppula A, Gunnaiah R. Nanotechnology for the detection of plant pathogens. *Plant Nano Biology*. 2022;Volume 2, November 2022, 100018. doi: <https://doi.org/10.1016/j.plana.2022.100018>.
113. Joshi D, Diggavi S, Chaisson MJP, Kannan S. HQAlign: Aligning nanopore reads for SV detection using current-level modeling. *ArXiv*. 2023. Epub 20230110. PubMed PMID: 36713252; PubMed Central PMCID: PMC9882582.
114. Stoler N, Nekrutenko A. Sequencing error profiles of Illumina sequencing instruments. *NAR Genom Bioinform*. 2021;3(1):lqab019. Epub 20210327. doi: 10.1093/nargab/lqab019. PubMed PMID: 33817639; PubMed Central PMCID: PMC8002175.
115. Choi Y, Chan AP, Kirkness E, Telenti A, Schork NJ. Comparison of phasing strategies for whole human genomes. *PLoS Genet*. 2018;14(4):e1007308. Epub 20180405. doi: 10.1371/journal.pgen.1007308. PubMed PMID: 29621242; PubMed Central PMCID: PMC5903673.
116. Li T, Ball D, Zhao J, Murray RM, Liu X, Sham PC, et al. Family-based linkage disequilibrium mapping using SNP marker haplotypes: application to a potential locus for schizophrenia at chromosome 22q11. *Mol Psychiatry*. 2000;5(1):77-84. doi: 10.1038/sj.mp.4000638. PubMed PMID: 10673772.
117. Ott A, Schnable JC, Yeh CT, Wu L, Liu C, Hu HC, et al. Linked read technology for assembling large complex and polyploid genomes. *BMC Genomics*. 2018;19(1):651. Epub 20180904. doi: 10.1186/s12864-018-5040-z. PubMed PMID: 30180802; PubMed Central PMCID: PMC6122573.
118. Mostovoy Y, Levy-Sakin M, Lam J, Lam ET, Hastie AR, Marks P, et al. A hybrid approach for de novo human genome sequence assembly and phasing. *Nat Methods*. 2016;13(7):587-90. Epub 20160509. doi: 10.1038/nmeth.3865. PubMed PMID: 27159086; PubMed Central PMCID: PMC4927370.
119. Kitzman JO. Haplotypes drop by drop. *Nat Biotechnol*. 2016;34(3):296-8. doi: 10.1038/nbt.3500. PubMed PMID: 26963554.
120. Roy J. Carver Biotechnology Center - Applications. Available from: <https://biotech.illinois.edu/htdna/applications>.
121. Huang SW, Hung SJ, Wang JR. Application of deep sequencing methods for inferring viral population diversity. *J Virol Methods*. 2019;266:95-102. Epub 20190125. doi: 10.1016/j.jviromet.2019.01.013. PubMed PMID: 30690049.
122. Nkone P, Loubser S, Quinn TC, Redd AD, Ismail A, Tiemessen CT, et al. Deep sequencing of the HIV-1 polymerase gene for characterisation of cytotoxic T-lymphocyte epitopes during early and chronic disease stages. *Virol J*. 2022;19(1):56. Epub 20220328. doi: 10.1186/s12985-022-01772-8. PubMed PMID: 35346259; PubMed Central PMCID: PMC8959563.

123. Brunstein J. In-depth coverage: some useful NGS terms. *MLO Med Lab Obs.* 2014;46(11):32-3. PubMed PMID: 25626214.
124. Jiang Y, Jiang Y, Wang S, Zhang Q, Ding X. Optimal sequencing depth design for whole genome re-sequencing in pigs. *BMC Bioinformatics.* 2019;20(1):556. Epub 20191108. doi: 10.1186/s12859-019-3164-z. PubMed PMID: 31703550; PubMed Central PMCID: PMC6839175.
125. Melas M, Mathew MT, Mori M, Jayaraman V, Wilson SA, Martin C, et al. Somatic variation as an incidental finding in the pediatric next-generation sequencing era. *Cold Spring Harb Mol Case Stud.* 2021;7(6). Epub 20211209. doi: 10.1101/mcs.a006135. PubMed PMID: 34716204; PubMed Central PMCID: PMC8751410.
126. Grobe N, Cherif A, Wang X, Dong Z, Kotanko P. Sample pooling: burden or solution? *Clin Microbiol Infect.* 2021;27(9):1212-20. Epub 20210418. doi: 10.1016/j.cmi.2021.04.007. PubMed PMID: 33878507.
127. Brynildsrud O. COVID-19 prevalence estimation by random sampling in population - optimal sample pooling under varying assumptions about true prevalence. *BMC Med Res Methodol.* 2020;20(1):196. Epub 20200723. doi: 10.1186/s12874-020-01081-0. PubMed PMID: 32703158; PubMed Central PMCID: PMC7376319.
128. Phatarfod RM, Sudbury A. The use of a square array scheme in blood testing. *Stat Med.* 1994;13(22):2337-43. doi: 10.1002/sim.4780132205. PubMed PMID: 7855467.
129. Roth WK. History and Future of Nucleic Acid Amplification Technology Blood Donor Testing. *Transfus Med Hemother.* 2019;46(2):67-75. Epub 20190205. doi: 10.1159/000496749. PubMed PMID: 31191192; PubMed Central PMCID: PMC6514489.
130. Schottstedt V, Tuma W, Bunger G, Lefevre H. PCR for HBV, HCV and HIV-1 experiences and first results from a routine screening programme in a large blood transfusion service. *Biologicals.* 1998;26(2):101-4. doi: 10.1006/biol.1998.0144. PubMed PMID: 9811513.
131. Olsson C, Waldenstrom E, Westermarck K, Landegre U, Syvanen AC. Determination of the frequencies of ten allelic variants of the Wilson disease gene (ATP7B), in pooled DNA samples. *Eur J Hum Genet.* 2000;8(12):933-8. doi: 10.1038/sj.ejhg.5200566. PubMed PMID: 11175281.
132. Wang T, Lin CY, Rohan TE, Ye K. Resequencing of pooled DNA for detecting disease associations with rare variants. *Genet Epidemiol.* 2010;34(5):492-501. doi: 10.1002/gepi.20502. PubMed PMID: 20578089; PubMed Central PMCID: PMC4096227.
133. Lee JS, Choi M, Yan X, Lifton RP, Zhao H. On optimal pooling designs to identify rare variants through massive resequencing. *Genet Epidemiol.* 2011;35(3):139-47. Epub 20110119. doi: 10.1002/gepi.20561. PubMed PMID: 21254222; PubMed Central PMCID: PMC3176340.
134. He D, Zaitlen N, Pasaniuc B, Eskin E, Halperin E. Genotyping common and rare variation using overlapping pool sequencing. *BMC Bioinformatics.* 2011;12 Suppl 6(Suppl

6):S2. Epub 20110728. doi: 10.1186/1471-2105-12-S6-S2. PubMed PMID: 21989232; PubMed Central PMCID: PMC3194190.

135. Lin WY, Zhang B, Yi N, Gao G, Liu N. Evaluation of pooled association tests for rare variant identification. *BMC Proc.* 2011;5 Suppl 9(Suppl 9):S118. Epub 20111129. doi: 10.1186/1753-6561-5-S9-S118. PubMed PMID: 22373333; PubMed Central PMCID: PMC3287842.

136. Golan D, Erlich Y, Rosset S. Weighted pooling--practical and cost-effective techniques for pooled high-throughput sequencing. *Bioinformatics.* 2012;28(12):i197-206. doi: 10.1093/bioinformatics/bts208. PubMed PMID: 22689761; PubMed Central PMCID: PMC3371840.

137. Ramsey DM, Futschik A. DNA pooling and statistical tests for the detection of single nucleotide polymorphisms. *Stat Appl Genet Mol Biol.* 2012;11(5):Article 1. Epub 20120925. doi: 10.1515/1544-6115.1763. PubMed PMID: 23023700.

138. Wang W, Yin X, Soo Pyon Y, Hayes M, Li J. Rare variant discovery and calling by sequencing pooled samples with overlaps. *Bioinformatics.* 2013;29(1):29-38. Epub 20121027. doi: 10.1093/bioinformatics/bts645. PubMed PMID: 23104896; PubMed Central PMCID: PMC3530907.

139. Zakharov S, Salim A, Thalamuthu A. Comparison of similarity-based tests and pooling strategies for rare variants. *BMC Genomics.* 2013;14:50. Epub 20130124. doi: 10.1186/1471-2164-14-50. PubMed PMID: 23343094; PubMed Central PMCID: PMC3600007.

140. Cao CC, Li C, Huang Z, Ma X, Sun X. Identifying rare variants with optimal depth of coverage and cost-effective overlapping pool sequencing. *Genet Epidemiol.* 2013;37(8):820-30. Epub 20131028. doi: 10.1002/gepi.21769. PubMed PMID: 24166758.

141. Cao CC, Li C, Sun X. Quantitative group testing-based overlapping pool sequencing to identify rare variant carriers. *BMC Bioinformatics.* 2014;15:195. Epub 20140617. doi: 10.1186/1471-2105-15-195. PubMed PMID: 24934981; PubMed Central PMCID: PMC4229885.

142. Malde K. Estimating the information value of polymorphic sites using pooled sequences. *BMC Genomics.* 2014;15 Suppl 6(Suppl 6):S20. Epub 20141017. doi: 10.1186/1471-2164-15-S6-S20. PubMed PMID: 25571927; PubMed Central PMCID: PMC4239578.

143. Guan W, Li C. Design of DNA pooling to allow incorporation of covariates in rare variants analysis. *PLoS One.* 2014;9(12):e114523. Epub 20141208. doi: 10.1371/journal.pone.0114523. PubMed PMID: 25485788; PubMed Central PMCID: PMC4259344.

144. Bansal V. A statistical method for the detection of variants from next-generation resequencing of DNA pools. *Bioinformatics.* 2010;26(12):i318-24. doi:

10.1093/bioinformatics/btq214. PubMed PMID: 20529923; PubMed Central PMCID: PMC2881398.

145. Kim SY, Li Y, Guo Y, Li R, Holmkvist J, Hansen T, et al. Design of association studies with pooled or un-pooled next-generation sequencing data. *Genet Epidemiol.* 2010;34(5):479-91. doi: 10.1002/gepi.20501. PubMed PMID: 20552648; PubMed Central PMCID: PMC5001557.

146. Niranjana TS, Adamczyk A, Bravo HC, Taub MA, Wheelan SJ, Irizarry R, et al. Effective detection of rare variants in pooled DNA samples using Cross-pool tailcurve analysis. *Genome Biol.* 2011;12(9):R93. Epub 20110928. doi: 10.1186/gb-2011-12-9-r93. PubMed PMID: 21955804; PubMed Central PMCID: PMC3308056.

147. Koboldt DC, Chen K, Wylie T, Larson DE, McLellan MD, Mardis ER, et al. VarScan: variant detection in massively parallel sequencing of individual and pooled samples. *Bioinformatics.* 2009;25(17):2283-5. Epub 20090619. doi: 10.1093/bioinformatics/btp373. PubMed PMID: 19542151; PubMed Central PMCID: PMC2734323.

148. Druley TE, Vallania FL, Wegner DJ, Varley KE, Knowles OL, Bonds JA, et al. Quantification of rare allelic variants from pooled genomic DNA. *Nat Methods.* 2009;6(4):263-5. Epub 20090301. doi: 10.1038/nmeth.1307. PubMed PMID: 19252504; PubMed Central PMCID: PMC2776647.

149. Out AA, van Minderhout IJ, Goeman JJ, Ariyurek Y, Ossowski S, Schneeberger K, et al. Deep sequencing to reveal new variants in pooled DNA samples. *Hum Mutat.* 2009;30(12):1703-12. doi: 10.1002/humu.21122. PubMed PMID: 19842214.

150. Lupton MK, Proitsi P, Danillidou M, Tsolaki M, Hamilton G, Wroe R, et al. Deep sequencing of the Nicastrin gene in pooled DNA, the identification of genetic variants that affect risk of Alzheimer's disease. *PLoS One.* 2011;6(2):e17298. Epub 20110225. doi: 10.1371/journal.pone.0017298. PubMed PMID: 21364883; PubMed Central PMCID: PMC3045431.

151. Marroni F, Pinosio S, Di Centa E, Jurman I, Boerjan W, Felice N, et al. Large-scale detection of rare variants via pooled multiplexed next-generation sequencing: towards next-generation Ecotilling. *Plant J.* 2011;67(4):736-45. Epub 20110711. doi: 10.1111/j.1365-313X.2011.04627.x. PubMed PMID: 21554453.

152. Jin SC, Pastor P, Cooper B, Cervantes S, Benitez BA, Razquin C, et al. Pooled-DNA sequencing identifies novel causative variants in PSEN1, GRN and MAPT in a clinical early-onset and familial Alzheimer's disease Ibero-American cohort. *Alzheimers Res Ther.* 2012;4(4):34. Epub 20120820. doi: 10.1186/alzrt137. PubMed PMID: 22906081; PubMed Central PMCID: PMC3506948.

153. Ramos E, Levinson BT, Chasnoff S, Hughes A, Young AL, Thornton K, et al. Population-based rare variant detection via pooled exome or custom hybridization capture with or without individual indexing. *BMC Genomics.* 2012;13:683. Epub 20121206. doi:

10.1186/1471-2164-13-683. PubMed PMID: 23216810; PubMed Central PMCID: PMC3534616.

154. Chen CT, McDavid AN, Kahsai OJ, Zebari AS, Carlson CS. Efficient identification of rare variants in large populations: deep re-sequencing the CRP locus in the CARDIA study. *Nucleic Acids Res.* 2013;41(7):e85. Epub 20130213. doi: 10.1093/nar/gkt092. PubMed PMID: 23408856; PubMed Central PMCID: PMC3627584.

155. Cheeseman IH, McDew-White M, Phyo AP, Sriprawat K, Nosten F, Anderson TJ. Pooled sequencing and rare variant association tests for identifying the determinants of emerging drug resistance in malaria parasites. *Mol Biol Evol.* 2015;32(4):1080-90. Epub 20141221. doi: 10.1093/molbev/msu397. PubMed PMID: 25534029; PubMed Central PMCID: PMC4379400.

156. Anand S, Mangano E, Barizzzone N, Bordoni R, Sorosina M, Clarelli F, et al. Next Generation Sequencing of Pooled Samples: Guideline for Variants' Filtering. *Sci Rep.* 2016;6:33735. Epub 20160927. doi: 10.1038/srep33735. PubMed PMID: 27670852; PubMed Central PMCID: PMC5037392.

157. Jiao H, Kulyte A, Naslund E, Thorell A, Gerdhem P, Kere J, et al. Whole-Exome Sequencing Suggests LAMB3 as a Susceptibility Gene for Morbid Obesity. *Diabetes.* 2016;65(10):2980-9. Epub 20160718. doi: 10.2337/db16-0522. PubMed PMID: 27431458.

158. Diez-Fairen M, Benitez BA, Ortega-Cubero S, Lorenzo-Betancor O, Cruchaga C, Lorenzo E, et al. Pooled-DNA target sequencing of Parkinson genes reveals novel phenotypic associations in Spanish population. *Neurobiol Aging.* 2018;70:325 e1- e5. Epub 20180514. doi: 10.1016/j.neurobiolaging.2018.05.008. PubMed PMID: 29887346.

159. Mitsui J, Fukuda Y, Azuma K, Tozaki H, Ishiura H, Takahashi Y, et al. Multiplexed resequencing analysis to identify rare variants in pooled DNA with barcode indexing using next-generation sequencer. *J Hum Genet.* 2010;55(7):448-55. Epub 20100520. doi: 10.1038/jhg.2010.46. PubMed PMID: 20485446.

160. Harakalova M, Nijman IJ, Medic J, Mokry M, Renkens I, Blankensteijn JD, et al. Genomic DNA pooling strategy for next-generation sequencing-based rare variant discovery in abdominal aortic aneurysm regions of interest-challenges and limitations. *J Cardiovasc Transl Res.* 2011;4(3):271-80. Epub 20110301. doi: 10.1007/s12265-011-9263-5. PubMed PMID: 21360310; PubMed Central PMCID: PMC3099005.

161. Wang SR, Carmichael H, Andrew SF, Miller TC, Moon JE, Derr MA, et al. Large-scale pooled next-generation sequencing of 1077 genes to identify genetic causes of short stature. *J Clin Endocrinol Metab.* 2013;98(8):E1428-37. Epub 20130614. doi: 10.1210/jc.2013-1534. PubMed PMID: 23771920; PubMed Central PMCID: PMC3733853.

162. Zuzarte PC, Denroche RE, Fehringer G, Katzov-Eckert H, Hung RJ, McPherson JD. A two-dimensional pooling strategy for rare variant detection on next-generation sequencing platforms. *PLoS One.* 2014;9(4):e93455. Epub 20140411. doi: 10.1371/journal.pone.0093455. PubMed PMID: 24728235; PubMed Central PMCID: PMC3984111.

163. Pihlstrom L, Rengmark A, Bjornara KA, Toft M. Effective variant detection by targeted deep sequencing of DNA pools: an example from Parkinson's disease. *Ann Hum Genet.* 2014;78(3):243-52. Epub 20140324. doi: 10.1111/ahg.12060. PubMed PMID: 24660942.
164. Gui H, Bao JY, Tang CS, So MT, Ngo DN, Tran AQ, et al. Targeted next-generation sequencing on Hirschsprung disease: a pilot study exploits DNA pooling. *Ann Hum Genet.* 2014;78(5):381-7. Epub 20140620. doi: 10.1111/ahg.12076. PubMed PMID: 24947032.
165. Ferraro MB, Savarese M, Di Fruscio G, Nigro V, Guarracino MR. Prediction of rare single-nucleotide causative mutations for muscular diseases in pooled next-generation sequencing experiments. *J Comput Biol.* 2014;21(9):665-75. Epub 20140716. doi: 10.1089/cmb.2014.0037. PubMed PMID: 25029289.
166. Wang J, Skoog T, Einarsdottir E, Kaartokallio T, Laivuori H, Grauers A, et al. Investigation of rare and low-frequency variants using high-throughput sequencing with pooled DNA samples. *Sci Rep.* 2016;6:33256. Epub 20160916. doi: 10.1038/srep33256. PubMed PMID: 27633116; PubMed Central PMCID: PMC5025741.
167. Suszynska M, Ratajska M, Galka-Marciniak P, Ryszkowska A, Wydra D, Debniak J, et al. Variant Identification in BARD1, PRDM9, RCC1, and RECQL in Patients with Ovarian Cancer by Targeted Next-generation Sequencing of DNA Pools. *Cancer Prev Res (Phila).* 2022;15(3):151-60. doi: 10.1158/1940-6207.CAPR-21-0295. PubMed PMID: 34906988.
168. Gomez J, Reguero JR, Moris C, Alvarez V, Coto E. Non optical semi-conductor next generation sequencing of the main cardiac QT-interval duration genes in pooled DNA samples. *J Cardiovasc Transl Res.* 2014;7(1):133-7. Epub 20131105. doi: 10.1007/s12265-013-9516-6. PubMed PMID: 24190697.
169. Sobczyk J, Pyne MT, Barker A, Mayer J, Hanson KE, Samore MH, et al. Efficient and effective single-step screening of individual samples for SARS-CoV-2 RNA using multi-dimensional pooling and Bayesian inference. *J R Soc Interface.* 2021;18(179):20210155. Epub 20210616. doi: 10.1098/rsif.2021.0155. PubMed PMID: 34129787; PubMed Central PMCID: PMC8205536.
170. Turgeon RD, Barry AR, Pearson GJ. Familial hypercholesterolemia: Review of diagnosis, screening, and treatment. *Can Fam Physician.* 2016;62(1):32-7. PubMed PMID: 26796832; PubMed Central PMCID: PMC4721838.
171. Bouhairie VE, Goldberg AC. Familial hypercholesterolemia. *Cardiol Clin.* 2015;33(2):169-79. doi: 10.1016/j.ccl.2015.01.001. PubMed PMID: 25939291; PubMed Central PMCID: PMC4472364.
172. Defesche JC. Defining the challenges of FH screening for familial hypercholesterolemia. *J Clin Lipidol.* 2010;4(5):338-41. doi: 10.1016/j.jacl.2010.08.022. PubMed PMID: 21122674.

173. Brice P, Burton H, Edwards CW, Humphries SE, Aitman TJ. Familial hypercholesterolaemia: a pressing issue for European health care. *Atherosclerosis*. 2013;231(2):223-6. Epub 20131002. doi: 10.1016/j.atherosclerosis.2013.09.019. PubMed PMID: 24267231.
174. Civeira F, International Panel on Management of Familial H. Guidelines for the diagnosis and management of heterozygous familial hypercholesterolemia. *Atherosclerosis*. 2004;173(1):55-68. doi: 10.1016/j.atherosclerosis.2003.11.010. PubMed PMID: 15177124.
175. Zubieliene K, Valteryte G, Jonaitiene N, Zaliaduonyte D, Zabiela V. Familial Hypercholesterolemia and Its Current Diagnostics and Treatment Possibilities: A Literature Analysis. *Medicina (Kaunas)*. 2022;58(11). Epub 20221117. doi: 10.3390/medicina58111665. PubMed PMID: 36422206; PubMed Central PMCID: PMC9692978.
176. Yuan G, Wang J, Hegele RA. Heterozygous familial hypercholesterolemia: an underrecognized cause of early cardiovascular disease. *CMAJ*. 2006;174(8):1124-9. doi: 10.1503/cmaj.051313. PubMed PMID: 16606962; PubMed Central PMCID: PMC1421462.
177. Goldberg AC, Hopkins PN, Toth PP, Ballantyne CM, Rader DJ, Robinson JG, et al. Familial hypercholesterolemia: screening, diagnosis and management of pediatric and adult patients: clinical guidance from the National Lipid Association Expert Panel on Familial Hypercholesterolemia. *J Clin Lipidol*. 2011;5(3):133-40. Epub 20110311. doi: 10.1016/j.jacl.2011.03.001. PubMed PMID: 21600517.
178. Almostadoctor - Dyslipidaemia. Available from: <https://almostadoctor.co.uk/encyclopedia/dyslipidaemia>.
179. Thomas H, Diamond J, Vieco A, Chaudhuri S, Shinnar E, Cromer S, et al. Global Atlas of Cardiovascular Disease 2000-2016: The Path to Prevention and Control. *Glob Heart*. 2018;13(3):143-63. Epub 20181006. doi: 10.1016/j.gheart.2018.09.511. PubMed PMID: 30301680.
180. Frostegard J. Immunity, atherosclerosis and cardiovascular disease. *BMC Med*. 2013;11:117. Epub 20130501. doi: 10.1186/1741-7015-11-117. PubMed PMID: 23635324; PubMed Central PMCID: PMC3658954.
181. Wikipedia - Atherosclerosis. Available from: <https://en.wikipedia.org/wiki/Atherosclerosis>.
182. Lind L. Circulating markers of inflammation and atherosclerosis. *Atherosclerosis*. 2003;169(2):203-14. doi: 10.1016/s0021-9150(03)00012-1. PubMed PMID: 12921971.
183. What Is Atherosclerosis? Available from: <https://www.nhlbi.nih.gov/health/atherosclerosis>.
184. Lee Y, Siddiqui WJ. Cholesterol Levels2022.

185. Kwiterovich PO, Jr. Recognition and management of dyslipidemia in children and adolescents. *J Clin Endocrinol Metab.* 2008;93(11):4200-9. Epub 20080812. doi: 10.1210/jc.2008-1270. PubMed PMID: 18697860.
186. Singh S, Bittner V. Familial hypercholesterolemia--epidemiology, diagnosis, and screening. *Curr Atheroscler Rep.* 2015;17(2):482. doi: 10.1007/s11883-014-0482-5. PubMed PMID: 25612857.
187. Familial hypercholesterolaemia (FH) : report of a second WHO consultation, Geneva, 4 September 1998. 1999. Available from: <https://apps.who.int/iris/handle/10665/66346>.
188. Hobbs HH, Brown MS, Goldstein JL. Molecular genetics of the LDL receptor gene in familial hypercholesterolemia. *Hum Mutat.* 1992;1(6):445-66. doi: 10.1002/humu.1380010602. PubMed PMID: 1301956.
189. HYPERCHOLESTEROLEMIA, FAMILIAL, 2; FHCL2. Available from: <https://www.omim.org/entry/144010>.
190. HYPERCHOLESTEROLEMIA, FAMILIAL, 4; FHCL4. Available from: <https://www.omim.org/entry/603813>.
191. HYPERCHOLESTEROLEMIA, FAMILIAL, 3; FHCL3. Available from: <https://www.omim.org/entry/603776>.
192. Watts GF, Sullivan DR, Poplawski N, van Bockxmeer F, Hamilton-Craig I, Clifton PM, et al. Familial hypercholesterolaemia: a model of care for Australasia. *Atheroscler Suppl.* 2011;12(2):221-63. Epub 20110913. doi: 10.1016/j.atherosclerosissup.2011.06.001. PubMed PMID: 21917530.
193. LOW DENSITY LIPOPROTEIN RECEPTOR; LDLR. Available from: <https://www.omim.org/entry/606945>.
194. Brown MS, Goldstein JL. Receptor-mediated endocytosis: insights from the lipoprotein receptor system. *Proc Natl Acad Sci U S A.* 1979;76(7):3330-7. doi: 10.1073/pnas.76.7.3330. PubMed PMID: 226968; PubMed Central PMCID: PMC383819.
195. Brown MS, Goldstein JL. Analysis of a mutant strain of human fibroblasts with a defect in the internalization of receptor-bound low density lipoprotein. *Cell.* 1976;9(4 PT 2):663-74. doi: 10.1016/0092-8674(76)90130-6. PubMed PMID: 189940.
196. APOLIPOPROTEIN B; APOB. Available from: <https://www.omim.org/entry/107730>.
197. Soria LF, Ludwig EH, Clarke HR, Vega GL, Grundy SM, McCarthy BJ. Association between a specific apolipoprotein B mutation and familial defective apolipoprotein B-100. *Proc Natl Acad Sci U S A.* 1989;86(2):587-91. doi: 10.1073/pnas.86.2.587. PubMed PMID: 2563166; PubMed Central PMCID: PMC286517.

198. Awan Z, Baass A, Genest J. Proprotein convertase subtilisin/kexin type 9 (PCSK9): lessons learned from patients with hypercholesterolemia. *Clin Chem*. 2014;60(11):1380-9. Epub 20140923. doi: 10.1373/clinchem.2014.225946. PubMed PMID: 25248569.
199. PROPROTEIN CONVERTASE, SUBTILISIN/KEXIN-TYPE, 9; PCSK9. Available from: <https://www.omim.org/entry/607786>.
200. Abifadel M, Varret M, Rabes JP, Allard D, Ouguerram K, Devillers M, et al. Mutations in PCSK9 cause autosomal dominant hypercholesterolemia. *Nat Genet*. 2003;34(2):154-6. doi: 10.1038/ng1161. PubMed PMID: 12730697.
201. Cohen J, Pertsemlidis A, Kotowski IK, Graham R, Garcia CK, Hobbs HH. Low LDL cholesterol in individuals of African descent resulting from frequent nonsense mutations in PCSK9. *Nat Genet*. 2005;37(2):161-5. Epub 20050116. doi: 10.1038/ng1509. PubMed PMID: 15654334.
202. McDonagh M, Peterson K, Holzhammer B, Fazio S. A Systematic Review of PCSK9 Inhibitors Alirocumab and Evolocumab. *J Manag Care Spec Pharm*. 2016;22(6):641-53q. doi: 10.18553/jmcp.2016.22.6.641. PubMed PMID: 27231792.
203. LOW DENSITY LIPOPROTEIN RECEPTOR ADAPTOR PROTEIN 1; LDLRAP1. Available from: <https://www.omim.org/entry/605747>.
204. Canizales-Quinteros S, Aguilar-Salinas CA, Huertas-Vazquez A, Ordonez-Sanchez ML, Rodriguez-Torres M, Venturas-Gallegos JL, et al. A novel ARH splice site mutation in a Mexican kindred with autosomal recessive hypercholesterolemia. *Hum Genet*. 2005;116(1-2):114-20. Epub 20041117. doi: 10.1007/s00439-004-1192-9. PubMed PMID: 15599766.
205. Sanchez-Hernandez RM, Prieto-Matos P, Civeira F, Lafuente EE, Vargas MF, Real JT, et al. Autosomal recessive hypercholesterolemia in Spain. *Atherosclerosis*. 2018;269:1-5. Epub 20171206. doi: 10.1016/j.atherosclerosis.2017.12.006. PubMed PMID: 29245109.
206. APOLIPOPROTEIN A-II; APOA2. Available from: <https://www.omim.org/entry/107670>.
207. Takada D, Emi M, Ezura Y, Nobe Y, Kawamura K, Iino Y, et al. Interaction between the LDL-receptor gene bearing a novel mutation and a variant in the apolipoprotein A-II promoter: molecular study in a 1135-member familial hypercholesterolemia kindred. *J Hum Genet*. 2002;47(12):656-64. doi: 10.1007/s100380200101. PubMed PMID: 12522687.
208. INTER-ALPHA-TRYPSIN INHIBITOR, HEAVY CHAIN 4; ITIH4. Available from: <https://www.omim.org/entry/600564>.
209. Fujita Y, Ezura Y, Emi M, Sato K, Takada D, Iino Y, et al. Hypercholesterolemia associated with splice-junction variation of inter-alpha-trypsin inhibitor heavy chain 4 (ITIH4) gene. *J Hum Genet*. 2004;49(1):24-8. Epub 20031206. doi: 10.1007/s10038-003-0101-8. PubMed PMID: 14661079.

210. GROWTH HORMONE RECEPTOR; GHR. Available from: <https://www.omim.org/entry/600946>.
211. Takada D, Ezura Y, Ono S, Iino Y, Katayama Y, Xin Y, et al. Growth hormone receptor variant (L526I) modifies plasma HDL cholesterol phenotype in familial hypercholesterolemia: intra-familial association study in an eight-generation hyperlipidemic kindred. *Am J Med Genet A*. 2003;121A(2):136-40. doi: 10.1002/ajmg.a.20172. PubMed PMID: 12910492.
212. Amselem S, Duquesnoy P, Attree O, Novelli G, Bousnina S, Postel-Vinay MC, et al. Laron dwarfism and mutations of the growth hormone-receptor gene. *N Engl J Med*. 1989;321(15):989-95. doi: 10.1056/NEJM198910123211501. PubMed PMID: 2779634.
213. Amselem S, Sobrier ML, Duquesnoy P, Rappaport R, Postel-Vinay MC, Gournelen M, et al. Recurrent nonsense mutations in the growth hormone receptor from patients with Laron dwarfism. *J Clin Invest*. 1991;87(3):1098-102. doi: 10.1172/JCI115071. PubMed PMID: 1999489; PubMed Central PMCID: PMC329906.
214. Caglar C, Friedman J. Restriction of food intake by PPP1R17-expressing neurons in the DMH. *Proc Natl Acad Sci U S A*. 2021;118(13). doi: 10.1073/pnas.2100194118. PubMed PMID: 33753517; PubMed Central PMCID: PMC8020659.
215. PROTEIN PHOSPHATASE 1, REGULATORY SUBUNIT 17; PPP1R17. Available from: <https://www.omim.org/entry/604088>.
216. Ono S, Ezura Y, Emi M, Fujita Y, Takada D, Sato K, et al. A promoter SNP (-1323T>C) in G-substrate gene (GSBS) correlates with hypercholesterolemia. *J Hum Genet*. 2003;48(9):447-50. Epub 20030903. doi: 10.1007/s10038-003-0055-x. PubMed PMID: 12955585.
217. LIPASE A, LYSOSOMAL ACID; LIPA. Available from: <https://www.omim.org/entry/613497>.
218. Reiner Z, Guardamagna O, Nair D, Soran H, Hovingh K, Bertolini S, et al. Lysosomal acid lipase deficiency--an under-recognized cause of dyslipidaemia and liver dysfunction. *Atherosclerosis*. 2014;235(1):21-30. Epub 20140415. doi: 10.1016/j.atherosclerosis.2014.04.003. PubMed PMID: 24792990.
219. Sjouke B, Defesche JC, de Randamie JSE, Wiegman A, Fouchier SW, Hovingh GK. Sequencing for LIPA mutations in patients with a clinical diagnosis of familial hypercholesterolemia. *Atherosclerosis*. 2016;251:263-5. Epub 20160709. doi: 10.1016/j.atherosclerosis.2016.07.008. PubMed PMID: 27423329.
220. Muntoni S, Wiebusch H, Jansen-Rust M, Rust S, Schulte H, Berger K, et al. Heterozygosity for lysosomal acid lipase E8SJM mutation and serum lipid concentrations. *Nutr Metab Cardiovasc Dis*. 2013;23(8):732-6. Epub 20120712. doi: 10.1016/j.numecd.2012.05.009. PubMed PMID: 22795295.
221. Van der Auwera GA, Carneiro MO, Hartl C, Poplin R, Del Angel G, Levy-Moonshine A, et al. From FastQ data to high confidence variant calls: the Genome Analysis

Toolkit best practices pipeline. *Curr Protoc Bioinformatics*. 2013;43(1110):11 0 1- 0 33. doi: 10.1002/0471250953.bi1110s43. PubMed PMID: 25431634; PubMed Central PMCID: PMC4243306.

222. The International Genome Sample Resource. Available from: <https://www.internationalgenome.org/>.

223. Mao Y, Zhang G. A complete, telomere-to-telomere human genome sequence presents new opportunities for evolutionary genomics. *Nat Methods*. 2022;19(6):635-8. doi: 10.1038/s41592-022-01512-4. PubMed PMID: 35689027.

224. Sakuragi S, Moriguchi J, Ohashi F, Ikeda M. Reference value and annual trend of white blood cell counts among adult Japanese population. *Environ Health Prev Med*. 2013;18(2):143-50. Epub 20120930. doi: 10.1007/s12199-012-0304-8. PubMed PMID: 23054993; PubMed Central PMCID: PMC3590320.

225. Matsuda T, Bebenek K, Masutani C, Rogozin IB, Hanaoka F, Kunkel TA. Error rate and specificity of human and murine DNA polymerase ϵ . *J Mol Biol*. 2001;312(2):335-46. doi: 10.1006/jmbi.2001.4937. PubMed PMID: 11554790.

226. DNA Replication and Causes of Mutation. Available from: <https://www.nature.com/scitable/topicpage/dna-replication-and-causes-of-mutation-409/>.

227. Shestak AG, Bukaeva AA, Saber S, Zaklyazminskaya EV. Allelic Dropout Is a Common Phenomenon That Reduces the Diagnostic Yield of PCR-Based Sequencing of Targeted Gene Panels. *Front Genet*. 2021;12:620337. Epub 20210201. doi: 10.3389/fgene.2021.620337. PubMed PMID: 33633783; PubMed Central PMCID: PMC7901947.

228. Hirschhorn JW, Avery A, Schandl CA. Managing a PCR Contamination Event in a Molecular Pathology Laboratory. *Methods Mol Biol*. 2023;2621:15-26. doi: 10.1007/978-1-0716-2950-5_2. PubMed PMID: 37041437.

229. Longo MC, Berninger MS, Hartley JL. Use of uracil DNA glycosylase to control carry-over contamination in polymerase chain reactions. *Gene*. 1990;93(1):125-8. doi: 10.1016/0378-1119(90)90145-h. PubMed PMID: 2227421.

230. Illumina Inc. TruSight™ Oncology UMI Reagents. Available from: <https://www.illumina.com/content/dam/illumina-marketing/documents/products/datasheets/trusight-oncology-umi-reagents-datasheet-1000000050425.pdf>.

231. Crowgey EL, Mahajan N, Wong WH, Gopalakrishnapillai A, Barwe SP, Kolb EA, et al. Error-corrected sequencing strategies enable comprehensive detection of leukemic mutations relevant for diagnosis and minimal residual disease monitoring. *BMC Med Genomics*. 2020;13(1):32. Epub 20200304. doi: 10.1186/s12920-020-0671-8. PubMed PMID: 32131829; PubMed Central PMCID: PMC7057603.

232. Arslan S, Garcia FJ, Guo M, Kellinger MW, Kruglyak S, LeVieux JA, et al. Sequencing by avidity enables high accuracy with low reagent consumption. Nat Biotechnol. 2023. Epub 20230525. doi: 10.1038/s41587-023-01750-7. PubMed PMID: 37231263.
233. PacBio Inc. SBB sequencing. Available from: <https://www.pacb.com/blog/sbb-sequencing/>.
234. Nishimura H, Kakizaki I, Muta T, Sasaki N, Pu PX, Yamashita T, et al. cDNA and deduced amino acid sequence of human PK-120, a plasma kallikrein-sensitive glycoprotein. FEBS Lett. 1995;357(2):207-11. doi: 10.1016/0014-5793(94)01364-7. PubMed PMID: 7805892.



EKLER

EK 1: Çalışmada kullanılan primerlerin dizileri ve elde edilen ampliconların boyları

Boş hücreler havuzların mutasyon tespit edilmeyen genlerine aittir. 6 geninin hiçbirisinde mutasyon tespit edilmeyen 8 adet negatif havuz yeşil ile işaretlenmiştir. Hiçbir havuzun APOA2 geninde mutasyon tespit edilmemiştir

Gen	Ekzon	Primer adı	Primer dizisi	Amplicon boyu (bp)
LIPA	2	2F	TTCACCTTTTACTTGCCCTAAATCTGG	313
		2R	GTATGGGTATGGCTTCATGTAGCTCAC	
	3	3F	AAGCCTGGAGAACATAGTTTATCTGCTC	347
		3R	AACCTCAACACAGTTAGTGCTTTCGTC	
	4	4F	ATGGTCCAGTTTCACTGCTACCTTGC	503
		4R	GAGTTACCACCTATCTACCTCTTCTGCTG	
	5	5F	TGTTGTTTCTGGACTCACTGCTGG	509
		5R	TCTCTCACATCCCTATCTTGCTTCATC	
	6	6F	GTGGTATTGTTTGCCTGGGTCTCAG	540
		6R	TTTTCTCCAGACTCTCCTTGCTCTC	
	7	7F	ATGTCCACTGGTTGCCATTCTCTCC	391
		7R	TTAGAGTTCTGATGAGGTCATTCCAC	
	8	8F	ATGGCTCTAGTTTTAGTGCTTTGAAGG	286
		8R	CCAGATCAGATTTGTAAGCAGGTTGTC	
LDLRAP1	1	1F	AGGGCTTGCGATTTGCTATCTTCTG	894
		1R	GGGCACCCGGAGTCAAAAGTTGG	
	2-3	2F	TAGGAAAGAAGGCTGGTGAGAGCTG	1521
		3R	CCAGATACCCAGAGCTGAGCTGC	
	4	4F	AAGTGCCCTGCTTCCAGTCTTTACC	457
		4R	AGTAATGGTCACCTTTGGGTCTGC	
	5-7	5F	AGACCTGGAGAAGAACAGCCTTTGC	1708
		7R	AAGAGGCCAAGCATTGTGAGGAGAC	
	8	8F	CCATAGAAGTCATCAGGTCCAATGC	703
		8R	AGGAAATCAACTTGAGCTGCAGTGTG	
	9	9F	GGAGGTGCTTTGATCTGAGGTTACTCC	437
		9R	GCTCTGATCTTCAGGCTGACAATGC	
APOA2	1-4	1F	CCCATCCATAGTCCCTGTCACCTG	1353
		4R	CAGTGGGTGTTCTAGAGGCCAGC	
ITI4	1	1F	TAGATCCCAATTTCCAGAACCTGC	500
		1R	GTCAGCATAAATCATTCTCCTGAGGTC	

	2	2F	AATGTTACATATGAGATGCCTCACTG	574
		2R	AACACCTGGACAGCACTTAGGGATG	
	3-6	3F	CCTTCGTCCCTTCTATAAAATGAATGC	1721
		6R	AACTAAATCACGTGTTTTCTGTCTGTGC	
	7-11	7F	TTTGATGGTTGTCAGGTTTATGATGTC	2007
		11R	CCCTGTCTGACTGCAACCTGAGAG	
	12-17	12F	TCTCATCCCTGTATGCTCTGTGACTG	2100
		17R	ATGATCCCAGGTTGAACTGACCCAC	
	18-20	18F	GGTCCCTTGAAGTGTCTGATCC	984
		20R	TCCTCCTCTAGGCTCACCTCCAC	
GHR	2	2F	CGGGAAGTAGAATTTATTACAACCTGC	496
		2R	GCCAAAGTTGACTGACTACTGCATTCC	
	3	3F	ATATCCAACCTGCCTTCCCATTGTG	234
		3R	ACATCTCCTCCCAGTCTCCAATTCC	
	4	4F	GGCTTCATCCCAGTAGTTTCTTCACAC	322
		4R	CACTTCCATTCCCACAGTACATCTGAC	
	5	5F	AGTACCAAACGGCCTCGTAGCAGTC	538
		5R	CTTGTTCAACAGCTGATCCCATCTC	
	6	6F	AGAGCAAAAAGAAATATACAACATGGCAG	576
		6R	AGCATGAACATGTGGTATTAGGTCAAAGG	
	7	7F	AACTATCTGGAAGACTAATGCACAGTTATAGC	700
		7R	CACCAAATTCCTGTCTCTAAACTTCCAC	
	8	8F	TGCATTCTGTTTCAGTGTTATCAAGAG	560
		8R	CTTGTTGAGGACTTGAGGGAAAGAGC	
	9-10	9F	TTTCTTTCTTTCAAATACTAAATCTGGCTG	1966
		10R	ACACCCACAATATTAAGACAATCACTGC	
PPP1R17	ex2	2F	TCTCTGAGAGAGGGCAAATATTGGAC	473
		2R	CCTCTCTCTGACTTGCCTGCAAACC	
	ex3	3F	CATCAAGTGCAATTAGGAACCAAACAG	358
		3R	GGGAGTGACAAATTACTTTAAAATACCTG	
	ex4	4F	TGGTGGCCATAATACTAAATTCCTCAG	635
		4R	GTAATGCTCTGACCCTTTCCCATCC	
	ex5	5F	CTAAGAGGTTTCTGGGTTGCCATGC	355
		5R	AAAGAGGTGACTGGCAATTTGATGG	

EK 2: 2D08 analiz sonuçları dizin tablosu.

D129V het D129V	2	3	C167R het C167R	5	6	7	8	2D08x01	C222X	257	258	259	260	261	262	263	D139N het D139N	2D08x33	C222R
9	10	11	12	13	14	15	16	2D08x02		265	266	267	268	269	270	271	272	2D08x34	
17	18	19	20	21	22	23	24	2D08x03		273	274	C222R het	276	277	278	279	280	2D08x35	
25	26	27	28	29	C222X het	31	32	2D08x04		281	282	283	284	285	286	287	D139N het D139N	2D08x36	
33	34	35	36	37	38	IVS4+2T>C IVS4+2T>C	40	2D08x05		289	290	291	292	C173Y het C173Y	294	295	296	2D08x37	
41	42	43	44	C222R hom C222R	46	47	48	2D08x06		297	298	299	300	301	C222R het	303	304	2D08x38	
49	50	51	52	53	54	55	56	2D08x07		305	306	307	308	309	310	311	312	2D08x39	
57	58	59	60	C222X het	62	63	64	2D08x08		313	314	315	316	317	318	319	320	2D08x40	
2D08y01	2D08y02	2D08y03	2D08y04	2D08y05	2D08y06	2D08y07	2D08y08			2D08y33	2D08y34	2D08y35	2D08y36	2D08y37	2D08y38	2D08y39	2D08y40		
65	66	P181R het P181R	68	69	C222R het	71	72	2D08x09	IVS4+2T>C C222R C167W	D139N het D139N	322	323	324	325	326	327	328	2D08x41	
73	D129V het D129V	75	C222R het	77	C167W het	IVS4+2T>C het	80	2D08x10		D139N het D139N	330	331	332	D221N het D221N	334	335	336	2D08x42	
81	D139N het D139N	S177P het S177P	84	85	86	87	88	2D08x11		N169N het N169N	338	339	340	341	342	343	344	2D08x43	
89	90	91	92	93	94	95	96	2D08x12		345	346	347	348	349	350	351	352	2D08x44	
97	98	99	c.502delG het c.502delG	101	Y188X het Y188X	103	C167W het	2D08x13		c.506delA het c.506delA	354	355	356	357	358	359	360	2D08x45	
105	106	107	IVS4+2T>C het	109	110	111	112	2D08x14		361	362	363	364	365	C173Y het C173Y	367	368	2D08x46	
113	114	115	116	117	118	119	120	2D08x15		369	370	371	372	373	374	375	376	2D08x47	
121	IVS4+2T>C het	123	124	125	126	127	128	2D08x16		377	378	379	380	381	382	383	384	2D08x48	
2D08y09	2D08y10	2D08y11	2D08y12	2D08y13	2D08y14	2D08y15	2D08y16			2D08y41	2D08y42	2D08y43	2D08y44	2D08y45	2D08y46	2D08y47	2D08y48		
129	130	131	132	133	134	135	136	2D08x17	G207G	385	386	387	388	389	390	391	392	2D08x49	
137	138	139	140	141	142	143	144	2D08x18		393	394	395	396	397	398	399	400	2D08x50	
145	146	D129V het D129V	148	149	150	151	152	2D08x19		401	402	C231X hom C231X	404	405	406	407	c.372delG het c.372delG	2D08x51	
153	154	G207G hom	156	157	158	159	160	2D08x20		409	410	411	412	413	414	415	416	2D08x52	
161	G207G hom	163	164	165	Y188X het Y188X	167	168	2D08x21		417	418	419	S177L het S177L	421	422	423	424	2D08x53	
169	170	171	172	173	174	175	176	2D08x22		425	426	427	428	429	430	431	432	2D08x54	
177	178	179	180	181	182	183	184	2D08x23		433	D136Y het D136Y	435	436	437	438	C173Y het C173Y	440	2D08x55	
185	186	187	188	189	190	S177P het S177P	192	2D08x24		441	442	443	444	445	446	447	448	2D08x56	
2D08y17	2D08y18	2D08y19	2D08y20	2D08y21	2D08y22	2D08y23	2D08y24			2D08y49	2D08y50	2D08y51	2D08y52	2D08y53	2D08y54	2D08y55	2D08y56		
193	194	195	196	197	198	199	200	2D08x25		449	450	451	452	453	454	455	456	2D08x57	D168E
201	C222R het C222R	203	204	D118Y hom D118Y	206	207	208	2D08x26		457	458	459	460	461	462	463	464	2D08x58	
209	210	211	212	213	214	215	216	2D08x27		465	466	467	468	469	470	471	472	2D08x59	
c.502delG het c.502delG	218	219	D129V het D129V	221	222	223	224	2D08x28		473	S177P hom S177P	475	476	D168E het	478	479	480	2D08x60	
225	226	227	228	229	230	231	232	2D08x29		481	482	483	D168E het	485	486	487	488	2D08x61	
233	234	235	236	237	238	239	240	2D08x30		489	490	491	492	493	494	495	496	2D08x62	
241	242	243	C160R het C160R	245	246	247	248	2D08x31		497	498	499	D139N het D139N	501	502	503	504	2D08x63	
249	250	251	252	253	254	255	256	2D08x32		505	506	507	D139N het D139N	509	510	511	C173Y het C173Y	2D08x64	
2D08y25	2D08y26	2D08y27	2D08y28	2D08y29	2D08y30	2D08y31	2D08y32			2D08y57	2D08y58	2D08y59	2D08y60	2D08y61	2D08y62	2D08y63	2D08y64		

EK 3: 2D16 analiz sonuçları dizin tablosu

D129V het	2	3	C167R het C167R	5	6	7	8	129	130	131	132	133	134	135	136	2D16x01
9	10	11	12	p.153_155delFQC hom p.153_155delFQC	14	15	16	137	138	139	140	141	142	143	144	2D16x02
17	18	19	20	21	22	23	24	145	146	D129V het	148	149	150	151	152	2D16x03
25	26	27	28	29	C222X het	31	32	153	154	G207G hom	156	157	158	159	160	2D16x04
33	34	35	36	37	38	IVS4+2T>C het	40	161	G207G hom	163	164	165	Y188X het	167	168	2D16x05
41	42	43	44	C222R hom	46	47	48	169	170	171	172	173	174	175	176	2D16x06
49	50	51	52	53	54	55	56	177	178	179	180	181	182	183	184	2D16x07
57	58	59	60	C222X het	62	63	64	185	186	187	188	189	190	S177P het	192	2D16x08
65	66	P181R het P181R	68	69	C222R het	71	72	193	194	195	196	197	198	199	200	2D16x09
73	D129V het	75	C222R het	77	C167W het	IVS4+2T>C het	80	201	C222R het	203	204	D118Y hom D118Y	206	207	208	2D16x10
81	D139N het	S177P het	84	85	86	87	88	209	210	211	212	213	214	215	216	2D16x11
89	90	91	92	93	94	95	96	c.502delG het	218	219	D129V het	221	222	223	224	2D16x12
97	98	99	c.502delG het	101	Y188X het	103	C167W het	225	226	227	228	229	230	231	232	2D16x13
105	106	107	IVS4+2T>C het	109	110	111	112	233	234	235	236	237	238	239	240	2D16x14
113	114	115	116	c.340-344delTTTCG het c.340-344delTTTCG	118	119	120	241	242	243	C160R het C160R	245	c.340-344delTTTCG het c.340-344delTTTCG	247	248	2D16x15
121	IVS4+2T>C het	123	124	125	126	127	128	249	250	251	252	253	254	255	256	2D16x16
2D16y01	2D16y02	2D16y03	2D16y04	2D16y05	2D16y06	2D16y07	2D16y08	2D16y09	2D16y10	2D16y11	2D16y12	2D16y13	2D16y14	2D16y15	2D16y16	
D129V	IVS4+2T>C	S177P	C167W	c.502delG	Y188X	C222X	C222R	G207G								

257	258	259	260	261	262	263	D139N het	385	386	387	388	389	390	391	392	2D16x17
265	266	267	268	269	270	271	272	393	394	395	396	397	398	399	400	2D16x18
273	274	C222R het	276	277	278	279	280	401	402	C231X hom C231X	404	405	406	407	c.372delG het c.372delG	2D16x19
281	282	283	284	285	286	287	D139N het	409	410	411	412	413	414	415	416	2D16x20
289	290	291	292	C173Y het	294	295	296	417	418	419	S177L het S177L	421	422	423	424	2D16x21
297	298	299	300	301	C222R het	303	304	425	426	427	428	429	430	431	432	2D16x22
305	306	307	308	309	310	311	312	433	D136Y het	435	436	437	438	C173Y het	440	2D16x23
313	314	315	316	317	318	319	320	441	442	443	444	445	446	447	448	2D16x24
D139N het	322	323	324	325	326	327	328	449	450	451	452	453	454	455	456	2D16x25
D139N het	330	331	332	D221N het D221N	334	335	336	457	458	459	460	461	462	463	464	2D16x26
N169N het N169N	338	339	340	341	342	343	344	465	466	467	468	469	470	471	472	2D16x27
345	346	347	348	349	350	351	352	473	S177P hom S177P	475	476	D168E het	478	479	480	2D16x28
c.506delA het c.506delA	354	355	356	357	358	359	360	481	482	483	D168E het	485	486	487	488	2D16x29
361	362	363	364	365	C173Y het	367	368	489	490	491	492	493	494	495	496	2D16x30
369	370	371	372	373	374	375	376	497	498	499	D139N het	501	502	503	504	2D16x31
377	378	379	380	381	382	383	384	505	506	507	D139N het	509	510	511	C173Y het	2D16x32
2D16y17	2D16y18	2D16y19	2D16y20	2D16y21	2D16y22	2D16y23	2D16y24	2D16y25	2D16y26	2D16y27	2D16y28	2D16y29	2D16y30	2D16y31	2D16y32	
C173Y	C222R	D139N	D168E													

EK 4: 2D24 analiz sonuçları dizin tablosu

D129V het D129V	2	3	C167R het	5	6	7	8	257	258	259	260	261	262	263	D139N het	193	194	195	196	197	198	199	200
9	10	11	12	p.153_155delFQC	14	15	16	265	266	267	268	269	270	271	272	201	C222R het C222R	203	204	D118Y het D118Y	206	207	208
17	18	19	20	21	22	23	24	273	274	C222R het C222R	276	277	278	279	280	209	210	211	212	213	214	215	216
25	26	27	28	29	C222X het C222X	31	32	281	282	283	284	285	286	287	D139N het c.502delG	502delG het c.502delG	218	219	D129V het D129V	221	222	223	224
33	34	35	36	37	38	39	40	289	290	291	292	C173Y het C173Y	294	295	296	225	226	227	228	229	230	231	232
41	42	43	44	C222R hom C222R	46	47	48	297	298	299	300	301	C222R het	303	304	233	234	235	236	237	238	239	240
49	50	51	52	53	54	55	56	305	306	307	308	309	310	311	312	241	242	243	C160R het	245	246delTTT c.340_344delTTTCG	247	248
57	58	59	60	C222X het C222X	62	63	64	313	314	315	316	317	318	319	320	249	250	251	252	253	254	255	256
65	66	P181R het P181R	68	69	C222R het C222R	71	72	D139N het D139N	322	323	324	325	326	327	328	449	450	451	452	453	454	455	456
73	D129V het D129V	75	C222R het C222R	77	C167W het C167W	79	80	D139N het c.506delA	330	331	332	D221N het	334	335	336	457	458	459	460	461	462	463	464
81	D139N het	S177P het S177P	84	85	86	87	88	N160N het	338	339	340	341	342	343	344	465	466	467	468	469	470	471	472
89	90	91	92	93	94	95	96	345	346	347	348	349	350	351	352	473	S177P hom S177P	475	476	D168E het	478	479	480
97	98	99	S02delG het	101	Y188X het Y188X	103	C167W het C167W	S06delA het c.506delA	354	355	356	357	358	359	360	481	482	483	D168E het D168E	485	486	487	488
105	106	107	Y54+2T>C het IV54+2T>C	109	110	111	112	361	362	363	364	365	C173Y het C173Y	367	368	489	490	491	492	493	494	495	496
113	114	115	116	344delTTTCG c.340-344delTTTCG	118	119	120	369	370	371	372	373	374	375	376	497	498	499	D139N het D139N	501	502	503	504
121	Y54+2T>C het IV54+2T>C	123	124	125	126	127	128	377	378	379	380	381	382	383	384	505	506	507	D139N het D139N	509	510	511	C173Y het C173Y
129	130	131	132	133	134	135	136	385	386	387	388	389	390	391	392	D129V het D129V	2	3	C167R het C167R	5	6	7	8
137	138	139	140	141	142	143	144	393	394	395	396	397	398	399	400	9	10	11	12	155delFQC p.153_155delFQC	14	15	16
145	146	D129V het D129V	148	149	150	151	152	401	402	C231X hom C231X	404	405	406	407	372delG het	17	18	19	20	21	22	23	24
153	154	G207G hom G207G	156	157	158	159	160	409	410	411	412	413	414	415	416	25	26	27	28	29	C222X het C222X	31	32
161	G207G hom G207G	163	164	165	Y188X het Y188X	167	168	417	418	419	S177L het	421	422	423	424	33	34	35	36	37	38	39	40
169	170	171	172	173	174	175	176	425	426	427	428	429	430	431	432	41	42	43	44	C222R hom C222R	46	47	48
177	178	179	180	181	182	183	184	433	D136V het	435	436	437	438	C173Y het C173Y	440	49	50	51	52	53	54	55	56
185	186	187	188	189	190	S177P het S177P	192	441	442	443	444	445	446	447	448	57	58	59	60	C222X het C222X	62	63	64
2024y01	2024y02	2024y03	2024y04	2024y05	2024y06	2024y07	2024y08	2024y09	2024y10	2024y11	2024y12	2024y13	2024y14	2024y15	2024y16	2024y17	2024y18	2024y19	2024y20	2024y21	2024y22	2024y23	2024y24

EK 5: 2D32 analiz sonuçları dizin tablosu

2032a01	2032a02	2032a03	2032a04	2032a05	2032a06	2032a07	2032a08	2032a09	2032a10	2032a11	2032a12	2032a13	2032a14	2032a15	2032a16	2032a17	2032a18	2032a19	2032a20	2032a21	2032a22	2032a23	2032a24	2032a25	2032a26	2032a27	2032a28	2032a29	2032a30	2032a31	2032a32		
D129V hnt G207G	9	10	11	12	p.155delATC	14	15	16	265	266	267	268	269	270	271	272	137	138	139	140	141	142	143	144	393	394	395	396	397	398	399	400	
17	18	19	20	21	C222X hnt C222X	23	24	273	274	C222R hnt C222R	276	277	278	279	280	145	146	D129V hnt G207G	148	D129V hnt G207G	156	157	158	159	160	409	410	411	412	413	414	415	
25	26	27	28	29	C222X hnt C222X	31	32	281	282	283	284	285	286	287	D139N hnt G207G	153	154	G207G hnt G207G	163	164	165	166	167	168	417	418	419	420	421	422	423	424	
33	34	35	36	37	38	p.54-27<C-1 V54-27<C	40	289	290	291	292	C173Y hnt C173Y	294	295	296	161	G207G hnt G207G	170	171	172	173	174	175	176	425	426	427	428	429	430	431	432	
41	42	43	44	C222R hnt C222R	46	47	48	297	298	299	300	301	C222R hnt C222R	303	304	169	170	171	172	173	174	175	176	425	426	427	428	429	430	431	432		
49	50	51	52	53	54	55	56	305	306	307	308	309	310	311	312	177	178	179	180	181	182	183	184	433	D139N hnt G136Y	435	436	437	438	C173Y hnt C173Y	440		
57	58	59	60	C222X hnt C222X	62	63	64	313	314	315	316	317	318	319	320	185	186	187	188	189	190	S177P hnt S177P	192	441	442	443	444	445	446	447	448		
65	66	D139N hnt G136Y	68	69	C222R hnt C222R	71	72	D139N hnt G136Y	322	323	324	325	326	327	328	193	194	195	196	197	198	199	200	449	450	451	452	453	454	455	456		
73	D139N hnt G136Y	75	C167N hnt G136Y	77	C167N hnt G136Y	80	81	330	331	332	C221X hnt C221X	334	335	336	201	C221X hnt C221X	203	204	205	206	207	208	457	458	459	460	461	462	463	464			
81	D139N hnt G136Y	84	85	86	87	88	89	338	339	340	341	342	343	344	209	210	211	212	213	214	215	216	465	466	467	468	469	470	471	472			
97	98	99	S02delATC c.502delATC	101	Y188X hnt Y188X	103	C167N hnt G136Y	354	355	356	357	358	359	360	225	226	227	228	229	230	231	232	481	482	483	D139N hnt G136Y	485	486	487	488			
105	106	107	S4-27<C-1 V54-27<C	109	110	111	112	361	362	363	364	365	C173Y hnt C173Y	367	368	233	234	235	236	237	238	239	240	489	490	491	492	493	494	495	496		
113	114	115	116	S4delATC c.3-40-34delATC	118	119	120	369	370	371	372	373	374	375	376	241	242	243	C167N hnt G136Y	245	S4delATC c.3-40-34delATC	247	248	497	498	499	D139N hnt G136Y	501	502	503	504		
121	S4-27<C-1 V54-27<C	123	124	125	126	127	128	377	378	379	380	381	382	383	384	249	250	251	252	253	254	255	256	505	506	507	D139N hnt G136Y	509	510	511	C173Y hnt C173Y		
129	130	131	132	133	134	135	136	385	386	387	388	389	390	391	392	259	260	261	262	263	264	265	266	327	328	65	66	D139N hnt G136Y	68	69	C222R hnt C222R	70	
137	138	139	140	141	142	143	144	393	394	395	396	397	398	399	400	267	268	269	270	271	272	273	274	336	337	73	D139N hnt G136Y	75	D139N hnt G136Y	76	77	C167N hnt G136Y	78
145	146	D129V hnt G207G	148	149	150	151	152	401	402	C221X hnt C221X	404	405	406	407	408	275	276	277	278	279	280	281	282	344	345	81	D139N hnt G136Y	83	84	85	86	87	88
153	154	G207G hnt G207G	156	157	158	159	160	409	410	411	412	413	414	415	416	283	284	285	286	287	288	289	290	351	352	89	90	91	92	93	94	95	96
161	G207G hnt G207G	163	164	165	Y188X hnt Y188X	167	168	417	418	419	S177P hnt S177P	421	422	423	424	291	292	293	294	295	296	297	298	360	361	97	98	99	S02delATC c.502delATC	101	Y188X hnt Y188X	103	C167N hnt G136Y
169	170	171	172	173	174	175	176	425	426	427	428	429	430	431	432	361	362	363	364	365	C173Y hnt C173Y	367	368	105	106	107	p.54-27<C-1 V54-27<C	109	110	111	112		
177	178	179	180	181	182	183	184	433	D139N hnt G136Y	435	436	437	438	C173Y hnt C173Y	440	369	370	371	372	373	374	375	376	113	114	115	116	S4delATC c.3-40-34delATC	118	119	120		
185	186	187	188	189	190	S177P hnt S177P	192	441	442	443	444	445	446	447	448	377	378	379	380	381	382	383	384	121	p.54-27<C-1 V54-27<C	123	124	125	126	127	128		
193	194	195	196	197	198	199	200	449	450	451	452	453	454	455	456	297	298	299	300	301	302	303	304	257	258	259	260	261	262	263	264		
201	C222R hnt C222R	203	204	D1187 hnt D1187	206	207	208	457	458	459	460	461	462	463	464	265	266	267	268	269	270	271	272	9	10	11	12	p.153-155delATC	14	15	16		
209	210	211	212	213	214	215	216	465	466	467	468	469	470	471	472	273	274	C222R hnt C222R	276	277	278	279	280	17	18	19	20	21	22	23	24		
S02delATC c.502delATC	218	219	D129V hnt G207G	221	222	223	224	473	S177P hnt S177P	475	476	D139N hnt G136Y	478	479	480	281	282	283	284	285	286	287	D139N hnt G136Y	25	26	27	28	29	C222R hnt C222R	31	32		
225	226	227	228	229	230	231	232	481	482	483	D139N hnt G136Y	485	486	487	488	289	290	291	292	C173Y hnt C173Y	294	295	296	33	34	35	36	37	38	p.54-27<C-1 V54-27<C	40		
233	234	235	236	237	238	239	240	489	490	491	492	493	494	495	496	297	298	299	300	301	C222R hnt C222R	303	304	41	42	43	44	C222R hnt C222R	46	47	48		
241	242	243	C160R hnt C160R	245	S4delATC c.3-40-34delATC	247	248	497	498	499	D139N hnt G136Y	501	502	503	504	305	306	307	308	309	310	311	312	49	50	51	52	53	54	55	56		
249	250	251	252	253	254	255	256	505	506	507	D139N hnt G136Y	509	510	511	C173Y hnt C173Y	313	314	315	316	317	318	319	320	57	58	59	60	C222R hnt C222R	62	63	64		

EK 6: Tekrarlanabilirlik çalışması sonuçları

Örnek	Havuz	MAF	Ortalama	SD	%SD
FH0403 D129V het	2D-08_X01	0,200	0,136	0,034	25,1
	2D-08_Y01	0,150			
	2D-08_X10	0,140			
	2D-08_Y10	0,150			
	2D-08_X19	0,120			
	2D-08_Y19	0,070			
	2D-08_X28	0,120			
	2D-08_Y28	0,140			
FH06019 C173Y het	2D-08_X37	0,110	0,109	0,011	9,7
	2D-08_Y37	0,100			
	2D-08_X46	0,110			
	2D-08_Y46	0,120			
	2D-08_X55	0,100			
	2D-08_Y55	0,120			
	2D-08_X64	0,120			
	2D-08_Y64	0,090			
FH0403 D129V het	2D-16_X01	0,070	0,071	0,016	22,7
	2D-16_Y01	0,100			
	2D-16_X03	0,080			
	2D-16_Y02	0,090			
	2D-16_X10	0,060			
	2D-16_Y11	0,060			
	2D-16_X12	0,060			
	2D-16_Y12	0,050			
FH06019 C173Y het	2D-16_X21	0,060	0,063	0,022	35,8
	2D-16_Y21	0,100			
	2D-16_X23	0,013			
	2D-16_Y22	0,070			
	2D-16_X30	0,060			
	2D-16_Y31	0,070			
	2D-16_X32	0,070			
	2D-16_Y32	0,060			
FH0403 D129V het	2D-24_X01	0,038	0,030	0,008	27,3
	2D-24_Y01	0,027			
	2D-24_X04	0,023			
	2D-24_Y02	0,037			
	2D-24_X10	0,018			
	2D-24_Y03	0,032			
	2D-24_X17	0,034			
	2D-24_Y17	0,016			
	2D-24_X19	0,035			

	2D-24_Y20	0,041			
FH06019 C173Y het	2D-24_X05	0,028	0,028	0,011	38,3
	2D-24_Y13	0,028			
	2D-24_X14	0,013			
	2D-24_Y14	0,028			
	2D-24_X16	0,049			
	2D-24_Y15	0,027			
	2D-24_X23	0,015			
	2D-24_Y24	0,038			



EK 7: FH hasta grubu için tespit edilen 8 adet negatif havuz, 72 adet pozitif havuz, bulunan mutasyonlar ve frekansları

Boş hücreler havuzların mutasyon tespit edilmeyen genlerine aittir. 6 geninin hiçbirisinde mutasyon tespit edilmeyen 8 adet negatif havuz yeşil ile işaretlenmiştir. Hiçbir havuzun APOA2 geninde mutasyon tespit edilmemiştir.

Havuz	LIPA	LDLRAP1	APOA2	ITIH4	GHR	PPP1R17
LX1-1		Q136X; 0,52				
LX1-2	I252L; 0,17	Q136X; 0,66		R56X; 0,09	R229H; 0,12	
LX1-3						
LX1-4	Q298Q; 0,09	I74M; 0,12 IVS3+1G>T; 0,15 R238W; 0,10		E392E; 0,06	L7L; 0,07	
LX1-5						
LX1-6		R238W; 0,10			A613A; 0,08 L622L; 0,07	
LX1-7				IVS12+5G>A; 0,05		
LX1-8		Q136X; 0,53 IVS4-1G>A; 0,16				
LY1-1				E392E; 0,04	L7L; 0,04	
LY1-2		Q136X; 0,17 IVS4-1G>A; 0,25 R238W; 0,09		IVS12+5G>A; 0,05	A613A; 0,11 L622L; 0,11	
LY1-3		Q136X; 0,27 IVS3+1G>T; 0,15				
LY1-4	I252L; 0,11 Q298Q; 0,06	Q136X; 0,38 I74M; 0,09 R238W; 0,13				
LY1-5		Q136X; 0,35				
LY1-6	I252L; 0,11	Q136X; 0,37		R56X; 0,09		
LY1-7		Q136X; 0,31				
LY1-8		Q136X; 0,17			R229H; 0,12	
LX2-1						
LX2-2				E332K; 0,08	M530T; 0,12	
LX2-3						V143M; 0,13
LX2-4		R238W; 0,07				
LX2-5		Q136X; 0,39				
LX2-6		Q136X; 0,20 S202H 0,09 E217X; 0,11				
LX2-7		I74M; 0,13 R238W; 0,15 V290V; 0,08				

LX2-8		Q136X; 0,29				
LY2-1		Q136X; 0,38				
LY2-2					M530T; 0,09	
LY2-3						
LY2-4		S202H 0,12				
LY2-5		I74M; 0,08 R238W; 0,07		E332K; 0,09		
LY2-6						V143M; 0,07
LY2-7		Q136X; 0,23 R238W; 0,11 V290V; 0,06				
LY2-8		Q136X; 0,33 E217X; 0,15				
LX3-1						
LX3-2				D87D; 0,06	R229H; 0,08	
LX3-3		R238W; 0,07				
LX3-4	I252L; 0,09					
LX3-5				G249R; 0,07 N333N; 0,06		
LX3-6	I252L; 0,12	S202H; 0,06 R238W; 0,10				
LX3-7		S202H; 0,08		V327L; 0,08		
LX3-8		R238W; 0,06				
LY3-1				V327L; 0,07		
LY3-2				G249R; 0,08	R229H; 0,09	
LY3-3		S202H; 0,08				
LY3-4						
LY3-5		R238W; 0,16		D87D; 0,06		
LY3-6	I252L; 0,20					
LY3-7		S202H; 0,06		N333N; 0,04		
LY3-8		R238W; 0,07				
LX4-1						
LX4-2		Q136X; 0,17		T832M; 0,09		
LX4-3					D564E; 0,13	
LX4-4		R238W; 0,16			R179C; 0,06 V499M; 0,06	
LX4-5				E332K; 0,11	R179C; 0,11	
LX4-6		R238W; 0,13 c.246delG; 0,17				
LX4-7		R238W; 0,06 Q86X; 0,08		D892D; 0,08	V499M; 0,07	
LX4-8	I252L; 0,09				R179C; 0,08	
LY4-1		Q86X; 0,11		D892D; 0,10		
LY4-2		R238W; 0,06		T832M; 0,08	R179C; 0,05	
LY4-3					R179C; 0,08	

LY4-4	I252L; 0,11	Q136X; 0,16				
LY4-5		R238W; 0,16			V499M; 0,04	
LY4-6				E332K; 0,07	R179C; 0,07	
LY4-7		R238W; 0,03			D564E; 0,04	
LY4-8		c.246delG; 0,20			V499M; 0,10	
LX5-1				L150L; 0,07 G249R; 0,03		
LX5-2				E332K; 0,11 M358I; 0,06		
LX5-3		R238W; 0,09		G418S; 0,05		
LX5-4				T896T; 0,10		
LX5-5				R689C; 0,06 W850C; 0,06		
LX5-6		R238W; 0,05				
LX5-7		R238W; 0,06		T586T; 0,13 P919P; 0,06		
LX5-8		R238W; 0,09 E217X; 0,08		P919P; 0,06	E62K; 0,06	
LY5-1				T896T; 0,09		
LY5-2				M358I; 0,07 T586T; 0,10		
LY5-3		R238W; 0,08		P919P; 0,04		
LY5-4		R238W; 0,06		G418S; 0,05	E62K; 0,08	
LY5-5						
LY5-6				E332K; 0,10 P919P; 0,06		
LY5-7		R238W; 0,06		G249R; 0,03 R689C; 0,05 W850C; 0,07		
LY5-8		R238W; 0,15		L150L; 0,08		

EK 8: Tespit edilen mutasyon çeşitleri

Gen	Mutasyon	cDNA pozisyonu	hg19 pozisyonu	dbSNP kodu	MAF	ClinVar*
LIPA	I252L	c.754A>T	chr10:90,983,509	rs147493628	<0,01	İY
LIPA	Q298Q	c.894G>A	chr10:90,982,268	rs116928232	<0,01	P
LDLRAP1	I74M	c.222C>G	chr1:25,880,546	rs199630684	<0,01	-
LDLRAP1	c.246delG	c.246delG	chr1:25,881,365	-	-	-
LDLRAP1	Q86X	c.256C>T	chr1:25,881,375	rs997683431	-	-
LDLRAP1	IVS3+1G>T	c.344+1G>T	chr1:25,881,464	-	-	-
LDLRAP1	Q136X	c.406C>T	chr1:25,883,705	rs121908325	<0,01	P
LDLRAP1	IVS4-1G>A	c.460-1G>A	chr1:25,889,134	rs2124688583	-	P
LDLRAP1	S202H	c.604_605delTCinsCA	chr1:25,889,632	rs386629678	-	İY
LDLRAP1	E217X	c.649G>T	chr1:25,890,184	rs2044402416	-	P
LDLRAP1	R238W	c.712C>T	chr1:25,890,247	rs41291058	0,02	İ
LDLRAP1	V290V	c.870C>T	chr1:25,893,426	rs200433841	<0,01	İY
ITIH4	R56X	c.166C>T	chr3:52,863,220	rs761970932	<0,01	-
ITIH4	D87D	c.261T>C	chr3:52,861,204	rs751642445	<0,01	-
ITIH4	L150L	c.450C>T	chr3:52,860,876	rs143427404	<0,01	İ
ITIH4	G249R	c.745G>C	chr3:52,859,916	rs139719930	<0,01	-
ITIH4	V327L	c.979G>T	chr3:52,858,479	rs1343545265	<0,01	-
ITIH4	E332K	c.994G>A	chr3:52,858,464	rs772851128	<0,01	-
ITIH4	N333N	c.999C>T	chr3:52,858,459	rs142248896	<0,01	-
ITIH4	M358I	c.1074G>A	chr3:52,858,303	rs749483845	<0,01	-
ITIH4	E392E	c.1176G>A	chr3:52,858,016	rs374402919	<0,01	-
ITIH4	G418S	c.1252G>A	chr3:52,857,940	rs141154056	<0,01	-
ITIH4	IVS12+5G>A	c.1679+5G>A	chr3:52,855,002	-	-	-
ITIH4	R689C	c.2065C>T	chr3:52,853,421	rs546117123	<0,01	-
ITIH4	T586T	c.1758G>A	chr3:52,854,723	rs768999504	<0,01	-
ITIH4	T832M	c.2495C>T	chr3:52,848,356	rs150977188	<0,01	-
ITIH4	W850C	c.2550G>C	chr3:52,848,301	rs148943741	<0,01	-
ITIH4	D892D	c.2676C>T	chr3:52,848,038	rs750458578	<0,01	-
ITIH4	T896T	c.2688G>A	chr3:52,848,026	rs532679730	<0,01	-
ITIH4	P919P	c.2757G>A	chr3:52,847,473	rs78780447	0,01	-
GHR	L7L	c.21G>A	chr5:42,565,997	rs756572945	<0,01	-
GHR	E62K	c.184G>A	chr5:42,689,039	rs121909361	<0,01	P
GHR	R179C	c.535C>T	chr5:42,700,021	rs121909362	<0,01	B
GHR	R229H	c.686G>A	chr5:42,711,376	rs6177	<0,01	B
GHR	V499M	c.1495G>A	chr5:42,719,104	rs373984725	<0,01	B
GHR	M530T	c.1589T>C	chr5:42,719,198	rs148677728	<0,01	B
GHR	D564E	c.1692C>A	chr5:42,719,301	rs1176485395	<0,01	-
GHR	A613A	c.1839C>A	chr5:42,719,448	rs1220068702	<0,01	-
GHR	L622L	c.1866C>T	chr5:42,719,475	rs772856718	<0,01	İY
PPP1R17	V143M	c.427G>A	chr7:31,746,856	rs367827815	<0,01	-

* İ: İyimsir, İY: İyimsire yakın, P: Patojenik, B: Belirsiz öneme sahip

EK 9: Pozitif bulunan örnekler ve genotipleri

Bir mutasyon taşıyan örnekler (64 adet):

Örnek kodu	Gen	Mutasyon	Zigozite
L005	LDLRAP1	Q136X	homozigot
L006	LDLRAP1	Q136X	homozigot
L008	LDLRAP1	Q136X	homozigot
L010	LDLRAP1	Q136X	homozigot
L011	LDLRAP1	Q136X	homozigot
L013	LDLRAP1	Q136X	homozigot
L015	LDLRAP1	Q136X	homozigot
L016	GHR	R229H	heterozigot
L027	LDLRAP1	IVS3+1G>T	homozigot
L050	ITIH4	IVS12+5G>A	heterozigot
L058	LDLRAP1	IVS4-1G>A	homozigot
L059	LDLRAP1	Q136X	homozigot
L060	LDLRAP1	Q136X	homozigot
L061	LDLRAP1	Q136X	homozigot
L063	LDLRAP1	Q136X	homozigot
L064	LDLRAP1	Q136X	homozigot
L074	GHR	M530T	heterozigot
L077	ITIH4	E332K	heterozigot
L086	PPP1R17	V143M	heterozigot
L095	LDLRAP1	R238W	heterozigot
L097	LDLRAP1	Q136X	homozigot
L104	LDLRAP1	Q136X	homozigot
L105	LDLRAP1	Q136X	homozigot
L108	LDLRAP1	S202H	heterozigot
L112	LDLRAP1	E217X	heterozigot
L119	LDLRAP1	V290V	heterozigot
L127	LDLRAP1	Q136X	homozigot
L138	GHR	R229H	heterozigot
L141	ITIH4	D87D	heterozigot
L149	LDLRAP1	R238W	heterozigot
L158	LIPA	I252L	heterozigot
L162	ITIH4	G249R	heterozigot
L167	ITIH4	N333N	heterozigot
L174	LIPA	I252L	heterozigot
L175	LDLRAP1	S202H	heterozigot
L176	LDLRAP1	R238W	heterozigot
L177	ITIH4	V327L	heterozigot
L179	LDLRAP1	S202H	heterozigot
L189	LDLRAP1	R238W	heterozigot

L202	ITIH4	T832M	heterozigot
L204	LDLRAP1	Q136X	homozigot
L215	GHR	D564E	heterozigot
L218	GHR	R179C	heterozigot
L223	LDLRAP1	R238W	heterozigot
L237	LDLRAP1	R238W	heterozigot
L240	LDLRAP1	c.246delG	homozigot
L242	LDLRAP1	R238W	heterozigot
L248	GHR	V499M	heterozigot
L251	GHR	R179C	heterozigot
L252	LIPA	I252L	heterozigot
L263	ITIH4	G249R	heterozigot
L264	ITIH4	L150L	heterozigot
L266	ITIH4	M358I	heterozigot
L270	ITIH4	E332K	heterozigot
L275	LDLRAP1	R238W	heterozigot
L276	ITIH4	G418S	heterozigot
L281	ITIH4	T896T	heterozigot
L303	LDLRAP1	R238W	heterozigot
L306	ITIH4	T586T	heterozigot
L307	ITIH4	P919P	heterozigot
L312	LDLRAP1	R238W	heterozigot
L318	ITIH4	P919P	heterozigot
L319	LDLRAP1	E217X	heterozigot
L320	LDLRAP1	R238W	heterozigot

İki mutasyon taşıyan örnekler (8 adet):

Örnek kodu	Gen	Mutasyon	Zigozite
L012	LIPA	I252L	heterozigot
	LDLRAP1	Q136X	homozigot
L025	ITIH4	E392E	heterozigot
	GHR	L7L	heterozigot
L117	LDLRAP1	I74M	homozigot
	LDLRAP1	R238W	homozigot
L221	LDLRAP1	R238W	heterozigot
	GHR	V499M	heterozigot
L230	ITIH4	E332K	heterozigot
	GHR	R179C	heterozigot
L241	LDLRAP1	Q86X	heterozigot
	ITIH4	D892D	heterozigot
L295	ITIH4	R689C	heterozigot
	ITIH4	W850C	heterozigot
L316	LDLRAP1	R238W	heterozigot
	GHR	E62K	heterozigot

Üç mutasyon taşıyan örnekler (3 adet):

Örnek kodu	Gen	Mutasyon	Zigozite
L014	LIPA	I252L	heterozigot
	LDLRAP1	Q136X	homozigot
	ITIH4	R56X	heterozigot
L028	LIPA	Q298Q	heterozigot
	LDLRAP1	I74M	homozigot
	LDLRAP1	R238W	homozigot
L042	LDLRAP1	R238W	heterozigot
	GHR	A613A	heterozigot
	GHR	L622L	heterozigot

ÖZGEÇMİŞ

Adı Soyadı: Haldun Doğan

Doğum Yeri:

Doğum Tarihi:

Medeni Hali:

Yabancı Dili: İngilizce

Eğitim Durumu

Lise: Ankara Fen Lisesi (1994-1996)

Bursa Nilüfer Lisesi (1996-1997)

Lisans: Ortadoğu Teknik Üniversitesi, Fen-Edebiyat Fakültesi

Moleküler Biyoloji ve Genetik Bölümü (1997-2002)

Yüksek Lisans: Ankara Üniversitesi, Biyoteknoloji Enstitüsü (2005-2009)

Önlisans: Anadolu Üniversitesi, Açıköğretim Fakültesi

Sağlık Kurumları İşletmeciliği (2016-2018)

Yayınlar ve Bildiriler

Murat Buyuktepe, Ibrahim Kaplan, Eyup Bayatli, Haldun Dogan, Hasan Caglar Ugur. Significance of O6-methyl guanine methyltransferase promoter methylation in high grade glioma patients: optimal cutoff point, CpG locus, and genetic assay. J Neurooncol. 2023 Aug;164(1):171-177. DOI: 10.1007/s11060-023-04397-1.

Soydan E., Sahin U., Ceylaner S., gökmen a., Dogan H., Kurdal Okcu M., ilhan o., Arslan Ö., Ozcan M. Allojeneik Hematopoetik Kök Hücre Nakli Sonrası Aseptik Nekroz Gelişen Olgularda Vitamin D Reseptör Gen Polimorfizmi. Türkiye'de Lösemi Lenfoma Miyelom Araştırmaları. Cilt numarası: 7 Sayı: 2. ss.51-56. DOI: 10.5578/llm.20229802. 2023.

Rabia Miray Kisl  Ekinci, Sibel Balci, Haldun Dogan, Serdar Ceylaner, Celal Varan, Sevcan Erdem, Fatma Coban, Atil Bisgin. Camptodactyly-Arthropathy-Coxa Vara-Pericarditis Syndrome Resembling Juvenile Idiopathic Arthritis: A Single-Center Experience from Southern Turkey. Mol Syndromol. 2021 Apr;12(2):112-117. doi: 10.1159/000513111.

       S, Acipayam C, Yenmi  Inan MN, Do an H. Iron Refractory Iron Deficiency Anemia Due to 374 Base Pairs Deletion in the TMPRSS6 Gene Journal of Pediatric Hematology/Oncology. 41(5):e333-e335, JUL, 2019.

Kenan  stanbulu, Nilg n K ksal, Merih  etinkaya, Hilal  zkan, Tahsin Yakut, Mutlu Karkucak, Haldun Do an. The potential utility of real-time PCR of the 16S-rRNA gene in the diagnosis of neonatal sepsis. The Turkish Journal of Pediatrics 2019; 61: 493-499 DOI:10.24953/turkjped.2019.04.004.

Ayşe Tuba Altug, Asli Senol, Ozlem Nasibe Ozkepir, Haldun Dogan, Serdar Ceylaner, Erhan Ozdiler. Evaluation of MSX1 gene as the common candidate gene of nonsyndromic congenital hypodontia and cleft lip and palate. Journal of Cleft Lip Palate and Craniofacial Anomalies, DOI:10.4103/jclpca.jclpca_87_17, 2018.

Dilek Dilli, Nazan Neslihan Do an, Mehmet  ah  pek, Yunus  avuş, Serdar Ceylaner, Haldun Do an, Arzu Dursun, Tuncay K          , Ay eg l Zenciro lu. MaFOS-GDM trial: Maternal fish oil supplementation in women with gestational diabetes and cord blood DNA methylation at insulin like growth factor-1 (IGF-1) gene. Clin Nutr ESPEN. 2018 Feb;23:73-78. doi: 10.1016/j.clnesp.2017.12.006. Epub 2017 Dec 30.

Utku Tanto lu, Seher Y ksel, Cihangir Akyol, Haldun Do an, N khet Kutlay,       Kuzu, Hilal  zda , Mehmet Ayhan Kuzu. Use of immunohistochemical versus microsatellite analyses as markers for colorectal cancer. Turkish Journal of Biochemistry, Volume 43, Issue 2, Pages 134-141, ISSN (Online) 1303-829X. 2017.

Mehmet Boyraz, Ediz Ye      , Fatih Ezg , Aysun Bideci, Haldun Do an, Korkut Ulucan, Peyami Cinaz. Effect of Cytokine Signaling 3 Gene Polymorphisms in Childhood Obesity. J Clin Res Pediatr Endocrinol 2016;8(4):452-460, 2016.

Ayşe Tuba Altuğ, Aslı Şenol, Özlem Özkepir, Haldun Doğan, Serdar Ceylaner, Gülay Ceylaner, F. Erhan Özdiler. Sendromik Olmayan Konjenital Maksiller Lateral Kesici Diş Eksikliği ve Dudak Damak Yarıklarının Ortak Genetik Alt Yapısı. Türkiye Klinikleri J Orthod-Special Topics. 2016;2(3):15-22.

Sebahat Yılmaz Ağıladioğlu, Zehra Aycan, Semra Çetinkaya, Veysel Nijat Baş, Aşan Önder, Havva Nur Peltek Kendirci, Haldun Doğan and Serdar Ceylaner. Maturity onset diabetes of youth (MODY) in Turkish children: sequence analysis of 11 causative genes by next generation sequencing. J Pediatr Endocrinol Metab. 2016 Apr;29(4):487-96. DOI: 10.1515/jpem-2015-0039.

Özgür Aldemir, Kadri Karaer, Cengiz Çevik, Haldun Doğan, Cumali Gökçe. A Novel Mutation of SCL26A4 Gene in Turkish Family with Pendred Syndrome. Mustafa Kemal Üniversitesi Tıp Dergisi. 2015;C:6 S:22.

Nejat Akar, Serdar Ceylaner, Haldun Doğan. A Novel Beta Globin Gene Polymorphism in Turkish Population: IVS-II 706 G>A. The Journal of Pediatric Research 2015;2(2):112-3.

Ezgu F, Hasanoglu A, Okur I, Biberoglu G, Tumer L, Eminoglu T, Dogan H. Rapid screening of 10 common mutations in Turkish Gaucher patients using electronic DNA microarray. Blood Cells Mol Dis. 2008 Mar-Apr;40(2):246-7. Epub 2007 Nov 26.

Posterler

Özgür Aldemir, Haldun Doğan, Kadri Karaer, Cengiz Çevik. Yeni Nesil Dizi Analizi ile İşitme Kayıplı Hastalarda Sık Görülen Sağırılık Genlerine Ait Mutasyonların Taranması. 11. Ulusal Tıbbi Genetik Kongresi. P-361. 2014.

Özlem Nesibe Özkepir, Ayşe Tuba Altuğ, Haldun Doğan, Serdar Ceylaner, Gülay Ceylaner. Sendromik Olmayan Hipodonti Olgularında PAX9 Genindeki Mutasyon ve Polimorfizmlerin Değerlendirilmesi. 14. Uluslararası Türk Ortodonti Derneği Sempozyumu. P-97. 2015.

İrem Kaplan, Deniz Kör, Berna Şeker Yılmaz, Derya Bulut, Aysun Karabay Bayazıt, Haldun Doğan, Ali Anarat, Halise Neslihan Önenli Mungan. Böbrek Transplantasyonu Yapılan Sistinozisli İki Olgu Sunumu. 5. Uluslararası Katılımlı Lizozomal Hastalıklar Kongresi. P42. 2016.

Davetli Konuşmacı

Tıbbi Genetik Derneği, Preimplantasyon Genetik Tanı Online Toplantı. PGT Uygulamalarında SNP Belirteçleri. 50 dk. 24.04.2021.

Ankara Üniversitesi Biyomedikal Mühendisliği Bölümü (EMB-TECH), Online Kurs. PCR Teknik Eğitimi. 80 dk. 14.11.2020.

Illumina, Kanser Genetiğinde Yeni Nesil Uygulamalar Sempozyumu. Türkiye’den Son Kullanıcı Deneyim Paylaşımı: Nextera XT ve Trusight One Workflow. 30 dk. 19 Eylül 2018.

Biyoinformatik Forumu, PHI Tech Bioinformatics, TÜBİTAK – Marmara Teknokent ve Gebze Teknik Üniversitesi (GTÜ). Panel: Kanser Çalışmalarında DNA Dizileme. 35 dk. 26 Mayıs 2017.

Tıbbi Genetik Derneği, Online Genetik Okulu. Yeni Nesil Dizi Analizinde IGV Programının Kullanımı. 60 dk. 28.03.2017.

2. Illumina Yeni Nesil Dizileme Çalıştayı. End-user talks: Rare diseases. 20 dk. 03-05 Mayıs 2016.

Tıbbi Genetik Derneği, Online Kurs. Sanger DNA Dizi Analizi, Sorunlar Çözüm Önerileri. 80 dk. 01.10.2015.

3rd Course in Next Generation Sequencing, European School of Genetic Medicine, İstanbul Üniversitesi, Tıbbi Genetik Derneği. Nadir Hastalıklarda Yeni Nesil Dizi Analizi Uygulamalarına Klinik Yaklaşım: Avantaj ve Sorunlar. 40 dk. 07-10 Mayıs 2014.

1. Uluslararası Gen-Arel Moleküler Biyoloji ve Genetik Öğrenci Kongresi, Arel Üniversitesi. Genetics Centers: Where Routine Laboratory Meets Scientific Study. 30 dk. 10-12 Mayıs 2013.

İş Tecrübesi

Kurumu: Montwell Ltd. Şti.

Görevi: Saha Uygulama Uzmanı

Yılları: 2005-2007

Kurumu: Genlab Ltd. Şti.

Görevi: Saha Uygulama Koordinatörü

Yılları: 2007-2009

Kurumu: İntergen Genetik Hastalıklar Tanı ve Uygulama Merkezi AŞ.

Görevi: Moleküler Laboratuvar Direktörü

Yılları: 2009-



TEZDEN ÇIKAN YAYINLAR



















