



T.C.  
EGE ÜNİVERSİTESİ  
Sağlık Bilimleri Enstitüsü



**NANOPORE SEKANSLAMA TEKNOLOJİSİ  
KULLANILARAK, KİŞİ TANIMLAMA KİTİ  
GELİŞTİRİLMESİ**

**Doktora Tezi**

Bilçay AKGÜN

Çocuk Sağlığı ve Hastalıkları Anabilim Dalı

İzmir  
2019

T.C.  
EGE ÜNİVERSİTESİ  
Sağlık Bilimleri Enstitüsü

**NANOPORE SEKANS LAMA TEKNOLOJİSİ  
KULLANILARAK, KİŞİ TANIMLAMA KİTİ  
GELİŞTİRİLMESİ**

**Doktora Tezi**

Bilçay AKGÜN

Danışman  
Prof. Dr. Ferda ÖZKINAY

Çocuk Sağlığı ve Hastalıkları Anabilim Dalı  
Genetik Doktora Programı

İzmir  
2019

### Tez Değerlendirme Kurulu Üyeleri

(Adı Soyadı)

(İmza)

Başkan : Prof.Dr. Ferda Özkınay

(Danışman)

Üye : Prof.Dr. Özgür Çoğulu

Üye : Prof.Dr. Hülseyin Onay

Üye : Doç.Dr. Tufan Çankaya

Üye : Doç.Dr. Aslı Subaşıoğlu



The image shows four handwritten signatures in blue ink, each placed on a horizontal dotted line. The signatures are: 1. Prof. Dr. Ferda Özkınay (top), 2. Prof. Dr. Özgür Çoğulu, 3. Prof. Dr. Hülseyin Onay, and 4. Doç. Dr. Tufan Çankaya. The signature of Doç. Dr. Aslı Subaşıoğlu is not clearly visible.

Doktora Tezinin kabul edildiği tarih: 26.08.2019

## Önsöz

Tıp fakültesi ikinci sınıfta karar verdiğim genetik bilimcisi olma hayalimi, akademisyenlik seviyesine taşıma yolunda en büyük adımı bugün itibariyle atmış bulunmaktayım. Dört sene önce biyoinformatik alanına hakim bir genetik hekimi olmak hedefiyle başladığım doktora serüvenimin sonuna gelip baktığımda, bu hedefimi büyük ölçüde gerçekleştirdiğimi görmenin derin bir gurur ve motivasyonunu yaşamaktayım. Bu süreçte en çok ilgimi çeken konu, şüphesiz ki tezimin de başrol oyuncusu olan nanopore sekanslama biyoinformatiği oldu. Prof. Dr. Hüseyin Onay hocam sayesinde, bu teknolojiyi çok erken bir dönemde tanıma ve uygulama fırsatı buldum. Bu teknolojinin ilk cihazı olan Minion neredeyse piyasaya çıkar çıkmaz öğrenmeye başladığım için, gerek yazılımsal gerekse donanımsal olarak yeniliklere hakim bir şekilde kendimi sürekli güncelleyerek ilerleyebildim. Tez konumun nanopore teknolojisiyle ilgili olması için danışman hocam Prof. Dr. Ferda Özkınay'a danıştığımda ise son derece olumlu karşıladı ve bu teknolojiyi ve biyoinformatiğini detaylıca öğrenmem için her zaman yardımcı oldu. Tez sürecim, hocalarımla ve benim de teknolojiye hakimiyetimiz sayesinde sorunsuzca ilerledi. Bu tez çalışması sonunda ortaya çıkan prototipin daha büyük ölçekli çalışmalara öncülük edeceğine inanıyorum. Bu süreçte, güncel dizi analizi teknolojileri uygulaması ve biyoinformatiği konularına hakim olup ilerlemem için desteklerini ve yardımlarını esirgemeyip her zaman yanımda olan hocalarıma, çalışma arkadaşlarıma ve 2211/C Yurt İçi Öncelikli Alanlar Doktora Burs Programı'yla burs desteği sağlayan TÜBİTAK'a sonsuz teşekkürlerimi sunuyorum.

İzmir, 07.08.2019

Bilçay AKGÜN

## Özet

### **Nanopore sekanslama teknolojisi kullanılarak, kişi tanımlama kiti geliştirilmesi**

**Amaç:** İnsan Deoksiribo Nükleik Asidi (DNA) bilinen en stabil biyolojik materyallerden bir tanesi olup, kişiye özel bir çok tanımlama bölgesi içermektedir. Son yıllarda, DNA üzerindeki belirli Tek Nükleotid Polimorfizm'lerinin (Single Nucleotide Polymorphism-SNP) genotiplendirilmesi, özellikle kimliklendirme çalışmalarında giderek yaygınlaşan ve yüksek ayırım gücü sunan bir yöntemdir. Genotiplendirme işlemi, yeni nesil dizi analizi teknolojisindeki hızlı ilerlemeler sayesinde giderek pratikleşmiş olsa da bu cihazların birçoğunun uzun sürede sonuç vermesi ve büyük veri analizleri gerektirmesi, özellikle hızlı kimliklendirme gereken durumlarda büyük bir dezavantaj oluşturmaktadır. Bununla birlikte, 4. nesil olarak geçen Oxford Nanopore teknolojisini kullanan bir cihazın (Minion) uzun okumalar yapması, hızlı analiz ve küçük boyutları sayesinde sahada kullanıma imkan sağlaması, bu dezavantajları ortadan kaldırmıştır. Bu projede; Oxford Nanopore sekanslama platformuna uyum sağlayacak, teknolojik ve pratik SNP temelli bir kişi tanımlama prototipi geliştirilmesi amaçlanmıştır.

**Gereç ve yöntem:** Ticari DNA kullanılarak, literatürde kişi tanımlama için sıklıkla kullanılan 52 SNP bölgesi multipleks Polimeraz Zincir Reaksiyonu (Polymerase Chain Reaction-PCR) metodu ile amplifiye edilmiştir. Bu ampliconlar Oxford Nanopore Minion cihazı ile sekanslanmıştır. Sekans verilerine göre özgün bir biyoinformatik iş akışı tasarlanmıştır.

**Bulgular:** Üç ticari DNA örneği üzerinden yapılan çalışmalarda Nanopore verilerinin 52 SNP bölgesinde de tam olarak doğru sonuç verdiği gözlenmiştir. Toplam süre 1 saatin altında tamamlanmıştır.

**Sonuç:** Bu tez sonucunda 52 adet SNP bölgesi Oxford Nanopore Minion cihazıyla başarılı bir şekilde sekanslanmıştır. Aynı zamanda SNP temelli kişi tanımlama için dünya üzerinde erişilmiş en hızlı süre elde edilmiştir.

**Anahtar Kelimeler;** SNP; kişi tanımlama; nanopore

## Abstract

### **Development of human identification kit using nanopore sequencing technology**

**Objective:** Human DNA is one of the most stable biological materials known and contains a number of individual identification sites. In recent years, genotyping of certain SNP on DNA has become increasingly common, especially in identification studies, and offers high discrimination power. Although the genotyping process has become increasingly practical due to the rapid advances in next generation sequencing technology, many of these devices have long-lasting results and require large data analyzes, especially in cases where rapid identification is required. However, the Minion device, which is defined as the 4th generation sequence device, has a small size and enables rapid analysis, thus eliminating these disadvantages. In this thesis; it is aimed to develop a technological and practical SNP based human identification prototype that will adapt to the Oxford Nanopore sequencing platform.

**Materials and methods:** Using commercial DNA, 52 SNP regions frequently used for identification in the literature were amplified by multiplex PCR method. These amplicons were sequenced with the Oxford Nanopore Minion device. A unique bioinformatics workflow is designed based on sequence data.

**Results:** Nanopore data were found to be completely accurate in 52 SNP regions in three commercial DNA samples. Total time was completed under 1 hour.

**Conclusion:** As a result of this thesis, 52 SNP sites were sequenced successfully with Oxford Nanopore Minion device. At the same time, the fastest time has been achieved in the world for SNP based human identification.

**Keywords;** SNP; human identification; nanopore

## İçindekiler

Önsöz.....	II
Özet .....	III
Abstrast.....	IV
İçindekiler .....	V
Tablolar Dizini .....	VII
Şekiller Dizini .....	VIII
Grafikler Dizini.....	IX
Kısaltma Listesi.....	X
1. Giriş.....	1
1.1. Araştırmanın Problemi .....	1
1.2. Araştırmanın Sorusu .....	2
1.3. Araştırmanın Hipotezleri.....	3
1.4. Araştırmanın Varsayımları .....	3
1.5. Araştırmanın Sınırlılıkları.....	3
1.6. Araştırmanın Amacı.....	3
2. Genel Bilgiler.....	5
2.1. Polimorfik alanlar .....	5
2.1.1. VNTR.....	5
2.1.2. STR.....	5
2.1.3. SNP.....	6
2.2. Sekans teknolojileri .....	8
2.2.1. Yeni nesil dizi analizi:.....	8
2.2.2. Oxford Nanopore Teknolojisi:.....	9
2.3. Biyoinformatik .....	11
2.3.1. Basecalling .....	11
2.3.2. Haritalama ve hizalama .....	11
2.3.3. Varyant çağırma .....	11
3. Gereç ve Yöntem.....	12
3.1. Araştırmanın tipi .....	12
3.2. Araştırmanın yeri ve zamanı .....	12
3.3. Araştırmanın örnekleme.....	12
3.4. Moleküler çalışmalar .....	12

3.4.1. SNP bölgelerinin seçimi.....	12
3.4.2. Primer Dizaynı .....	12
3.4.3. PCR çalışmaları .....	14
3.4.3.1. Standart PCR çalışmaları .....	14
3.4.3.2. Süper hızlı PCR çalışmaları.....	16
3.4.4. Sekanslama çalışmaları .....	16
3.4.4.1. Yeni nesil dizi analizi .....	16
3.4.4.2. Nanopore ile sekanslama.....	19
3.4.5. Biyoinformatik .....	22
3.4.5.1. Yeni nesil dizi analizi informatiği.....	22
3.4.5.2. Nanopore sekanslama informatiği .....	22
4. Bulgular .....	27
4.1. Yeni nesil dizi analizi bulguları .....	27
4.2. Nanopore sekanslama bulguları.....	28
4.3. Biyoinformatik iş akışı bulguları.....	34
4.3.1. Basecalling bulguları .....	34
4.3.2. Haritalama ve hizalama bulguları .....	35
4.3.3. Varyant çağırma bulguları .....	35
5. Tartışma.....	40
6. Sonuç ve Öneriler .....	45
7. Kaynaklar .....	46
Ekler .....	54
Teşekkür .....	55
Özgeçmiş.....	57

## Tablolar Dizini

<b>Tablo 1.</b> SNPforID konsorsiyumu tarafından belirlenen 52 SNP bölgesi ve allel bilgileri.....	7
<b>Tablo 2.</b> Çalışmamızda kullanılan primer dizileri.....	13
<b>Tablo 3.</b> M13 primer dizileri.....	14
<b>Tablo 4.</b> İlk PCR işlemi kondisyon ve koşulları .....	15
<b>Tablo 5.</b> Multipleks PCR işlemi için kullanılan kondisyon ve koşullar .....	16
<b>Tablo 6.</b> Yeni nesil dizi analizi çalışması için kütüphane hazırlığında PCR amplifikasyonu basamağı inkübasyon koşulları .....	18
<b>Tablo 7.</b> Üç farklı ticari DNA'nın 52 SNP için yeni nesil dizi analizi genotiplendirmesi .....	27
<b>Tablo 8.</b> DNA-1'e ait nanopore sekanslama verilerinin metrik değerleri.....	29
<b>Tablo 9.</b> DNA-1'e ait nanopore sekanslama 5. dakika verilerinin genotiplendirme okuma derinliği bulguları .....	29
<b>Tablo 10.</b> DNA-1'e ait nanopore sekanslama 10. dakika verilerinin genotiplendirme okuma derinliği bulguları .....	30
<b>Tablo 11.</b> DNA-1'e ait nanopore sekanslama 15. dakika verilerinin genotiplendirme okuma derinliği bulguları .....	31
<b>Tablo 12.</b> DNA-2'ye ait nanopore sekanslama 5. dakika verilerinin genotiplendirme okuma derinliği bulguları .....	32
<b>Tablo 13.</b> DNA-3'e ait nanopore sekanslama 5. dakika verilerinin genotiplendirme okuma derinliği bulguları .....	33
<b>Tablo 14.</b> Üç farklı DNA'ya ait nanopore sekanslama 5., 10. ve 15. Dakika verilerindeki en düşük ve ortalama okuma derinliği sayıları.....	34
<b>Tablo 15.</b> DNA-1'e ait nanopore sekaslama 5., 10. ve 15. Dakika verilerinin Guppy programıyla basecalling işlemi süreleri .....	35
<b>Tablo 16.</b> DNA-1'e ait nanopore sekanslama verilerinin bcfTools programının farklı modifikasyonlarıyla varyant çağırma süreleri .....	36
<b>Tablo 17.</b> Yazılan scriptle, tüm nanopore sekans verilerinin tekrar değerlendirilmesi sonucu elde edilen süreler.....	39

## Şekiller Dizini

Şekil 1. Oxford Nanopore FlowCell'ine örnek ve buffer yükleme aşamaları.....	21
Şekil 2. İnsan mitokondriyal kromozomuna ait referans dizi örneği .....	23
Şekil 3. Oluşturduğumuz fasta formatındaki referansa ait SNP ve dizi örnekleri (Kırmızı ok – genotiplendirmek istediğimiz SNP'lere ait baz) .....	23
Şekil 4. HDFView programıyla açılmış fast5 dosyası görüntüsü .....	24
Şekil 5. rs10495407 SNP'inin IGV programı görüntüsü .....	25
Şekil 6. Philisa cihazına ait program ile oluşturulan PCR süresi ve koşullarının görüntüsü .....	28
Şekil 7. Varyant çağırma programları için oluşturulan manifest dosyası görüntüsü.	36
Şekil 8. Tüm programların tek bir komutla çalışmasını sağlayacak özel olarak yazılmış bash script'in ilk bölümü .....	37
Şekil 9. Tüm programların tek bir komutla çalışmasını sağlayacak özel olarak yazılmış bash script'in ikinci bölümü .....	38

## Grafikler Dizini

Şekil tablosu ögesi bulunamadı.



## Kısaltma Listesi

DNA	: Deoksiribo nükleik asit
SNP	: Tek nükleotid polimorfizmi
PCR	: Polimeraz zincir reaksiyonu
RFLP	: Restriksiyon fragman uzunluk polimorfizmi
VNTR	: Değişken sayılı ardışık tekrarlar
STR	: Kısa ardışık tekrarlar
RNA	: Ribo nükleik asit
FBI	: Federal soruşturma bürosu
CODIS	: Combined DNA index system
CNV	: Kopya sayısı değişikliği
$\alpha$ -HL	: Alfa hemolisin
ddH <sub>2</sub> O	: Didistile su
TBE	: Tris/Borate/EDTA
DMSO <sub>4</sub>	: Dimetil sülfat
MgCl <sub>2</sub>	: Magnezyum
UCSC	: University of California, Santa Cruz

## 1. Giriş

### 1.1. Araştırmanın Problemi

İnsan DNA'sı kullanılarak kimlik belirleme çalışmaları; adli olaylardan, paternite belirlenmesine kadar birçok alanda yaygın olarak kullanılan ve sonuçları itibariyle büyük önemi olan çalışmalardır. Bu çalışmalar, Alec Jeffreys ve arkadaşlarının "DNA profillemesi" ya da "DNA fingerprinting" olarak bilinen tekniği tanımlayıp, adli olgularda başarılı bir şekilde uygulamalarıyla başlamış ve günümüze kadar artan bir ivmeyle devam etmiştir (Jeffreys, Brookfield ve Semeonoff, 1985). DNA profillemesinde geçmişten günümüze dek çeşitli yöntemler kullanılmıştır. Bunlar başlıca Restriksiyon Fragman Uzunluk Polimorfizmi (Restriksiyon Fragment Length Polymorphism-RFLP) ve PCR temelli yöntemlerdir. PCR temelli yöntemleri de temel olarak Değişken Sayılı Ardışık Tekrarlar (Variable Number of Tandem Repeat-VNTR), Kısa Ardışık Tekrarlar (Short Tandem Repeat-STR) ve SNP analizi olarak üçe ayırmak mümkündür (Primorac ve Schanfield, 2014).

RFLP yönteminde DNA, restriksiyon enzimleri aracılığıyla fragmanlara bölünmekte ve bu fragmanlar genellikle agaroz jel üzerinde elektroforez işlemine tabii tutularak görüntülenmektedir. Bu işlemin amacına göre de sonuçlar analiz edilmektedir (Mittal, Chaturvedi ve Tulsyan, 2013). RFLP, adli tıptaki DNA uygulamalarında kullanılan ilk yöntem olarak bilinmektedir. (Kashyap, Sitalaximi, Chattopadhyay ve Trivedi, 2017). PCR ise Mullis tarafından keşfedilen ve çok kısa bir süre içerisinde DNA'daki istenilen bölgeleri amplifiye etmekte kullanılan, günümüzdeki moleküler biyoloji çalışmalarının temelini oluşturan yöntemdir (Mullis, 1990). PCR'ın keşfinden sonra diğer birçok alanda olduğu gibi DNA profillemesi konusundaki keşifler de giderek hızlanmıştır. DNA profillemeye çalışmalarında şu ana kadar VNTR, STR ve SNP bölgeleri analiz edilerek kimliklendirmeler yapılmıştır (Schneider, 2012). VNTR'ler genellikle kromozomların subtelomerik bölgelerinde bulunan 10 ila 100 baz çifti arasında değişen dizilerin tekrarlarından oluşan polimorfik bölgeleridir. VNTR profillemesi genel olarak zordur ve yerini STR'lara bırakmıştır. STR'lar ise 2 ile 6 bazlık dizilerin tekrarından oluşan mikrosatellit bölgeleridir (Reid, Storts, Riss ve Minor, 2004). VNTR'lere göre daha kısa olup daha az tekrar sayısına sahip oldukları için analizleri daha kolay bir şekilde analiz edilebilmektedir. Günümüzde DNA profillemesinde altın standart olarak çeşitli STR bölgeleri analiz edilmektedir (J. M. J.

M. Butler, 2005). STR analizi, Sanger sekanslama olarak da bilinen kapiller elektroforez cihazları aracılığıyla yapılmaktadır. Ancak oldukça konvansiyonel ve en eski otomatize sekanslama yöntemi olan kapiller elektroforez işlemiyle STR analizi hem uzun sürmektedir hem de sık sık işlemin tekrarlanması gerekmektedir (Gammon ve Mayers, 2019). STR bölgeleri kısa tekrar bölgelerinden oluştuğu için yeni nesil dizi analizi cihazlarıyla başarılı bir şekilde dizilenememektedirler. Tüm bu dezavantajlardan dolayı DNA profillemesinde STR analizi yerini yavaş yavaş SNP analizine bırakmaya başlamıştır. SNP'ler genomda oldukça yaygın olarak görülen tek nükleotid polimorfizmlerdir ve STR'lara kıyasla oldukça düşük mutasyona uğrama oranına sahiptirler ayrıca çok daha kolay bir şekilde analiz edilebilmektedirler (Daniel ve Walsh, 2006).

Yeni nesil dizi analizi teknolojisindeki hızlı ilerlemeler, yeni teknolojilerin uygulama alanına girmesiyle devam etmektedir. Özellikle 4. nesil dizi analizi olarak tanımlanan "Oxford Nanopore" teknolojisi, yeni nesil dizi analizi teknolojisinin getirdiği, aynı anda çok sayıda bölgenin analizi, baz başına maliyet düşüklüğü gibi tüm faydaların yanısıra uzun okumalar yapması, hızlı analiz imkanı ve özellikle küçük boyutları sayesinde sahada, laboratuvar dışında ve olgu başında kullanıma imkan sağlamaktadır. Nanopore sekanslama teknolojisi kullanılarak, tek bir DNA veya Ribo Nükleik Asit (Ribonucleic Acid-RNA) molekülü, PCR amplifikasyonu veya örneğin kimyasal etiketlenmesine gerek olmadan dizilenebilmektedir (Feng, Zhang, Ying, Wang ve Du, 2015). Bu teknoloji şu ana dek viral patojenlerin hızlı bir şekilde tanımlanması, Ebola'nın izlenmesi, gıda güvenliğinin izlenmesi, insan genom dizilenmesi, bitki genomu dizilenmesi, antibiyotik direncinin izlenmesi ve haplotipleme gibi pek çok alanda kullanılmıştır ve yapılan bu çalışmalar teknolojinin avantajlarının kimliklendirmede de kullanılabileceğini net bir şekilde ortaya koymaktadır (Belser ve diğerleri, 2018; Miten Jain ve diğerleri, 2018; Quick ve diğerleri, 2016; Schmidt ve diğerleri, 2017).

## **1.2. Araştırmanın Sorusu**

DNA profillemeye, saha içerisinde uygulanabilen, doğru ve hızlı sonuç veren yöntem nedir?

### **1.3. Araştırmanın Hipotezleri**

Nanopore sekanslama teknolojisi, insan DNA'sından aynı anda çok sayıda SNP genotiplendirmesi için uygun bir teknolojidir.

Nanopore sekanslama teknolojisi SNP temelli DNA profillemeye saha üzerinde doğru ve en hızlı sonucu verecek olan teknolojidir.

### **1.4. Araştırmanın Varsayımları**

Araştırmada, izole edilmiş şekilde satın alınmış DNA'lar kullanılmıştır. Bizim de benzer kalitede bir DNA izolasyonu yapacağımız varsayılarak deneylere devam edilmiştir.

### **1.5. Araştırmanın Sınırlılıkları**

Araştırmamızda yalnızca 3 farklı DNA ile genotiplendirme işlemi yapılmıştır. Prototipten kit haline geçiş aşamasında, örnek sayısını artırarak temsil gücünü ölçmek gerekmektedir.

Genomik materyal olarak ticari olarak satın alınan DNA'lar kullanılmıştır. Devam projesinde, kan, tükürük, semen gibi çeşitli biyolojik materyallerden izolasyon yapıp, kitin zorlu koşullardaki başarısını da göstermek gerekmektedir.

Ortaya çıkan prototip teorik olarak sahada sonuç vermeye uygun gözükse de, tüm deneyler laboratuvar ortamında yapılmıştır. Kitleştirme aşamasında saha üzerinde de denemeler yapılmalıdır.

### **1.6. Araştırmanın Amacı**

Bu tezde, nanopore teknolojisi kullanılarak dünyanın en hızlı ve laboratuvar dışında uygulanabilir SNP tabanlı kişi tanımlama kiti prototipi geliştirilmesi amaçlanmıştır.

Bu amaç doğrultusunda 52 SNP (rs1490413, rs876724, rs1357617, rs2046361, rs717302, rs1029047, rs917118, rs763869, rs1015250, rs735155, rs901398, rs2107612, rs1886510, rs1454361, rs2016276, rs729172, rs740910, rs1493232, rs719366, rs1031825, rs722098, rs733164, rs826472, rs2831700, rs873196, rs1382387, rs2111980, rs2056277, rs1024116, rs727811, rs1413212, rs938283, rs1979255, rs1463729, rs2076848, rs1355366, rs907100, rs354439, rs2040411, rs737681, rs2830795, rs251934, rs914165, rs10495407, rs1360288, rs964681, rs1005533, rs8037429, rs891700, rs1335873, rs1028528, rs1528460) analiz edebilen bir prototip hazırlanacaktır. Bu prototipin, laboratuvar ortamı veya donanımlı teknik personel gereksiz saha içerisinde dakikalar içinde sonuç verebilecek olma

potansiyeli nedeniyle yalnızca tıbbi ya da adli vakalarda deęil, kesin kimliklendirme gereken tüm durumlarda kullanım potansiyeli olacaęı öngörülmektedir.



## 2. Genel Bilgiler

### 2.1. Polimorfik alanlar

Geçmişten günümüze, DNA profillemesinde insan genomunda bulunan polimorfik bölgelerden yararlanılmıştır. Bu polimorfik alanlar; VNTR, STR ve SNP'lerdir (Primorac ve Schanfield, 2014).

#### 2.1.1. VNTR

VNTR'ler ilk olarak 1980 yılında tanımlanmış, DNA üzerinde yer alan ardışık tekrar bölgeleridir (Wyman ve White, 1980). Bu bölgeler 6 ya da daha yüksek sayıda nükleotidin ardışık olarak çok sayıda tekrar etmesiyle oluşmaktadır (Näslund ve diğerleri, 2005). VNTR'lar, DNA profillemesinde ilk kullanılan yüksek polimorfik alanlardır (Jeffreys ve diğerleri, 1985). VNTR analizi için ilk etapta RFLP yöntemi kullanılmıştır. RFLP sonucunda oluşan fragmanlar genellikle agaroz jel elektroforezi kullanılarak büyüklüğüne göre ayrılmaktadır ve ardından Southern Blot yöntemi ile hedeflenen VNTR bölgelerinin tespiti yoluyla kişiye DNA profillemesi saptanabilmektedir. Elde edilen bant paterni kişiye özgün olsa da yorumlaması ve uygulaması oldukça zor olan yöntemlerdir. Bu zorlukların yanısıra degrade örneklerde VNTR ile ayırım neredeyse imkasız bir hal almaktadır.

#### 2.1.2. STR

VNTR'dan farklı olarak STR lokusları, daha kısa tekrar dizilerine (2-7bp) sahip olup, bu sayede degrade örneklerde daha dirençli aynı zamanda daha yüksek duyarlılık sağlamaktadır. STR tiplendirme artık adli tıp laboratuvarlarında kullanılan primer yöntem olmuştur. Tetramerik ve pentamerik tekrar dizileri en yaygın kullanılan STR tipleri olup, tek bir PCR reaksiyonunda 16'ya kadar lokus analize girebilmektedir. Diğer PCR temelli metodlar ile karşılaştırıldığında, STR analizi daha hızlı sonuç vermektedir. STR bölgelerinin de çok yüksek polimorfizm göstermesi, ayırım gücünün yükselmesini sağlamaktadır (Tracey, 2001).

The Federal Soruşturma Bürosu (Federal Bureau of Investigation - FBI) laboratuvarı, 1997'de ulusal bir çalışma başlatarak, Combined DNA Index System (CODIS) olarak isimlendirilen çekirdek STR lokus seti oluşturmuştur (Budowle, Moretti, Baumstark, Defenbaugh ve Keys, 1999). Seçilen 13 lokus CSF1PO, FGA, TH01, TPOX, VWA,

D3S1358, D5S818, D7S820, D8S1179, D13S317, D16S539, D18S51 ve D21S11 belirteçlerinden oluşmaktadır. Genel olarak, 15 STR lokusu kullanarak gerçekleştirilen bir DNA profillemeye, yaklaşık  $4 \times 10^{-18}$ 'lik bir duyarlılıkta sonuç verebilmektedir (Marjanovic ve diğerleri, 2006). Daha da kolaylık sağlamak amacıyla, birçok laboratuvar, 16'ya kadar STR lokusunu çoğaltma kapasitesine sahip kitler kullanmaktadır. STR'lar günümüzde hala DNA profillemeye altın standart bir yöntemdir (J. M. J. M. Butler, 2005).

STR temelli DNA profillemenin de dezavantajları vardır. Karışık örneklerin yorumu, özellikle "stutter" nedeniyle komplike bir husustur (Brookes, Bright, Harbison ve Buckleton, 2012). PCR artefaktı, esas allelden 1 tekrar daha kısa "yanlış" allellere yol açabilmektedir. Fragman boyutlarında küçük değişikliklere yol açan tam olmayan tekrar ünitleri içeren alleller görüldüğünde, başka bir durum meydana gelmektedir. Primer bağlanma bölgesi mutasyonları bulunan örneklerde o belirteç kullanılamaz. Bu durumlar, kötü pik dengesi veya heterozigot lokusta "allel dropout" a sebep olur (Heinrich, Müller, Rand, Brinkmann ve Hohoff, 2004). Aynı zamanda STR'lar kapiller elektroforez cihazlarıyla analiz edilebilmektedir (J. M. Butler, Buel, Crivellente ve McCord, 2004). Bu cihazlar Sanger sekanslama teknolojisi olarak da bilinen birinci nesil dizi analizi cihazlarıdır. Sekans başı maliyet yüksekliğinin yanısıra yüksek tekrar oranı gibi sorunlar da teşkil etmektedirler.

### **2.1.3. SNP**

SNP'ler genom üzerinde yer alan ve en sık karşılaşılan tek nükleotid değişiklikleridir. Büyük çoğunlukla bialleliktirler. Genomda yaklaşık 1000 bazda 1 görülürler (Nussbaum, 2016). SNP haplotiplerinin %80'i tüm etnik gruplarda görülmektedir ancak yaklaşık %8'inin popülasyona özgü olduğu bildirilmiştir (Patil ve diğerleri, 2001). dbSNP veritabanında şu ana dek 2 milyarın üstünde SNP bildirilmiştir. Bu rakam 2006 yılında 10 milyon civarındayken, 13 sene içerisinde 200 kat artmıştır (Sherry, 2001).

Son yıllarda kişi tanımlamada SNP kullanımı giderek yaygınlaşmaya başlamıştır. Bunun sebeplerinden ilki yeni nesil dizi analizi teknolojilerinin giderek yaygınlaşması ve SNP genotiplendirmesini oldukça kolay bir hale getirmesidir. Bir diğeri SNP'ler yalnızca kişi ayırmakla kalmayıp ırk, coğrafya tayini gibi bilgiler sunmasıdır. Son

olarak da degrade olmuş DNA'da diğer polimorfik belirteçlere kıyasla çok daha kolay bir şekilde genotiplendirme yapılabilir (Primorac ve Schanfield, 2014).

SNPforID konsorsiyumu 2003 yılında kurulmuş ve STR'lara göre daha kolay genotiplendirilebilir ve yüksek ayırım gücüne sahip, tablo 1'de belirtilen 52 otozomal SNP belirlemiştir (Sanchez ve diğerleri, 2006). Bu konsorsiyumun ortaya çıkışında, STR'ların degrade DNA genotiplendirmesindeki başarısızlığı ve SNP'lerin 500 pg'a kadar DNA ile genotiplendirebilme yeteneği rol oynamıştır. 52 SNP kullanılarak yapılan genotiplendirmenin ayırım gücünün  $5 \times 10^{-19}$  olduğu tespit edilmiştir. Bu rakam, laboratuvarlarda yaygın olarak kullanılan 15 belirteçli STR kitlerinin ayırım gücünden yaklaşık 10 kat daha yüksektir.

**Tablo 1.** SNPforID konsorsiyumu tarafından belirlenen 52 SNP bölgesi ve allel bilgileri

SNP	Referans allel	Alternatif allel	SNP	Referans allel	Alternatif allel
rs1005533	G	A	rs2111980	T	C
rs1015250	G	C	rs251934	A	G
rs1024116	C	T	rs2830795	A	G
rs1028528	A	G	rs2831700	A	G
rs1029047	T	A	rs354439	A	T
rs1031825	A	C	rs717302	G	A
rs10495407	G	A	rs719366	G	A
rs1335873	T	A	rs722098	A	G
rs1355366	T	C	rs727811	G	T
rs1357617	A	T	rs729172	G	T
rs1360288	C	T	rs733164	G	A
rs1382387	C	A	rs735155	C	T
rs1413212	T	C	rs737681	T	C
rs1454361	T	A	rs740910	A	G
rs1463729	C	T	rs763869	G	A
rs1490413	G	A	rs8037429	C	T
rs1493232	C	A	rs826472	T	C
rs1528460	C	T	rs873196	C	T

rs1886510	G	A	rs876724	C	T
rs1979255	C	G	rs891700	A	G
rs2016276	T	C	rs901398	C	T
rs2040411	G	A	rs907100	G	C
rs2046361	T	A	rs914165	G	A
rs2056277	C	T	rs917118	C	T
rs2076848	A	T	rs938283	T	C
rs2107612	G	A	rs964681	T	C

## 2.2. Sekans teknolojileri

### 2.2.1. Yeni nesil dizi analizi:

DNA dizi analizi tek nükleotid varyasyonları, kopya sayı değişikliği (copy number variation-CNV) ve insersiyon/delesyon mutasyonları gibi genetik değişiklikleri moleküler düzeyde saptamayı sağlayan oldukça önemli bir yöntemdir. ‘National Human Genome Research Institute’ 2004 yılında “1000 Genom” projesini başlatmış ve insan genomunu 1000 dolardan daha az maliyet ile dizileyebilecek teknolojilerin geliştirilmesini hedeflemiştir (1000 Genomes Project Consortium ve diğerleri, 2015). Bu sayede “1000 Genom” projesinin başlangıcından günümüze kadar yeni nesil dizileme yöntemleri geliştirilmiştir.

İkinci nesil dizi analizi olarak da adlandırılan yeni nesil dizi analizi yüzbinlerce hatta milyonlarca okumayı aynı anda yapabilen bir teknolojidir. yeni nesil dizi analizi ile aynı anda birden çok hastada pek çok sayıda genin moleküler analizi yapılabilmektedir. Günümüzde var olan bütün platformlarda yeni nesil dizi analizi; kütüphane hazırlığı, amplifikasyon, dizileme reaksiyonu ve veri analizi olmak üzere 4 basamaktan oluşmaktadır (Buermans ve den Dunnen, 2014). Kütüphane hazırlığı sırasında genomik materyal dizileme reaksiyonuna uygun hale getirilmek amacıyla rastgele olarak 35-500 bp uzunluğunda parçalara ayrılır ve DNA fragmentlerinin uçlarına uygun adaptörler takılır. Kütüphane hazırlığı yaklaşık 24-36 saat kadar süre gerektirmektedir. Kütüphane hazırlığı sonrasında, kullanılan platforma göre farklı yöntemler kullanılarak (bridge amplifikasyon, emülsiyon PCR) genomik materyal amplifiye edilir. Bir flow-cell’e optimal kütüphane malzemesi yüklenmesi ile yaklaşık

800-1000 K/m<sup>2</sup> kadar küme elde edilir. Daha sonra dizi analizi basamağı yine kullanılan platforma bağı olarak sentez temelli veya ligasyon temelli olarak yapılır (Buermans ve den Dunnen, 2014; Metzker, 2010). Günümüzde yeni nesil dizi analizi araştırma laboratuvarlarında ve klinik tanı merkezlerinde yaygın olarak kullanıma girmiştir. Yeni nesil dizi analizi ile tüm genom, tüm ekzom veya hedeflenmiş gen analizleri yapılabilmektedir.

Tüm bu avantajlarına rağmen yeni nesil dizi analizi yönteminin bazı eksiklikleri bulunmaktadır. Özellikle okuma derinliğinin düşük olduğu durumlarda sekans hataları ve yanlış hizalama ile sonuçlanabilmektedir. Kısa okumalar yapması sebebiyle yeni nesil dizi analizi ile büyük delesyon veya insersiyon mutasyonlarının, üçlü nükleotid tekrar bölgelerinin ve bazı CNV'lerin saptanması mümkün değildir (Swanson, Ramos ve Snyder, 2014). Bu durum yeni nesil dizi analizi ile STR temelli DNA profillemesi yaklaşımının sağlanmasına engel oluşturmaktadır. Ayrıca yeni nesil dizi analizi ile başına okuma ve analiz maliyeti Sanger dizi analizine göre düşük olsa da analiz için gerekli cihazın ilk kurulumu oldukça maliyetlidir. Örnek başına maliyeti düşürmek için olguların DNA örneklerini biriktirdikten sonra analiz önerilmektedir, bu da zaman kaybına yol açmaktadır (Y. Wang, Yang ve Wang, 2014). Bu sebeple yeni nesil dizi analizi kullanabilen genetik tanı merkezi sayısı oldukça azdır ve pek çok merkezde genetik hastalıkların tanısında Sanger dizi analizi kullanılmaktadır. Sanger dizi analizi tek nükleotid varyasyonları ve küçük delesyon/insersiyon tipi mutasyonlarda halen altın standart bir yöntem olsa da analiz basamakları oldukça zahmetli ve zaman alıcıdır. Ayrıca analiz başına gereken maliyet de yeni nesil dizi analizine göre daha yüksektir.

### **2.2.2. Oxford Nanopore Teknolojisi:**

Yeni nesil dizi analizi teknolojisindeki hızlı ilerlemeler, yeni teknolojilerin uygulama alanına girmesiyle devam etmektedir. 4. nesil dizi analizi olarak adlandırılan Oxford Nanopore teknolojisi tek DNA molekülünü real-time dizileme yapan ucuz ve ultra-hızlı bir teknolojidir (Lu, Giordano ve Ning, 2016; Y. Wang ve diğerleri, 2014). DNA molekülünün nano boyuttaki porlardan geçişi sırasında nükleotidlerin oluşturduğu elektrofizyolojik değişiklikler analiz edilerek DNA dizisi ortaya çıkartılır (Fyta, 2015). Bu teknoloji ile tek bir molekül çift zincir veya tek zincir DNA veya RNA, PCR amplifikasyonuna, kimyasal etiketlenmesine, DNA sentezine veya ligasyona

gereksinim duymadan dizilenebilmektedir. Oxford Nanopore teknolojisi ile kütüphane hazırlığı ve sekanslama çok kısa sürede ve pratik olarak yapılabilmektedir. Ayrıca PCR basamağının da ortadan kalkması ile dizi analizi süresi günlerden saatlere düşmektedir. Teknolojinin diğer çok önemli bir avantajı da çok uzun okumalar (104–106 bin baz) yapılabilmesidir (Lu ve diğerleri, 2016; Y. Wang ve diğerleri, 2014).

Oxford Nanopore Minion cihazı, ilk kez 2014 yılında kullanıma sunulmuş olan, 90 gr ağırlığında ve 10x3x2 cm boyutlarında, günümüzde var olan en küçük dizi analizi cihazıdır (Lu ve diğerleri, 2016). Herhangi bir bilgisayarın standart USB 3.0 çıkışına takılarak kullanılmaktadır. Günümüzde maliyeti yaklaşık 1000 dolar civarındadır. Bu kadar ucuz, küçük ve pratik olması her türlü laboratuvar koşullarında, laboratuvar dışında veya direk olgu başında son derece hızlı sürede analiz olanağı sağlamaktadır.

Oxford Nanopore Teknolojisi ile dizi analizi, üzerinde nano boyutta porlar bulunan bir membran aracılığıyla yapılmaktadır. Bu membran elektrolitler içeren bir sıvıyı cis ve trans kompartmanlarına ayırır. Bu ortama bir voltaj uygulandığında pordan geçen iyonik akım standart elektrofizyolojik teknikler ile ölçülebilmektedir. Nanopordan geçen iyonik akım membran üzerinden geçen voltajın kontrolü ile yapılmaktadır. Ortama uygulanan voltaj farkı sebebiyle ortaya çıkan iyonik akım negatif yüklü DNA molekülünün nanopordan geçişini sağlamaktadır. DNA molekülünün nanopordan geçişi sırasında akımda nükleotide özgü bir değişiklik oluşmaktadır. İyonik akım değişikliklerinin ölçülmesi ile DNA dizisi tanımlanmaktadır (Branton ve diğerleri, 2008; Fyta, 2015). Bu teknolojiye yeni nesil dizi analizinden farklı olarak nükleotidlerin floresan boya ile etiketlenmesi gerekmemektedir. Nanoporlar biyolojik ve sentetik olmak üzere iki gruba ayrılır.  $\alpha$ - haemolysin ( $\alpha$ -HL) ilk tanımlanan ve en çok kullanılan biyolojik nanopordur. Biyolojik nanoporlar por boyut ve yapısının sabit olması ve stabil olmamaları nedeniyle dizi analizinde başarısızlığa yol açmışlardır. Bu sebeple sentetik por içeren katı ve plastik membranlar geliştirilmiştir (Feng ve diğerleri, 2015).

Oxford Nanopore Teknolojisi'nde, özel tasarlanmış kitler sayesinde kütüphane hazırlık süreci çok basittir ve 10 dakika kadar kısa bir sürede tamamlanabilmektedir. MinION cihazı bilgisayara bağlandıktan sonra software programları ile bağlantı kurularak \*.fast5 formatında dizi analizi verisi oluşturmaktadır. Bu teknoloji ile dizi analizi ve veri oluşumu eş zamanlı olarak gerçekleşmesi sayesinde analiz için istenilen dizi boyutuna ve okuma derinliğine ulaşıldığında dizi analizi durdurulabilmektedir. Bu

özellik teknolojinin en önemli avantajlarından bir diğeridir. Bu teknolojiyle uzun okumalar “real-time” olarak yapılabilmesine rağmen teknolojinin en büyük açığı özellikle homopolimer bölgelerindeki hata oranlarının yüksek olmasıdır. Ancak son zamanlarda bu hata oranı Oxford’un geliştirdiği 1D2 kitleri, okuma derinliğinin artırılması ve biyoinformatik algoritmalarla hata oranı optimal seviyelere indirilmiştir (Y. Wang ve diğerleri, 2014).

## **2.3. Biyoinformatik**

### **2.3.1. Basecalling**

Yeni nesil dizi teknolojilerinin tümü, kimyasal, donanımsal ve optik sensörlerin karmaşık bir etkileşimi içindedirler. Sensörlerden gelen kompleks verilerin, nükleotitlerle eşleştirilmesi işlemi basecalling olarak adlandırılmaktadır. Her sekans platformu farklı bir donanımsal mekanizma ve sensör yapısına sahip olduğu için ortaya çıkan ilk veriler de farklı olmaktadır. Bunun için farklı platformlar için farklı programlar geliştirilmiştir (Ledergerber ve Dessimoz, 2011). Bu işlem sonucunda genellikle okumalardaki nükleotid dizileri ve bunlara ait kalite skorlarını içeren \*.fastq formatında dosyalar oluşmaktadır.

### **2.3.2. Haritalama ve hizalama**

Terminolojik olarak haritalama ve hizalama aynı anlamda kullanılsa da aralarında küçük nüanslar vardır. Haritalama, sekans platformundan elde edilen herbir okumanın, referans genom ile karşılaştırma işlemidir. Hizalama ise haritalama işleminin daha ayrıntılı bir şekilde yapılmış halidir. Hizalama sonunda okumalara ait kromozomal lokasyon bilgileri elde edilmektedir. Yapılan çalışmanın türüne göre farklı algoritmalar içeren farklı programlar kullanılmaktadır.

### **2.3.3. Varyant çağırma**

Yeni nesil dizi analizi sonucunda ortaya çıkan okumalardaki tek nükleotid varyantlarının liste haline getirilmesi işlemi olarak tanımlanabilir. Bu işlem sonucunda yalnızca referans genomla eşleşmeyen varyantlar elde edilebileceği gibi, lokasyonu verilen herhangi bir nükleotide ait genotip, okuma derinliği, kalite skoru gibi bilgileri de içerebilmektedir.

### **3. Gereç ve Yöntem**

#### **3.1. Araştırmanın tipi**

Araştırmamız, DNA profillemede yeni bir yöntem deneme ve varolan yöntemi geliştirme niteliği taşıdığından metodolojik araştırmadır.

#### **3.2. Araştırmanın yeri ve zamanı**

Araştırmamızın, ön çalışmaları 2018 yılı Haziran ayında başlamış olup, 26 Nisan 2019'da Ege Üniversitesi Bilimsel Araştırma Projeleri Genel Müdürlüğü tarafından desteklenmeye başlamıştır. Destek sürecinden sonraki ıslak laboratuvar çalışmaları Ege Üniversitesi Tıp Fakültesi Tıbbi Genetik Anabilim Dalı'nda gerçekleştirilmiştir.

#### **3.3. Araştırmanın örnekleme**

Araştırmamızda, Coriell Enstitüsü'nden satın alınan 3 farklı ticari DNA örneği kullanılmıştır. Bu açıdan etik kurul gerekliliği bulunmamaktadır.

#### **3.4. Moleküler çalışmalar**

##### **3.4.1. SNP bölgelerinin seçimi**

Çalışmamızda, SNPforID konsorsiyumu tarafından belirlenen 52 SNP bölgesi kullanılmıştır (Amigo, Phillips, Lareu ve Carracedo, 2008; Sanchez ve diğerleri, 2006).

##### **3.4.2. Primer Dizaynı**

Her bir SNP için The University of California Santa Cruz (UCSC) veritabanı (James Kent ve diğerleri, 2002) ve Primer3 (Koressaar ve Remm, 2007; Untergasser ve diğerleri, 2012) yazılımı kullanılarak primer dizaynları yapılmıştır . Primer seçim parametleri; primer bağlanma sıcaklığı optimum 60,0 °C olacak şekilde 59,0 °C ile 61,0 °C arasında, primer uzunlukları 18-30 nükleotid aralığında ve primer GC yüzdesi yaklaşık %50 civarında olacak şekilde belirlenmiştir. Ürün boyutu aralığı 200 baz çifti civarında olacak şekilde dizaynlar gerçekleştirilmiştir. Bu büyüklüğün seçilmesinin sebebi sonraki aşamalarda multipleks PCR denemeleri sırasında, süreyi ve siklus sayısını minimumda tutabilmektir. Primer dizileri tablo 2'de belirtilmiştir.

**Tablo 2.** Çalışmamızda kullanılan primer dizileri

SNP	Forward	Reverse
rs1490413	TTGGCTATTTGGAGATTGGTG	GTTGGTGGAAGGGACTGAGA
rs876724	TGCCCTCACAGTCTCAGCTA	ACCAATCCAGGGGGACTAGA
rs1357617	TGTGTCTCAAACGCCATCAG	AAACCTGGGGCTTTTGACCTC
rs2046361	AAGCATTCACTACTGGCTCA	TTTCTGCCAAATGTATCCTTACC
rs717302	GGCATATTATAAATAAACCGAGAAAAA	ATGTGGGCAAGCAATCTAGC
rs1029047	GGAAATTTGTAGATCATTCTGGGTA	CACTGAAGTTTCAGGGCTCA
rs917118	CAGAATCCAGCTTCTCCAC	CTTGAGCAAAGGCCTTGAGA
rs763869	TTGAACAACACAACGACAAATG	AGGGACACAGCCAAACCATA
rs1015250	TCATCCCCATGTGCTAACAA	ACCGACATGGGAAATGTCA
rs735155	CCCTCTCCACGTATTTGA	GCTCCAGAAGGGACCTAACC
rs901398	CAAGAAAGAGGCCCTCACAC	CAACGCCACCATCATAACAGA
rs2107612	GGTGAAGTCTGGATTTTCTTTTG	AGGGAATTCGAAATCTGAAGG
rs1886510	TTCTGCATTTTCTCCCTGCT	AACCTCCAGTCTCCAAACC
rs1454361	TGGGTGAGGCCTGAGTTATC	AGCCACGTTGGGATCTTATG
rs2016276	ATCGCAGTGAGAAAGGGATG	TCTGGCTCACTGGTTGTTTG
rs729172	GCCTAAGCCAGCGTGTAGTA	CAATCACGAATACCACCGTTT
rs740910	TGACACATCCATGGTATGGAATA	TGCATCCTCACTTCCCTTTC
rs1493232	GGAACCTTCTAAATCTCCTTGATG	TGGAAAGCAGCATAATTGTCA
rs719366	TCGGGCAACAAAGTTAGACC	AGTCACTTCTCGGCAGCATT
rs1031825	TTGTGTTCTTCTCCGTGAGC	TGTCTCGAGGGTCTTTTCTCA
rs722098	CCTTAGGAATCCCCGTTTAC	CCCACAGATAGGGGGAAAAAT
rs733164	GGGCACAGCTTCTTGACTCT	GGCCAGCTGTGTAAATGTT
rs826472	TGTGTTCCGATTTGGTAAACTC	CCAGCAAAAACCTCTTTTCTCC
rs2831700	GGCAGTGAGGAATAGCTGATG	TTTTTCCAGCCCTGCTACC
rs873196	ATGCCACTCACCTCTCCAAC	TGTGCCTGGCTAGGAAAGTT
rs1382387	TTGCACGGTTCTGTGTAACAA	TTCTTCCCTTTTGTAGAATTGA
rs2111980	GTAACCGGAGGCTCTTCCTT	GCCACAGGGTGACACTTCTT
rs2056277	CAGACAAACACCTCCCTGGT	AAATGATTGGGCCCTGTAT
rs1024116	TCAATCCCACTCTTCCCTTT	TGCAGCATTTTGCTTGAAGT
rs727811	GCACCTGTATTAGTGGCCTTG	CCAGAATGAGCCAAAGTATGTTT
rs1413212	GCAAAAAGATTGGACAGGCTAA	GTTTTGCAGGCCGACTTATT
rs938283	GAATCACAGGAAGTTGCAAAAA	CCCCTTGATCTGTGCATGT
rs1979255	TAGCATGGGGGTTTTTGAAG	CCAGTTCAGAGACTATGGATGGTA
rs1463729	ACATCTGCACGTGCTTTGTGTC	TCTGGGAGCCTGTAGGAGAA
rs2076848	ATGCCATTTTGAAGCCAAAC	TGGCTTAAATCAATGGGTCTTC
rs1355366	GAACCCAAGAGAGCCAAGAA	AAATGCTCCTGCACCAAAAT
rs907100	CCTCCTGCCAGAACCACTAA	CGAAAAGGTGAAACCAGTCC
rs354439	TGATAAGAGGTGCTTCGGAAA	TTTCTAGTAAAGCTCTGGCATGAA
rs2040411	TCACATTGCTCTGTTCATCTGC	TCTGTGATGCATAGGGTCCA
rs737681	GCCCTGAGACCTCTCTCCAT	CCACCGTCGCTTTATTCTGT

rs2830795	ATAGGAAGGCTGCAGGTTGA	GGGAGGTTGTCATCAGAGA
rs251934	CATTGGCTGTGATAGCTCCA	TGACCCTTGCAGAGAACTGA
rs914165	ATGGCCCCGATACAATGAT	ATGAACAAGGACCCAGCTT
rs10495407	ATGCCTGCAGGTTTGGTATC	TTGGTGGTCTCTTCACATGG
rs1360288	ACGCTGGTCACTCCAGCTAC	TCTGAGAAGCAGGGAGGATG
rs964681	AAACGGGGAGCTCTGTCTG	AGCCGGGGACTTTGACTATT
rs1005533	TTCACAGGGTTTCAAAGGAA	AAGGGAAGCCCTTTGAGC
rs8037429	GCCAAGGCATCTGAAATAGG	CATTATTTGCTGAAAAGTGCTACG
rs891700	TTTGTACACTTTTTTCAGAGGTGGT	GAGCTTTGGCAACAGGGTAG
rs1335873	TCTGTCCCCAACAGAAGTT	GCCATGAACCTTTGGACCTA
rs1028528	CGAGTGCCATTGAATCATTTT	TGCTGTGGAGGAGAAGGAAG
rs1528460	TTCTCAACAACCTCCCCCTTG	ATTATGTTGGGATGGGGTGA

Tüm forward ve reverse primerlerin 5' uçlarına tablo 3'te belirtilen M13 kuyruk primeri eklenmiştir (Oetting ve diğerleri, 1995). Bu işlemin esas amacı multipleks PCR işlemi sırasında, primerlerin bağlanma sıcaklık dereceleri arasındaki farklılıkların SNP bölgeleri arasında amplifikasyon dengesizliği yaratmasının önünü geçmektir.

**Tablo 3.** M13 primer dizileri

**M13**

Forward	GTAAAACGACGGCCAG
Reverse	CAGGAAACAGCTATGAC

Primerler liyofilize stok halinde elimize ulaştıktan sonra, üreticinin önerdiği koşullarda 100 mM'lık konsantrasyon elde edilecek şekilde ddH<sub>2</sub>O ile sulandırılmıştır. Sonrasında 0,2 ml'lik tüplerde 10 µl stok primeri ve 90 µl ddH<sub>2</sub>O birleştirilerek 10 mM konsantrasyonda 52 primer çifti oluşturulmuştur.

**3.4.3. PCR çalışmaları**

**3.4.3.1. Standart PCR çalışmaları**

Öncelikle 52 çift primer tek tek tablo 4'teki kondisyon ve koşullarla, standart bir 96 kuyucuklu PCR cihazında amplifiye edilmiş ve agaroz jel elektroforezi uygulanmıştır.

**Tablo 4.** İlk PCR işlemi kondisyon ve koşulları

PCR kondisyonları		PCR koşulları		
Kimyasal	Miktar	Sıcaklık	Süre	Döngü Sayısı
Maxima Hot Start PCR Master Mix	12,5 µl	94 °C	5 dakika	1
ddH <sub>2</sub> O	5 µl	94 °C	30 saniye	
Primer - Forward	1 µl	60 °C	30 saniye	35
Primer - Reverse	1 µl	72 °C	60 saniye	
DNA	100 ng	72 °C	7 dakika	1

#### Agaroz jel elektroforezi

1. Cam beher içine, 4 gr agaroz ve 200 ml 0.5X Tris/Borate/EDTA (TBE) Buffer konulup mikrodalga fırında agaroz saydam bir görünüm alıncaya kadar ısıtılmıştır.
2. Sonrasında üzerine dikkatli bir şekilde 32 µl etidyum bromür eklenerek karıştırılmıştır ve ürünlerin yükleneceği kuyucukların oluşması için taraklar yerleştirilmiştir.
3. Jel soğumaya bırakılmıştır.
4. Parafilm üzerinde 5 µl Orange G ve 5 µl PCR ürünleri karıştırılarak kuyucuklara yüklenmiştir.
5. Amplifiye olan ürünlerin uzunluklarını değerlendirmek için ilk ve son kuyucuklara DNA ladder yüklenmiştir.
6. Örnekler 110 Volt akımda 45 dakika yürütüldükten sonra jel görüntüleme sistemi ile görüntüler alınmıştır.

Agaroz jel elektroforeziyle bantlar konfirme edildikten sonra 26 SNP'e ait forward ve reverse primerleri, her birinden 5 µl eklenerek 0,5 ml'lik tüplerde birleştirilmiştir. Bu işlemin sonunda 26 primer çiftini barındıran 2 ayrı tüpte primer miksleri (miks 1 – ilk 26 primer çifti, miks 2 – ikinci 26 primer çifti) oluşturulmuştur. Bu aşamadan sonra tablo 5'teki kondisyon ve koşullarla, 0.2 ml'lik tüpler içerisinde PCR işlemi

gerçekleştirilmiştir. Elde edilen PCR ürünleri, Illumina NextSeq cihazıyla sekanslanmak üzere, -20 °C'lik buzdolabında saklanmıştır.

**Tablo 5.** Multipleks PCR işlemi için kullanılan kondisyon ve koşullar

PCR kondisyonları		PCR koşulları		
Kimyasal	Miktar	Sıcaklık	Süre	Döngü Sayısı
KAPA2G Fast Multiplex Mix	12,5 µl	94 °C	60 saniye	1
ddH <sub>2</sub> O	5 µl	94 °C	5 saniye	
DMSO4	1 µl	60 °C	30 saniye	30
MgCl <sub>2</sub>	1,5 µl	72 °C	60 saniye	
Primer Miks DNA	5 µl 100 ng	72 °C	2 dakika	1

#### 3.4.3.2. Süper hızlı PCR çalışmaları

Bu basamakta PCR cihazı olarak Philisa® Thermal Cycler kullanılmıştır. Bu cihaz standart olarak 1,5 saat süren PCR süresini 20 dakikanın altına düşüren bir mekanik tasarıma sahiptir (Geyer ve Hanson, 2013). Tablo 5'teki kondisyon ve koşullarla, 0.2 ml'lik tüpler içerisinde bu cihazla da PCR işlemi gerçekleştirilmiştir. Elde edilen PCR ürünleri, Illumina NextSeq cihazıyla sekanslanmak üzere, -20 °C'lik buzdolabında saklanmıştır.

#### 3.4.4. Sekanslama çalışmaları

##### 3.4.4.1. Yeni nesil dizi analizi

Bu basamakta, hem 3 farklı ticari DNA örneğinin DNA profillerini belirlemek, hem de önceki basamaklardaki multipleks PCR çalışmalarının tüm SNP'ler için amplifikasyon başarısını ölçmek amaçlanmıştır. Nextera XT kütüphane hazırlık hiti ve NextSeq Reagent v2.5 kartuşu kullanılarak, Illumina NextSeq 550 cihazında sekanslama işlemi gerçekleştirilmiştir. Aşağıdaki protokol kullanılmıştır:

##### Fermentas ile pürifikasyon

1. 1,5 ml'lik eppendorf tüpüne birleştirilmiş 60 µl ampikon eklenmiştir.

2. Bunun üstüne, 60 µl Binding Buffer ve 60 µl izopropanol eklenmiştir.
3. Bu karışım pipetaj yapılarak spin kolona aktarılmıştır.
4. 14000 rpm'de 1 dakika santrifüj yapılmıştır.
5. Ardından spin kolonun dışındaki tüp atılarak, 2ml'lik yeni toplama tüpü takılmıştır.
6. Spin kolona 700 µl Wash Buffer eklenmiştir.
7. 14000 rpm'de 1 dakika santrifüj yapılmıştır.
8. Spin kolonun dışındaki tüp atılarak, 2ml'lik yeni toplama tüpü takılmıştır.
9. 14000 rpm'de 1 dakika santrifüj yapılmıştır.
10. Spin kolonun dışındaki tüp atılarak ve spin kolon 1,5 ml'lik steril Eppendorf tüpü üstüne yerleştirilmiştir.
11. Spin kolona 50 µl Elution Buffer eklenmiştir.
12. Sonrasında 14000 rpm'de 1 dakika santrifüj edilip ve spin kolon atılmıştır.
13. Bu işlem sonrasında pürifiye PCR amplikonları elde edilmiştir.
14. Konsantasyon ölçümü Qubit cihazı ile yapılmıştır.

#### DNA tegmentasyonu

1. Her örnek için 1 ng birleştirilmiş pürifiye amplikon, total volümü 5 µl olacak şekilde nuclease free su ile tamamlanarak, 0.2 ml'lik eppendorf tüpe alınmıştır.
2. Üzerine 5 µl Tagment DNA Buffer ve 5 µl Amplicon Tagment Mix eklenip 5 kez pipetaj yapılmıştır.
3. 280 g hızda 20°C'de 1 dakika santrifüj edilmiştir.
4. Daha sonra PCR tüpleri thermocyclera yerleştirilerek 55 °C'de 5 dakika inkübasyon sağlanıp ve sonrasında 10 °C'ye soğutulmuştur.

#### Tegmente DNA'nın temizlenmesi

1. İnkübasyon işleminden sonra örneklere 5 µl Neutraline Tagment Buffer eklenip 10 kez pipetaj yapılmıştır.
2. Oda ısısında 5 dakika inkübe edilmiştir.

### PCR amplifikasyonu

1. Önceki basamakta elde edilen 20 µl ürüne 15 µl Nextera PCR Master Mix, 5 µl Index1 ve 5 µl Index 2 eklenmiştir.
2. 280 xg hızda 20 °C’de 1 dakika santrifüj edilmiştir.
3. Daha sonra PCR tüpleri PCR cihazına yerleştirilerek tablo 6’da belirtilen koşullarda inkübasyon sağlanmıştır.

**Tablo 6.** Yeni nesil dizi analizi çalışması için kütüphane hazırlığında PCR amplifikasyonu basamağı inkübasyon koşulları

Sıcaklık	Süre	Döngü Sayısı
72 °C	3 dakika	1
95 °C	30 saniye	
95 °C	10 saniye	
55 °C	30 saniye	12
72 °C	30 saniye	
72 °C	5 dakika	1
10 °C	∞	

### PCR temizleme

1. Ürün thermocyclerdan çıkarılarak 280 xg hızda 20 °C’de 1 dakika santrifüj edilmiştir.
2. Üzerine 90 µl AMPure XP Bead eklenip ve 10 kez pipetaj yapılmıştır.
3. Oda ısısında 5 dakika inkübe edilmiştir.
4. Manyetik standde 2 dakika bekletilmiştir.
5. Supernatant kısmı atılmıştır.
6. Manyetik standden kaldırmadan 190 µl taze hazırlanmış %80’lik etanol eklenerek ve 30 saniye beklenilmiştir.
7. Supernatant kısmı atılmıştır.
8. 6 ve 7. basamaklar tekrarlanmıştır.

9. Manyetik stand'tan kaldırmadan oda ısısında 15 dakika kurumaya bırakılmıştır.
10. Manyetik stand'tan kaldırılıp 52,5 µl Resuspension Buffer eklenmiş ve 10 kez pipetaj yapılmıştır.
11. 2 dakika oda sıcaklığında inkübe edilmiştir.
12. Manyetik stand'a alınıp 2 dakika inkübe edilmiştir.
13. 40 µl supernatant yeni bir tüpe 0.2ml'lik Eppendorf tüpe'e alınmıştır.
14. Bütün ürünlerin DNA konsantrasyonu Qubit DNA konsantrasyon ölçme cihazı ile ölçülmüştür.
15. Her bir pool 100 ng DNA içerecek şekilde birleştirilmiştir.
16. Birleştirilen pool'ların DNA konsantrasyonu tekrar Qubit DNA konsantrasyon ölçme cihazı ile ölçülmüştür.
17. Son üründen 15 ng DNA içerecek şekilde uygun koşullarda kite yüklenmiştir.
18. Kartuş NextSeq Illumina cihazına yerleştirilip uygun program seçilerek çalıştırılmıştır.

#### **3.4.4.2. Nanopore ile sekanslama**

Bu basamakta Oxford Nanopore Minion cihazı kullanılarak sekanslama işlemi gerçekleştirilmiştir. Kit olarak ise SQK-RAD004 kodlu hızlı kütüphane hazırlık kiti kullanılmıştır. Üreticinin önerdiği protokol ile yaklaşık 10 dakikada kütüphane hazırlığı tamamlanmaktadır ve sekans işlemine geçilmektedir. Minion cihazı, herhangi bir bilgisayara USB 3.0 noktası aracılığıyla bağlanarak çalışmaktadır ve öncesinde herhangi bir hazırlık aşaması yoktur. Örnekler yüklendikten yaklaşık 2 dakika sonra real-time olarak veri akışı başlamaktadır. Aşağıdaki protokol izlenmiştir:

##### **Kütüphane hazırlığı**

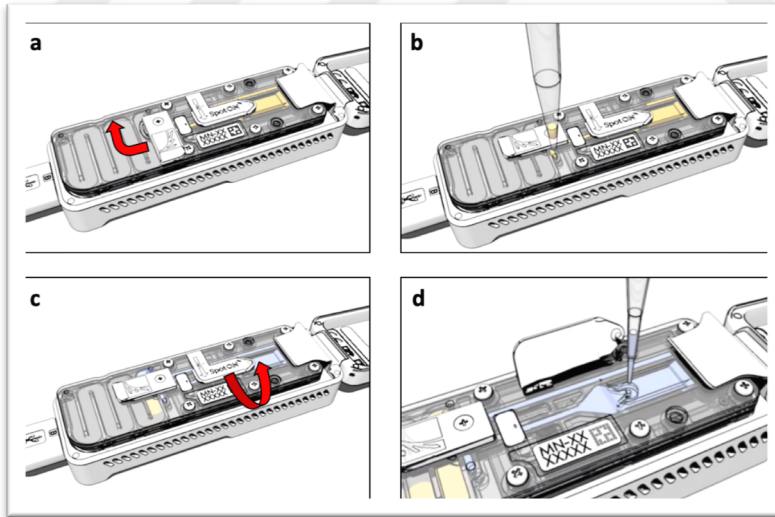
1. Multipleks PCR sonrası 0,2 ml'lik iki ayrı primer miksiyle amplifiye edilen ampikonlar tek bir tüpte birleştirilip ve yaklaşık 50 µl volüme sahip 52 SNP'e ait ampikon miksi oluşturulmuştur.
2. Yaklaşık 50 µl volüme sahip 52 SNP'e ait ampikonları içeren bu miksten 5 µl, 1,5 ml'lik Eppendorf DNA LoBind tüpüne alınmıştır.

3. Total volüm 7,5 µl olacak şekilde nuclease-free su ile tamamlanıp ve pipetaj yapılmıştır.
4. Ardından tüp birkaç saniyeliğine santrifüj edilmiştir.
5. Sequencing Buffer, Loading Beads, Flush Buffer, Flush Tether, Fragmentation Mix ve Rapid Adapter kimyasalları, buzdolabından çıkartılıp ve oda sıcaklığında erimek üzere bırakılmıştır.
6. Sonrasında bu kimyasallar birkaç saniyeliğine santrifüjlenip ve pipetaj yapılmıştır.
7. Hemen ardından buz kalıbının üstüne alınmıştır.
8. 5 µl genomik DNA miksi 0,2 ml'lik PCR tüpüne transfer edilmiştir.
9. Tüpün içine 2.5 µl Fragmentation Mix eklenmiştir.
10. Karışıma pipetaj yapılmıştır.
11. Sonrasında thermal cycler cihazına alınıp ve 30°C'de 1 dakika, hemen ardından 80°C'de 1 dakika inkübe edilmiştir.
12. Hızlıca pipetaj yapıldıktan sonra buz kalının üstüne alınmıştır.
13. Bu karışımdan 10 µl alınıp ve 0,2 ml'lik eppendorf tüpüne aktarılmıştır.
14. Üstüne 1 µl Rapid Adapter eklenmiştir.
15. Pipetaj yapıldıktan sonra 5 dakikalığına oda sıcaklığında inkübasyona bırakılmıştır.

#### Flow Cell hazırlığı ve yükleme

1. MinION sekans cihazı, bilgisayara bağlanıp ve kapağı açılarak üstüne FlowCell yerleştirilmiştir.
2. FlowCell'in yükleme portu kapağı saat yönünde 90° döndürülüp ve yükleme portunun açılması sağlanmıştır (Şekil 1a).
3. 1000 µl'lik bir pipet 200 µl'ye ayarlanmıştır.
4. Bu pipet aracılığıyla yükleme portundan baloncuklar ve küçük bir miktar buffer çekilmiştir (Şekil 1b).
5. Flush Tether tüpünün içine doğrudan 30 µl Flush Buffer eklenip ve pipetaj yapılmıştır.

6. Ardından bu karışımın 800 µl'si, baloncuk oluşmamasına dikkat edilerek yavaş bir şekilde yükleme portunun içine yüklenilmiştir.
7. 5 dakika oda sıcaklığında inkübasyona bırakılmıştır.
8. Sonrasında kütüphane hazırlığı basamağındaki 11 µl'lik amplikonları içeren karışımın üstüne; 34 µl Sequencing Buffer, 25,5 µl Loading Beads ve 4,5 µl ddH<sub>2</sub>O eklenmiştir.
9. FlowCell üstündeki SpotON örnek portu kapağı yukarı kaldırılıp ve portun açılması sağlanmıştır. (Şekil 1c).
10. 5. basamakta geriye kalan 200 µl'lik Flush Tether ve Flush Buffer karışımı yavaşça yükleme portunun içine yüklenmiştir.
11. Hemen ardından 8. basamaktaki 75 µl'lik miks damla damla olacak şekilde SpotON örnek portuna yüklenilmiştir (Şekil 1d).
12. Bu işlemin ardından önce SpotON örnek portu kapağı, sonra ise yükleme portu kapağı dikkatlice kapatılmıştır.



**Şekil 1.** Oxford Nanopore FlowCell'ine örnek ve buffer yükleme aşamaları

#### Sekanslama işleminin başlatılması

1. MinION cihazı USB 3.0 destekli bağlantı noktası aracılığıyla bilgisayara bağlanılmıştır.

2. Sonrasında MinKNOW uygulaması açılmıştır.
3. FlowCell ID ve kullanılan kit seçimleri yapılmıştır (SQK-RAD004).
4. Sonrasında sekanslama işlemi başlatılmıştır.

### **3.4.5. Biyoinformatik**

#### **3.4.5.1. Yeni nesil dizi analizi informatiği**

##### Basecalling

Illumina NextSeq cihazından çıkan ham datalar analiz edilmiştir. NextSeq cihazında oluşan \*.bcl formatındaki ham veriler, bcl2fastq2 programı aracılığıyla \*.fastq formatına dönüştürülmüştür (Illumina, 2018).

##### Haritalama ve Hizalama

BWA programı MEM opsiyonu kullanılarak kullanılarak hg19 insan referans genomuna göre haritalanmıştır (Li ve Durbin, 2010). Haritalama sonrasında ortaya çıkan \*.sam formatındaki dosyalar SamTools programı aracılığıyla \*.bam formatına dönüştürülüp indekslenmiştir (Li ve diğerleri, 2009).

Ardından bu dosyalar IGV programı açılarak görüntülenmiştir (Robinson, Thorvaldsdóttir, Wenger, Zehir ve Mesirov, 2017). Sonrasında herbir SNP'e ait kromozomal lokasyonlar girilerek SNP'lerin genotip bilgileri çıkarılmıştır.

#### **3.4.5.2. Nanopore sekanslama informatiği**

##### Referans oluşturma

Biyoinformatik analizi süresinin olabildiğince kısa olması asıl hedefimiz olduğu için, cihazdan çıkacak olan okumaları 3 milyar bazlı insan genomuna haritalayıp vakit kaybetmek yerine yeni ve prototipe özgün bir \*.fasta formatında referans oluşturulmuştur. Standart referans genomlar genel olarak şekil 2'deki formattadır.

```

>chrM
GATCACAGGTCTATCACCTATTAACCACTCACGGGAGCTCTCCATGCAT
TTGGTATTTTCGTCTGGGGGGTGTGCACGCGATAGCATTGCGAGACGCTG
GAGCCGGAGCACCTATGTGCGAGTATCTGTCTTTGATTCTGCCTCATT
CTATTATTTATCGCACCTACGTTCAATATTACAGGCGAACATACCTACTA
AAGTGTGTTAATTAATTAATGCTTGTAGGACATAATAATAACAATTGAAT
GTCTGCACAGCCGCTTCCACACAGACATCATAACAAAAAATTTCCACCA
AACCCCCCCTCCCCCGCTTCTGGCCACAGCACTTAAACACATCTCTGC
CAAACCCCAAAAACAAGAACCCTAACACCAGCCTAACCCAGATTTCAAAT

```

**Şekil 2.** İnsan mitokondriyal kromozomuna ait referans dizi örneği

Çalışmamızda benzer formatta bir referans oluşturulmuştur. Kromozom isimleri yerine SNP'lerin dbSNP veritabanının yer alan rs kodları yazılmıştır. Ardından UCSC veritabanından, SNP'in bulunduğu noktanın 5' ve 3' uçlarından 100'er baz uzaklığı kapsayacak şekilde diziler elde edilmiştir. Toplamda her bir SNP'e ait 201 bazlık bir sanal referans oluşturulmuştur. Bu referansı, şekil 3'teki formatta 52 kromozomlu ve her bir kromozomu 201 baz içeren bir sanal genom olarak tanımlamak mümkündür.

```

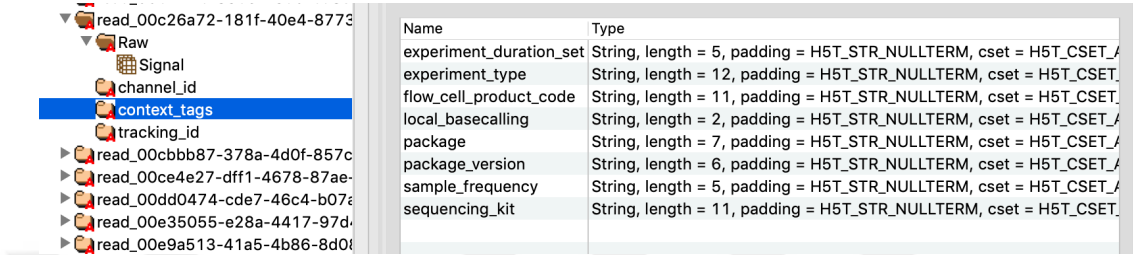
>rs1490413
attaagcacatataaaatgctcagcacagaaataaaagctctttgtttct
gacagctcagaactgcctggtgtggactgggctgatgtgggttctttgca
Gaaactggctggcctcagagcagggacactagtgagaagtgtctggttcag
gtgtggacagcccagagaggcttctcagtccctccaccaacatggggga
t
>rs876724
acagctttgccctcacagtctcagctaccctgggaaagagtttgccaggg
gaatcatcatgcaggctccatTTTTATACCCTGCactgaagtataagta
CattttttgtcacactctgctaactgcctgctcatagatattcaaatTTA
gtgatgtagacagactcctaacttctcatggTTTAAATgTTTAAACAA
c

```

**Şekil 3.** Oluşturduğumuz fasta formatındaki referansa ait SNP ve dizi örnekleri (Kırmızı ok – genotiplendirmek istediğimiz SNP'lere ait baz)

Nanopore sekans cihazı, diğer bilinen yeni nesil dizi analizi cihazlarından farklı olarak \*.fast5 uzantılı ham veriler oluşturmaktadır. Sekanslama işleminin 3. dakikasından itibaren cihaz ham veriler üretmeye başlamaktadır ve sekans süresi boyunca veri akışı devam etmektedir. Fast5 uzantısı, aslında HDF5 data modeli şeklinde olan bir çeşit hiyerarşik veri formatıdır (Lu ve diğerleri, 2016). Bu veri saklama şekli, son derece büyük ve kompleks verilerin daha kolay bir şekilde yönetilebilmesini sağlamaktadır.

Diğer yeni nesil dizi analizi platformları sadece dizi bilgisi ve bu diziye ait kalite skorlarını içeren \*.fastq formatında veriler oluşturulmuşken, Nanopore'dan elde edilen \*.fast5 formatındaki veriler, sekans işlemine ait, pikoamper düzeyindeki sinyal değişikliklerine kadar bir çok metrik parametreyi de içermektedir. Bu veriler HDFView programıyla açılıp, şekil 4'teki şekilde içerik bilgisine erişilebilmektedir.



Name	Type
experiment_duration_set	String, length = 5, padding = H5T_STR_NULLTERM, cset = H5T_CSET_A
experiment_type	String, length = 12, padding = H5T_STR_NULLTERM, cset = H5T_CSET_A
flow_cell_product_code	String, length = 11, padding = H5T_STR_NULLTERM, cset = H5T_CSET_A
local_basecalling	String, length = 2, padding = H5T_STR_NULLTERM, cset = H5T_CSET_A
package	String, length = 7, padding = H5T_STR_NULLTERM, cset = H5T_CSET_A
package_version	String, length = 6, padding = H5T_STR_NULLTERM, cset = H5T_CSET_A
sample_frequency	String, length = 5, padding = H5T_STR_NULLTERM, cset = H5T_CSET_A
sequencing_kit	String, length = 11, padding = H5T_STR_NULLTERM, cset = H5T_CSET_A

Şekil 4. HDFView programıyla açılmış fast5 dosyası görüntüsü

#### Baz çağırma (Basecalling)

Nanopore özelinde bakıldığında, ard arda kaydedilen her bir pordaki pikoamper düzeyinde değişiklikler, baz çağırma işlemiyle nükleotidlerle eşleştirilmiştir. Bu işlem için birden çok sayıda program mevcuttur. Bu işlem için Scrappie, Guppy ve Flappie kullanılmıştır. Bu programların herbirinin parametreleri modifiye edilerek çok sayıda deneme yapılmış ve optimal sonucu en kısa sürede veren program-parametre kombinasyonu seçilmiştir. Bu işlem sonucunda \*.fastq formatında veriler elde edilmiştir.

#### Haritalama ve hizalama

Bu basamakta, nanopore sekanslama informatiğinde haritalamada en sık kullanılan 3 program seçilerek denemeler yapılmıştır. Bunlar, BWA, minimap2 ve graphMap programlarıdır. BWA ve minimap2 birbirine yakın algoritmalara sahip olup, lineer referansa göre haritalama yapsalar da, graphMap bunlardan farklı olarak grafik tabanlı bir referansa göre haritalama da yapabilmektedir. Grafik tabanlı haritalamanın özellikle uzun okumalarda, haritalama kalitesini artırarak hata oranını düşürdüğü bilinmektedir (Sović ve diğerleri, 2016). \*.fastq formatındaki veriler bu programlar aracılığıyla başta oluşturulan sanal referansa göre haritalanıp \*.sam formatında dosyalar elde edilmiştir. Ardından bu dosyalar SamTools programı ile \*.bam

formatına dönüştürülmüştür. Sonrasında \*.bam dosyası yine SamTools aracılığıyla indekslenip, \*.bam.bai formatında indeks dosyası elde edilmiştir.

Önceki basamakta oluşan \*.bam dosyası confirmasyon amacıyla IGV programıyla görüntülenmiştir. Bunun için öncelikle ilk basamakta oluşturulan referans dosyası, IGV'nin “create \*.genome file” seçeneği aracılığıyla programa tanıtılmıştır. Sonrasında ise kromozom arama boşluğuna her bir SNP'e ait rs kodları yazılarak şekil 5'teki gibi tek tek görüntülenmiştir.



Şekil 5. rs10495407 SNP'inin IGV programı görüntüsü

### Varyant çağırma

Bu basamakta, haritalanmış \*.bam formatındaki dosya, bcfTools, GATK ve freebayes programları kullanılarak \*.vcf formatındaki dosya elde edilmiştir. Genel olarak varyant çağırma işleminde yalnızca referans ile uyumsuz varyantlar elde edilmektedir. Ancak bizim çalışmamızda mutasyon tespiti değil, her bir SNP'e ait genotiplendirme bilgilendirmesi istenildiğinden programların parametreleri, referansla uyumlu da olsa uyumsuz da olsa tüm bilgileri verecek şekilde configure edilmiştir. Bunun için programların “pile up” modülleri kullanılmıştır. Bu sayede olabildiğince düşük kriterlere göre filtrelenen baz bilgilerini içeren \*.vcf dosyaları elde edilmiştir. Ayrıca

programların insersiyon ve delesyon analiziyle vakit kaybetmelerini önlemek için, yalnızca SNP analizine yönelik modifikasyonlar yapılmıştır.

### Script oluşturma

Bu aşamadan sonra elde edilen \*.vcf dosyasına, yalnızca her bir SNP'e ait, isim genotiplendirme bilgisi, okuma derinliği ve okuma yüzdeleri bilgisini ekrana getirecek şekilde Python dilinde özel olarak yazılmış bir kod uygulanmıştır.

Tüm bu işlemlerin sonunda doğru sonucu en hızlı sürede veren program kombinasyonu seçilmiştir. Bu programlar tek bir komutla çalışacak şekilde UNIX tabanlı platformlara uyumlu özel bir bash script yazılmıştır. Bu sayede programları tek tek çalıştırmakla vakit kaybedilmemiş olacak ve nihai sonucu hızlı bir şekilde ulaşılabacaktır.

## 4. Bulgular

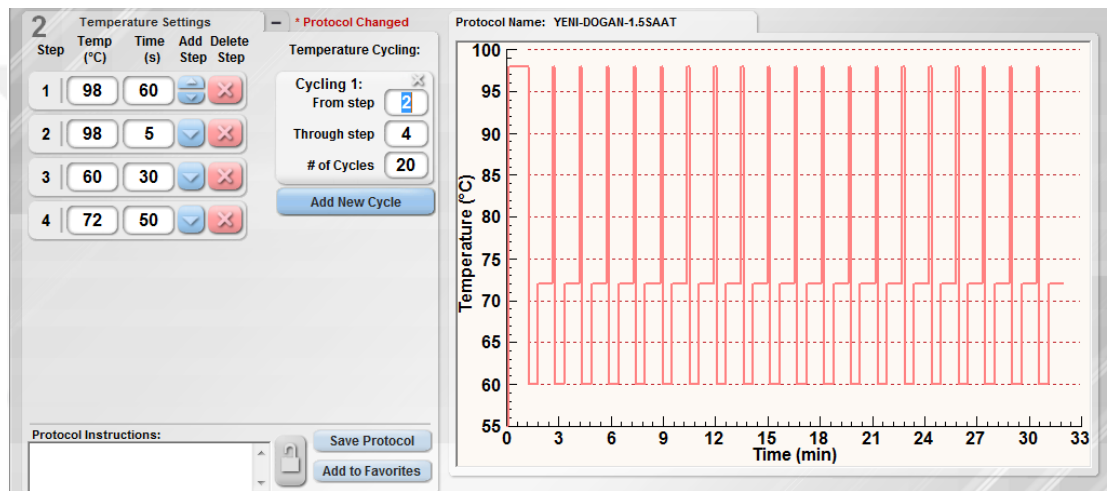
### 4.1. Yeni nesil dizi analizi bulguları

Üç farklı ticari DNA örneğini 52 SNP bölgesini için genotiplendirmek ve multipleks PCR başarısını ölçmek amacıyla NextSeq platformu ile gerçekleştirilen sekanslama işleminde, SNP'lerin herbiri tüm multipleks denemelerinde başarılı bir şekilde analiz edilmiştir. Tüm SNP'lerin okuma derinliklerinin 500'ün üzerinde olduğu görülmüştür. Üç farklı ticari DNA'nın genotiplendirmesi tablo 7'deki şekildedir.

**Tablo 7.** Üç farklı ticari DNA'nın 52 SNP için yeni nesil dizi analizi genotiplendirmesi

SNP	DNA-1	DNA-2	DNA-3	SNP	DNA-1	DNA-2	DNA-3
rs1005533	G/A	G/G	G/A	rs2111980	T/T	T/C	T/C
rs1015250	G/G	G/G	G/C	rs251934	G/G	A/G	A/A
rs1024116	C/C	C/T	C/T	rs2830795	A/G	A/G	A/G
rs1028528	A/A	A/G	A/G	rs2831700	A/A	A/G	A/A
rs1029047	T/A	A/A	T/T	rs354439	A/A	A/A	A/T
rs1031825	A/A	A/A	A/A	rs717302	A/A	A/A	G/A
rs10495407	G/A	G/G	G/G	rs719366	A/A	G/A	A/A
rs1335873	T/A	T/A	A/A	rs722098	A/A	A/G	A/G
rs1355366	T/C	T/T	T/T	rs727811	G/G	G/T	G/G
rs1357617	T/T	A/T	A/T	rs729172	G/G	G/T	G/T
rs1360288	C/C	C/T	C/T	rs733164	A/A	G/A	G/A
rs1382387	A/A	C/A	A/A	rs735155	C/C	C/C	C/T
rs1413212	T/C	C/C	T/C	rs737681	T/C	T/T	T/C
rs1454361	A/A	T/A	T/T	rs740910	A/G	A/G	A/G
rs1463729	C/C	C/C	C/T	rs763869	G/A	G/A	G/G
rs1490413	G/A	G/A	G/A	rs8037429	C/C	C/T	C/T
rs1493232	C/A	A/A	C/C	rs826472	C/C	T/C	C/C
rs1528460	T/T	C/T	T/T	rs873196	C/T	T/T	C/T
rs1886510	G/A	G/G	G/G	rs876724	C/T	C/C	C/T
rs1979255	C/G	C/G	C/G	rs891700	A/G	A/G	A/A
rs2016276	T/C	T/T	T/T	rs901398	T/T	T/T	C/T
rs2040411	G/A	G/A	A/A	rs907100	G/C	G/C	C/C
rs2046361	T/T	T/A	T/A	rs914165	A/A	G/A	G/A
rs2056277	C/T	C/C	C/T	rs917118	T/T	C/T	C/T
rs2076848	A/T	A/T	A/A	rs938283	T/T	T/T	T/T

Yeni nesil dizi analizin sonuçlarının, hem standart hem süper hızlı PCR cihazı ile gerçekleştiren tüm multipleks PCR denemelerinde tam amplifikasyon başarısını doğrulamasıyla, sonraki aşamalara en kısa sürede PCR'ı tamamlanan ampliconlar seçilerek devam edilmiştir. Bu PCR, Philisa cihazı kullanılarak toplam 33 dakikada tamamlanmış olup, cihaza özgü yazılım aracılığıyla oluşturulan döngü sayısı ve koşullar şekil 6'daki gibidir.



Şekil 6. Philisa cihazına ait program ile oluşturulan PCR süresi ve koşullarının görüntüsü

#### 4.2. Nanopore sekanslama bulguları

Oxford Nanopore Minion cihazı ile gerçekleştirilen sekans işlemi başladıktan hemen sonra 5. dakika, 10. dakika ve 15. dakikalarda oluşan ham datalar ayrı ayrı analiz edilmiştir. DNA-1'e ait 5., 10., ve 15. dakika ham datalarının metrik değerleri tablo 8'deki gibidir.

**Tablo 8.** DNA-1'e ait nanopore sekanslama verilerinin metrik deęerleri

Süre	Basecalling Filtresi	Okuma Sayısı	Okunan baz sayısı	Okuma uzunluęu ortalaması	Okuma uzunluęu N50	Ortalama kalite skoru
5 dakika	Toplam	12.000	5.101.403	269	363	8,58
	Filtrelenmiř	10.214	4.285.634	270	342	8,82
10 dakika	Toplam	24.000	10.408.440	270	402	8,55
	Filtrelenmiř	20.423	8.746.153	270	369	8,79
15 dakika	Toplam	40.000	17.796.360	271	465	8,55
	Filtrelenmiř	33.911	14.893.268	272	423	8,79

DNA-1'e ait 5., 10., ve 15. dakikalara ait genotiplendirme ve her bir SNP için okuma derinlięi sonuçları sırasıyla tablo 9, tablo 10 ve tablo 11'deki gibidir.

**Tablo 9.** DNA-1'e ait nanopore sekanslama 5. dakika verilerinin genotiplendirme okuma derinlięi bulguları

SNP	Okuma Derinlięi	Genotip	SNP	Okuma Derinlięi	Genotip
rs1005533	66	G/A (%59/%41)	rs2111980	130	T/T (%88)
rs1015250	139	G/G (%98)	rs251934	146	G/G (%87)
rs1024116	70	C/C (%91)	rs2830795	158	A/G (%58/%42)
rs1028528	93	A/A (%91)	rs2831700	161	A/A (%94)
rs1029047	51	T/A (%57/%42)	rs354439	168	A/A (%100)
rs1031825	215	A/A (%86)	rs717302	128	A/A (%87)
rs10495407	150	G/A (%62/%37)	rs719366	61	A/A (%92)
rs1335873	174	T/A (%55/%44)	rs722098	159	A/A (%97)
rs1355366	77	T/C (%33/%66)	rs727811	162	G/G (%96)
rs1357617	41	T/T (%97)	rs729172	158	G/G (%93)
rs1360288	85	C/C (%89)	rs733164	83	A/A (%94)
rs1382387	88	A/A (%96)	rs735155	110	C/C (%99)
rs1413212	220	T/C (%51/%49)	rs737681	153	T/C (%60/%39)
rs1454361	141	A/A (%95)	rs740910	267	A/G (%48/%52)
rs1463729	136	C/C (%87)	rs763869	182	G/A (%61/%38)
rs1490413	48	G/A (%64/%35)	rs8037429	71	C/C (%86)
rs1493232	155	C/A (%33/%66)	rs826472	98	C/C (%88)
rs1528460	178	T/T (%95)	rs873196	133	C/T (%36/%64)
rs1886510	232	G/A (%58/%42)	rs876724	53	C/T (%63/%36)

rs1979255	128	C/G (%62/%38)	rs891700	135	A/G (%52/%47)
rs2016276	321	T/C (%64/%35)	rs901398	139	T/T (%87)
rs2040411	217	G/A (%46/%53)	rs907100	279	G/C (%56/%44)
rs2046361	47	T/T (%86)	rs914165	131	A/A (%90)
rs2056277	232	C/T (%50/%50)	rs917118	145	T/T (%98)
rs2076848	196	A/T (%47/%53)	rs938283	100	T/T (%94)
rs2107612	78	A/A (%92)	rs964681	100	C/C (%87)

**Tablo 10.** DNA-1'e ait nanopore sekanslama 10. dakika verilerinin genotiplendirme okuma derinliği bulguları

SNP	Okuma Derinliği	Genotip	SNP	Okuma Derinliği	Genotip
rs1005533	126	G/A (%59/%41)	rs2111980	257	T/T (%88)
rs1015250	281	G/G (%98)	rs251934	276	G/G (%87)
rs1024116	145	C/C (%90)	rs2830795	320	A/G (%58/%41)
rs1028528	192	A/A (%95)	rs2831700	295	A/A (%96)
rs1029047	104	T/A (%60/%40)	rs354439	323	A/A (%99)
rs1031825	429	A/A (%91)	rs717302	278	A/A (%89)
rs10495407	318	G/A (%61/%38)	rs719366	119	A/A (%92)
rs1335873	347	T/A (%56/%44)	rs722098	314	A/A (%97)
rs1355366	155	T/C (%37/%63)	rs727811	316	G/G (%96)
rs1357617	85	T/T (%97)	rs729172	355	G/G (%94)
rs1360288	180	C/C (%94)	rs733164	152	A/A (%95)
rs1382387	185	A/A (%95)	rs735155	207	C/C (%99)
rs1413212	412	T/C (%54/%45)	rs737681	300	T/C (%59/%40)
rs1454361	259	A/A (%93)	rs740910	535	A/G (%47/%52)
rs1463729	287	C/C (%90)	rs763869	343	G/A (%60/%40)
rs1490413	122	G/A (%62/%37)	rs8037429	144	C/C (%86)
rs1493232	311	C/A (%43/%56)	rs826472	186	C/C (%88)
rs1528460	324	T/T (%95)	rs873196	292	C/T (%36/%64)
rs1886510	454	G/A (%55/%44)	rs876724	107	C/T (%62/%37)
rs1979255	250	C/G (%63/%35)	rs891700	306	A/G (%54/%45)
rs2016276	605	T/C (%62/%37)	rs901398	303	T/T (%88)
rs2040411	424	G/A (%46/%53)	rs907100	537	G/C (%56/%44)
rs2046361	95	T/T (%90)	rs914165	269	A/A (%90)
rs2056277	484	C/T (%52/%47)	rs917118	292	T/T (%98)
rs2076848	367	A/T (%50/%50)	rs938283	190	T/T (%96)
rs2107612	161	A/A (%91)	rs964681	201	C/C (%88)

**Tablo 11.** DNA-1'e ait nanopore sekanslama 15. dakika verilerinin genotiplendirme okuma derinliği bulguları

SNP	Okuma Derinliği	Genotip	SNP	Okuma Derinliği	Genotip
rs1005533	208	G/A (%59/%40)	rs2111980	427	T/T (%88)
rs1015250	464	G/G (%96)	rs251934	490	G/G (%87)
rs1024116	245	C/C (%90)	rs2830795	542	A/G (%57/%42)
rs1028528	315	A/A (%94)	rs2831700	472	A/A (%96)
rs1029047	187	T/A (%54/%45)	rs354439	572	A/A (%98)
rs1031825	732	A/A (%93)	rs717302	468	A/A (%89)
rs10495407	532	G/A (%60/%39)	rs719366	189	A/A (%89)
rs1335873	557	T/A (%55/%44)	rs722098	493	A/A (%97)
rs1355366	281	T/C (%39/%61)	rs727811	501	G/G (%94)
rs1357617	129	T/T (%96)	rs729172	574	G/G (%93)
rs1360288	285	C/C (%96)	rs733164	249	A/A (%92)
rs1382387	287	A/A (%96)	rs735155	360	C/C (%97)
rs1413212	668	T/C (%52/%47)	rs737681	505	T/C (%61/%38)
rs1454361	418	A/A (%93)	rs740910	881	A/G (%48/%52)
rs1463729	438	C/C (%91)	rs763869	572	G/A (%60/%40)
rs1490413	187	G/A (%59/%40)	rs8037429	227	C/C (%88)
rs1493232	545	C/A (%42/%57)	rs826472	317	C/C (%88)
rs1528460	538	T/T (%97)	rs873196	481	C/T (%37/%63)
rs1886510	745	G/A (%56/%43)	rs876724	200	C/T (%62/%38)
rs1979255	444	C/G (%61/%38)	rs891700	487	A/G (%55/%44)
rs2016276	954	T/C (%61/%38)	rs901398	504	T/T (%90)
rs2040411	687	G/A (%46/%53)	rs907100	856	G/C (%57/%43)
rs2046361	143	T/T (%93)	rs914165	458	A/A (%92)
rs2056277	824	C/T (%53/%46)	rs917118	489	T/T (%97)
rs2076848	610	A/T (%50/%49)	rs938283	341	T/T (%94)
rs2107612	221	A/A (%95)	rs964681	327	C/C (%91)

Nanopore sekanslama sonuçları, yeni nesil dizi analizi basamağında ortaya çıkan DNA-1'e ait genotiplendirme sonuçlarıyla karşılaştırıldığında, 5., 10. ve 15. dakika sonunda elde edilen verilerin tümünün %100'lük bir başarıyla genotiplendirmeyi yaptığını göstermiştir. DNA-1'e ait bu sonuçlar 5. dakika sonunda ortaya çıkan ham datanın bile genotiplendirme için son derece yeterli olduğunu göstermektedir. DNA-2

ve DNA-3 ile yapılan nanopore sekanslama çalışmalarında da benzer şekilde 5. Dakika sonunda tam doğrulukla genotiplendirmelerin yapıldığı gösterilmiştir. DNA-2 ve DNA-3'ün 5. dakika nanopore sekanslama verileri sırasıyla tablo 12 ve tablo 13'teki şekildedir.

**Tablo 12.** DNA-2'ye ait nanopore sekanslama 5. dakika verilerinin genotiplendirme okuma derinliği bulguları

SNP	Okuma Derinliği	Genotip	SNP	Okuma Derinliği	Genotip
rs1005533	103	G/G (%95)	rs2111980	194	T/C (%52/%48)
rs1015250	188	G/G (%89)	rs251934	126	A/G (%39/%59)
rs1024116	133	C/T (%49/%41)	rs2830795	151	A/G (%54/%46)
rs1028528	169	A/G (%55/%44)	rs2831700	127	A/G (%60/%39)
rs1029047	113	A/A (%94)	rs354439	53	A/A (%91)
rs1031825	203	A/A (%97)	rs717302	169	A/A (%96)
rs10495407	68	G/G (%87)	rs719366	199	G/A (%52/%47)
rs1335873	118	T/A (%61/%39)	rs722098	140	A/G (%63/%36)
rs1355366	166	T/T (%97)	rs727811	135	G/T (%47/%52)
rs1357617	105	A/T (%36/%63)	rs729172	201	G/T (%51/%49)
rs1360288	194	C/T (%58/%41)	rs733164	174	G/A (%56/%43)
rs1382387	186	C/A (%57/%42)	rs735155	109	C/C (%87)
rs1413212	136	C/C (%97)	rs737681	164	T/T (%93)
rs1454361	92	T/A (%50/%49)	rs740910	181	A/G (%38/%61)
rs1463729	122	C/C (%90)	rs763869	318	G/A (%53/%46)
rs1490413	142	G/A (%62/%37)	rs8037429	102	C/T (%45/%54)
rs1493232	274	A/A (%93)	rs826472	142	T/C (%54/%43)
rs1528460	163	C/T (%46/%54)	rs873196	116	T/T (%92)
rs1886510	172	G/G (%97)	rs876724	122	C/C (%88)
rs1979255	133	C/G (%59/%40)	rs891700	187	A/G (%61/%38)
rs2016276	206	T/T (%64/%35)	rs901398	110	T/T (%90)
rs2040411	170	G/A (%48/%51)	rs907100	169	G/C (%59/%41)
rs2046361	154	T/A (%61/%39)	rs914165	121	G/A (%49/%50)
rs2056277	249	C/C (%93)	rs917118	122	C/T (%60/%40)
rs2076848	111	A/T (%58/%41)	rs938283	146	T/T (%95)
rs2107612	146	A/A (%89)	rs964681	104	T/C (%57/%41)

**Tablo 13.** DNA-3'e ait nanopore sekanslama 5. dakika verilerinin genotiplendirme okuma derinliği bulguları

SNP	Okuma Derinliği	Genotip	SNP	Okuma Derinliği	Genotip
rs1005533	168	G/A (%43/%56)	rs2111980	158	T/C (%63/%36)
rs1015250	132	G/C (%54/%45)	rs251934	131	A/A (%89)
rs1024116	152	C/T (%38/%61)	rs2830795	110	A/G (%48/%51)
rs1028528	103	A/G (%41/%58)	rs2831700	142	A/A (%96)
rs1029047	221	T/T (%87)	rs354439	149	A/T (%53/%46)
rs1031825	144	A/A (%93)	rs717302	208	G/A (%49/%50)
rs10495407	110	G/G (%99)	rs719366	103	A/A (%97)
rs1335873	92	A/A (%90)	rs722098	127	A/G (%61/%39)
rs1355366	133	T/T (%93)	rs727811	199	G/G (%88)
rs1357617	177	A/T (%51/%49)	rs729172	164	G/T (%89)
rs1360288	101	C/T (%60/%39)	rs733164	185	G/A (%52/%47)
rs1382387	174	A/A (%92)	rs735155	116	C/T (%59/%40)
rs1413212	153	T/C (%62/%38)	rs737681	137	T/C (%44/%56)
rs1454361	170	T/T (%88)	rs740910	218	A/G (%50/%49)
rs1463729	128	C/T (%46/%54)	rs763869	179	G/G (%91)
rs1490413	188	G/A (%47/%52)	rs8037429	111	C/T (%64/%35)
rs1493232	125	C/C (%91)	rs826472	120	C/C (%94)
rs1528460	140	T/T (%90)	rs873196	102	C/T (%45/%54)
rs1886510	101	G/G (%95)	rs876724	130	C/T (%57/%43)
rs1979255	156	C/G (%52/%48)	rs891700	124	A/A (%98)
rs2016276	149	T/T (%56/%43)	rs901398	117	C/T (%53/%46)
rs2040411	53	A/A (%94)	rs907100	231	C/C (%88)
rs2046361	113	T/A (%61/%38)	rs914165	141	G/A (%47/%51)
rs2056277	279	C/T (%54/%44)	rs917118	103	C/T (%57/%42)
rs2076848	79	A/A (%86)	rs938283	86	T/T (%92)
rs2107612	166	A/G (%58/%42)	rs964681	130	T/C (%52/%47)

Üç farklı DNA örneğine ait, 5., 10. ve 15. dakika sonunda elde edilen verilerde en düşük ve ortalama okuma derinliği verileri tablo 14'te verilmiştir.

**Tablo 14.** Üç farklı DNA'ya ait nanopore sekanslama 5., 10. ve 15. Dakika verilerindeki en düşük ve ortalama okuma derinliği sayıları

		5. dakika	10. dakika	15. dakika
<b>DNA-1</b>	En düşük okuma derinliği	41	85	129
	Ortalama okuma derinliği	149	279	476
<b>DNA-2</b>	En düşük okuma derinliği	68	157	245
	Ortalama okuma derinliği	152	308	483
<b>DNA-3</b>	En düşük okuma derinliği	53	114	189
	Ortalama okuma derinliği	143	299	470

Bu sonuçlar, Oxford Nanopore Minion cihazının 5 dakikalık sekans verilerinin 52 SNP bölgesinin analizi için yeterli olduğunu göstermektedir.

### 4.3. Biyoinformatik iş akışı bulguları

Çalışmamızda farklı basamaklarda çeşitli biyoinformatik araç ve programlar, parametreleri modifiye edilerek çok sayıda deneme yapılmıştır. Bu denemeler sonucunda optimal sonucu en kısa sürede veren programların kombinasyonu seçilmiştir.

#### 4.3.1. Basecalling bulguları

Üç farklı programla çok sayıda basecalling denemesi yapılmıştır. En ideal program seçilirken süre, okuma sayısı ve doğru sonuç verebilme yeteneği değerlendirilmiştir. Üç programın da standart parametreler kullanılarak verdiği genotiplendirme sonuçların aynı olduğu ve yeni nesil dizi analizi verileriyle tam olarak örtüştüğü görülmüştür. Ancak süreler arasında anlamlı farklılık tespit edilmiştir. Tüm denemelerde en hızlı ve en yüksek okuma sayısı veren programın Oxford Nanopore Technologies ekibinin yazdığı Guppy programı olduğu görülmüştür. Bundan dolayı, sonraki süre denemelerine bu programla devam edilmiştir. Bu programa özel olarak

atanan bir \*.cfg dosyasıyla tablo 15'deki şekilde yaklaşık 6 kat daha hızlı sonuç verdiği ce okuma sayıları bir miktar düşse de genotiplendirme sonuçlarının etkilenmediği görülmüştür.

**Tablo 15.** DNA-1'e ait nanopore sekaslama 5., 10. ve 15. Dakika verilerinin Guppy programıyla basecalling işlemi süreleri

<b>Basecalling süreleri</b>			
	<b>5. dakika verisi</b>	<b>10. dakika verisi</b>	<b>15. dakika verisi</b>
<b>Standart</b>	13 dakika	32 dakika	51 dakika
	58 saniye	8 saniye	22 saniye
<b>Hızlı modda</b>	2 dakika	5 dakika	8 dakika
	19 saniye	12 saniye	14 saniye

#### **4.3.2. Haritalama ve hizalama bulguları**

Bu basamakta kullanılan 3 program da, yaklaşık 10 kb (52SNP x 201 baz) nükleotid içeren çalışmamıza özgün referans sayesinde beklenilenden çok daha hızlı sonuçlar vermişlerdir. Okuma derinliği ve genotiplendirme sonuçları arasında da hiçbir fark bulunamamıştır. Bu sebeplerden dolayı haritalama basamağı için yeni nesil dizi analizi informatiğinde en bilinen ve köklü uygulama olan BWA programı seçilmiştir.

Hizalamada ise son derece hızlı ve etkili sonuç veren SamTools kullanılmıştır. Haritalama ve hizalamanın toplam süresi, 5. dakika, 10. dakika ve 15 dakika verileri için son derece kısa olup, 2 saniyenin altında olarak bulunmuştur.

#### **4.3.3. Varyant çağırma bulguları**

Bu basamakta denenilen 3 farklı program ve süreleri arasında anlamlı farklılıklar bulunmuştur. FreeBayes ve GATK parametre modifikasyonlarına rağmen, bcfTools'un modifikasyonsuz halinden daha uzun sürede sonuç verdikleri için esas olarak bcfTools seçilerek devam edilmiştir. İlk etapta karşılaşılan süreler beklenilenin çok üstünde olmuştur. Bundan dolayı programın tüm değiştirilebilir konfigürasyonları detaylıca incelenmiştir. Programın manfeste göre analiz seçeneği görülünce, yalnızca SNP'in olduğu noktaya dair varyant bilgisini elde etmek amacıyla \*.bed formatında

şekil 7'deki manifest dosyası oluşturulmuştur. Bu manifest sayesinde program \*.bam dosyasındaki tüm nükleotidlere ait varyant bilgilerini analiz etmekle vakit kaybetmeyerek, yalnızca SNP'in olduğu 101. baza ait bilgileri çıkaracaklardır. 101. baz seçilmesinin nedeni, başta oluşturulan referansta her bir SNP'in 101. baz pozisyonunda yer almasıdır. Bu işlem sonrasında sürelerde dramatik bir düşüş yaşanmıştır. Ancak yine de, farklı parametreleri modifiye ederek yapılan çeşitli denemeler sonucunda 2 saniyenin altına düşülebilmektedir. Bu süreler tablo 16'da verilmiştir.

```
rs1490413 100 101 SNP_1
rs876724 100 101 SNP_2
rs1357617 100 101 SNP_3
rs2046361 100 101 SNP_4
rs717302 100 101 SNP_5
rs1029047 100 101 SNP_6
rs917118 100 101 SNP_7
rs763869 100 101 SNP_8
rs1015250 100 101 SNP_9
rs735155 100 101 SNP_10
```

Şekil 7. Varyant çağırma programları için oluşturulan manifest dosyası görüntüsü

**Tablo 16.** DNA-1'e ait nanopore sekanslama verilerinin bcfTools programının farklı modifikasyonlarıyla varyant çağırma süreleri

<b>bcfTools ile varyant çağırma süreleri</b>			
	<b>5. dakika verisi</b>	<b>10. dakika verisi</b>	<b>15. dakika verisi</b>
<b>Standart</b>	1 dakika	2 dakika	4 dakika
	58 saniye	46 saniye	8 saniye
<b>Manifestli</b>	19 saniye	27 saniye	41 saniye
<b>Manifest + Modifikasyon</b>	<2 saniye	<2 saniye	<2 saniye

Son olarak seçilen programlar için tek bir komutla çalışabilen ve içinde Python diliyle yazılmış olan özel kodları da barındıran şekil 8 ve şekil 9'daki bash script yazılmıştır.

```

#!/usr/local/bin/bash
#Bilcag Akgun - Thesis Script

VERSION=1.0

#####

read -p "Input directory:" selected_dir
read -p "Output directory:" output_dir

START_DATE=`date`

echo "Job started at: $START_DATE"

SECONDS=0

/Users/bilcag/GitHub/ont-guppy-cpu/bin/guppy_basecaller -c
dna_r9.4.1_450bps_fast.cfg \
--cpu_threads_per_caller 8 -q 0 --input_path $selected_dir/* --save_path
$output_dir -r

cd $output_dir

cat *.fastq > merged.fastq

cat $output_dir/merged.fastq | bwa mem -t 8 -a
/Users/bilcag/Dropbox/Tez/bin/SNP_52.fasta - | samtools sort -O BAM -@4 -o
$output_dir/52plex_sorted.bam -

samtools index $output_dir/52plex_sorted.bam

bcftools mpileup -d 100000 --thread 8 -B -q 0 -Q 0 -R
/Users/bilcag/Dropbox/Tez/bin/SNPforID.bed -f
/Users/bilcag/Dropbox/Tez/bin/SNP_52.fasta $output_dir/52plex_sorted.bam |
bcftools call -V indels -m - > $output_dir/52plex.vcf

python <<END_OF_PYTHON
#!/anaconda2/bin/python python

import tkinter
from tkinter import filedialog

outDir = filedialog.askdirectory(initialdir="/", title='Please select a
directory')

snpData = {}

with open(outDir+'/52plex.vcf') as f:
    for l in f:
        if l.startswith('#'):
            continue

            snp, pos, id, ref, alt, qual, filter, info, d, dd = l.split()

```

**Şekil 8.** Tüm programların tek bir komutla çalışmasını sağlayacak özel olarak yazılmış bash script'in ilk bölümü

```

        if int(pos) != 101:
            continue

        par = {}
        for p in info.split(';'):
            pv = p.split('=')
            par[pv[0]] = pv[1]

        snpData[snp] = {'pos': pos, 'ref': ref, 'alt': alt, 'qual': qual,
            'filter': filter, 'info': par}

print('Got data for {} SNPs:'.format(len(snpData)))

with open(outDir + '/ttt_profile.csv', 'w') as f:
    f.write('snp, coverage, ref_allele, ref_percent, alt_allele,
        alt_percent, genotype\n')

    for s in sorted(snpData.keys()):
        totalDepth = int(snpData[s]['info']['DP'])
        depthList = [int(d) for d in snpData[s]['info']['DP4'].split(',')]
        refDepth = sum(depthList[0:2])
        altDepth = sum(depthList[2:4])

        if refDepth > altDepth and altDepth/refDepth < 0.1:
            genotype = snpData[s]['ref'] + snpData[s]['ref']
        elif altDepth > refDepth and refDepth/altDepth < 0.1:
            genotype = snpData[s]['alt'] + snpData[s]['alt']
        else:
            genotype = snpData[s]['ref'] + snpData[s]['alt']

        if snpData[s]['alt'] == '.':
            f.write(','.join([s, str(totalDepth), snpData[s]['ref'],
                '{:.1f}'.format(100*refDepth/totalDepth), ',', ',',
                snpData[s]['ref']+snpData[s]['ref']] + '\n')
            print(' {} ({} ) {} ( {:.1f} % )'.format(s, totalDepth,
                snpData[s]['ref'], 100*refDepth/totalDepth))
        else:
            f.write(','.join([s, str(totalDepth), snpData[s]['ref'],
                '{:.1f}'.format(100*refDepth/totalDepth), snpData[s]['alt'],
                '{:.1f}'.format(100*altDepth/totalDepth), genotype] + '\n')
            print(' {} ({} ) {} ( {:.1f} % ) {} ( {:.1f} % )'.format(s,
                totalDepth, snpData[s]['ref'], 100*refDepth/totalDepth,
                snpData[s]['alt'], 100*altDepth/totalDepth))

END_OF_PYTHON

END_DATE=`date`

echo "Job ending at: $END_DATE"

ELAPSED="Elapsed: $((($SECONDS / 3600))hrs $((($SECONDS / 60) % 60))min
    $((($SECONDS % 60))sec"

echo "$ELAPSED"

```

**Şekil 9.** Tüm programların tek bir komutla çalışmasını sağlayacak özel olarak yazılmış bash script'in ikinci bölümü

Bu özel scriptle, elimizdeki tüm ham veriler için informatik işlemler baştan tekrar edilmiş ve tablo 15'teki süreler elde edilmiştir. Bu süreler çalışmamızın biyoinformatik analiz basamağında elde edilen en kısa sürelerdir.

**Tablo 17.** Yazılan scriptle, tüm nanopore sekans verilerinin tekrar değerlendirilmesi sonucu elde edilen süreler

<b>Biyoinformatik analiz süresi</b>			
	<b>DNA-1</b>	<b>DNA-2</b>	<b>DNA-3</b>
<b>5. dakika</b>	2 dakika	2 dakika	2 dakika
<b>ham verisi</b>	27 saniye	44 saniye	16 saniye
<b>10. dakika</b>	5 dakika	5 dakika	5 dakika
<b>ham verisi</b>	13 saniye	34 saniye	4 saniye
<b>15. dakika</b>	8 dakika	8 dakika	8 dakika
<b>ham verisi</b>	27 saniye	58 saniye	11 saniye

Tüm sonuçlar bütün olarak değerlendirildiğinde; multipleks PCR, kütüphane hazırlığı, nanopore sekanslama ve biyoinformatik analizin yaklaşık olarak 53 dakikada tamamlandığı ve üç farklı DNA için genotiplendirme sonuçlarının yeni nesil dizi analizi sonuçlarıyla tam olarak örtüştüğü görülmektedir.

## 5. Tartışma

Kiři tanımlama, DNA üzerindeki yüksek polimorfik alanların farklı yöntemlerle analiz edilip deęerlendirilmesiyle, iki bireyi dünyadaki en yüksek duyarlılıkla birbirinden ayrılmasıdır. Adli ya da tıbbi perspektifin dıřında, yüksek güvenlik önlemi gerektiren her durumda uygulanıp, sonuç verebilme potansiyeli taşıyan bir moleküler laboratuvar uygulamasıdır. Kiři tanımlama için halihazırda uygulanan yöntemlerin en büyük handikapı büyük laboratuvarlara, pahalı ekipmanlara ihtiyaç duyulması ve yeterince hızlı sonuç verilememesidir. Çalışmamızda bu dezavantajlara büyük ölçüde çözüm sunacak bir prototip geliştirilmiştir. Bu prototipin temelini oluşturan sekanslama teknolojisi ve seçilen polimorfik bölgeler farklı açılardan tartışılacaktır.

STR'ler özellikle 1990'lardan itibaren, yüksek derecede polimorfik yapıları nedeniyle kiři tanımlama için tercih edilen markerlardır (Jobling ve Gill, 2004). Kapiller elektroforez temelli STR analizi kiři tanımlamada Dünya çapında en sık kullanılan yöntem olsa da, SNP temelli DNA profillemesi son on yılda giderek önem kazanmaya başlamıştır. İnsan genom projesinin tamamlanmasından sonra SNP'lerle ilgili veritabanları oluşmaya başlamıştır (Belmont ve dięerleri, 2005). Aynı zamanda bir dönem yaygın olarak yapılan GWAS çalışmaları hem SNP'ler ile ilgili bu verileri büyütmiş hem de popülasyondaki SNP dağılımıyla ilgili fikir sahibi olmamızı sağlamıştır (Choudhury ve dięerleri, 2014). SNP'ler STR'lara kıyasla hem düşük mutasyon oranına sahiptirler hem de biallelik yapılarından dolayı bir çok otomatize sekans platformunda oldukça kolay bir şekilde sekanslanabilirler (Børsting, Mogensen ve Morling, 2013). SNP'ler aynı zamanda atasal coęrafya hakkında da bilgiler vermektedir bu açıdan STR'lardan çok daha esnektir (Q. Wang ve dięerleri, 2016). Bu bilgiler doęrultusunda çalışmamızda dünya üzerinde hala altın standart olan STR temelli genotiplendirmedense, daha az yaygın ancak gelecek vadeden SNP temelli bir yöntem seçilmiştir. Geliştirilen prototip sadece SNP genotiplendirmesiyle kalmayıp çok daha bilgi verici olabilecektir. Dięer bir açıdan, SNP genotiplendirmesi sadece adli tıp alanında deęil, tıbbın dięer birçok alanında da aktif olarak kullanılmaktadır. Bunlar başlıca kanser, farmakogenetik, kardiyovasküler yatkınlıklar gibi alanlardır (Kim ve Misra, 2007). Ayrıca talasemi, kistik fibrozis gibi tek gen hastalıklarının tanısında da ilk basamak test olarak, sık görülen tek nükleotid mutasyonları reverse dot blot yöntemiyle incelenmektedir. Çalışmamızın bulgularına bakıldığında 2 part

halinde multipleks olarak amplifiye edilen 52 SNP başarılı bir şekilde genotiplendirilmiştir. Bu prototip yalnızca kişi tanımlamada değil, primer değişiklikleriyle spesifik tek nükleotid değişikliklerinin incelendiği tüm alanlarda uygulanabilir bir potansiyel taşımaktadır.

SNP'lerin analizi için çeşitli birçok yöntem geliştirilmiştir. Bunlar, enzim kesimi, hibridasyon, primer uzaması, SNaPShot, kütle spektrometresi ve microarray gibi çeşitli yöntemlerdir (Cho ve diğerleri, 2014). Bu yöntemler genellikle tek seferde az sayıda SNP'i analiz edebilmekte ya da microarray gibi SNP sayısı arttığında ise yüksek maliyet ve laboratuvar iş gücüyle karşımıza çıkmaktadır. Güncel bir teknoloji olan yeni nesil dizi analizi teknolojileri yüksek verimlilik ve örnek başı düşük maliyet gibi avantajlarıyla, SNP'in genotiplendirmesinde önceki yöntemlerin dezavantajlarını ortadan kaldırmıştır (Weber-Lehmann ve diğerleri, 2014). Ancak yeni nesil dizi analizinin maliyet etkin bir şekilde kullanılabilmesi için örnek sayısını yüksek tutmak gerekmektedir. Bu da küçük laboratuvarlarda bu teknolojiyi kullanışsız kılmaktadır. Oxford Nanopore teknolojisi en güncel sekans teknolojisi olup, görece düşük bir maliyetle, laboratuvarlara ihtiyaç duymaksızın dizi analizi yapabilmektedir (Ip ve diğerleri, 2015). Teknolojinin satışa sunulduğu son birkaç yıldan bu yana en hızlı büyüyen sekanslama teknolojisi olarak karşımıza çıkmaktadır (Miten Jain, Olsen, Paten ve Akeson, 2016). Çalışmamızda bütün bu sebeplerden dolayı Oxford Nanopore teknolojisi tercih edilmiştir. Kullandığımız flowcell'ler özel bir protokolle yıkanarak iki kez daha kullanma imkanı sunmasına rağmen, toplam maliyet açısından STR genotiplendirmesinin hala gerisindedir. Bununla birlikte, Oxford Nanopore Technologies yakın zamanda tek kullanımlık, daha küçük kapasiteli ve çok daha uygun fiyatlı Flongle isimli flowcell'leri satışa sunacağını duyurmuştur. Bu durum da, çok yakın gelecekte maliyet olarak da STR genotiplendirmesinin önüne geçilebileceğini öngördürtmektedir.

Oxford Nanopore teknolojisi esasen mikrobiyoloji çalışmalarında aktif ve efektif olarak kullanılmaktadır (Tyler ve diğerleri, 2018). Bunun sebebi de teknoloji ilk çıktığında yüksek hata oranının tür belirleme ya da genome-assembly çalışmalarını etkileyecek ölçüde sorun yaratmasa da, söz konusu insan genomu olunca klinik rapor vermede yetersiz kalmasıdır. Ancak geçtiğimiz yıllar içerisinde bu hata oranı hem geliştirilen yazılımlarla hem de kitteki donanımsal modifikasyonlarla giderek düşürülmüştür (M. Jain ve diğerleri, 2017; Rang, Kloosterman ve de Ridder, 2018).

Şu ana denk insan genomu kullanılarak yapılmış onlarca çalışma mevcuttur. Bunlar Y kromozomu sekanslaması, *GBA* geni dizi analizi, hemofili hastalığında moleküler tanı, genomda metilasyon saptanması gibi güncel çalışmalardır (Chatron ve diğerleri, 2019; Kuderna ve diğerleri, 2019; Leija-Salazar ve diğerleri, 2019; Simpson ve diğerleri, 2017). Aynı zamanda bu teknoloji kullanılarak yapılan iki adet kişi tanımlama çalışması mevcuttur. Bunlardan ilki bizim çalışmamızda da olduğu gibi 52 SNP'lik bir yöntemsel çalışma diğeri ise triallelik genotiplendirme çalışmasıdır (Cornelis ve diğerleri, 2019; Cornelis, Gansemans, Deleye, Deforce ve Van Nieuwerburgh, 2017). Bu çalışmaların ilkinde, yaklaşık 50 bazlık amplikonlar ligazla randomize şekilde uç uca eklenerek daha uzun ve nanopore sekanslama teknolojisi açısından daha makul olarak kabul edilebilecek amplikon uzunluğu elde edilmiş sonrasında sekanslanmıştır. Ancak çalışmamızda buna gerek duyulmamıştır çünkü çalışmamızın en büyük farkı hıza ve sahada uygulanabilirliğe odaklanmasıdır. Çalışmamızda SQK-RAD004 kodlu hızlı kütüphane hazırlık kiti kullanılmıştır. Bu hızlı kütüphane hazırlık kiti, yalnızca 30kb ve üstü amplikonlar için önerilmektedir. Bunu sebebi de bu hızlı kütüphane hazırlık kiti, amplikonları transpozaz'lar aracılığıyla motor proteine bağladığı için ancak 30kb ve üstündeki fragmanları efektif bir şekilde motor proteine bağlanabileceği öngörülmektedir (Mitsuhashi ve diğerleri, 2017). Çalışmamızın yürütücülerinden Prof. Dr. Hüseyin Onay ve ekibinin gerçekleştirdiği farklı projelerde küçük amplikonların, doğru konsantrasyonda kullanılarak bu hızlı kütüphane hazırlık kitiyle dahi başarılı bir şekilde sekanslanabildiği ispatlanmıştır. Çalışmamız bu açıdan diğer iki çalışmadan ayrılmaktadır.

Oxford Nanopore Minion cihazını son derece gelişmiş bir ampermetre olarak tanımlamak yanlış olmayacaktır. Her bir pordan geçen nükleotide göre pikoamper seviyesinde akım farklılıklar gelmekte ve bunlar da ham verilerde yer almaktadır (Y. Wang ve diğerleri, 2014). Biyoinformatik işlemlerde ilk basamak olarak basecalling adı verilen pikoamper düzeyindeki farklılıkları barından \*.fast5 formatındaki bu ham verileri, nükleotidlerle eşleştirmek gerekmektedir. Bununla ilgili hem Oxford Nanopore Technologies şirketi hem de biyoinformatikçiler ve yazılımcılar şu ana dek çeşitli programlar geliştirmişlerdir. Wick ve arkadaşlarının son derece güncel olan yayınında Oxford Nanopore Technologies tarafından geliştirilen 4 farklı basecalling programı detaylıca karşılaştırılmıştır (Wick, Judd ve Holt, 2019). Bu programlar içerisinde hem performans hem de yüksek doğruluk bakımından en idealinin Guppy

olduğu belirtilmiştir. Bizim çalışmamızda da 3 farklı basecalling programıyla denemeler yapılmış ve benzer şekilde gerek hız gerekse doğru genotiplendirme bakımından Guppy'nin diğer programlardan çok daha önde olduğu tespit edilmiştir. Aynı zamanda Wick ve arkadaşlarının çalışmalarındaki bir önemli nokta da konsensus sekans doğruluğunun 99.94%'lere kadar çıktığını tespit etmeleri olmuştur. Bu da son dört yıl içerisinde bu teknolojiye büyük gelişimi farklı bir açıdan göstermektedir.

Nanopore sekanslama teknolojisi, basecalling adımıdan sonra diğer yeni nesil teknolojilerinden farklı olarak daha kompleks biyoinformatik analizlere ihtiyaç duymaktadır. Bunun nedeni de yer yer %20'lere varabilen hata oranıdır (Ashton ve diğerleri, 2015). Ancak bu verilerin bir çoğu geçmiş yıllarda ve metagenomik çalışmalardan elde edilmiştir (Loit ve diğerleri, 2019). Bu durum biyoinformatik olarak yazılımları güncellemeyi mecbur kılmıştır ve hata oranını düşürmek için birçok farklı yazılım ortaya çıkmıştır. Bu yazılımlar hibrid sekanslama, assembly, grafik temelli haritalama gibi farklı algoritmalarla hata oranını düşürmektedirler (Goodwin ve diğerleri, 2015; Sović ve diğerleri, 2016; Szalay ve Golovchenko, 2015). Ancak bunların en büyük dezavantajı iyi donanımdaki bilgisayarlarda bile onlarca dakikadan saatlere varan sonuç verme süreleridir. Nanopore çalışmalarıyla ilgili geçmiş tecrübemiz, tek nükleotid varyantlarından ziyade insersiyon ve delesyonlarda cihazın daha çok yanıldığını bize göstermiştir. Esas amacımız da SNP araştırmak ve analiz süresini minimumda tutmak olduğu için biyoinformatik iş akışımıza herhangi bir hata düzeltme yazılımı eklenmemiştir. Sonuçlarımıza baktığımızda ise hata düzeltme yazılımları kullanılmadan bile 3 farklı DNA örneği için de 2 ile 3 dakika arasında %100 doğruluk ile biyoinformatik analiz gerçekleştiği görülmüştür.

Biyoinformatiğin en temel basamaklarından olan haritalama, kabaca referans bir diziyle cihazdan çıkan okumaları eşleştirme olarak tanımlansa da, nanopore teknolojisinde diğer yeni nesil dizi analizi teknolojilerinden daha farklı bir yere sahiptir. Nanopore sekanslama teknolojisinin hata profili ve homojen olmayan hata dağılımı göz önüne alınarak, var olan haritalama programları güncellenmiş ya da bu teknoloji için yeni programlar yazılmıştır (Pavlovikj, Moriyama ve Deogun, 2017). Haritalama programların algoritmaları genel anlamda farklılık gösterdiği için, farklı programlar farklı sonuçlar verebilmektedir. Hatta bu programların seçimine bağlı olarak yanlış haritalanmalar da ortaya çıkabilmektedir (Smith, Xuan ve Zhang, 2008). Çalışmamızda kullanılan BWA, Minimap ve GraphMap programlar literatürde

nanopore teknolojisi ile ilgili alıřmalarda sıklıkla karřımıza ıkmaktadır (Senol Cali, Kim, Ghose, Alkan ve Mutlu, 2018). Farklı yayınlar Minimap'in hız anlamında diđer iki programda önde olduđunu gösterse de (Li, 2018), bizim alıřmamızda zaten 5 dakikalık ham datalar 10 kb gibi oldukça küçük bir alana haritalacađı için 3 program arasında gerek süre gerekse haritalama kalitesi bakımından anlamlı bir farklılık bulunmamıştır.

DNA profillemeye Dünyadaki en büyük otör kabul edilen John Butler, řu an bukkal sürüntüden yapılan genotiplendirmenin en iyi kořullarda dahi 90 dakika civarında sonuç verdiđini belirtmiş ve genotiplendirmede geleceđin çok daha hızlı ve daha yüksek sensitiviteyle sonuç veren yöntem ve platformlara dođru evrileceđini vurgulamıştır (J. M. Butler, 2015). Bizim alıřmamız da aslında bunun başarılı bir örneđini teşkil etmektedir. alıřmamızda toplamda 53 dakikada sonuca varılarak gerek hız gerekse yüksek duyarlılık anlamında literatüre göre çok daha başarılı sonuçlar elde edilmiştir. Ancak ticari olarak satın alınan DNA örneklerinin kullanılması nedeniyle izolasyon basamađı atlanmıştır. alıřmamızın öncü olacađı devam projelerinde farklı vücut sıvılarından hem izolasyonla hem de izolasyonsuz denemeler yapıp, o kořullar altında da genotiplendirme başarısının teyit edilmesi gerekmektedir.

## 6. Sonuç ve Öneriler

1. Çalışmamızda, kişi tanımlamada en yaygın kullanılan 52 farklı SNP, Oxford Nanopore Minion cihazı kullanılarak tam doğrulukla genotiplendirilmiştir.
2. PCR, kütüphane hazırlığı, sekanslama ve biyoinformatik analiz basamakları toplamı 53 dakika sürmüştür. Bu süre, DNA profillemede SNP temelli genotiplendirmede dünya üzerinde ulaşılmış en kısa süredir. Bu süreyi elde etmek için süper hızlı PCR cihazı, hızlı kütüphane hazırlık kiti, Oxford Nanopore sekanslama teknolojisi ve modifiye edilmiş biyoinformatik iş akışı kullanılmıştır.
3. Çalışmamızda kullanılan biyoinformatik iş akışı, tek bir komutla, herhangi bir ekstra işleme gerek olmaksızın birkaç dakikalık bir sürede sonuç verebilmektedir. Bu açıdan, uygulayacak olan kişinin herhangi bir komut yazma becerisine sahip olma ya da sonuçları tek tek analiz etme gerekliliği ortadan kalkmaktadır.
4. Geliştirdiğimiz SNP temelli kişi tanımlama kiti prototipi, sahada ve tecrübesi olmayan bir teknik personel tarafından bile kolayca uygulanabilir bir potansiyelindedir. Aynı zamanda, sonraki büyük ölçekli çalışmalar için öncü olma niteliği taşımaktadır.

## 7. Kaynaklar

- 1000 Genomes Project Consortium, Auton, A., Brooks, L. D., Durbin, R. M., Garrison, E. P., Kang, H. M., ... Abecasis, G. R. (2015). A global reference for human genetic variation. *Nature*. doi:10.1038/nature15393
- Amigo, J., Phillips, C., Lareu, M. ve Carracedo, Á. (2008). The SNPforID browser: An online tool for query and display of frequency data from the SNPforID project. *International Journal of Legal Medicine*. doi:10.1007/s00414-008-0233-7
- Ashton, P. M., Nair, S., Dallman, T., Rubino, S., Rabsch, W., Mwaigwisya, S., ... O'Grady, J. (2015). MinION nanopore sequencing identifies the position and structure of a bacterial antibiotic resistance island. *Nature Biotechnology*. doi:10.1038/nbt.3103
- Belmont, J. W., Boudreau, A., Leal, S. M., Hardenbol, P., Pasternak, S., Wheeler, D. A., ... Stewart, J. (2005). A haplotype map of the human genome. *Nature*. doi:10.1038/nature04226
- Belser, C., Istace, B., Denis, E., Dubarry, M., Baurens, F. C., Falentin, C., ... Aury, J. M. (2018). Chromosome-scale assemblies of plant genomes using nanopore long reads and optical maps. *Nature Plants*. doi:10.1038/s41477-018-0289-4
- Børsting, C., Mogensen, H. S. ve Morling, N. (2013). Forensic genetic SNP typing of low-template DNA and highly degraded DNA from crime case samples. *Forensic Science International: Genetics*. doi:10.1016/j.fsigen.2013.02.004
- Branton, D., Deamer, D. W., Marziali, A., Bayley, H., Benner, S. A., Butler, T., ... Schloss, J. A. (2008). The potential and challenges of nanopore sequencing. *Nature Biotechnology*. doi:10.1038/nbt.1495
- Brookes, C., Bright, J. A., Harbison, S. ve Buckleton, J. (2012). Characterising stutter in forensic STR multiplexes. *Forensic Science International: Genetics*. doi:10.1016/j.fsigen.2011.02.001
- Budowle, B., Moretti, T. R., Baumstark, A. L., Defenbaugh, D. A. ve Keys, K. M. (1999). Population data on the thirteen CODIS core short tandem repeat loci in African Americans, U.S. Caucasians, Hispanics, Bahamians, Jamaicans, and Trinidadians. *Journal of Forensic Sciences*, 44(6), 1277–1286. doi:10.1520/JFS14601J
- Buermans, H. P. J. ve den Dunnen, J. T. (2014). Next generation sequencing technology: Advances and applications. *Biochimica et Biophysica Acta* -

*Molecular Basis of Disease*. doi:10.1016/j.bbadis.2014.06.015

Butler, J. M. (2015). The future of forensic DNA analysis. *Philosophical Transactions of the Royal Society B: Biological Sciences*. doi:10.1098/rstb.2014.0252

Butler, J. M., Buel, E., Crivellente, F. ve McCord, B. R. (2004). Forensic DNA typing by capillary electrophoresis using the ABI Prism 310 and 3100 genetic analyzers for STR analysis. *ELECTROPHORESIS*. doi:10.1002/elps.200305822

Butler, J. M. J. M. (2005). *Forensic DNA typing: biology, technology, and genetics of STR markers*. *Chemistry & biodiversity*. doi:10.1007/s13398-014-0173-7.2

Chatron, N., Schluth-Bolard, C., Frétigny, M., Labalme, A., Vilchez, G., Castet, S. M., ... Jourdy, Y. (2019). Severe hemophilia A caused by an unbalanced chromosomal rearrangement identified using nanopore sequencing. *Journal of Thrombosis and Haemostasis*. doi:10.1111/jth.14460

Cho, S., Yu, H. J., Han, J., Kim, Y., Lee, J. ve Lee, S. D. (2014). Forensic application of SNP-based resequencing array for individual identification. *Forensic Science International: Genetics*. doi:10.1016/j.fsigen.2014.07.005

Choudhury, A., Hazelhurst, S., Meintjes, A., Achinike-Oduaran, O., Aron, S., Gamiieldien, J., ... Ramsay, M. (2014). Population-specific common SNPs reflect demographic histories and highlight regions of genomic plasticity with functional relevance. *BMC Genomics*. doi:10.1186/1471-2164-15-437

Cornelis, S., Gansemans, Y., Deleye, L., Deforce, D. ve Van Nieuwerburgh, F. (2017). Forensic SNP Genotyping using Nanopore MinION Sequencing. *Scientific Reports*. doi:10.1038/srep41759

Cornelis, S., Gansemans, Y., Vander Plaetsen, A. S., Weymaere, J., Willems, S., Deforce, D. ve Van Nieuwerburgh, F. (2019). Forensic tri-allelic SNP genotyping using nanopore sequencing. *Forensic Science International: Genetics*. doi:10.1016/j.fsigen.2018.11.012

Daniel, R. ve Walsh, S. J. (2006). The Continuing Evolution of Forensic DNA Profiling - From STRS to SNPS. *Australian Journal of Forensic Sciences*. doi:10.1080/00450610609410633

Feng, Y., Zhang, Y., Ying, C., Wang, D. ve Du, C. (2015). Nanopore-based fourth-generation DNA sequencing technology. *Genomics, Proteomics and Bioinformatics*. doi:10.1016/j.gpb.2015.01.009

Fyta, M. (2015). Threading DNA through nanopores for biosensing applications.

*Journal of Physics Condensed Matter*. doi:10.1088/0953-8984/27/27/273101

- Gammon, K. ve Mayers, C. (2019). Optimisation of rapid STR analysis using a standard DNA forensic pipeline, (March), 1–4. <https://www.biorxiv.org/content/biorxiv/early/2019/03/12/575357.full.pdf> adresinden erişildi.
- Geyer, C. N. ve Hanson, N. D. (2013). Rapid screening of transformants using the Streck Philisa® thermal cycler. *BioTechniques*. doi:10.2144/000114106
- Goodwin, S., Gurtowski, J., Ethe-Sayers, S., Deshpande, P., Schatz, M. C. ve McCombie, W. R. (2015). Oxford Nanopore sequencing, hybrid error correction, and de novo assembly of a eukaryotic genome. *Genome Research*. doi:10.1101/gr.191395.115
- Heinrich, M., Müller, M., Rand, S., Brinkmann, B. ve Hohoff, C. (2004). Allelic drop-out in the STR system ACTBP2 (SE33) as a result of mutations in the primer binding region. *International Journal of Legal Medicine*. doi:10.1007/s00414-004-0473-0
- Illumina. (2018). Illumina Indexed Sequencing Guide 02/2018. *Illumina*. doi:10.1264/jsme2.ME14112
- Ip, C. L. C., Loose, M., Tyson, J. R., de Cesare, M., Brown, B. L., Jain, M., ... Olsen, H. E. (2015). MinION Analysis and Reference Consortium: Phase 1 data release and analysis. *F1000Research*. doi:10.12688/f1000research.7201.1
- Jain, M., Tyson, J. R., Loose, M., Ip, C. L. C., Eccles, D. A., O’Grady, J., ... Olsen, H. E. (2017). MinION Analysis and Reference Consortium: Phase 2 data release and analysis of R9.0 chemistry. *F1000Research*. doi:10.12688/f1000research.11354.1
- Jain, Miten, Koren, S., Miga, K. H., Quick, J., Rand, A. C., Sasani, T. A., ... Loose, M. (2018). Nanopore sequencing and assembly of a human genome with ultra-long reads. *Nature Biotechnology*. doi:10.1038/nbt.4060
- Jain, Miten, Olsen, H. E., Paten, B. ve Akeson, M. (2016). The Oxford Nanopore MinION: Delivery of nanopore sequencing to the genomics community. *Genome Biology*. doi:10.1186/s13059-016-1103-01
- James Kent, W., Sugnet, C. W., Furey, T. S., Roskin, K. M., Pringle, T. H., Zahler, A. M. ve Haussler, D. (2002). The human genome browser at UCSC. *Genome Research*. doi:10.1101/gr.229102. Article published online before print in May

2002

- Jeffreys, A. J., Brookfield, J. F. Y. ve Semeonoff, R. (1985). Positive identification of an immigration test-case using human DNA fingerprints. *Nature*. doi:10.1038/317818a0
- Jobling, M. A. ve Gill, P. (2004). Encoded evidence: DNA in forensic analysis. *Nature Reviews Genetics*. doi:10.1038/nrg1455
- Kashyap, V. K., Sitalaximi, T., Chattopadhyay, P. ve Trivedi, R. (2017). DNA Profiling Technologies in Forensic Analysis. *International Journal of Human Genetics*. doi:10.1080/09723757.2004.11885864
- Kim, S. ve Misra, A. (2007). SNP Genotyping: Technologies and Biomedical Applications. *Annual Review of Biomedical Engineering*. doi:10.1146/annurev.bioeng.9.060906.152037
- Koressaar, T. ve Remm, M. (2007). Enhancements and modifications of primer design program Primer3. *Bioinformatics*. doi:10.1093/bioinformatics/btm091
- Kuderna, L. F. K., Lizano, E., Julià, E., Gomez-Garrido, J., Serres-Armero, A., Kuhlwil, M., ... Marques-Bonet, T. (2019). Selective single molecule sequencing and assembly of a human Y chromosome of African origin. *Nature Communications*. doi:10.1038/s41467-018-07885-5
- Ledergerber, C. ve Dessimoz, C. (2011). Base-calling for next-generation sequencing platforms. *Briefings in Bioinformatics*. doi:10.1093/bib/bbq077
- Leija-Salazar, M., Sedlazeck, F. J., Toffoli, M., Mullin, S., Mokretar, K., Athanasopoulou, M., ... Proukakis, C. (2019). Evaluation of the detection of GBA missense mutations and other variants using the Oxford Nanopore MinION. *Molecular Genetics and Genomic Medicine*. doi:10.1002/mgg3.564
- Li, H. (2018). Minimap2: Pairwise alignment for nucleotide sequences. *Bioinformatics*. doi:10.1093/bioinformatics/bty191
- Li, H. ve Durbin, R. (2010). Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics*. doi:10.1093/bioinformatics/btp698
- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., ... Durbin, R. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics*. doi:10.1093/bioinformatics/btp352
- Loit, K., Adamson, K., Bahram, M., Puusepp, R., Anslan, S., Kiiker, R., ... Tedersoo, L. (2019). Relative performance of Oxford Nanopore MinION vs. Pacific

- Biosciences Sequel third-generation sequencing platforms in identification of agricultural and forest pathogens. *bioRxiv*. doi:10.1101/592972
- Lu, H., Giordano, F. ve Ning, Z. (2016). Oxford Nanopore MinION Sequencing and Genome Assembly. *Genomics, Proteomics and Bioinformatics*. doi:10.1016/j.gpb.2016.05.004
- Marjanovic, D., Bakal, N., Pojskic, N., Kapur, L., Drobnic, K., Primorac, D., ... Hadziselmovic, R. (2006). Allele frequencies for 15 short tandem repeat loci in a representative sample of Bosnians and Herzegovinians. *Forensic Science International*. doi:10.1016/j.forsciint.2005.01.006
- Metzker, M. L. (2010). Sequencing technologies the next generation. *Nature Reviews Genetics*. doi:10.1038/nrg2626
- Mitsuhashi, S., Kryukov, K., Nakagawa, S., Takeuchi, J. S., Shiraishi, Y., Asano, K. ve Imanishi, T. (2017). A portable system for rapid bacterial composition analysis using a nanopore-based sequencer and laptop computer. *Scientific Reports*. doi:10.1038/s41598-017-05772-5
- Mittal, B., Chaturvedi, P. ve Tulsyan, S. (2013). Restriction Fragment Length Polymorphism. *Brenner's Encyclopedia of Genetics: Second Edition* içinde . doi:10.1016/B978-0-12-374984-0.01314-0
- Mullis, K. B. (1990). The unusual origin of the polymerase chain reaction. *Scientific American*. doi:10.1038/scientificamerican0490-56
- Näslund, K., Saetre, P., Von Salomé, J., Bergström, T. F., Jareborg, N. ve Jazin, E. (2005). Genome-wide prediction of human VNTRs. *Genomics*. doi:10.1016/j.ygeno.2004.10.009
- Nussbaum, R. (2016). Human Genetic Diversity: Mutation and Polymorphism. *Thompson & Thompson Genetics in Medicine* içinde .
- Oetting, W. S., Lee, H. K., Flanders, D. J., Wiesner, G. L., Sellers, T. A. ve King, R. A. (1995). Linkage analysis with multiplexed short tandem repeat polymorphisms using infrared fluorescence and M13 tailed primers. *Genomics*. doi:10.1006/geno.1995.1264
- Patil, N., Berno, A. J., Hinds, D. A., Barrett, W. A., Doshi, J. M., Hacker, C. R., ... Cox, D. R. (2001). Blocks of limited haplotype diversity revealed by high-resolution scanning of human chromosome 21. *Science*. doi:10.1126/science.1065573

- Pavlovikj, N., Moriyama, E. N. ve Deogun, J. S. (2017). Comparative analysis of alignment tools for nanopore reads. *Proceedings - 2017 IEEE International Conference on Bioinformatics and Biomedicine, BIBM 2017* içinde . doi:10.1109/BIBM.2017.8217645
- Primorac, D. ve Schanfield, M. (2014). Forensic DNA Applications: An Interdisciplinary Perspective – a new book in forensic science. *Croatian Medical Journal*. doi:10.3325/cmj.2014.55.434
- Quick, J., Loman, N. J., Duraffour, S., Simpson, J. T., Severi, E., Cowley, L., ... Carroll, M. W. (2016). Real-time, portable genome sequencing for Ebola surveillance. *Nature*. doi:10.1038/nature16996
- Rang, F. J., Kloosterman, W. P. ve de Ridder, J. (2018). From squiggle to basepair: Computational approaches for improving nanopore sequencing read accuracy. *Genome Biology*. doi:10.1186/s13059-018-1462-9
- Reid, Y., Storts, D., Riss, T. ve Minor, L. (2004). *Authentication of Human Cell Lines by STR DNA Profiling Analysis. Assay Guidance Manual*.
- Robinson, J. T., Thorvaldsdóttir, H., Wenger, A. M., Zehir, A. ve Mesirov, J. P. (2017). Variant review with the integrative genomics viewer. *Cancer Research*. doi:10.1158/0008-5472.CAN-17-0337
- Sanchez, J. J., Phillips, C., Børsting, C., Balogh, K., Bogus, M., Fondevila, M., ... Morling, N. (2006). A multiplex assay with 52 single nucleotide polymorphisms for human identification. *Electrophoresis*. doi:10.1002/elps.200500671
- Schmidt, K., Mwaigwisya, S., Crossman, L. C., Doumith, M., Munroe, D., Pires, C., ... Livermore, D. M. (2017). Identification of bacterial pathogens and antimicrobial resistance directly from clinical urines by nanopore-based metagenomic sequencing. *Journal of Antimicrobial Chemotherapy*. doi:10.1093/jac/dkw397
- Schneider, P. M. (2012). Beyond STRs: The role of diallelic markers in forensic genetics. *Transfusion Medicine and Hemotherapy*. doi:10.1159/000339139
- Senol Cali, D., Kim, J. S., Ghose, S., Alkan, C. ve Mutlu, O. (2018). Nanopore sequencing technology and tools for genome assembly: computational analysis of the current state, bottlenecks and future directions. *Briefings in Bioinformatics*. doi:10.1093/bib/bby017
- Sherry, S. T. (2001). dbSNP: the NCBI database of genetic variation. *Nucleic Acids*

*Research*. doi:10.1093/nar/29.1.308

- Simpson, J. T., Workman, R. E., Zuzarte, P. C., David, M., Dursi, L. J. ve Timp, W. (2017). Detecting DNA cytosine methylation using nanopore sequencing. *Nature Methods*. doi:10.1038/nmeth.4184
- Smith, A. D., Xuan, Z. ve Zhang, M. Q. (2008). Using quality scores and longer reads improves accuracy of Solexa read mapping. *BMC Bioinformatics*. doi:10.1186/1471-2105-9-128
- Sović, I., Šikić, M., Wilm, A., Fenlon, S. N., Chen, S. ve Nagarajan, N. (2016). Fast and sensitive mapping of nanopore sequencing reads with GraphMap. *Nature Communications*. doi:10.1038/ncomms11307
- Swanson, A., Ramos, E. ve Snyder, H. (2014). Next generation sequencing is the impetus for the next generation of laboratory-based genetic counselors. *Journal of Genetic Counseling*. doi:10.1007/s10897-013-9684-1
- Szalay, T. ve Golovchenko, J. A. (2015). De novo sequencing and variant calling with nanopores using PoreSeq. *Nature Biotechnology*. doi:10.1038/nbt.3360
- Tracey, M. (2001). Short tandem repeat-based identification of individuals and parents. *Croatian medical journal*.
- Tyler, A. D., Mataseje, L., Urfano, C. J., Schmidt, L., Antonation, K. S., Mulvey, M. R. ve Corbett, C. R. (2018). Evaluation of Oxford Nanopore's MinION Sequencing Device for Microbial Whole Genome Sequencing Applications. *Scientific Reports*, 8(1), 10931. doi:10.1038/s41598-018-29334-5
- Untergasser, A., Cutcutache, I., Koressaar, T., Ye, J., Faircloth, B. C., Remm, M. ve Rozen, S. G. (2012). Primer3-new capabilities and interfaces. *Nucleic Acids Research*. doi:10.1093/nar/gks596
- Wang, Q., Fu, L., Zhang, X., Dai, X., Bai, M., Fu, G., ... Li, S. (2016). Expansion of a SNaPshot assay to a 55-SNP multiplex: Assay enhancements, validation, and power in forensic science. *Electrophoresis*. doi:10.1002/elps.201500353
- Wang, Y., Yang, Q. ve Wang, Z. (2014). The evolution of nanopore sequencing. *Frontiers in Genetics*. doi:10.3389/fgene.2014.00449
- Weber-Lehmann, J., Schilling, E., Gradl, G., Richter, D. C., Wiehler, J. ve Rolf, B. (2014). Finding the needle in the haystack: Differentiating "identical" twins in paternity testing and forensics by ultra-deep next generation sequencing. *Forensic Science International: Genetics*. doi:10.1016/j.fsigen.2013.10.015

Wick, R. R., Judd, L. M. ve Holt, K. E. (2019). Performance of neural network basecalling tools for Oxford Nanopore sequencing. *Genome Biology*. doi:10.1186/s13059-019-1727-y

Wyman, A. R. ve White, R. (1980). A highly polymorphic locus in human DNA. *Proceedings of the National Academy of Sciences*. doi:10.1073/pnas.77.11.6754



## **Ekler**

“Nanopore sekanslama teknolojisi kullanılarak, kiři tanımlama kiti geliştirilmesi” başlıklı çalışmamız, TGA-2019-20451 proje kodu ile Ege Üniversitesi Bilimsel Araştırma Projeleri Koordinatörlüğü tarafından desteklenmiştir.



## Teşekkür

Doktora eğitim süresince, engin bilgi ve deneyimlerini her fırsatta paylaşan, çalışma disipliniyle bizlere örnek olan, motivasyonu ve desteği hiç eksilmeyen, bilimi ve bilimi sevmeyi bana öğreten, kapısı bizlere hep açık olan ve en önemlisi ekip kavramının akademik sınırları aşan bir aile ortamı olduğunu gösteren, bu süreçte öğrencisi olmaktan onur duyduğum danışman hocam Prof. Dr. Ferda Özkınay'a,

Hayatımda tanıdığım en çalışkan, neşeli ve gerçek bir bilim insanı olan, bilgi birikimiyle aydınlatan, başarılarıyla yol gösteren, motivasyonunu hiç kaybetmeyen, genetik serüvenimi başlatan ve beni bugünlere taşıyan, yanında bulunduğum her saniye vizyonumu genişleten, hem akademik hem de sosyal anlamda rol model olarak benimsediğim değerli hocam Prof. Dr. Hüseyin Onay'a,

Doktoraya başladığım ilk günden itibaren en zor zamanlarda bile güler yüzü hiç eksik olmayan, moleküler genetiği tüm detaylarıyla öğreten, kliniğin genetikteki önemini kavramamı sağlayan, biyoinformatik alanında bana her zaman destek olan, bilimsel başarılarının yanı sıra pür iyilik, naifliğiyle nasıl yaşanılmalı sorusuna cevap olan hocam Doç. Dr. Tahir Atik'e,

Doktora eğitimim boyunca poliklinik ve servis vizitleri sırasında bilgi ve tecrübelerini hem bilimsel hem de popüler genetik konularında paylaşan, Tıbbi Genetik Anabilim Dalı'nın kapılarının her zaman açık olduğunu vurgulayıp laboratuvarlarında çalışmamızı sağlayan, yayın ve kitap yazımlarında bizlere de yer veren değerli hocam Prof. Dr. Özgür Çoğulu'ya,

Dört senelik doktora süresince en çok birlikte çalıştığım, bilgilerini her zaman paylaşan, çalışkanlığı bana örnek olan, zor zamanlarda bile motive kalmamızı sağlayan, doktora öğrencisi olduğumu hissettirmeden sorumluluk bilinciyle işlere odaklanmamı sağlayan, neşesi hiç eksik olmayan uzmanım Uzm. Dr. Esra Işık'a,

Desteklerini hissettiğim, bilgi ve tecrübelerinden yararlandığım değerli hocalarım Prof. Dr. Haluk Akın, Doç. Dr. Emin Karaca, Doç. Dr. Asude Durmaz, Doç. Dr. Burak Durmaz ve Doç. Dr. Ayça Aykut'a,

İstanbul Acıbadem Üniversitesi'nde bir ay boyunca ekibinde yer almamı sağlayan, biyoinformatik ile ilgili temellerimi oluşturan, enerjisi ve yaklaşımıyla her zaman pozitif kalmanızı sağlayan değerli hocam Prof. Dr. Uğur Sezerman'a,

Tecrübeleriyle yanımda olan Uzm. Dr. Erhan Parıltay, Uzm. Dr. Aslı Ece Solmaz, Uzm. Dr. Melis Köse ve Uzm. Dr. Durdugül Ayyıldız'a,

Bu süreçte biyoinformatik konusunda beni her zaman teşvik eden, farklı alanlardaki hakimiyetlerimizi birleştirip birlikte yenilikçi çalışmalara imza attığımız, hem şu anki hem de gelecekte iş arkadaşlarım ve dostlarım Uzm. Dr. Hilmi Bolat ve Dr. Semih Aşıkocalı'ya,

İzmir'e geldiği günden beri hem çalışma hem de sosyal anlamda yanımda olan, kaliteli sohbeti ve konulara bakış perspektifi ufkumu genişleten arkadaşım Dr. Elif Uzay'a, Doktora eğitimim boyunca birlikte çalışmaktan büyük keyif aldığım hekim arkadaşlarım Uzm. Dr. Hasan Taşlıdere, Uzm. Dr. Ayşe Nur Kavasoğlu, Uzm. Dr. İpek Ceylan, Dr. Özgüç Semih Şimşir, Dr. Ayhan Keçelioğlu, Dr. Tuba Sözen Türk, Dr. Zehra Cengisiz, Dr. Yasemin Karaca, Dr. Gizem Kök ve Dr. Duygu Arıcan'a

Bu süreçte bana her zaman yardımcı olan biyolog, teknisyen, sekreter ve personel arkadaşlara,

“Nanopore sekanslama teknolojisi kullanılarak, kişi tanımlama kiti geliştirilmesi” isimli çalışmamızı, TGA-2019-20451 proje koduyla destekleyen Ege Üniversitesi Bilimsel Araştırma Projeleri Koordinatörlüğüne,

Biyoinformatik alanında 100/2000 doktora bursuyla beni destekleyen Yüksek Öğretim Kurulu'na,

Bu süreçte 2211/C Yurt İçi Öncelikli Alanlar Doktora Burs Programı'yla destek olan TÜBİTAK'a,

Doktora başlamaya beni teşvik eden, maddi ve manevi anlamda sonsuz destek olan, zor zamanlarda bile verdikleri güven ile dik kalmamı sağlayan ve bugünlere ulaşmamda en büyük emeğin sahipleri olan annem, babam ve anneannem'e,

En içten teşekkür, sevgi ve saygılarımı sunarım.

İzmir, 07.08.2019

Bilçay AKGÜN

## Özgeçmiş

### Eğitim

- Ege Üniversitesi Tıp Fakültesi 2009-2015
- Ege Üniversitesi Tıp Fakültesi Araştırma Eğitimi Programı 2011-2015
- Aristotle University of Thessaloniki Medical School 2014-2014

### Beceriler

#### Klinik Genetik

- 1000’i geçkin hastanın poliklinikte ayrıntılı dismorfolojik muayenesi ve aile ağacı
- Konsültasyon hekimliği
- Genetik servisi yatan hastalarından sorumlu asistan hekimlik

#### Moleküler Genetik Laboratuvarı

- DNA izolasyonu, RNA izolasyonu, Restriksiyon ve Ligasyon işlemleri, Pürifikasyon, PCR, Metilasyon Spesifik PCR (MS-PCR), qPCR, QF-PCR, Inverse Shifting(IS)-PCR, 1. (Sanger), 2. (Illumina platformları) ve 3. (Oxford Nanopore) Nesil Dizi Analizi

#### Sitogenetik Laboratuvarı

- Periferik kandan kromozomal karyotipleme

#### Biyoinformatik

- Birinci Nesil Dizi Analizi (Sanger)
- İkinci Nesil Dizi Analizi
- Üçüncü Nesil Dizi Analizi (Oxford Nanopore)

### Burslar

- 100/2000 Doktora Bursu (Biyoenformatik alanında) - Yüksek Öğretim Kurulu
- 2211/C Yurt İçi Öncelikli Alanlar Doktora Bursu - TÜBİTAK

### Projeler

## TÜBİTAK – 1511

- Fulminant Hepatite Sebep olan Viral Patojenlerin Tespiti için Nanopore Sekanslama Teknolojisi Temelli, Hızlı ve Hasta Başında Uygulanabilir Tanı Kiti Geliştirilmesi:HEPASEQ

### Ege Üniversitesi - Genel Araştırma Projesi

- Nanopore sekanslama teknolojisi kullanılarak, kişi tanımlama kiti geliştirilmesi

### Ege Üniversitesi - Genel Araştırma Projesi

- Türkiyede Rotavirüs ve Norovirüs Gastroenteritine Yatkınlıkta FUT2 Gen Polimorfizmlerinin Araştırılması

### Ege Üniversitesi - Genel Araştırma Projesi

- CRISPR-Cas9 Teknolojisi ile in vitro F8 Geninin Fonksiyonunun Bozulmasının Sağlanması

### Ege Üniversitesi – Tez Projesi

- Ağır Hemofili A Hastalarında İmmün Düzenleyici Sitokinleri Kodlayan Gen Polimorfizmlerinin İnhibitör Gelişimi İle İlişkisinin Belirlenmesi

### Ege Üniversitesi – Tez Projesi

- Alkalen Fasfataz Düzeyi Düşük Olan Ve Klinik Olarak Hipofosfatazya Düşünülen Hastalarda Alkalen Fosfataz (ALP) Geni Dizi Analizi İle Moleküler Tanımın Araştırılması

### Ege Üniversitesi – Tez Projesi

- Ağır Hemofili A Hastalarında İmmün Düzenleyici Sitokinleri Kodlayan Gen Polimorfizmlerinin İnhibitör Gelişimi İle İlişkisinin Belirlenmesi

### Ege Üniversitesi – Genel Araştırma Projesi

- Otizmlili Olgularda Moleküler Karyotipleme Yöntemi ile Genetik Etiyolojinin Aydınlatılması

## Ödüller

### Sözel Sunum Birinciliği

- Hemofili Vakalarla Eğitim Sempozyumu, 22-24 Şubat 2019, Antalya

- “Hemofili B Hastalığı Moleküler Analizinde Türkiye Deneyimi: F9 Gen Mutasyon Spektrumu ve Genotip-Fenotip İlişkisi”

#### Sözel Sunum Birinciliği

- Uluslararası Türkiye Hemofili Kongresi, 14-17 Nisan 2018, İstanbul
- “Hemofili A Hastalarında F8 Gen Mutasyon Spektrumu: 21 Yeni Mutasyon Tanımlanması”

#### Sözel Sunum Birinciliği

- HEmofili Vakalarla Eğitim Sempozyumu( HEVES ), 23-25 Şubat 2018, Antalya
- “Ağır Hemofili A Olgularında Yeni Nesil Dizi Analizi İle F8 Geni Büyük Delesyonlarının Belirlenmesi”

#### Sözel Sunum İkinciliği

- Hematolojik Genetik Sempozyumu, 14-16 Şubat 2018, İzmir
- “Hemofili A Hastalarında F8 Gen Mutasyon Spektrumu: 18 Yeni Mutasyon Tanımlanması”

#### Sözel Sunum Birinciliği

- Nörometabolik Dismorfoloji Sempozyumu, 10-12 Mart 2016, İstanbul
- “Otizmlili Olgularda Moleküler Karyotipleme Yöntemi ile Genetik Etiyolojinin Aydınlatılması”

#### Yayınlar

Ozbaran B, Akgun B, Kacamak D, Kose S, Kavasoglu A, Onay H. The Role of Molecular Karyotyping in Elucidating the Genetic Etiology of Autism. Turk Psikiyatri Derg. 2017 Fall;28(3):156-162. doi:10.5080/u18239.

Onay H, Kacamak D, Kavasoglu A, Akgun B, Yalcinli M, Kose S, Ozbaran B. Mutation Analysis of NRXN1 Gene in Autism Spectrum Disorders. Balkan J Med Genet. 2017 Mar 8;19(2):17-22. doi: 10.1515/bjmg-2016-0031.

Karaca I, Yilmaz SG, Palamar M, Onay H, Akgun B, Aytacoglu B, Aykut A, Ozkinay FF. Evaluation of CNTNAP2 gene rs2107856 polymorphism in Turkish population

with pseudoexfoliation syndrome. Int Ophthalmol. 2017 Dec 19. doi: 10.1007/s10792-017-0800-3.

Isik E, Onay H, Atik T, Akgun B, Cogulu O, Ozkinay F. Clinical and genetic features of L1 syndrome patients: Definition of two novel mutations. Clinical Neurology and Neurosurgery 2018; 172: 20–23. doi: 10.1016/j.clineuro.2018.06.007

Ozdemir TR, Karaca NE, Marshall JD, Kutukculer N, Aksu G, Ozgul RK, Ozanturk A, Isik E, Akgun B, Ozdemir HH, Darcan S, Ozkinay F, Cogulu O. Immunodeficiency in a Child with Alström Syndrome. Indian J Pediatr. 2018 Oct;85(10):924-926. doi: 10.1007/s12098-018-2740-y.

#### Kitap bölümleri

Cogulu O, Akgun B. Genetic Biomarkers in Pediatric Acute Myeloblastic Leukemia. Türkiye Klinikleri J Pediatr Sci 2016;12(4):90-100 (In Turkish)

Akgun B, Atik T. Trinükleotid Tekrar Hastalıkları. Tıbbi Genetik (Laboratuvar ve Klinik). Nobel Tıp Yayınları, 2017;393-397. ISBN:9786059215442

Akgun B, Ozkinay F. Genetik Hastalıklarda Tedavi. Tıbbi Genetik (Laboratuvar ve Klinik). Nobel Tıp Yayınları, 2017;s.313-317. ISBN:9786059215442

Atik T, Akgun B. Kalıtsal Cilt Hastalıkları (Genodermatozlar). Tıbbi Genetik (Laboratuvar ve Klinik). Nobel Tıp Yayınları, 2017;s.299-302. ISBN:9786059215442

#### Uluslararası kongre ve sempozyumlar

European Society of Human Genetics Conference (ESHG)

– 21-24 Mayıs 2016, Barcelona, İspanya

European Society of Human Genetics Conference (ESHG)

– 27-30 Mayıs 2017, Kopenhag, Danimarka

13th Balkan Congress of Human Genetics

– 17-20 Nisan 2019, Edirne, Türkiye

**İletişim için mail adresi:** bilcagakgun@hotmail.com