

**KARADENİZ TEKNİK ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ**



TRABZON



KARADENİZ TEKNİK ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ



Karadeniz Teknik Üniversitesi Fen Bilimleri Enstitüsünce

Unvanı Verilmesi İçin Kabul Edilen Tezdir.

Tezin Enstitüye Verildiği Tarih : / /

Tezin Savunma Tarihi : / /

Tez Danışmanı :

Trabzon

ÖNSÖZ

Çağımızın en önemli hastalıklarından biri olan kanserin erken tanı ve tedavisi için geliştirilen uygulamalara katkı sağlamak amacıyla yaptığım tez çalışmasının gerçekleştirilmesinde bilgi ve tecrübesini benimle paylaşan, çalışma boyunca desteğini eksik etmeyen kıymetli hocam Dr. Öğr. Üyesi İbrahim SAVRAN'a teşekkür ederim.

Çalışma süresince bilgi ve görüşlerinden faydalandığım Öğr. Gör. Gökhan DİLEK'e, İnş. Yük. Müh. Yusuf PARLAYAN'a ve iş dünyasındaki yol arkadaşlarım Phi Software ailesine teşekkür ederim.

Maddi ve manevi hiçbir fedakarlıktan kaçınmayan, aldığım her kararın en büyük destekçisi olan aileme ve hayatıma girdiği günden itibaren her anımı güzelleştiren, aynı zamanda çalışmamın tüm aşamalarında yanımda olan sevgili eşim Dt. Aysel TOPRAK AKBULUT'a sonsuz teşekkür ve sevgilerimi sunarım.

Safa AKBULUT
Trabzon 2022

TEZ ETİK BEYANNAMESİ

Yüksek Lisans Tezi olarak sunduğum “Genetik Verilerde Mutasyon Analizi ve Varyant Tespiti” isimli çalışmanın başından sonuna kadarki tüm süreçlerini danışman hocam Dr. Öğr. Üyesi İbrahim SAVRAN’ın sorumluluk ve gözetiminde tamamladığımı, verileri kendi imkanlarımla toplayarak, çalışmaları uygun ortam ve laboratuvarlarda yaptığımı, farklı kaynaklardan aldığım verileri metin içerisinde ve kaynaklar kısmında tam olarak gösterdiğimi, çalışma sürecinde bilimsel araştırma ve etik kurallara uygun olarak davrandığımı ve aksi hallerin tespit edilmesi halinde her sonucu kabul ettiğimi bildiririm.

21/02/2022

Safa AKBULUT

İÇİNDEKİLER

	<u>Sayfa No</u>
ÖNSÖZ.....	III
TEZ ETİK BEYANNAMESİ.....	IV
İÇİNDEKİLER.....	V
ÖZET	VII
SUMMARY	VIII
ŞEKİLLER DİZİNİ.....	IX
TABLolar DİZİNİ.....	XI
SEMBOLLER DİZİNİ	XII
1. GENEL BİLGİLER.....	1
1.1. Giriş.....	1
1.2. Temel Kavramlar.....	1
1.2.1. Nükleik Asitler	1
1.2.1.1. DNA	2
1.2.1.2. RNA	3
1.2.2. Hücre Siklusu	5
1.3. Mutasyon.....	8
1.3.1. Bulundukları Yere Göre Mutasyonlar.....	9
1.3.2. Oluşum Mekanizmasına Göre Mutasyonlar.....	9
1.3.3. Fenotipik Etkilerine Göre Mutasyonlar.....	9
1.3.4. Moleküler Değişiklik Tipine Göre Mutasyonlar	10
1.3.4.1. Nokta (Gen) Mutasyonları	10
1.3.4.1.1. Baz Çifti Değişikliği Mutasyonları (Substitution Mutations)	10
1.3.4.1.2. Sessiz, Yanlış Anlamlı ve Anlamsız Mutasyonlar	11
1.3.4.1.3. Çerçeve Kayması (Frameshift) Mutasyonları	12
1.3.4.2. Kromozom Mutasyonları	13
1.3.4.2.1. Sayısına Göre Kromozom Mutasyonları.....	13
1.3.4.2.2. Yapısına Göre Kromozom Mutasyonları	13
1.4. Karsinogenez (Kanser Oluşumu)	14
1.5. p53 Geni	17

1.6.	BRAF Geni.....	17
1.7.	DNA Dizileme.....	21
1.7.1.	İlk Nesil DNA Dizileme Yöntemleri	23
1.7.2.	İkinci Nesil DNA Dizileme Yöntemleri.....	24
1.7.3.	Yeni Nesil DNA Dizileme Yöntemleri	25
2.	YAPILAN ÇALIŞMALAR	45
2.1.	Yöntemler.....	45
2.1.1.	Verilerin Elde Edilmesi	45
2.1.1.1.	NCBI	45
2.1.1.2.	Fasta Formatı.....	45
2.1.2.	K-MER Analizi	27
2.1.3.	Mutasyonlu Gen Verisini Dizileme.....	28
2.1.4.	Kullanılan Veri Yapısı	29
2.1.4.1.	Hash Fonksiyonları	29
2.1.4.2.	Bloom Filter	30
2.1.5.	Mutasyon Algoritmaları	31
2.1.5.1.	Mutasyon Tespit Algoritması.....	31
2.1.5.2.	Mutasyon Türü Belirleme Algoritması	32
3.	BULGULAR	34
4.	TARTIŞMA	36
5.	SONUÇLAR	37
6.	ÖNERİLER	38
7.	KAYNAKLAR.....	40
	ÖZGEÇMİŞ.....	46

Yüksek Lisans

ÖZET

GENETİK VERİLERDE MUTASYON ANALİZİ VE VARYANT TESPİTİ

Safa AKBULUT

Karadeniz Teknik Üniversitesi
Fen Bilimleri Enstitüsü
Bilgisayar Mühendisliği Anabilim Dalı
Danışman: Dr. Öğr. Üyesi İbrahim SAVRAN
2022, 45 Sayfa

Tüm canlılar, genetik yapılarında nükleotit bazlarının özgün bir dizisini bulundurur. Mutasyon, bu dizilerde gerçekleşen kalıtsal değişimlerdir. Mutasyonlar, fenotipte pozitif veya negatif etkilere sebep olabilir ya da nötr kalarak herhangi bir değişiklik yaratmayabilir. Çağımızdaki en tehlikeli hastalıklardan biri olan kanser, genlerdeki mutasyonlar sebebiyle gerçekleşen kontrolsüz hücre büyümesidir ve hücresel düzeyde gerçekleşmektedir. Mutasyonlar, kanserin oluşum sebepleri hakkında bilgi verdiği için kanser araştırmalarında kritik bir öneme sahiptir. Gelişen teknolojiyle birlikte özel cihazlarla okunan DNA dizilerinde mutasyonlar tespit edilebilmektedir. Mutasyon tespit işlemleri, oldukça maliyetli ve uzun zaman alan uygulamalardır. Bu çalışmada nükleotit baz dizilerinde gerçekleşen küçük ölçekli mutasyonların tespiti için yazılımsal çözümler araştırılmıştır. Araştırmalar sonucunda yazılımsal deneyler yapılarak mutasyon tespitinin hızlı bir şekilde gerçekleştirilebilmesi için algoritmalar önerilmiştir. Çalışmada, sağlıklı gen verisi olarak NCBI (National Center for Biotechnology Information) üzerinden paylaşılan BRAF (B-Raf proto-oncogene) geni kullanılmıştır. Sağlıklı gen bilgisi, hash fonksiyonları kullanan Bloom Filter veri yapısında saklanmıştır. Hasta bireylere ait DNA dizileri, yeni nesil dizileme yöntemlerine uygun olacak şekilde fasta formatında okunmuş olup filtredeki sağlıklı veriler ile karşılaştırılarak oluşan küçük ölçekli mutasyonlar başarılı bir şekilde tespit edilmiştir.

Anahtar Kelimeler: Genetik, DNA, Mutasyon, Kanser, Braf Gen, Dizileme, K-mer, HashFonksiyonu, Bloom Filtresi

Master Thesis

SUMMARY

MUTATION ANALYSIS AND VARIANT DETECTION IN GENETIC DATA

Safa AKBULUT

Karadeniz Technical University
The Graduate School of Natural and Applied Sciences
Computer Engineering Graduate Program
Supervisor: Asst. Prof. Dr. İbrahim SAVRAN
2022, 45 Pages

All living things contain a unique sequence of nucleotide bases in their genetic makeup. Mutation is the genetic changes that occur in these sequences. Mutations may cause positive or negative effects in the phenotype, or they may remain neutral and not create any change. Cancer, one of the most dangerous diseases of our age, is uncontrolled cell growth due to mutations in genes and occurs at the cellular level. Mutations have a critical importance in cancer research as they provide information about the causes of cancer formation. With the developing technology, mutations in the DNA sequences can be detected by reading with special devices. Mutation detection processes are very costly applications which require long time. In this study, software solutions were investigated for the detection of small-scale mutations in nucleotide base sequences. Algorithms have been proposed for rapid mutation detection by performing software experiments based on the result of researches. In the study, BRAF (B-Raf proto-oncogene) gene shared via NCBI (National Center for Biotechnology Information) was used as healthy gene data. The healthy gene information was stored in the Bloom Filter data structure that use hash functions. The DNA sequences of the patients were read in fasta format in accordance with the new generation sequencing methods, and small-scale mutations were successfully detected by comparing them with the healthy data in the filter.

Keywords: Genetics, DNA, Mutation, Cancer, Braf Gene, Sequencing, K-mer, HashFunction, Bloom Filter

ŞEKİLLER DİZİNİ

	<u>Sayfa No</u>
Şekil 1.1. Nükleotit yapısı.	2
Şekil 1.2. DNA molekül yapısı.....	3
Şekil 1.3. Genetik bilgi aktarım aşamaları.	4
Şekil 1.4. tRNA'nın yapısı.	4
Şekil 1.5. Hücre siklusu.....	5
Şekil 1.6. Hücre siklusu kontrol noktaları	6
Şekil 1.7. DNA tamir mekanizması.....	8
Şekil 1.8. Nokta (gen) mutasyonu.	10
Şekil 1.9. Transisyonel mutasyon.....	10
Şekil 1.10. Transversiyonel mutasyon.....	11
Şekil 1.11. Sessiz, yanlış anlamlı ve anlamsız mutasyonlar.....	11
Şekil 1.12. Çerçeve kayması (insersiyon)	12
Şekil 1.13. Çerçeve kayması (delesyon).....	12
Şekil 1.14. Mutasyonların sınıflandırılması	14
Şekil 1.15. Karsinogenezin oluşumu.	16
Şekil 1.16. Ras / MAP kinaz yolu.	18
Şekil 1.17. BRAF V600 mutasyonu	18
Şekil 1.18. Erkeklerin tüm yaşlarda görülen kanser türü oranları, Türkiye	20
Şekil 1.19. Kadınların tüm yaşlarda görülen kanser türü oranları, Türkiye.	20
Şekil 1.20. Over kanseri, yaşa göre beş yıllık hayatta kalma (İngiltere, 2009 – 2013)... ..	21
Şekil 1.21. Kanser türlerinde BRAF gen mutasyonu oranı	21
Şekil 1.22. Kodon tablosu	23
Şekil 1.23. DNA dizileme sistemlerinin tarihsel gelişimi.	25
Şekil 2.1. Gen dizileme	29
Şekil 2.2. Hash fonksiyonu çalışma sistemi	30
Şekil 2.3. Bloom Filter çalışma yapısı.....	30
Şekil 2.4. Mutasyon tespit algoritmasının çalışma yapısı	32
Şekil 2.5. Mutasyon türlerinin oluşumu	33
Şekil 2.6. Mutasyon çeşidini belirleme	33

Şekil 3.1.	Hash fonksiyonu – çalışma zamanı grafiği	34
Şekil 3.2.	Hash fonksiyonu – FP oranı grafiği	35



TABLÖLAR DİZİNİ

	<u>Sayfa No</u>
Tablo 2.1. Örnek fasta formatı	27
Tablo 2.2. Alt diziler	27
Tablo 3.1. Filtre analizi	34
Tablo 3.2. FP olasılıkları	35
Tablo 4.1. Alt dizi karakteristiğine bağlı FP – zaman ilişkisi	36
Tablo 4.2. Hash fonksiyonu karakteristiğine bağlı FP – zaman ilişkisi	36

SEMBOLLER DİZİNİ

A	: Adenin
BRAF	: B-Raf proto-oncogene
C	: Sitozin
DNA	: Deoksiribonükleik asit
FP	: False Positive
G	: Guanin
GPU	: Grafik İşlem Birimi
K-mer	: Alt Dizi
M	: Mitoz
MAPK	: Mitogen-Activated Protein Kinases (Mitojen Aktive Edici Kinaz)
mRNA	: Mesajcı RNA
NCBI	: National Center for Biotechnology Information
p53	: Protein 53
PTC	: Papiller Tiroid Kanser
Raf	: V-raf Murine Sarcoma
RAS	: Reticular Activating System (Retiküler Aktive Sistemi)
Rb	: Retinoblastoma
RNA	: Ribonükleik asit
rRNA	: Ribozomal RNA
T	: Timin
TP53	: Tümör Protein 53
tRNA	: Taşıyıcı RNA
U	: Urasil

1. GENEL BİLGİLER

1.1. Giriş

Genetik verilerdeki küçük ölçekli mutasyonların tespit edilerek incelenmesi amacıyla yapılan çalışmada örnek DNA verileri üzerinde yazılımsal deneyler yapılarak başarılı sonuçlar alınması amaçlanmıştır. Mutasyon tespiti kanser ve benzeri genetik hastalıkların erken teşhisinde önemli rol oynamaktadır. Çalışma aşamalarında ilk olarak veri seti belirlenmiştir. BRAF mutasyonu kanser taramalarında erken tanı belirteci olarak kullanıldığından çalışma için uygun olarak seçilmiştir. NCBI gen bankası tarafından sağlanan sağlıklı ve mutasyonlu BRAF gen verileri, veri seti olarak kullanılmıştır. Mutasyonlu veriler evrensel bir gen dizilim dosyası olan fasta formatında tutulmuştur. Tespit işlemleri için, bir karşılaştırma veri yapısı olan özelleştirilmiş bir Bloom Filter kullanılmıştır. Bu sayede işlemlerin hızlandırılması amaçlanmıştır.

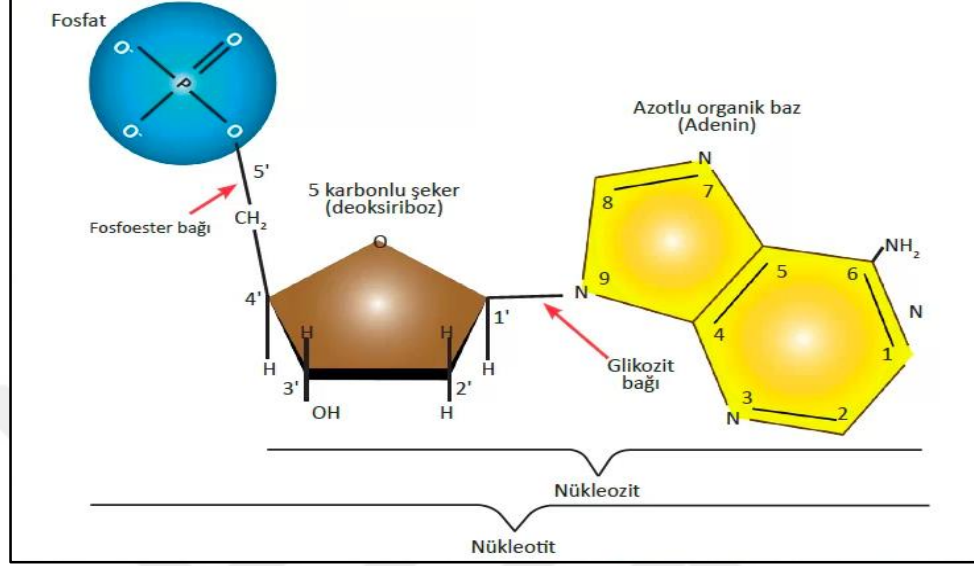
Çalışma kapsamında önerilecek yöntemin temel seviyede gen mutasyonlarının tespiti için yapılan biyoinformatik çalışmalara katkı sağlaması hedeflenmektedir.

1.2. Temel Kavramlar

1.2.1. Nükleik Asitler

Nükleik asitler canlı hücrelerinde kalıtımı sağlayan, metabolik olayların gerçekleşmesinde ve kontrolünde rol alan temel yapı taşı moleküllerdir. Temel görevi genetik bilgi aktarımını sağlamak olan nükleik asitler, nükleotit adı verilen basit birimlerden oluşmaktadır [1]. Nükleotitler, bir fosfat, bir pentoz (halkasal formdaki beş karbonlu şeker) ve bir azotlu organik bazdan oluşan kimyasal bileşiklerdir (Şekil 1.1). Nükleotitlerin yapısında bulunan A (Adenin) ve G (Guanin) bazları pürin; C (Sitozin), T (Timin) ve U (Urasil) bazları ise pirimidin bazları olarak adlandırılır. A, G ve C hem DNA hem de RNA'nın yapısında bulunur. T, yalnızca DNA'ya, U ise yalnızca RNA'ya özgüdür. Nükleotitler fosfodiester adı verilen kimyasal bağlar ile birbirlerine bağlanırlar. Nükleotitler

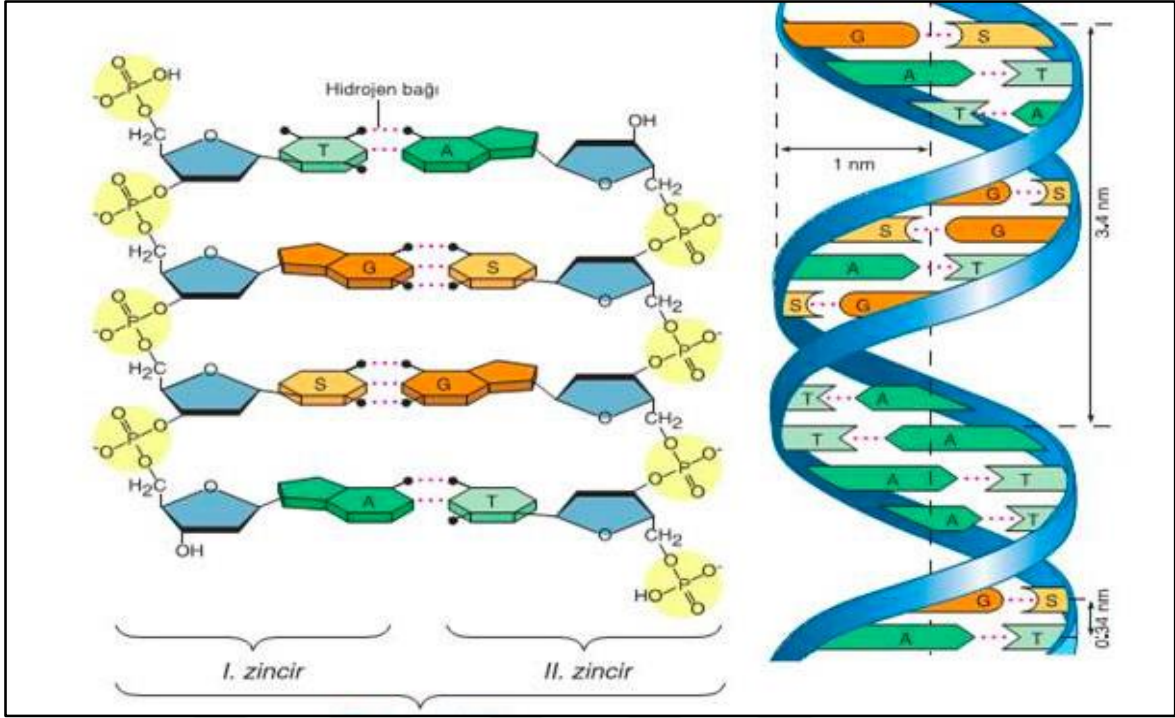
yapısında bulunan baz çeşidine göre isimlendirilir. Canlı hücrelerinde en yaygın olarak görülen nükleik asitler, DNA (Deokribonükleik asit) ve RNA (Ribonükleik asit)'dır [2].



Şekil 1.1. Nükleotit yapısı [3]

1.2.1.1. DNA

DNA, bir fosfat, bir azotlu baz ve bir deoksiribozdan oluşan kimyasal bileşiktir. DNA molekülü, çift zincirli sarmal bir yapıya sahiptir. Çift sarmal yapı, DNA'ya kendini eşleyebilme imkanı sağlamaktır [4]. DNA, yapısında bulunan A, T, G ve C bazlarının zayıf hidrojen bağlarıyla bağlanması ile oluşur [2]. Bu bağlanmada A ile G, T ile C eşleşmektedir. Moleküldeki A baz sayısı T baz sayısına, G baz sayısı ise C baz sayısına eşittir. A -T bağlantısında 2, G-C bağlantısında 3 adet zayıf hidrojen bağı kurulur. DNA bütün canlı hücrelerde bulunur ve kalıtsal bilgisiyi taşır. Bir türün farklı dokularındaki DNA'lar aynı baz yapısına sahiptir. DNA'nın baz içeriği türden türe farklılıklar gösterir. DNA, hücrelerde kromozom adı verilen yapıların içinde yer alır. DNA'nın anlamlı en küçük parçasına gen adı verilir [1].



Şekil 1.2. DNA molekül yapısı [5]

1.2.1.2. RNA

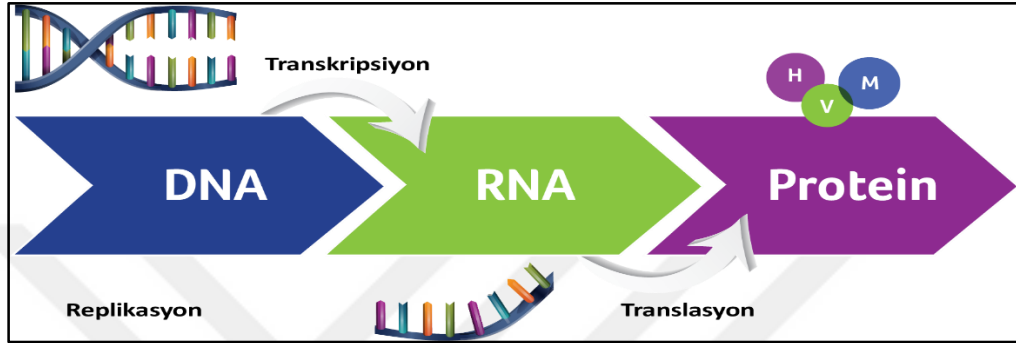
RNA, tek nükleotit zincirinden oluşan bir polinükleotittir. İçeriğindeki her bir nükleotit azotlu bir baz, bir fosfat ve bir riboz şekerden meydana gelir. RNA'da bulunan bazlar A, G, C ve U'dur. RNA genetik bilgi aktarımı işlemlerinde önemli görevlere sahiptir. Genetik bilginin aktarımı ve protein sentezinde görev alan üç çeşit RNA vardır.

Genetik bilginin aktarılması süreci replikasyon, transkripsiyon ve translasyon aşamalarından oluşmaktadır (Şekil 1.3). Bu süreçte mRNA (mesajcı RNA), tRNA (taşıyıcı RNA) ve rRNA (ribozomal RNA) olmak üzere başlıca üç çeşit RNA görev alır.

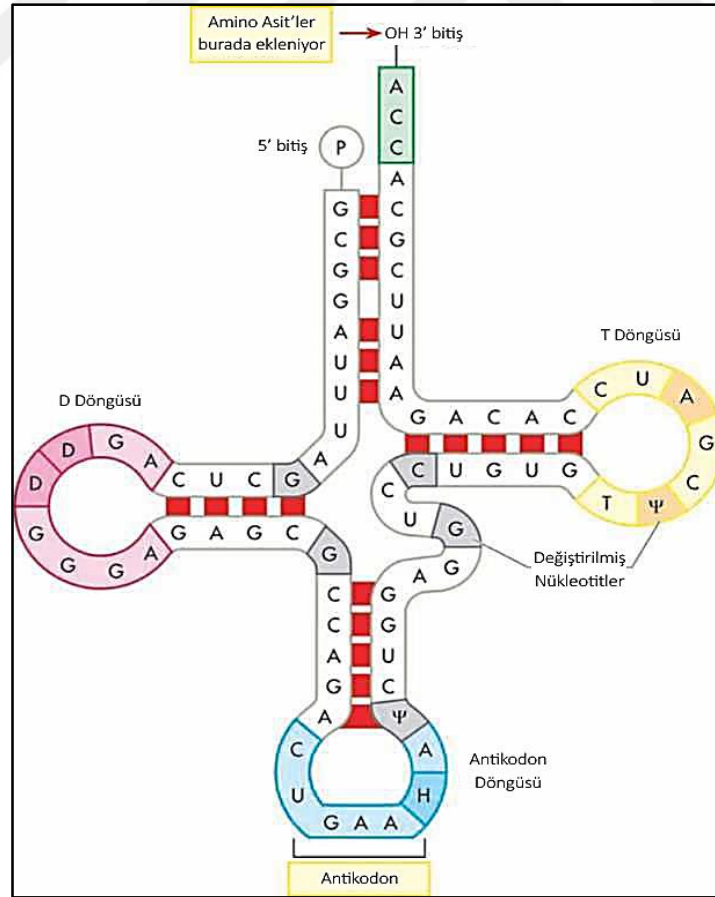
Replikasyon, DNA'nın kendini eşlemesine verilen addır. DNA replikasyonu gerçekleşeceği zaman çift sarmallı zincir yapı açılır. Açılan yapıdaki her bir sarmal DNA'nın kendini eşlemesi için bir şablon görevi görür. Eşleme tamamlandığında ilk DNA yapısının tamamen aynısı iki DNA molekülü elde edilmiş olur.

Transkripsiyon (yazılma), DNA şablonundaki nükleotit dizisine göre mRNA'nın sentezlenmesi işlemine verilen addır. Transkripsiyon hücre çekirdeğinde meydana gelir ve genetik bilgi mRNA şeklinde stoplazmaya taşınır.

Translasyon (okuma, çeviri) mRNA ile taşınan genetik bilginin tRNA tarafından okunarak protein sentezinin yapılması işlemidir. Translasyon ribozomda gerçekleşir. tRNA protein sentezi için gerekli olan aminoasitlerin taşınmasında görevlidir. Şekil 1.4'te gösterildiği gibi her bir tRNA'da aminositin bağlanacağı aminoasit bağlantı bölgesi ve kodonun bağlanacağı bir antikodon bölgesi bulunur. rRNA ribozomların yapısına katılır [4].



Şekil 1.3. Genetik bilgi aktarım aşamaları [6]



Şekil 1.4. tRNA'nın yapısı [7]

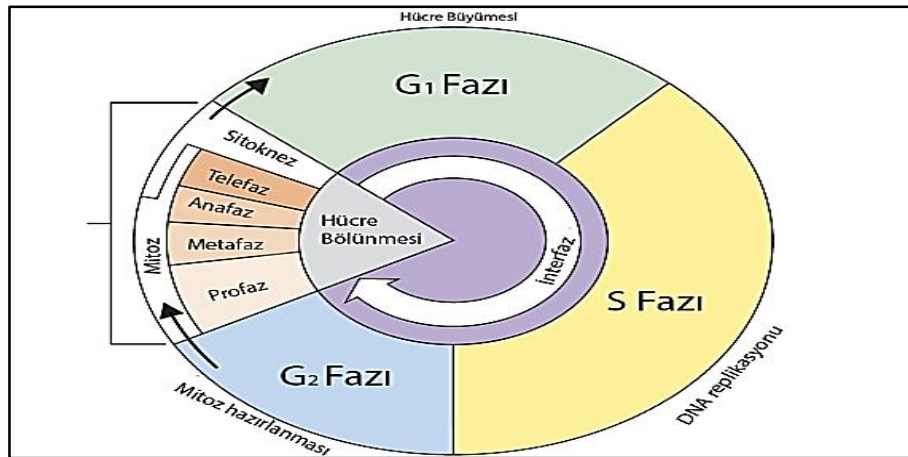
1.2.2. Hücre Siklusu

Canlılığın sürmesi için hücrelerin sürekli yenilenmesi gerekir. Bu yenilenme mekanizması bir yandan bölünüp çoğalırken bir yandan da ömrünü tamamlamış hücrelerin vücuttan atılması şeklinde olmaktadır. Genler tarafından kontrol edilen bu mekanizmada hücrelerin çoğalmasını sağlayan genler ve aşırı üremeyi engelleyen genler görev alır.

Devamlı bölünme gerçekleşen hücrelerde siklus G_1 , S ve G_2 fazlarının yer aldığı interfaz ve M (mitoz) safhalarından oluşmaktadır. (Şekil 1.5). Bu süreçte hücreler uyarılır ve büyürler. Hücreler bölünme sinyali almadığı süre boyunca istirahat fazı G_0 'da dururlar [8,9]. Hücre siklusunun %90'lık kısmını 16-24 saat süren interfaz oluşturur. Siklusun geriye kalan kısmını mitoz bölünme oluşturur ve 1-2 saat sürer.

Hücre büyümesi G_1 'de bulunan R point (kısıtlayıcı nokta) tarafından düzenlenir. Bu noktada siklus durur ya da tamamlanır [9]. Hücreler G_1 'de çevreden gelen sinyalleri alır ve büyüme başlangıcı için uyarı oluşturur. Bu aşamada DNA replikasyonu (sentez) için hazırlık yapılır fakat sentez gerçekleşmez. Protein ve RNA sentez işlemleri gerçekleşir. DNA sentezi ve kromozom duplikasyonu S fazında gerçekleşir.

Hücre G_2 fazında büyümesini sürdürür. RNA ve protein sentezi gerçekleşmeye devam eder. Hücre; P (profaz), M (metafaz), A (anafaz) ve T (telofaz)'den oluşan mitoz evresine hazırlanır. Telofazda stoplazmik bölünme tamamlanır. Bunu sonucunda genetik materyali aynı olan iki yeni hücre meydana gelmektedir [9,10,11,12]. Hücre siklusunda bulunan G_1/S , G_2/M ve metafaz-anafaz geçişlerindeki kontrol noktalarında siklusun devamına ya da tamamlanmasına karar verilir [9].

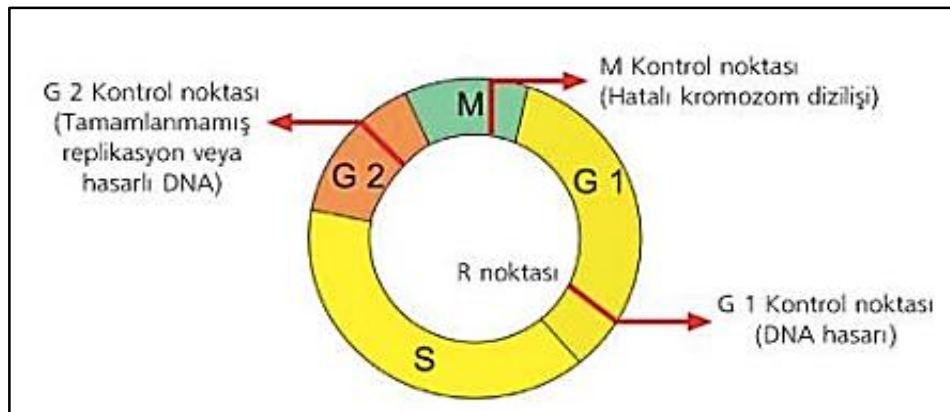


Şekil 1.5. Hücre siklusu [13]

G₁/S kontrol noktası: Restriksiyon noktası (kısıtlayıcı) olarak işlev görmektedir. Hücrenin G₁/S kontrol noktasında durmasını sağlayan p53 (Protein 53) genidir. DNA’da gerçekleşen bir hasar varsa hücre döngüsü bu kontrol noktasında durmaktadır. G₁ noktasında S fazına geçmeden önce hasarlı DNA’nın tamir edilmesine fırsat sağlanmaktadır (Şekil 1.2). Eğer onarım sağlanamazsa apoptozis (hücre ölümü) gerçekleşir. Bu fazda p53 geninin faaliyeti etkili olmaktadır. DNA hasarı olduğunda p53 üretimi artmaktadır. Artan p53, hücrenin G₁ fazında kalarak S fazına geçmesini engelleyen sinyaller oluşturmaktadır. İnsanlarda görülen kanserlerde p53’ü kodlayan genin çoğunlukla mutasyona uğramış olduğu tespit edilmiştir. Gerçekleşen mutasyonlar sonucunda p53 fonksiyon kaybına uğramakta ve hasar gören DNA’ya cevap olarak G₁’deki duraksama engellenmektedir [14]. Hasarlı DNA replike olarak onarılmadan yavru döllere aktarılmaktadır.

G₂/M kontrol noktası: Bu noktada DNA replikasyonu veya DNA sentezinin tamamlanması ile ilgili döngü ayarlanır. Replikasyonu gerçekleşmemiş veya hasar görmüş DNA burada algılanmakta ve döngüyü duraklatacak sinyaller oluşturulmaktadır. Siklus G₂ noktasında durursa hücre M aşamasına geçemez. Genomun replikasyonu tamamen ve kusursuz bir şekilde gerçekleşinceye kadar hücre G₂’de kalmaktadır. Replikasyon kusursuz bir şekilde gerçekleşince G₂’deki duraksama ortadan kalkmakta ve hücre M aşamasına geçmektedir. G₂/M kontrol noktasında meydana gelen bir defekt, kromozom anomalilerine sebep olmaktadır [15].

Mitoz (Metaz-anafaz) Kontrol Noktası: Bu noktada kromozomların iğ iplikleri üzerinde doğru konumlanıp konumlanmadığı kontrol edilmektedir. Mitotik iğ iplikleri üzerindeki bir veya daha fazla kromozomun eksik olduğu tespit edildiğinde kontrol noktası hücre döngüsünü durdurmaktadır.



Şekil 1.6. Hücre siklusu kontrol noktaları [16]

Hücre siklusunda bu fazlar sırasıyla olur. Bu fazlardan biri atlanıp diğerine geçilirse, genom bilgisi eksiksiz ve doğru kopyalanamayacağından hücrede hasar oluşabilir [14]. Kanser hücrelerindeki genetik instabilitenin en önemli nedenlerinden biri siklus kontrol noktalarında oluşan defektlerdir [15].

Herhangi bir etkenle hücre DNA'sında oluşan hasarın çeşidi ve boyutuna göre kontrol noktaları faz geçişlerini (G₁-S veya G₂-M) engeller (Şekil 1.2). DNA'da oluşan hasar DNA sentezini inhibe edebilir. Hücrelerde DNA replikasyonu gerçekleşmediğinde mitoz giriş, kinaz komplekslerinin inaktivasyonu ile engellenir [9].

Onkogenler ve tümör baskılayıcı genler, hücre siklusunda önemli bir yere sahip olan gen guruplarıdır [10].

Onkogenler: Kontrolünü kaybetmiş, protein kodlayan genlerdir. Doğrudan veya dolaylı olarak kanser gelişimini etkilemektedirler.

Proto-onkogenler: Artmış gen ekspresyonu ya da mutasyonlar sebebiyle onkogene dönüşebilme potansiyeline sahip olan normal genlerdir. Hücrenin büyüme ve bölünme aşamalarında kullanılan proteinlerini kodlar. Bu genlerde gerçekleşen mutasyonlar sonucunda tümör gelişebilmektedir [17].

Tümör Baskılayıcı Genler: Hücre büyümesini baskılayıcı etkileri ile kanser gelişimini baskılayan genlerdir. Bunlar p53 ve Rb (Retinoblastoma) genleridir. Tümör baskılayıcı genlerin kaybı kanser oluşumunda onkogen aktivasyonundan daha önemli bir role sahiptir. Normal bir hücrede DNA hasarı meydana geldiğinde p53 devreye girerek genomik kararlık oluşturur. Siklusu G₁ fazında inhibe eder ve hücreye tamir için gerekli olan zamanı sağlar. Hasar onarılamıyorsa hücreyi apoptozise götürür [9,14,18]. Kalıtsal mutasyon gerçekleşmiş DNA onarım genlerinde kanser ihtimali bulunmaktadır [11]. Rb, normal hücrelerde siklusu G₁'de inhibe eder. Bu genlerde gerçekleşen mutasyonlar hücre siklusunda inhibisyonu engeller ve normal olmayan hücre büyümesine sebep olurlar [17]. Ayrıca apoptozise karşı direnç gelişimine ve neoplastik farklılaşma ile kanserleşmenin başlamasına neden olurlar. p53 geni işlev kaybına uğrarsa hücre büyümesi kontrol edilemez. DNA tamiri olmaksızın hücre siklusu kontrolsüzce gerçeklemeye devam eder. Kanserogenezin sürmesi için, tümör baskılayıcı bir genin işlev kaybı çoğunlukla iki alelde de mutasyon gerçekleşmesini gerektirir [19].

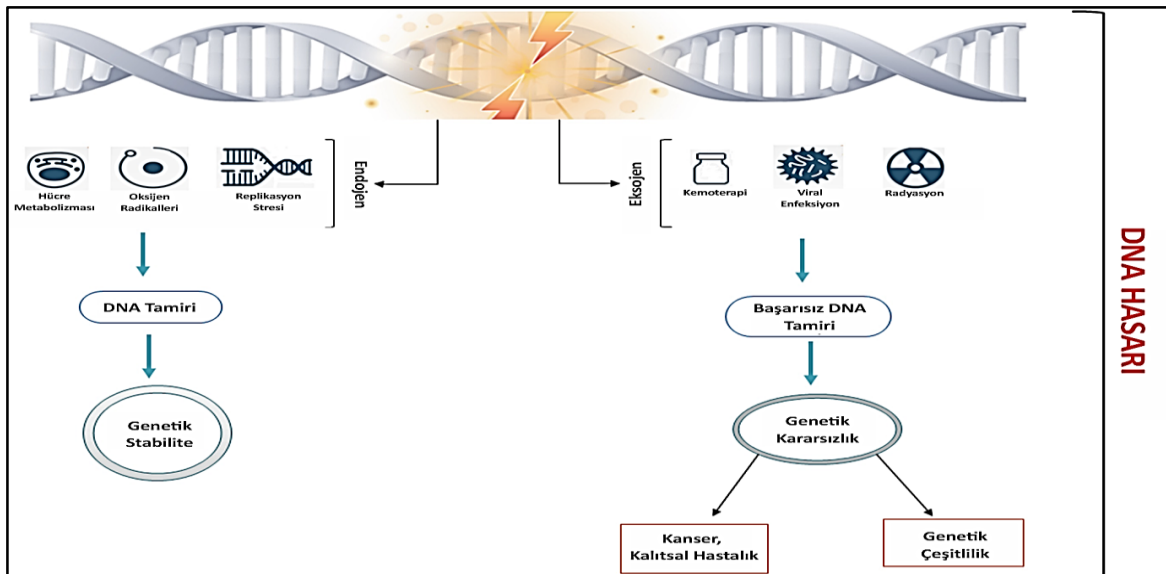
1.3 Mutasyon

Mutasyon, ilk olarak 1901 yılında bitkiler üzerinde yapılan çaprazlamalarda gözlemlenen bir varyasyonu isimlendirmek için kullanılmıştır. Mutasyon, bir canlının genetik materyalindeki nükleotit baz dizilerinde gerçekleşen kalıtsal değişikliklerdir. Bu değişiklikler somatik (vücut) hücrelerde ya da gamet (üreme) hücrelerinde olabilir. Mutasyonlar fenotipte pozitif veya negatif etkilere sebep olabilir ya da nötr kalarak fenotipte herhangi bir değişiklik yaratmayabilir. Mutasyonlar çoğunlukla DNA'nın replikasyonu ya da onarımı sırasında meydana gelen hatalarla ortaya çıkar. Genetik çeşitliliğin ana kaynağını oluşturan bu mutasyonlar genetik çalışmaların da temelini oluşturmaktadır [20]. Bir çok mutasyon sınıflandırma çeşidi bulunmaktadır.

Mutant: Mutasyona uğramış bir organizmaya mutant adı verilir. Mutant bir alel, değişmiş genetik bilgi içerir ve gen ürününü değiştirir.

Mutajen: Mutasyona neden olan etkene denir. Radyasyon, sıcaklık, X ışınları, bazı kimyasal maddeler, travma ve sigara katranı birer mutajendir.

Tüm organizmaların, hücreleri dışarıdan gelecek hasarlara karşı korumak için geliştirdiği bir DNA onarım mekanizması vardır. Bu mekanizma hücre ölümü, mutasyon, replikasyon hataları ve genomik kararsızlık gibi bir çok durumda devreye girer. DNA onarım mekanizmasının çalışmasındaki bir anormallik kansere ve yaşlanmaya yol açar (Şekil 1.7) [21]. Normal bir insan hücresinde replikasyon sırasında oluşan hata oranı 10^{-10} 'dur [22].



Şekil 1.7. DNA tamir mekanizması [23]

1.3.1. Bulundukları Yere Göre Mutasyonlar

Germ Hücre Mutasyonları: Gamet (üreme) hücrelerindeki mutasyon sonraki nesillere aktarılmaktadır. Bu mutasyonlar genetik hastalıkların, genetik çeşitliliğin ve evrimin temelini oluşturmaktadır.

Somatik Hücre Mutasyonları: Vücut hücrelerinde meydana gelir. Bu mutasyon, kansere neden olabildiği için önemlidir. Hücresel fonksiyon değişikliklerine, lokal hücre ölümüne ya da tümör oluşumuna yol açabilirler. Gelecek nesillere aktarılmazlar.

1.3.2. Oluşum Mekanizmasına Göre Mutasyonlar

Spontan Mutasyonlar: Herhangi bir mutajenik etken olmadan doğal olarak oluşan mutasyonlardır. Nükleotit dizilerinde rastgele değişiklikler oluşur. Replikasyon hatalarından ve baz modifikasyonlarından kaynaklanmaktadır.

Uyarılmış Mutasyonlar: Doğal ya da yapay ajanların etkilemesi ile oluşan mutasyonlardır. Radyasyon ve kimyasalların neden olduğu DNA hasarlarından oluşur. Mantar toksinleri, UV ışınları ve kozmik ışınlar bu mutasyona neden olan diğer ajanlardır.

Adaptif (Uyumsal) Mutasyon: Organizmanın, çevreden gelen belirli bir etkene uyum sağlamak amacıyla geliştirdiği mutasyon çeşididir [24].

1.3.3. Fenotipik Etkilerine Göre Mutasyonlar

Morfolojik Mutasyon: Organizmanın dış görünüşünü etkileyen mutasyonlardır.

İşlev Kaybı Mutasyonu: Gen ürününün işlevini kısmen ya da tamamen yok eden mutasyondur. Bu mutasyonlar genellikle resesif (çekinik)'tir. Çünkü bir genin işlev kaybını başka bir alel telafi edebilir [25,26].

İşlev Kazancı Mutasyonu: Gen ürününün yeni bir işlev kazanmasına sebep olan mutasyondur.

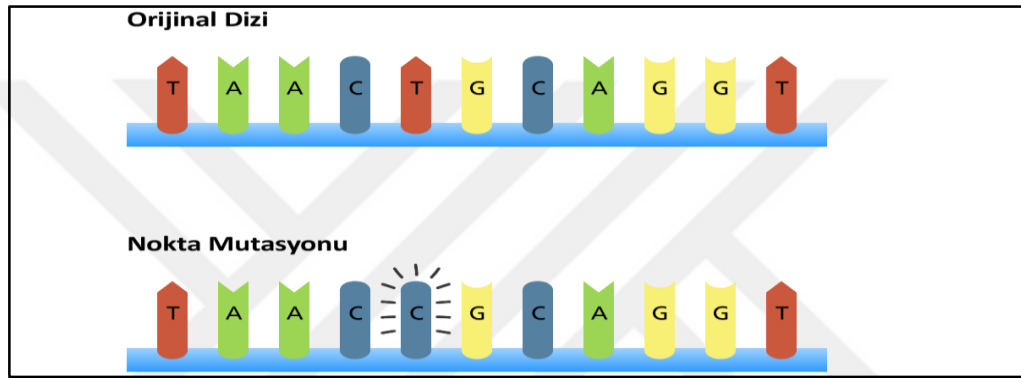
Nötral Mutasyon: Yeni gelen aminoasit öncekiyle aynı kimyasal özelliklere sahiptir. Genellikle gen ürününü etkilemez [27].

Letal (Öldürücü) Mutasyon: Hayati öneme sahip genlerde meydana gelerek ölüme yol açan mutasyon çeşididir.

1.3.4. Moleküler Değişiklik Tipine Göre Mutasyonlar

1.3.4.1. Nokta (Gen) Mutasyonları

Gen dizisinde bir veya birkaç nükleotitte meydana gelen değişikliklerdir. Bu değişiklikler sonraki nesillere aktarılabilir. Kendiliğinden ya da deneysel olarak meydana gelebilir. Kromozomların sayısında veya yapılarında herhangi bir değişiklik olmaz. Genelde replikasyon sırasında oluşur.



Şekil 1.8. Nokta (gen) mutasyonu [28]

1.3.4.1.1. Baz Çifti Değişikliği Mutasyonları (Substitution Mutations)

Transisyonel Mutasyon (Pürin-pürin veya pirimidin-pirimidin): Bir pürin diğer bir pürin ile veya bir pirimidin diğer bir pirimidin ile yer değiştirirse buna transisyon denir.

Normal gen dizisi (A-T)

T	G	T	G	A	A	A	A	T	T	C	T	A	G
A	C	A	C	T	T	T	T	A	A	G	A	T	C

Mutant gen dizisi (G-C)

T	G	T	G	A	A	G	A	T	T	C	T	A	G
A	C	A	C	T	T	C	T	A	A	G	A	T	C

Şekil 1.9. Transisyonel mutasyon

Transversiyonel Mutasyon (Pürin-pirimidin veya pirimidin-pürin): Bir pürin ve pirimidin karşılıklı olarak yer değiştirirse transversiyon olur.

Normal gen dizisi (C-G)

T	G	T	C	A	G	A	A	T	T	C	T	A	G
A	C	A	G	T	C	T	T	A	A	G	A	T	C

Mutant gen dizisi (G-)

T	G	T	G	A	G	A	A	T	T	C	T	A	G
A	C	A	C	T	C	T	T	A	A	G	A	T	C

Şekil 1.10. Transversiyonel mutasyon

1.3.4.1.2. Sessiz, Yanlış Anlamlı ve Anlamsız Mutasyonlar

Silent (Sessiz) Mutasyon: Gen ürünün yapı ve işleyişini değiştirmede için fenotipte gözlenebilir bir değişiklik yaratmazlar. Yeni oluşan baz, genin kodlayacağı aminoasiti değiştirmez. Örneğin UAC kodonunun UAU'ya dönüşümü bir silent mutasyondur. Çünkü üçüncü bazdaki C yerine U gelmesi kodlanacak aminoasitte bir değişiklik yaratmaz. Her ikisi de tirozin aminoasitini kodlar [4].

Missense (Yanlış Anlamlı) Mutasyon: Kodondaki anlamın değişmesi ile protein yapısına farklı bir aminoasit katılmasına sebep olan mutasyondur. Örneğin UAU kodunu UCU kodonuna dönüşebilir. Bu durumda protein yapısına tirozin yerine serin aminoasiti katılır [4]. Böylece olması gerekenden farklı bir protein meydana gelir.

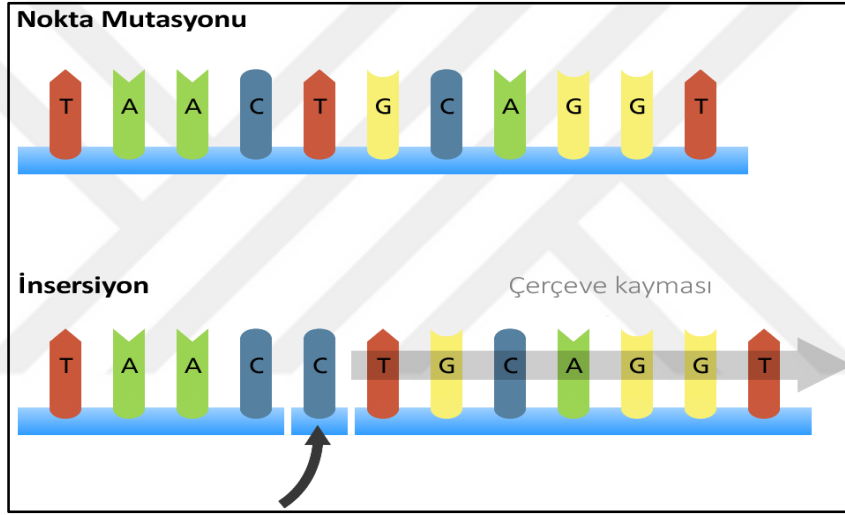
Nonsense (Anlamsız) Mutasyon: Bir kodonun hiçbir aminoasiti kodlamayıp sonlanma kodonu (UAA, UAG veya UGA) ya da başka birşeye dönüşmesine neden olan mutasyondur. Bu mutasyon zincirin erken sonlanmasına doğal olarak da daha kısa bir protein oluşmasına neden olur.

Nokta Mutasyonları				
Mutasyonsuz	Sessiz	Anlamsız	Yanlış Anlamlı	
			Konservatif	Anti-Konservatif
TTC	TTT	ATC	TCC	TGC
AAG	AAA	UAG	AGG	ACG
Lys	Lys	STOP	Arg	Thr

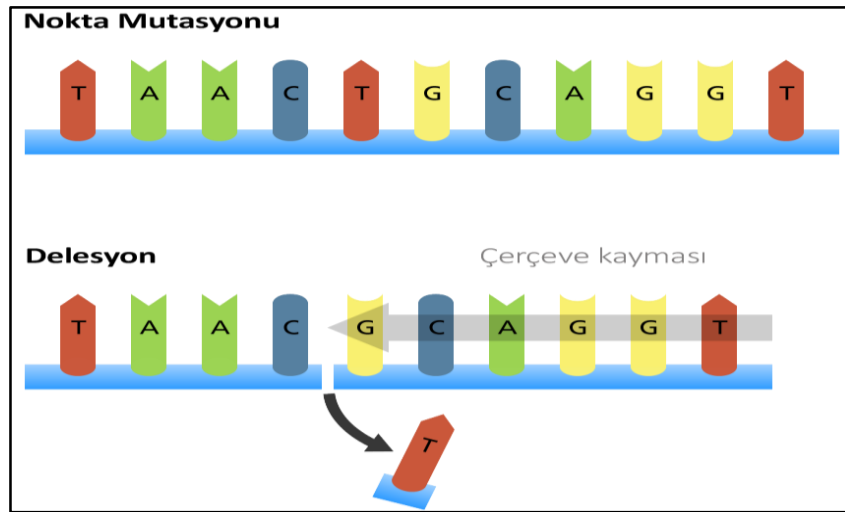
Şekil 1.11. Sessiz, yanlış anlamlı ve anlamsız mutasyonlar [29]

1.3.4.1.3. Çerçeve Kayması (Frameshift) Mutasyonları

Bir genin protein kodlayan bölgesinde, bir ya da daha fazla sayıda nükleotit girişi veya çıkışı ile kodonların değişmesine neden olan mutasyonlardır. Nükleotit girişi olduğunda insersiyon, çıkışı olduğunda ise delesyon olarak adlandırılmaktadır. Bu değişiklikle, okuma esnasında üçlü kodonlar ve tüm aminoasit dizisi aşağı doğru kayar. Bunun sonucunda normal proteinden apayrı yapıda işlevsiz bir protein oluşur. Bu mutasyonda bazen bir noktada sonlanma anlamına gelen bir kodon ortaya çıkabilir ve zincirin erken sonlanmasına neden olabilir [21, 22].



Şekil 1.12. Çerçeve kayması (insersiyon) [28]



Şekil 1.13. Çerçeve kayması (delesyon) [28]

1.3.4.2. Kromozom Mutasyonları

Normal (diploit) bir hücrede 46 adet kromozom bulunur. Bu kromozomlar 23 çift halinde düzenlenmiştir. Bu çiftleri oluşturan iki kromozomdan biri anneden diğeri babadan gelir. 23 adet kromozoma haploid (n) adı verilir. $2n$ kromozoma sahip olma durumuna diploit denilmektedir. Kromozom sayısının haploid ya da haploidin tam katları halinde artmasına euploidi denir. Kromozom mutasyonları mitoz ve mayoz bölünme sırasında oluşurlar. Kromozomlar her zaman düzenli ayrılmayabilir. Bu sebeple kromozomların yapısında veya sayısında değişiklikler olabilir. Kromozom sayısındaki bu değişiklikler heteroploidi olarak adlandırılmaktadır.

1.3.4.2.1. Sayısına Göre Kromozom Mutasyonları

Aneuploidi: Genomdaki bazı kromozomların sayıca eksilmesi ya da çoğalmasdır. Bu durumda haploidin tam katı oluşmaz. Bir kromozom fazla olduğunda trizomi, bir kromozom eksik olduğunda ise monozomi olarak adlandırılır.

- Monozomi: Diploid yapıdan bir kromozomun eksilmesidir.
- Trizomi: Diploid yapıya bir kromozomun eklenmesidir.

Poliploidi: $2n$ 'den fazla kromozom takımının bulunması durumuna denir. $3n$ triploidi, $4n$ tetraploidi olarak adlandırılır.

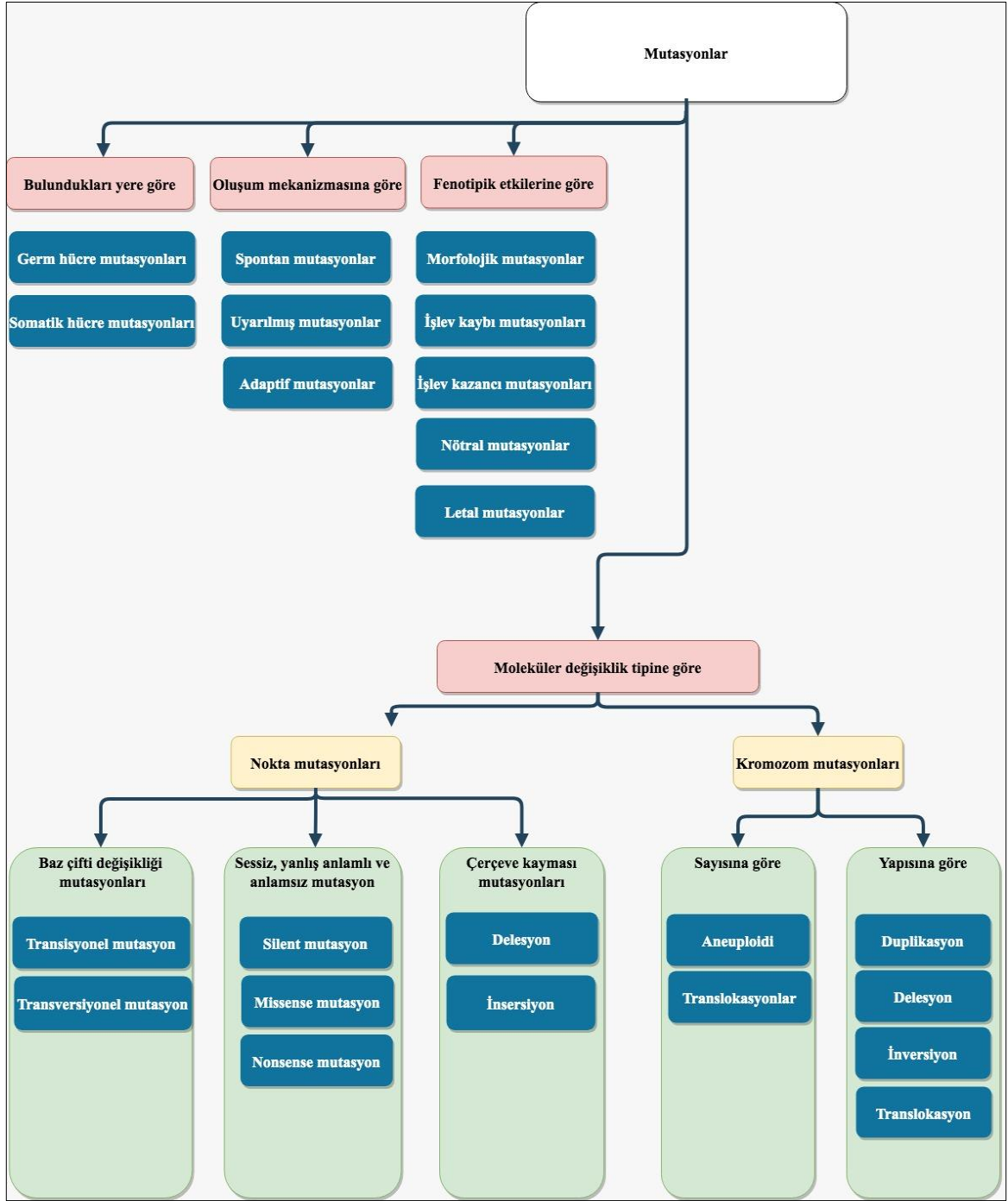
1.3.4.2.2. Yapısına Göre Kromozom Mutasyonları

Duplikasyon: Kromozomdaki bir segmentin çiftlenmesi olayına denir.

Delesyon: Kromozomun herhangi bir bölgesinde meydana gelen kopma ve ayrılmalardır.

İnversiyon: Kromozomun bir bölgesinin ters dönüp tekrar aynı kromozoma bağlanması durumudur.

Translokasyon: Kromozom segmentleri ve onların taşıdıkları gen zincirlerindeki yerleşimlerde görülen değişimlerdir.



Şekil 1.14. Mutasyonların sınıflandırılması

1.4. Karsinogenez (Kanser Oluşumu)

Hücreler endojen ve eksojen kaynaklı birçok etkene maruz kalır. Bu etkenler kanser olarak adlandırdığımız kontrolsüz hücre büyümesine sebep olurlar. Kanser ortaya çıkması

sadece kontrolsüz çoğalmaya bağlı değildir. Hücrede invazyon ve metastaz gibi diğer malignite özelliklerinin de bulunması gereklidir [30].

Tümör oluşumundaki iki temel aşama, neoplastik transformasyon ve neoplastik gelişim adımlarıdır. Normal hücrelerin neoplastik hücrelere dönüşümü olan neoplastik transformasyon, onkogenezis ve karsinogenezisin ilk basamağını oluşturur [31]. Neoplastik transformasyon, mutasyonların etkisiyle DNA replikasyon ve tamir mekanizmalarının düzgün çalışmaması sonucunda oluşur [11]. Mutasyonlar, proto-onkogenlerde, tümör baskılayıcı genler ile bu genleri düzenleyen genlerde ve büyüme kontrolü yapan genlerde oluşabilir [32,33]. Neoplastik gelişim ise kanser hücrelerinin büyüyüp gelişerek önce çevre dokulara sonrasında da uzak dokulara yayılarak büyümesidir. Sürekli büyüme ve çoğalma ile tümörde malignansiye varan değişiklikler meydana gelir [34].

Kanserler DNA dizisindeki birtakım anormalliklerle oluşur. Kanserlerin küçük bir bölümü kalıtsal nedenlerle olurken; %85-%90'lık kısmı canlı hücrelerin yaşamları süresince mutajenlere maruz kalması sonucunda hücre DNA'sında hatalar ve değişiklikler gerçekleşmesi ile ortaya çıkmaktadır [11].

Kanserin sebebinden bağımsız olarak hücreye ait genetik materyalde bozulma gerçekleşir. Tek bir gende mutasyon olabileceği gibi aynı anda birçok gende de mutasyon görülebilmektedir. Birden fazla gende aynı anda meydana gelen mutasyon hasarları, sadece bir gende oluşan mutasyona göre kanser oluşumunda daha fazla rol alır [11].

Gen yapısındaki bozulmalar ve değişiklikler normal hücrelerdeki yapım ve yıkım mekanizmasında bozulmalara neden olur. Hücrelerdeki bu biyolojik düzenin bozulması ve tümör oluşumuna neden olan genetik hasarların çeşitli nedenleri vardır. Bu nedenler aşağıda maddeler halinde belirtilmiştir.

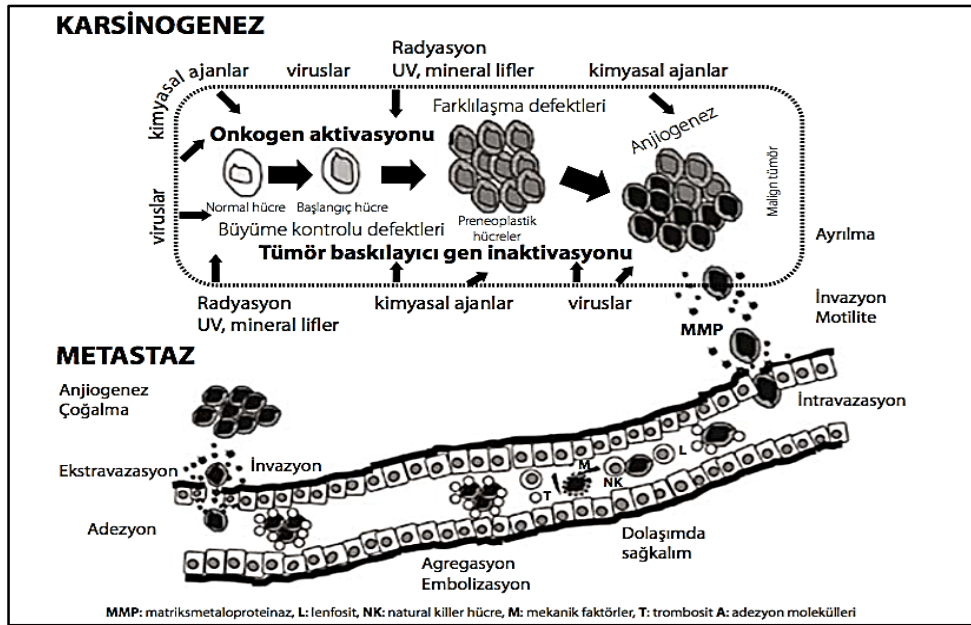
Kanser gelişimine sebep olan endojen faktörler;

- DNA'daki rastgele değişimler
- Oksidatif metabolizma
- Somatik rekombinasyon mekanizması (VDJ)

Kanser gelişimine sebep olan eksojen faktörler;

- Yağlı yiyecekler
- İlaçlar
- Sigara
- Alkol
- Benzen

- Kömür
- Kimyasal gazlar
- İyottan fakir besinler
- Radyasyon
- Güneş ışığı
- Isı
- Kronik travma
- Virüsler [35]
- EBV (Epstein Barr Virüsü)
- HCV (Hepatit C Virüsü)
- HBV (Hepatit B Virüsü)
- HPV (Human Papilloma Virüs)
- MCV (Merkel Hücreli Polyoma Virüsü)
- HTLV (İnsan T-lenfotropik Virus)
- KSHV (Kaposi Sarkom İlişkili Herpes Virus)
- Bakteriler [36]
- Helicobacter pylori
- Chlamydia trachomatis



Şekil 1.15. Karsinogenezin oluşumu [37]

1.5. p53 Geni

p53, 17p13.1 kromozom bölgesinde lokalize olan ve 393 aminoasitten oluşan bir proteini kodlayan gendir. Normal koşullarda hücrede düşük seviyededir ve normal hücre siklusunda görev almaz. Normal bir hücrede DNA hasarı, iğ iplik hasarı, hipoksi durumu veya onkogenlerle p53 seviyesi yükselir ve p53 aktivasyonu sağlanır.

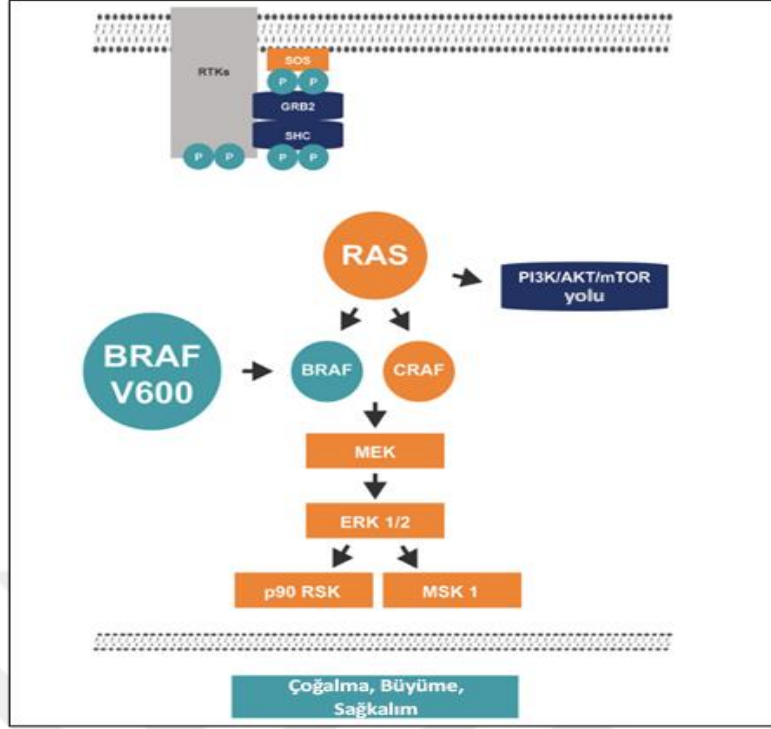
p53, herhangi bir hasar durumunda DNA onarımının yapılmasına zaman kazandırmak için hücre siklusunu G₁ fazında durdurur. p53 apoptoziste anahtar rol oynamanın yanı sıra farklılaşma ve yaşlanma olaylarında da görev alır [38].

p53 geni mutasyona uğramış hücrelerde tümör baskılayıcı genlerde kayıplar olur ve hücre bölünmeye devam eder. Bunun sonucunda da kanserleşme oluşur. İnsanlarda meydana gelen kanserler içerisinde görülme oranı en yüksek olan mutant gen p53'tür. Bu kanser çeşitlerinin %50'sinden fazlasında p53 mutasyonu olduğu bildirilmiştir [39,40]. Mutasyon olmayan alelde ise p53 kaybı söz konusudur.

p53'te görülen mutasyonlar; nokta mutasyonu, missense mutasyon, nonsense mutasyon ve silent mutasyon olarak rapor edilmiştir. Mutasyonlar çoğunlukla p53'ün DNA bağlanma bölgesinde görülmektedir [41].

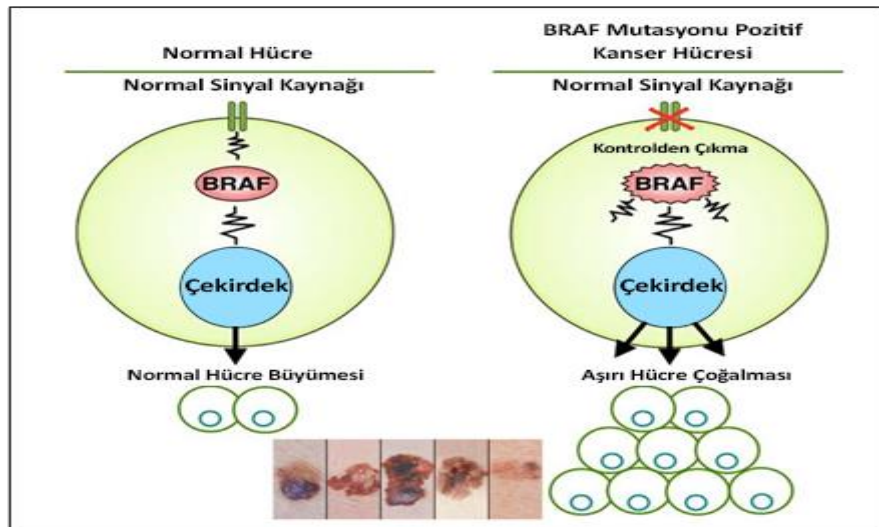
1.6. BRAF Geni

Onkogenler sınıfına ait olan BRAF, MAPK (mitojen aktive edici protein kinaz) mekanizması içerisinde bulunan ve hücre dışından gelen sinyallerin hücrenin çekirdeğine iletilmesini sağlayan proteinin üretim talimatlarını belirleyen gendir [42]. RAS (retiküler aktive sistemi)/MAPK yolu, hücrelerin büyümeyi, bölünmeyi (çoğalmayı), farklılaşmayı, hareketi(göç) ve apoptozu gerçekleştirmek için geçirdiği süreci düzenler. BRAF, 18 ekzondan oluşur. Reseptör tirozin kinaz yolunda RAS'tan sonra Raf (v-raf murine sarcoma) gelmektedir. Üç adet kinaz içerisinde en güçlü olan BRAF'tır [43].



Şekil 1.16. Ras / MAP kinaz yolu [44]

BRAF geninde oluşan mutasyonlar yoğunlukla 15.ekzonda bulunan 600. kodonda ve 11. ekzonda oluşur. Bu gendeki mutasyonlar en sık 600. kodonda görülür [45]. Bu mutasyonların BRAF üzerinde V600E aminoasit değişikliği yapması BRAF kinazda devamlı ve onkojenik aktivasyona neden olur [46]. BRAF mutasyonu tanı belirteci olarak kullanılmaktadır [47].



Şekil 1.17. BRAF V600 mutasyonu [48]

Braf genindeki mutasyonlar farklı hastalıklara ve malignitelere yatkınlık oluşturur. Over tümörleri kadınlarda en sık görülen kanser türlerinden biridir (Şekil 1.14). Kadın genital sistem kanserlerinin %30'unu oluştururlar [15]. Over kanserinin erken dönemde teşhis oranı düşük olduğundan hastalık teşhis sıklığı ile ölümler arasında orantılı olmayan bir risk söz konusudur. Over kanserleri, kadınlarda genital sistem kanserlerine bağlı ölümlerin %50'sini oluştururlar. Şekil 1.15'te 2009–2013 yılları arasında İngiltere'de yumurtalık kanseri tanısı konulan hastalar için yaşa bağlı hayatta kalma oranları verilmiştir. Over kanserlerinde nulliparite, aile öyküsü ve kalıtsal mutasyonlar risk faktörleri arasındadır ve tümör gelişine neden olurlar [49].

BRAF mutasyonları seröz borderline over tümörlerin üçte birinde, görülür. Daha az oranda da düşük dereceli seröz over tümörlerinde görülür. Düşük dereceli seröz kanserler KRAS, BRAF ve ERBB2 onkogenlerinde mutasyonlara sahiptir. Bu kanserler seröz borderline tümörlerden kademeli olarak gelişir [50]. BRAF ve KRAS seröz karsinomlarda tümör ilerlemesinde morfolojik bulgular gözükmeden çok önceki aşamalarda rol alır.

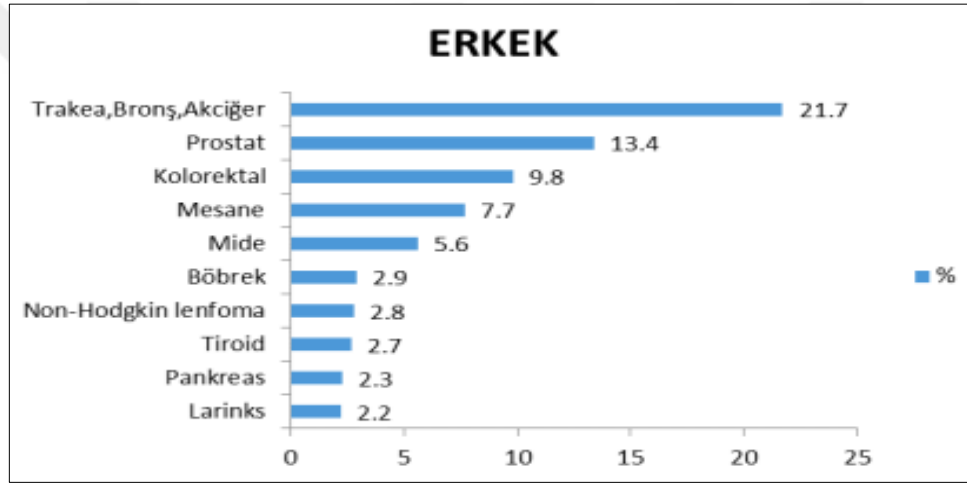
BRAF genindeki mutasyonlar malign melanomların üçte ikisinde görülürken, kolorektal ve overyane karsinomların daha az bir bölümünde görülür. Yüksek dereceli seröz over tümörlerinde ise TP53 (Tümör protein 53) mutasyonu görülür. Bu karsinomların %96'sının TP53 mutasyonuna sahip olduğu ifade edilmiştir [51]. Malign melanomalı olguların çoğunda BRAF geni V600E mutasyonu tanımlanmıştır. Kolorektal kanserler dünyada en yaygın görülen malignitelerden biridir (Şekil 1.13, Şekil 1.14). Kolorektal kanserlerde BRAF geninde çeşitli mutasyonlar gözlenmiştir. Kolorektal kanserlerin yaklaşık %9-11'inde BRAF mutasyonu görülmektedir. Kanser öncüsü bazı poliplerin içerisinde farklı oranlarda BRAF mutasyonu tespit edilmiştir. Tiroid kanseri endokrin maligniteler içerisinde en sık görülenidir ve tüm malignitelerin %1 kadarını oluşturur [52]. PTC (Papiller tiroid kanser), tüm tiroid malignitelerinin yaklaşık %85'ini oluşturmaktadır [53]. Sağ kalım oranı yüksek bir kanser çeşidi olsa bile hastalığın nüks etme olasılığı oldukça yüksektir. Nüks görülen hastaların küçük bir kısmı ölmektedir [54].

PTC, BRAF mutasyonlarının en sık görüldüğü hastalıklardan biridir [54]. PTC'de BRAF, RAS, RET, PAX8 gibi birçok gende mutasyonlar tespit edilmiştir [52]. PTC'deki en yaygın mutasyon BRAF V600E mutasyonudur. Bu mutasyon PTC'de iyi bir prognoz ve risk sınıflandırması elde etmek için önemli bir belirteç olarak ortaya çıkmıştır [55,56]. Nükseden PTC'de majör mutasyon BRAF mutasyonudur [57]. Yapılan çalışmalarda nükseden PTC'lerde BRAF mutasyonu daha yaygın olarak gözlenmiştir (nükseden tümörlerde %85, primer

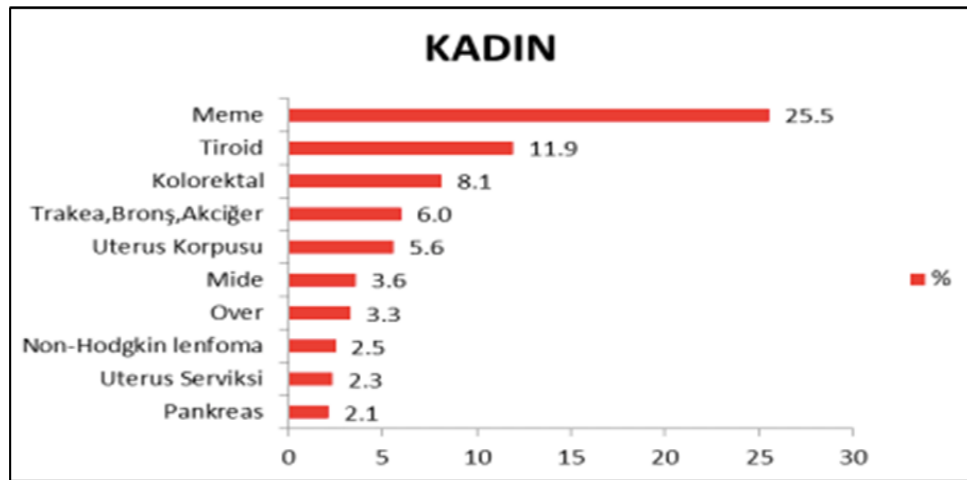
tümörlerde %65 BRAF mutasyonu pozitif olarak gözlenmiştir) [57,58]. BRAF mutasyonu, PTC nüks riskini belirlemek için preoperatif olarak kullanılabilecek faydalı bir göstergedir. Bu da kişiye özel tedavi seçenekleri oluşturulmasına yardımcı olabilir [54,59,60].

BRAF mutasyonunun daha kötü prognozlu, daha agresif bir hastalığın bağımsız bir belirteci olduğu gösterilmiştir [61]. Bu mutasyonun PTC için daha kötü bir prognoza neden olduğu tespit edilmiştir [62]. İyi huylu PTC’lerde BRAF mutasyonu gözlenmez [63].

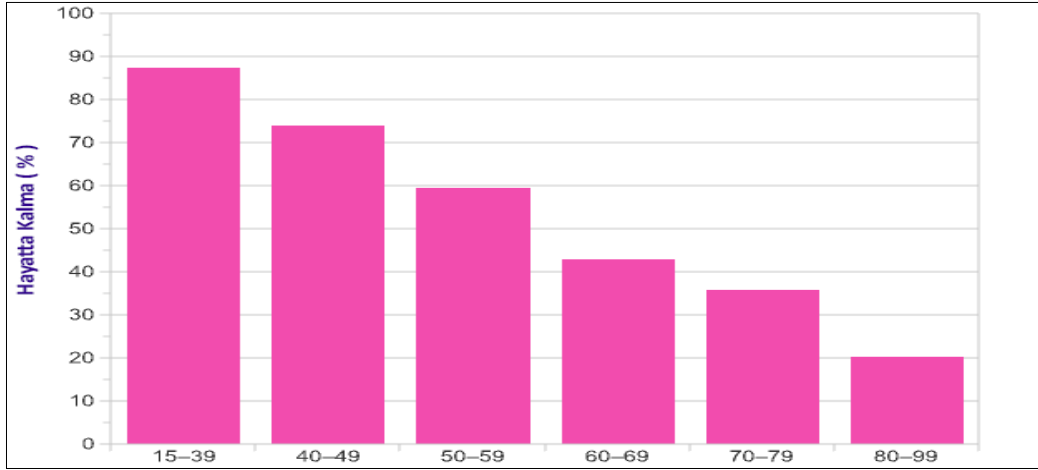
Diğer genetik değişimler gibi BRAF mutasyonu da bir hastalığın erken veya ilk dönemlerini gösterebilir. p53’te oluşan mutasyon gibi diğer bazı mutasyonlarda büyük ihtimalle geç dönemleri gösterir [64].



Şekil 1.18. Erkeklerin tüm yaşlarda görülen kanser türü oranları, Türkiye [65]



Şekil 1.19. Kadınların tüm yaşlarda görülen kanser türü oranları, Türkiye [65]



Şekil 1.20. Over kanseri, yaşa göre beş yıllık hayatta kalma (İngiltere, 2009 – 2013) [66]

2002 yılında Davies ve arkadaşları tarafından, kanser hastalıklarında BRAF gen mutasyonu varlığının araştırıldığı bir çalışma yayınlanmıştır [67]. Çalışmada 530 kanserli hücre ve 378 primer kanser örneği incelenmiştir. Kanser çeşitlerine göre BRAF gen mutasyonuna sahip olan hücre sayıları ve yüzdelik oranları Şekil 1.21’de gösterilmiştir.

Kanser türü	Örnek sayısı (adet)	Mutasyon (+)(adet)	Oran (%)
<i>Kanser Hücre serisi</i>			
Melanom	34	20	59
Kolorektal	40	7	18
Glioma	38	4	11
Akciğer	131	4	3
Sarkom	59	5	9
Over	26	1	4
Meme	45	1	2
Karaciğer	7	1	14
<i>Primer Kanser</i>			
Melanom	9	6	
Kolorektal	33	4	
Over	35	5	
Sarkom	182	1	

Şekil 1.21. Kanser türlerinde BRAF gen mutasyonu oranı

1.7. DNA Dizileme

DNA’daki nükleotit bazlarının kesin sırasının belirlenmesine DNA dizileme adı verilmektedir. İnsan genomundaki milyarlarca baz çiftinin dizilimindeki farklılıklar, her bireyin eşsiz genetik yapısını ortaya koymaktadır. Moleküler biyolojinin hızlı gelişimiyle

DNA çalışmaları ön plana çıkmış ve geçmişten günümüze kadar birçok DNA dizileme yöntemi geliştirilmiştir. Gelişen teknolojinin de yardımıyla çok sayıda çalışma yapılmakta ve yeni yöntemler ortaya koyulmaktadır. Her ne kadar çok sayıda yöntem keşfedilse de sadece belli sayıdaki yöntemler genel kabul görmekte ve uzun süre boyunca geçerli kalabilmektedir. Dizileme yöntemleri ile oluşturulan DNA dizileri, farklı organizmaların genomlarının kıyaslanması ve gen yapıları gibi konular hakkında bilgi sahibi olmamıza olanak sağlar [68]. Dizileme işlemleri, bir gen veya gen bölgesinde varyantlar veya mutasyonların tespiti, hastalıkların teşhis ve tedavisi gibi birçok amaç için kullanılmaktadır.

Gregor Mendel 1866 yılında bezelye bitkilerinde yaptığı melezleme çalışması ile genetiğin temelini atmıştır. 1869 yılında Fridedrich Miescher hücre çekirdeğinde asidik özellikli maddeler keşfetmiştir [61]. Avery ile arkadaşları gerçekleştirdikleri deneyler sonucunda genetik bilgileri DNA'nın aktardığını 1944'de kanıtlamıştır. Çift sarmal DNA yapısı Watson ve Crick tarafından 1953 yılında keşfetmiştir. Tüm bu genetik çalışmalar ve benzeri diğer araştırmalar nükleik asit dizileme sistemlerinin kökenini oluşturmuştur [68].

İlk DNA dizileri 1970'lerin başlarında elde edilmiştir. Dizileme yöntemleriyle ilgili yapılan çalışmalar yeni yöntemlerin geliştirilmesine ön ayak olmuştur.

Genetik kod, DNA ve RNA dizilerinde kodlanan genetik bilgileri hücreler yardımıyla proteinlere çeviren kuralların bütünüdür [69]. Bir organizmanın genomu DNA'sı içinde kaydedilmiştir. Genomun RNA veya protein kodlayan kısımları genleri oluşturmaktadır. Üç adet nükleotitin bir araya gelerek oluşturduğu yapıya kodon adı verilir [69]. Kod, kodonlar ile amino asitler arasındaki ilişkiyi tanımlar. Genellikle her bir kodon bir amino asidi kodlar. Genetik bir dizideki kod çevrilmeye başladığında nükleotitler baştan itibaren üçlü şekilde ayrılarak okunur. Çevirinin yapıldığı asıl çerçeve başlama kodonu tarafından belirlenir. Protein sentezini sonlandıran kodona durdurma (stop) kodonu adı verilir.

Nükleotitlerin üçlü dizilimi ile oluşan kodonlar $4^3 = 64$ farklı kombinasyonu mevcuttur (Şekil 1.16) [69]. Her bir kodon bir aminoasite veya bir stop (bitim) sinyaline karşılık gelmektedir.

		İkinci Baz Sırası				
		U	C	A	G	
Birinci Baz Sırası	U	UUU } Phe UUC } UUA } Leu UUG }	UCU } UCC } Ser UCA } UCG }	UAU } Tyr UAC } UAA Stop UAG Stop	UGU } Cys UGC } UGA Stop UGG Trp	U C A G
	C	CUU } CUC } Leu CUA } CUG }	CCU } CCC } Pro CCA } CCG }	CAU } His CAC } CAA Gln CAG }	CGU } CGC } Arg CGA } CGG }	U C A G
	A	AUU } AUC } Ile AUA } AUG Met	ACU } ACC } Thr ACA } ACG }	AAU } Asn AAC } AAA Lys AAG }	AGU } Ser AGC } AGA } Arg AGG }	U C A G
	G	GUU } GUC } Val GUA } GUG }	GCU } GCC } Ala GCA } GCG }	GAU } Asp GAC } GAA Glu GAG }	GGU } GGC } Gly GGA } GGG }	U C A G

Şekil 1.22. Kodon tablosu [70]

Geçmişten günümüze kadar yapılan çalışmalar neticesinde birçok DNA dizileme yöntemi ortaya çıkmıştır. Bu yöntemler üç ana kategoride toplanmıştır [71].

1.7.1. İlk Nesil DNA Dizileme Yöntemleri

Sanger Dideoksi Yöntemi (Zincir Sonlandırma): Uzun DNA parçalarını dizilemek için kullanılmaktadır. Zincir sonlandırma metodu olarak da adlandırılır. Sanger Deoksi yönteminde tek iplikli DNA şablonu, primer DNA, deoksinükleotitler ve zincir büyümesini sonlandıran modifiye edilmiş bazlar kullanılmaktadır. Doğru, hızlı ve ekonomik bir dizileme metodudur. Günümüzde popülerliğini yeni teknolojilere bırakmış olsa da hala bazı DNA dizileme uygulamalarında kullanılmaktadır. Güncel DNA Dizileme yöntemlerinde de kullanılan birçok temel kavramı tanımlamıştır [72].

Maxam-Gilbert Dizilemesi: DNA'nın kimyasallarla modifikasyonu ve elektroforezle belli bazlarda kesilmesine dayalı metottur. Sanger Dideoksi yönteminden birkaç yıl sonra bulunmasına rağmen ilk etapta Sanger yöntemine göre daha popüler olmuştur. Yöntemin karmaşıklığı ve zararlı kimyasalların kullanımı sebebiyle kullanışlı bir yöntem değildir. Sanger Dideoksi yönteminde yapılan iyileştirmeler sonucu yöntemin yaygınlaşmasıyla birlikte Maxam-Gilbert dizilemesinin popüleritesi azalmıştır.

Av Tüfeği Dizilemesi: Uzun DNA parçalarını dizilemek için kullanılan dizileme yöntemidir. Bu yöntemde büyük DNA parçası rast gele küçük parçalara ayrılır ve her bir parça klonlanır. Ayrılan küçük parçalarda kendi örtüşen kısımlarına bakılıp birleştirme

işlemi yapılır ve büyük dizilim elde edilir [73]. Bölünen parçaların bir genomdaki tüm bazları tamamen bulundurabilmesi için çok fazla sayıda klonlama işleminin yapılması gerekmektedir. Klonlama sayısına bağlı olarak elde edilen dizinin büyük bir kısmı tekrar eden fazlalık verilerden oluşmaktadır. Dizilim, tekrar eden dizi parçalarını doğru sıralamada yerleştiren bilgisayar programı yardımıyla analiz edilmektedir.

PCR Amplifikasyonu: DNA parçasının çok sayıda klonlanması temeline dayanan bir dizileme çeşididir. Uygun kimyasal ortamlar oluşturularak klonlanan parçaların başlangıç ve bitiş noktalarına yapay DNA parçaları eklenmektedir. Sonrasında çeşitli enzimler yardımıyla incelenecek kısım ayrıştırılmaktadır.

1.7.2. İkinci Nesil DNA Dizileme Yöntemleri

DNA dizilemelerinde kullanılan nesil kavramı, teknolojik gelişmeler sayesinde dizileme hızlarındaki yüksek artışa ve dizileme maliyetinin azalmasına bağlı olarak ortaya çıkmıştır. İkinci nesil dizilemelerin en önemli özelliği paralel yöntemler kullanarak çok fazla sayıda veri dizisini örnekleyerek sıralayabilmesidir. Yüksek hesaplama gücü sayesinde eski yöntemlere oranla 100 kat hızlı veri örnekleyebilmektedirler. En yaygın kullanılan yöntemler aşağıda belirtilmiştir [72].

454 Pirodizilemesi: Bu yöntemde, DNA 400-700 baz çifti uzunluğundaki tek sarmallı bölümlere parçalanır ve her bir parça adaptörlere bağlanır. PCR reaksiyonu ile DNA amplifiye edilerek birbirinin aynısı olan her bir DNA klonu adaptörlerde taşınır. Robotikler vasıtasıyla adaptörler, her biri sadece bir adaptör bulunduran milyonlarca kuyuya sahip fiber optik tablalara yerleştirilirler. Dört nükleotit belirli bir sırayla tabla üzerinden kayarlar. İçlerinden ışık saçan bir tanesi hangi bazın eklendiğini belirtir.

Illumina/Solexa Dizileme: Bu yöntem nükleotit zinciri sonlandırması ve DNA sentezi bakımından Sanger Dideoksi yöntemine benzemektedir. Sonlandırıcılar dideoksi yerine deoksi ribonükleotitlerdir ve tersine çevrilebilir şekilde eklenmektedir.

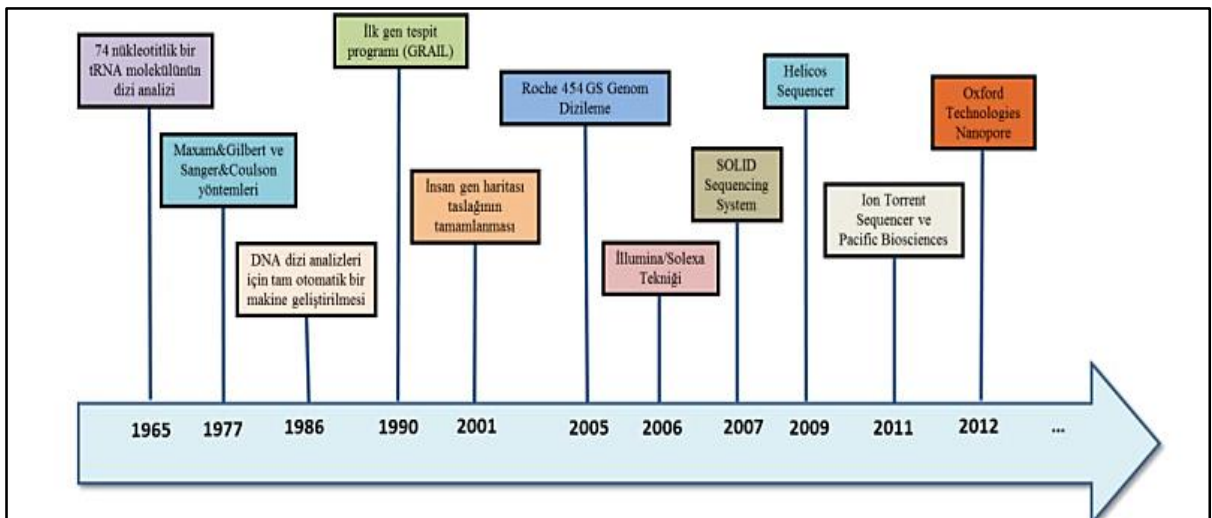
Solid Dizilemesi: Yapışma ile dizileme olarak da bilinmektedir. Hemen hemen her aşaması 454 Pirodizileme ile benzerlik göstermektedir. Büyük baz uzunluğu yerine çok daha kısa baz uzunluğu kullanır.

1.7.3. Yeni Nesil DNA Dizileme Yöntemleri

Üçüncü ve dördüncü nesil dizileme yöntemleri olarak karşımıza çıkan daha ucuz ve daha hızlı yöntemlerdir. Tek işlemede milyonlarca sayıda dizileme yapabilmektedirler [70]. Üçüncü nesil dizileme yöntemleri yaklaşım olarak mikroskopi ve nano-teknoloji olmak üzere iki temel üzerine oturtulmuştur. Bu yöntemlerin en karakteristik özelliği DNA'nın tek bir molekülünün dizilenmesidir. Dördüncü nesil dizileme yöntemlerini önceki yöntemlerden ayıran özellik ise optik algılamanın kullanılmamasıdır. Bu yöntemler ışık sonrası dizileme olarak da isimlendirilmektedir. DNA dizisine yeni bir deoksiribonükleotit eklendiği zaman protonların salınımı takip edilerek elektrot yardımıyla pH değerleri ölçülmektedir [72].

Kitlesel Paralel Dizileme: Paralel DNA dizileme tekniği kullanıldığı için dizilimde verim yüksektir [74]. Bu dizileme, DNA'nın parçalanması ve bu parçalara adaptör dizilerinin eklenmesiyle başlar. **Poloni Dizilemesi:** Yöntem, dizilemede genomun tekrar kullanılması metodunu benimsemiştir. Yeni nesil dizileme yöntemleri içerisinde büyük boyutlu genomu dizilemek için kullanılan ilk metotlardan biridir. Sanger yöntemine göre oldukça ucuzdur. **DNA Nanotop Dizilemesi:** DNA içerisinden alınan küçük parçalar çoğaltılarak DNA nanotopları meydana getirilir. Dönen çember ikileşmesi kullanılarak nanotoplar birbirlerine bağlanmaktadır. Bir organizmanın tüm genom diziliminin elde edilmesinde kullanılan yüksek verimliliğe sahip bir teknolojidir.

İyon Yarı İletken Dizilemesi: Işık kullanmayan dördüncü nesil yöntemdir. DNA sonuna eklenen protonların takibine ve ölçülen pH değerine bağlıdır.



Şekil 1.23. DNA dizileme sistemlerinin tarihsel gelişimi [68]

2. YAPILAN ÇALIŞMALAR

2.1. Yöntemler

Bu çalışmada, mutasyona sahip olan nükleotit baz dizilerinin tespiti ve gerçekleşen küçük ölçekli mutasyon çeşitlerinin belirlenebilmesi için yeni bir yöntem önerilmiştir. Yöntem uygulanırken kanser oluşumundaki etkisi göz önüne alınarak BRAF geni seçilmiştir. Fasta formatı şeklinde okunan mutasyonlu BRAF geni dizileri, sağlıklı BRAF geni ile karşılaştırılmıştır. Karşılaştırma yapılan gen dizileri belirli uzunluklardaki alt dizilere ayrılarak hash fonksiyonları kullanan Bloom Filter yapısında saklanmıştır. Bu sayede karşılaştırmanın tamamlanması ve mutasyonların tespitinin hızlandırılması amaçlanmıştır.

2.1.1. Verilerin Elde Edilmesi

Çalışmada kullanılan sağlıklı ve mutasyona uğramış DNA verileri, NCBI (National Center for Biotechnology Information) biyoteknoloji bilgi merkezinin depoladığı BRAF gen dizilerinden elde edilmiştir. Gen dizileri fasta formatına dönüştürülerek saklanmıştır.

2.1.1.1. NCBI

NCBI, ABD Ulusal Sağlık Enstitüsüne bağlı olarak genetik analizler üzerinde çalışmak amacıyla kurulmuştur. Günümüzde en önemli gen bankalarından biri olan NCBI'nin misyonu, sağlığı ve hastalığı kontrol eden temel moleküler ve genetik süreçlerin anlaşılmasına yardımcı olacak yeni bilgi teknolojileri geliştirmektir [75].

2.1.1.2. Fasta Formatı

Fasta formatı, baz çiftleri veya aminoasitlerin tek harfli kodlar kullanılarak temsil edildiği uluslararası standart haline gelmiş metin tabanlı bir formattır. İlk satır her zaman ">" karakteri ile başlar ve bir dizi tanımlamalar içerir. Tanımlama satırında dizilim ile ilgili

bilgiler bulunmaktadır. Satırdaki bilgilerin dizilime herhangi bir etkisi yoktur. Bu satırda hiç bir açıklama olmasa bile ">" karakteri olmalıdır. Sonraki satırda bazlardan oluşan veriler bulunur. Veri satırı tek satır olabileceği gibi çok fazla satırdan da oluşabilir. Uzun diziler genelde okunma yöntemine göre 60 veya 100 baz gibi belirli bir uzunluktan sonra alt satıra geçecek şekilde yazılmaktadır. Nükleotitler baş harfleri ile temsil edilir (Adenin → A). Dizide belirsiz olan nükleotitler için N harfi kullanılır [76, 77].

Tablo 2.1. Örnek fasta formatı [78]

>ref114_arb-silva hs-arb_silva-LUS-SSUref114
GCGGTGGCCATAGGGCCCTGTTGGTGGGGCATGACCTCCCGTGATGGGGCCC AGGGGGCTGGTGTGGCCGTGACTCTGGACGCAAGTTGGGCCATTCCCGTGG ATTACCCAGTTGTGGCCGCCCATGGGGAGCCTGGTGGTTGCTGGTGTGCCC TGTCTCGCTCCGCACCATCGTGTANNNNNNNNNNNNNNNNNNNNNNNNNNNNN NNNNNNN

2.1.2. K-MER Analizi

Çalışmada, K-mer dizileme yöntemi mutasyona uğramış nükleotitlerin tespit edilmesinde önemli rol oynamaktadır. Biyoinformatikte K-mer dizileri, biyolojik bir dizinin sabit bir uzunluğa sahip olan alt dizilere bölünerek analiz edilmesini sağlar. Yeni nesil DNA dizileme işlemlerinde K-mer yönteminden sıklıkla faydalanılır [79].

Örneğin 10 karakter uzunluklu 'CATAGTGCAC' DNA dizisinin sabit uzunluklu alt dizileri aşağıdaki tabloda belirtilmiştir.

Tablo 2.2. Alt diziler

k (uzunluk)	K-mer alt dizi elemanları	n (alt dizi sayısı)
3	CAT, ATA, TAG, AGT, GTC, TGC, GCA, CAC	8
4	CATA, ATAG, TAGT, AGTG, GTGC, TGCA, GCAC	7
5	CATAG, ATAGT, TAGTG, AGTGC, GTGCA, TGCAC	6
6	CATAGT, ATAGTC, TAGTGC, AGTGCA, GTGCAC	5
7	CATAGTG, ATAGTGC, TAGTGCA, AGTGCAC	4

Boyutu S olan bir DNA dizisinden k boyutlu n adet K-mer dizisi oluşur. Genel olarak K-mer dizileri için aşağıdaki denklemi kurabiliriz.

$$n = S - k + 1 \quad (2.1)$$

Denklem 3.1'in parametreleri aşağıdadır.

- n: Alt dizi sayısı
- S: DNA dizisinin uzunluğu
- K: Her bir alt dizi (K-mer) uzunluğu

2.1.3. Mutasyonlu Gen Verisini Dizileme

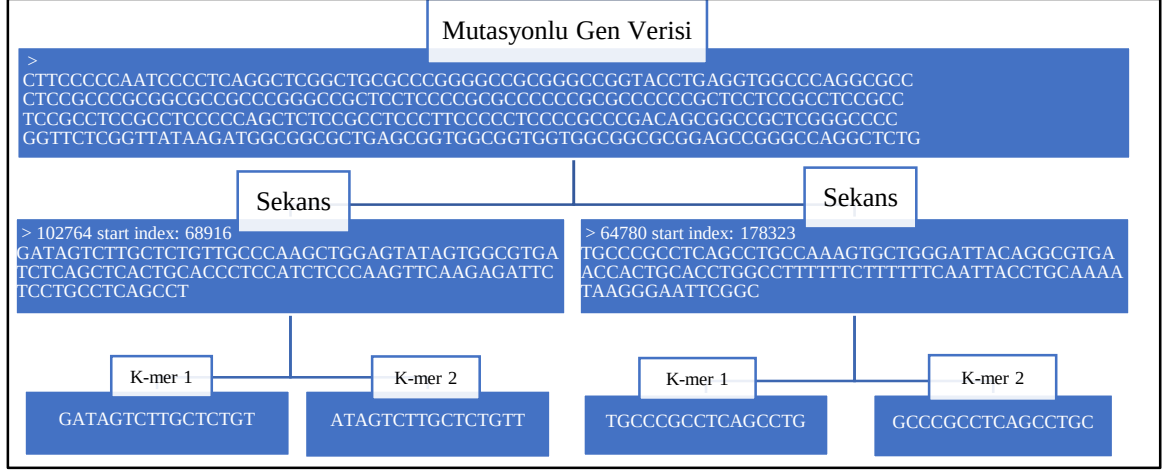
Günümüzde yüksek boyutlu DNA dizilerini okumak için kullanılan cihazlar veriyi parça parça ve birçok kez okuyarak elde ederler. Okuma sayısı, toplam veri boyutunun tek seferde okunan veri boyutuna oranının bir katsayı ile çarpılması sonucu belirlenir. Bu katsayı derinlik olarak ifade edilir. Bu hesaplamayı aşağıdaki denklem ile ifade edebiliriz.

$$n = T \times D / S \quad (2.2)$$

Denklem 3.2'in parametreleri aşağıdadır.

- n: Okunan dizi sayısı
- T: Toplam gen boyutu
- D: Derinlik
- S: Tek okumadaki sabit dizi boyutu

Bu çalışmada mutasyonlu DNA verileri, her bir okumada 100 nükleotit (S=100) okunacak şekilde ve 50 derinlikli okuma (D=50) yapılarak simüle edilmiştir. Her bir okuma sonucunda oluşan dizilere sekans adı verilmiştir. Okunan veriler fasta formatına uygun şekilde kaydedilmiştir.



Şekil 2.1. Gen dizileme

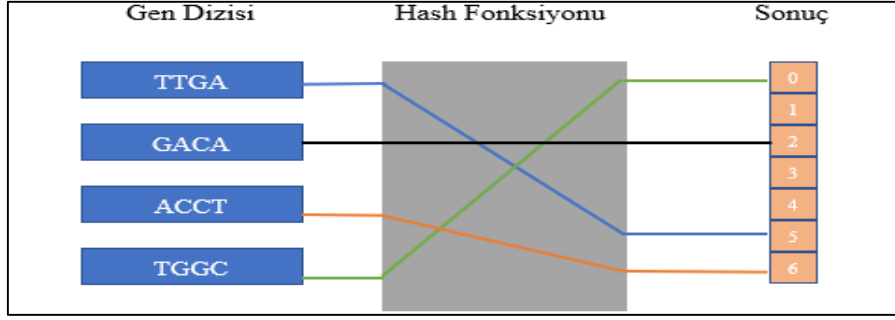
Şekil 2.1 de alt dizilere ait fasta formatındaki açıklama satırında, dizinin kaçınıcı alt dizi olduğu ve gen içinde hangi sırada olduğunu bilgisi tutulmaktadır.

2.1.4. Kullanılan Veri Yapısı

Çalışmada verilerimizi kıyaslama işleminde, aranan dizinin büyük veriler içerisinde olup olmadığını hızlı bir şekilde kontrol edebilmemizi sağlayan Bloom Filter kullanılmıştır.

2.1.4.1. Hash Fonksiyonları

Farklı boyutlara sahip her bir veri setini, her biri birbiri ile aynı boyutlu olan çıktılarına dönüştüren algoritmik bir yöntemdir. Aynı veri için her zaman aynı sonucu benzersiz olarak üretmektedir. Büyük veri tabloları içerisindeki verileri daha hızlı sorgulamak, karşılaştırma yapmayı kolaylaştırmak ve parola bilgisini gizlemek gibi işlemler için kullanılır. Hash fonksiyonları temel seviyede şifreleme aracı olarak kullanılır [80]. Şekil 2.2 de hash fonksiyonunun çalışması gösterilmiştir.

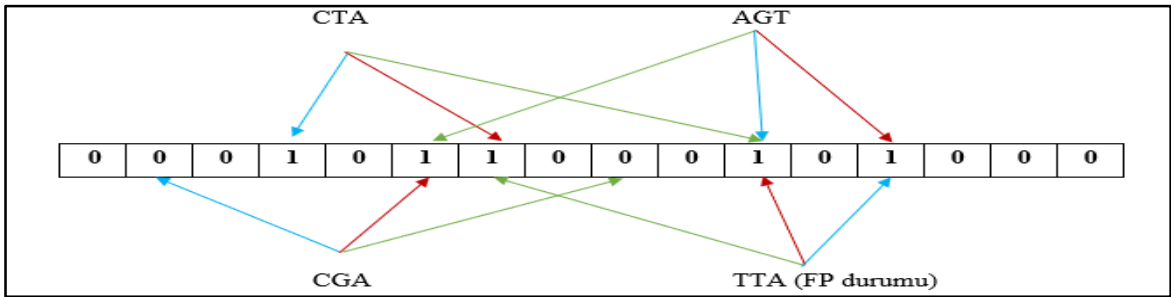


Şekil 2.2. Hash fonksiyonu çalışma sistemi

Belirli bir sayıdan fazla verinin aynı fonksiyona tabi tutulması sonucu fonksiyon farklı iki veri için aynı sonucu üretirse Çakışma (Collision) olur ve o fonksiyona artık güvenilemez [81]. Çalışmada kullanılan hash fonksiyonu sayısı verilerin saklandığı filtrenin boyutuna doğrudan etki etmektedir.

2.1.4.2. Bloom Filter

Bir öğe veya dizinin bir veri kümesi içinde bulunup bulunmadığını hızlı ve bellek açısından verimli bir şekilde bildirmek için tasarlanmış olasılıksal bir veri yapısıdır. Olasılıksal denmesinin sebebi, sorgulama yapıldığında filtreden elemanın filtrede bulunduğu cevabı gelse bile bunun kesin olmamasıdır. Ancak eleman filtrede bulunmuyorsa bu kesinlikle doğrudur.[82] Bloom Filter, başlangıçta tüm verilerin 0 olduğu bir bit dizisidir. Veriler, 0 ile filtrenin boyutu aralığında sonuç üreten bir veya daha fazla hash fonksiyonundan geçirilerek filtrenin hash sonucuna karşılık gelen indeksi 1 olarak setlenir. Bloom filter özellikle veri tabanı sorgularından önce kullanıldığında gereksiz disk erişiminin önüne geçebilmektedir.



Şekil 2.3. Bloom Filter çalışma yapısı

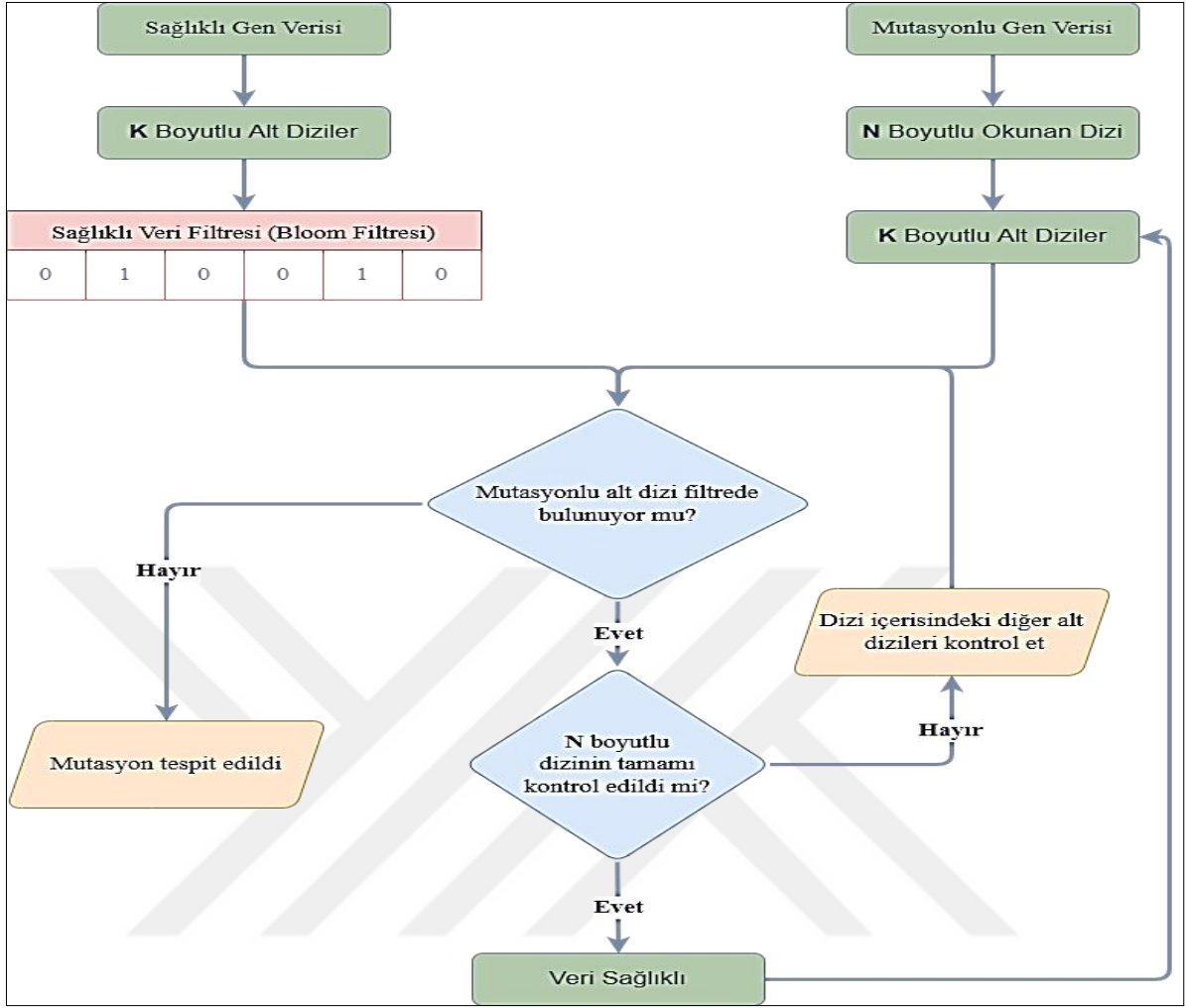
Örnek bir Bloom Filter şekil 2.3’de gösterilmiştir. İki farklı gen dizisi CTA ve AGT her biri farklı renkli oklar ile belirtilmiş üç adet hash fonksiyonundan geçirilerek filtre içerisinde konumlandırılmıştır. Yeni bir dizinin bu filtrede olup olmadığını sorgulamak için aynı hash fonksiyonları kullanılarak dizinin filtredeki karşılık düştüğü indekslere bakılır. Şekilde CGA dizisi sorgulandığında fonksiyon sonuçlarına karşılık setlenmemiş (0) indeksler olduğu için bu dizinin kesinlikle bu filtrede olmadığı söylenebilmektedir. TTA dizisinin fonksiyon çıktıları ise dizi filtrede olmamasına rağmen farklı genetik dizilerin hash fonksiyonlarının setlediği [80] indekslere karşılık geldiği için filtre bu diziyi içerisinde bulundurduğunu belirtir. Bu duruma FP (false positive) denir. Bloom filter dizisindeki öge sayısı arttıkça, FP olasılığı artar. Düzgün yapılandırma ile FP oranı %1’in altında tutulabilmektedir.[83]

2.1.5. Mutasyon Algoritmaları

2.1.5.1. Mutasyon Tespit Algoritması

Çalışmanın ilk aşamasında sağlıklı gen verileri sabit bir K-mer uzunluğu ve sabit bir hash fonksiyonu sayısı belirlenerek Bloom Filter’a eklenmiştir. K-mer uzunluğu ve hash fonksiyonu sayısının sabit tutulmasının nedeni mutasyonlu verileri filtreye eklerken de aynı değerlerin kullanılacak olmasıdır. Sabit boyutlu her bir okuma dizisinin alt dizileri bir döngüye sokularak sağlam verilerin olduğu filtre içerisinde aranır. Filtrede sorgulama yapıldığında filtre bu alt diziyi bulundurma ihtimali olduğunu (FP) veya kesin olarak bulundurmadığını cevaplar. Kesin olarak bulundurmuyor ise alt dizinin mutasyon barındırdığı bilgisine ulaşırız.

Mutasyonlu genden okunan alt dizilerin (K-mer) tek başına kontrol edilmesi doğru sonuç için yeterli değildir. Alt dizideki mutasyon sağlıklı gen içerisindeki doğru bir alt diziyi örtüşebilir. Bu durumun kontrolünü yapmak için mutasyonlu genden okunan alt diziyi bulunduran dizinin tamamı sorgudan geçirilerek doğrulanır.



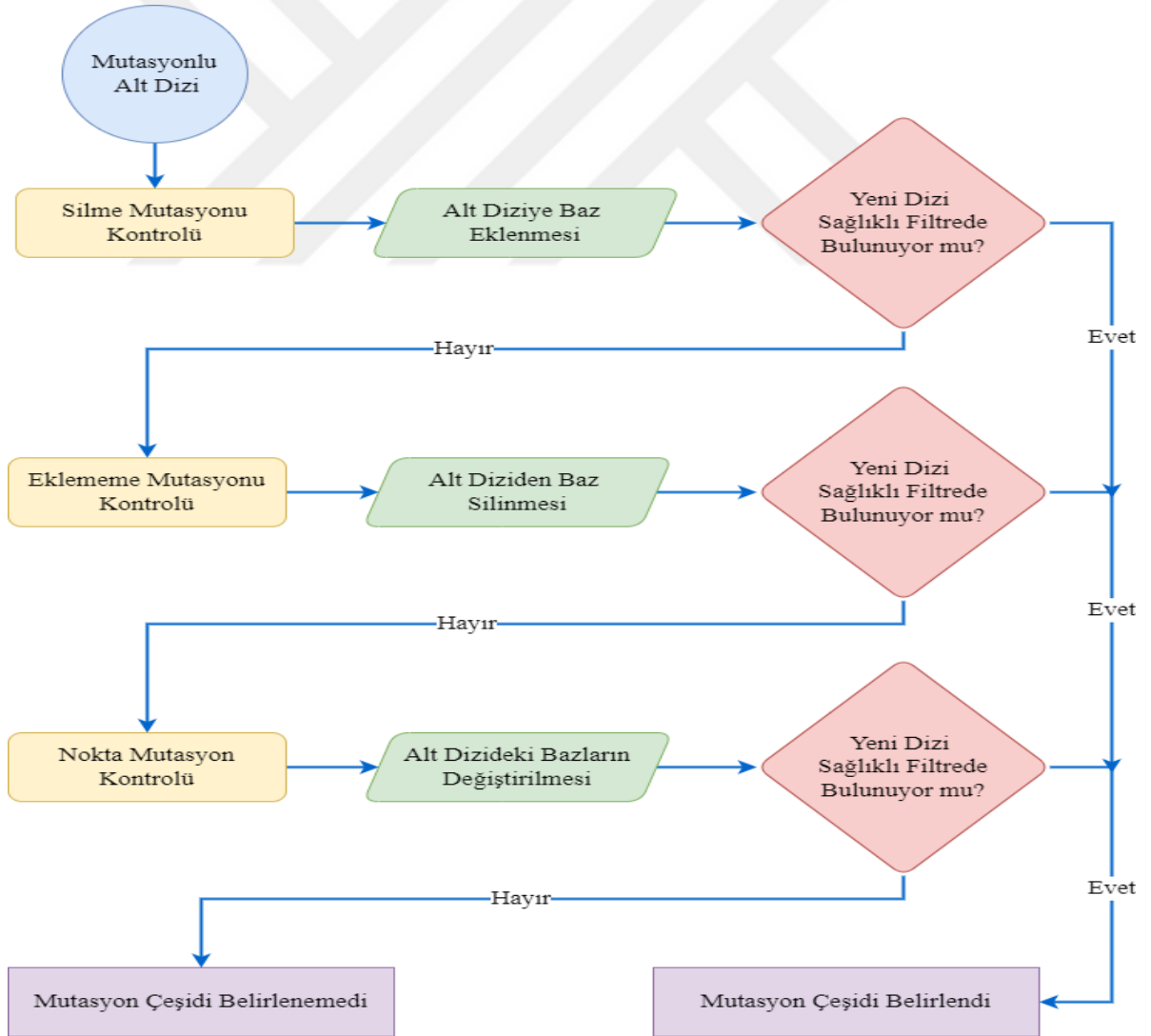
Şekil 2.4. Mutasyon tespit algoritmasının çalışma yapısı

2.1.5.2. Mutasyon Türü Belirleme Algoritması

Mutasyon tespit edilen K-mer alt dizisinde hangi mutasyon çeşidinin gerçekleşmiş olduğunu anlamak için çalışma yapılmıştır. Küçük bazlı mutasyon gerçekleşen alt diziye, geriye dönük olarak her veri için yer değiştirme, baz silme ve baz ekleme işlemleri uygulanarak yeni bir alt dizi oluşturulmuştur. Bu yeni K-mer, daha büyük boyutlu olan okuma dizisi içerisindeki mutasyonlu K-mer ile değiştirilmiş ve oluşan yeni dizinin sağlıklı gen filtresi içerisinde bulunup bulunmadığı kontrol edilmiştir. Çalışma içerisinde tespit edilebilen mutasyonların oluşumu Şekil 3.12 de gösterilmiştir.



Şekil 2.5. Mutasyon türlerinin oluşumu



Şekil 2.6. Mutasyon çeşidini belirleme

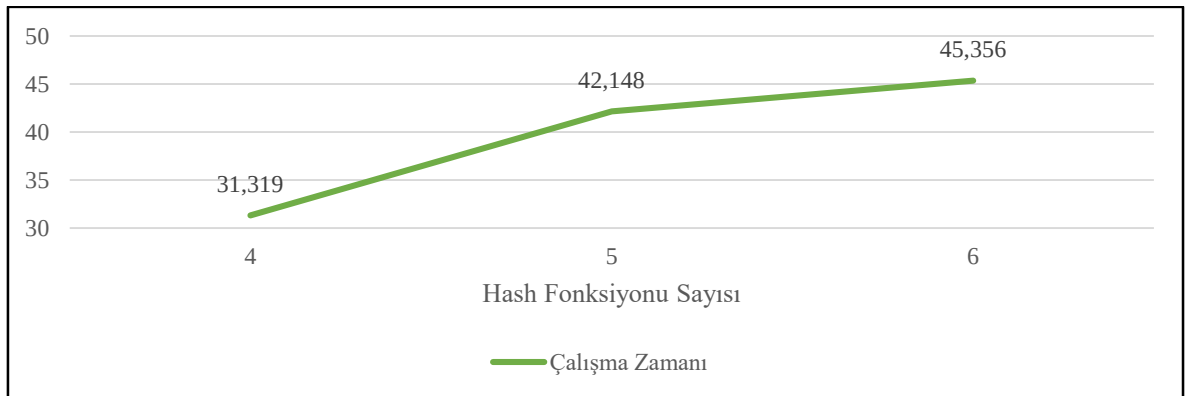
3. BULGULAR

Önerilen algoritmalar sağlıklı ve mutasyonlu iki farklı BRAF geni kullanılarak test edilmiştir. Farklı dizi boyutları ve hash fonksiyonu sayısına göre işlenen veriler Tablo 3.1 de karşılaştırılmıştır.

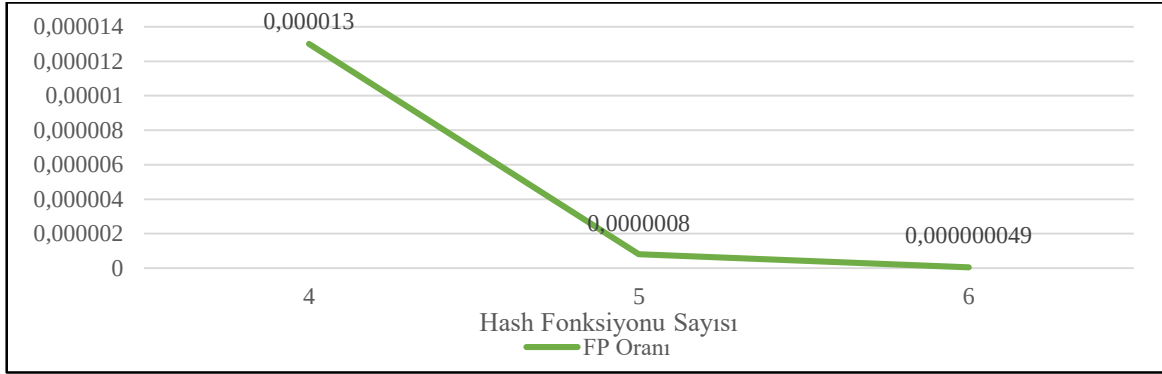
Tablo 3.1. Filtre analizi

Alt Dizi (K-mer) Boyutu	Hash Fonksiyonu Sayısı	Bloom Filtresi Boyutu	Çalışma Süresi (sn.)
16	4	13157376	31.319
16	5	16446720	42.148
16	6	19736064	45.356
20	4	16446400	31.273
20	5	20558000	39.044
20	6	24669600	45.007
24	4	19735296	29.792
24	5	24669120	38.971
24	6	29602944	43.469

Kullanılacak hash fonksiyonlarının sayısı arttıkça çalışma süresinin uzadığı görülmüştür. Bu olumsuzluğa rağmen hash fonksiyonu sayısının yüksek tutulması fonksiyon içerisinde oluşan çakışma (collision) hatasının filtreye etkisini azaltarak ve FP oranının düşük seviyede tutulmasını sağlamıştır. 16 boyutundaki alt dizi kullanıldığında hash fonksiyonunun çalışma süresine etkisi Şekil 3.1’de FP oranına etkisi de Şekil 3.2’de gösterilmiştir.



Şekil 3.1. Hash fonksiyonu – çalışma zamanı grafiği



Şekil 3.2. Hash fonksiyonu – FP oranı grafiği

Şekil 3.2’de gösterildiği gibi kullanılacak Bloom Filter’ın en az düzeyde FP oranına sahip olması için hash fonksiyonu sayısı 6 olarak seçilmiştir. FP oranı, filtre büyüklüğü, filtredeki eleman sayısı ve hash fonksiyonu sayılarına bağlı olarak aşağıdaki denklem ile belirlenmiştir.

$$p = (1 - e^{-kn/m})^k \quad [84] \quad (3.1)$$

Denklem 3.3’te kullanılan parametreler aşağıdadır:

- p: FP oranı
- k: Hash fonksiyonu sayısı
- n: Veri boyutu
- m: Bloom Filter boyutu

Çalışmada kullanılan alt dizi uzunluğu ve hash fonksiyonu sayısına göre hesaplanan FP olasılıkları Tablo 3.2 de belirtilmiştir.

Tablo 3.2. FP olasılıkları

Alt Dizi (K-mer) Boyutu	Hash Fonksiyonu Sayısı	FP Oranı
16	4	$1,3474 \times 10^{-5}$
16	5	$8,163 \times 10^{-7}$
16	6	$4,946 \times 10^{-8}$
20	4	$5,657 \times 10^{-6}$
20	5	$2,759 \times 10^{-7}$
20	6	$1,345 \times 10^{-8}$
24	4	$2,773 \times 10^{-6}$
24	5	$1,132 \times 10^{-7}$
24	6	$4,619 \times 10^{-9}$

4. TARTIŞMA

Çalışmada önerilen algoritmada, alt dizi uzunluğu ve hash fonksiyonu sayısının yüksek tutulması mutasyon tespiti adına daha kesin sonuçlar vermiştir. Bu veriler FP olasılığını azaltmış ancak tespit süresini uzatmıştır. Bu durumlar göz önüne alınarak çalışan algoritmada alt dizi uzunluğu için 16 ve hash fonksiyonu sayısı için 6 ideal değerler olarak belirlenmiştir. Farklı veri setleri kullanılması durumunda dizi boyutuna bağlı olarak alt dizi boyutu ve hash fonksiyonu sayısı değiştirilerek çalışma amacına uygun farklı hız ve performans sonuçları alınabilmektedir. Algoritmaya ait karakteristik değerler ile oluşan çalışma hızı ve FP ilişkileri Tablo 4.1 ve Tablo 4.2 de gösterilmiştir.

Tablo 4.1. Alt dizi karakteristiğine bağlı FP – zaman ilişkisi

Karakteristik	Karakteristik	FP Oranı	Çalışma Zamanı (sn)
H6	A16	$4,946 \times 10^{-8}$	45.356
H6	A20	$1,345 \times 10^{-8}$	45.007
H6	A24	$4,619 \times 10^{-9}$	43.469

Tablo 4.2. Hash fonksiyonu karakteristiğine bağlı FP – zaman ilişkisi

Karakteristik	Karakteristik	FP Oranı	Çalışma Zamanı (sn)
A16	H4	$1,3474 \times 10^{-5}$	31.319
A16	H5	$8,163 \times 10^{-7}$	42.148
A16	H6	$4,946 \times 10^{-8}$	45.356

Çalışmadaki algoritmalar, özellikleri aşağıda olan bilgisayarda çalıştırılmıştır.

- İşlemci: Intel® Core™ i7-6700HQ 2.60 GHz İşlemci
- İşletim Sistemi: Windows 10 Home Single Language 64-bit
- Ram: 24 GB DDR4
- Disk Kapasitesi: 256 GB SSD
- Ekran Kartı: NVIDIA GeForce GTX 950M

Algoritmanın çalıştırılacağı sistem özelliklerine göre hız ve performans tercihinine bağlı olarak Tablo 3.1, Tablo 4.1 ve Tablo 4.2 kontrol edilerek uygun karakteristikler seçilmelidir.

5. SONUÇLAR

- Bu çalışmada, mutasyona sahip olan nükleotit baz dizilerinin tespiti ve gerçekleşen mutasyon çeşitlerinin belirlenebilmesi için yeni bir yöntem önerilmiştir. Önerilerin algoritmaların veri seti üzerinde mutasyon olup olmadığını tespit edebilmesi ve mutasyon varsa hangi mutasyonun gerçekleştiğini belirleyebilmesi için yazılımsal deneyler gerçekleştirilmiştir. Deney sonucunda elde edilen veriler analiz edilerek en hızlı şekilde sonuca ulaşabilmek amaçlanmıştır. Çalışmada, sağlıklı gen verisi olarak Amerikan Gen Bankası NCBI üzerinden paylaşılan BRAF geni kullanılmıştır. Sağlıklı gen bilgisi hash fonksiyonları kullanan Bloom Filter veri yapısında saklanmıştır. Birden fazla hash fonksiyonu kullanımı ile algoritmanın güvenilirliği artırılmıştır. Hasta bireylere ait DNA dizileri, yeni nesil dizileme yöntemlerine uygun olacak şekilde fasta formatı halinde okunmuştur. Çalışma evrensel fasta formatı şeklinde hazırlanmış tüm verileri işleyebilmektedir. Okunarak dizi haline getirilmiş olan veriler filtrede bulunan sağlıklı veriler ile karşılaştırılarak verideki küçük ölçekli mutasyonlar başarılı bir şekilde tespit edilmiştir.
- Kullanılan Bloom filtresi ile gen verisinin sağlıklı olup olmadığı çok hızlı bir şekilde sorgulanmıştır. BRAF mutasyonundaki genetik değişiklikler bir hastalığın erken veya başlangıç safhalarını gösterebilmektedir. Bazı kanser çeşitlerinde erken tanı oranı düşük olduğundan tanı konulduğunda tedavi için geç kalınmaktadır. Mutasyonun erken evrede hızlı bir şekilde tespit edilmesi ile bu sorunun önüne geçilmesi amaçlanmaktadır.
- Önerilen çalışma ve algoritmaların, günümüzün en büyük sorunlarından olan kanser ve benzeri genetik hastalıklara sebebiyet veren gen mutasyonlarının tespitine ve bu alandaki biyoinformatik çalışmalara temel düzeyde katkı vermesi hedeflenmektedir.

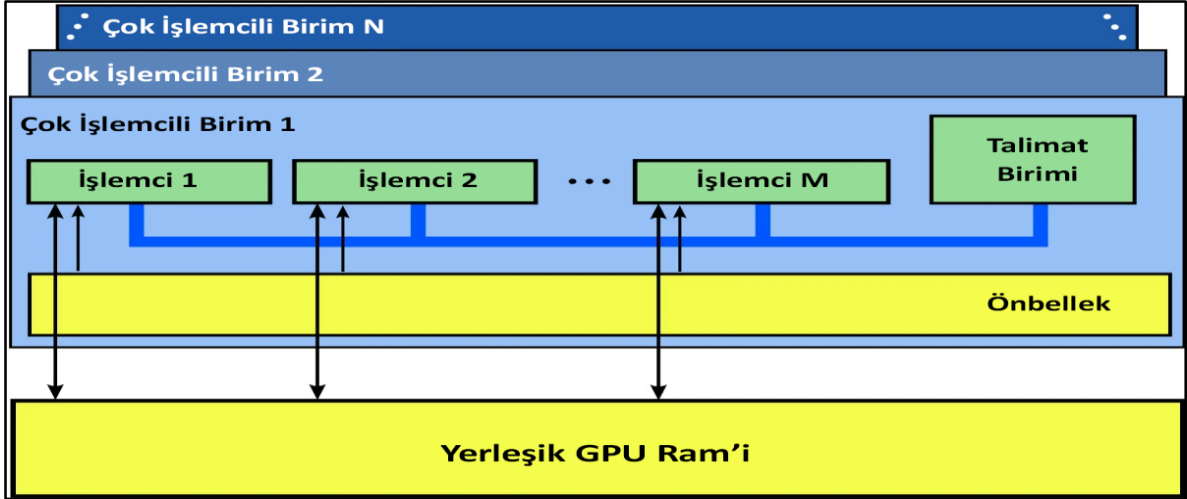
6. ÖNERİLER

Önerilen çalışmada mutasyon tespit algoritmasının başarılı sonuçlar verdiği gözlenmiştir. Sonraki aşamalarda algoritmanın hız performansının artırılması ve büyük ölçekli mutasyonların da tespit edilmesi amaçlanmaktadır. Çalışmanın geliştirilmesi adına planlanan çalışmalar maddeler halinde listelenmiştir.

- Mutasyon tespit algoritmasının, GPU (grafik işlem birimi) üzerinde çalışan CUDA ve benzeri kodlama dilleri kullanılarak kodlanması ile daha hızlı, daha başarılı sonuçlar alınması hedeflenmektedir.
- Gen verisi üzerinde çok fazla baz kaybı veya baz eklenmesi olan büyük ölçekli mutasyonlar tespit edilebilmektedir. Mevcut analizler ile bu mutasyon çeşitlerine ait kesin verilerin elde edilebilmesi için algoritmanın geliştirilmesi planlanmaktadır.
- Kullanılan Bloom Filter optimal düzeyde boyutlandırılarak algoritmanın çalışma performansının daha da artırılması amaçlanmaktadır.

Çalışmada önerilen çözüm yöntemlerinin temel seviyedeki biyoinformatik çalışmalara katkı sağlaması ve geleceğimiz için çok önemli olan bu alanda yeni algoritmik fikirler için bir başlangıç olması amaçlanmaktadır.

Gen dizilerindeki mutasyonların tespit edilebilmesi adına yapılan çalışmada geliştirilen güncel gen dizileme algoritmaları incelendiğinde GPU kullanımının yaygınlaştığı görülmektedir. GPU tabanlı bir algoritma olan MumberGPU yüksek hacimli gen dizilerini az maliyetli ve yüksek hızda dizileyebilmek için tasarlanmıştır. Dizi hizalaması için son ek ağacı (suffix arrays) veri yapısını kullanmaktadır. GPU üzerinde çalışan CUDA programlama dili ile kodlanmış olan program, aynı algoritma ile CPU (merkezi işlem birimi) üzerinde çalışan C dili kullanılarak geliştirilmiş programa göre 10 kat daha hızlı çalışmakta ve 3.5 kat daha fazla performans göstermektedir [85]. GPU'nun hızlı çalışmasını sağlayan mimarisi Şekil 6.1'de gösterilmiştir.



Şekil 6.1. GPU mimarisi [85]

Şekilde GPU'nun, her biri aynı komutu aynı anda yürüten birden çok (M) akış işlemcisi içeren birkaç (N) çok işlemcide nasıl organize edildiği gösterilmiştir. Her işlemci doku önbelleğine çok hızlı erişebilmektedir ancak yerleşik RAM'e okuma ve yazma işlemleri yüksek gecikme süresine sahiptir.

Mutasyon tespiti için incelenen bir diğer yöntem de MutScan algoritmasıdır. MutScan, FASTQ dosya formatı biçiminde saklanan verilerdeki hedef mutasyonları tespit etmeyi ve HTML tabanlı okuma yığın görselleştirmeleri oluşturarak mutasyonları doğrulamayı amaçlayan, hızlı çalışan bir araçtır. Mutasyon tespit işlemi için özelleştirilmiş hash fonksiyonları ve filtreler kullanılmaktadır [86].

Buna benzer mutasyon tespit uygulamaları ve çalışmada önerilen algoritmanın GPU tabanlı bir program haline getirilmesiyle güncel dizileme metotları ile uygun olarak çalışabileceği ve mutasyon tespit hızı ile performans seviyesinin daha üst bir noktaya ulaşabileceği düşünülmektedir.

7. KAYNAKLAR

1. <https://tr.wikipedia.org/wiki/DNA>. 23 Nisan 2021.
2. https://tr.wikipedia.org/wiki/N  kleik_asit. 20 Temmuz 2021.
3. <https://www.bilgial.com/page/34/>. 23 Haziran 2021.
4. Yıldırım, A., Bardakçı, F., Karataş, M. ve Tanyolaç, B., Molek  ler Biyoloji, İkinci Baskı, Nobel Yayın, Ankara, 2010.
5. <https://www.biyolojidersim.com/nukleik-asitlerin-yapisi-ve-replikasyon/>.23 Haziran 2021.
6. <https://pngset.com/download-free-png-unstn>. 13 Mart 2021.
7. <https://aibolita.com/mens-diseases/46714-protein-synthesis.html>. 18 Mart 2021
8. Vermeulen, K., VanBockstaele, D. R. ve Berneman, Z. N., The cell cycle: a review of regulation, deregulation and therapeutic targets in cancer, Cell Prolif, 36 (2003) 131-49.
9. Vermeulen K, Berneman Z. N.ve VanBockstaele D. R., Cell cycle and apoptosis, Cell Prolif, 36 (2003) 165-75.
10. Lodish, H., Berk, A, Zipursky, S. L., Matsudaira, P., Baltimore, D. ve Darnell, J. E., Molecular Cell Biology, fourth edition, WH Freeman and Co, New York, 2000.
11. Yokuş, B. ve   akır, D.   ., Kanser Biyokimyası, Dicle   niv. Vet. Fak. Dergisi, 1,2 (2012) 7-18.
12. Bellamy, COC., p53 and apoptosis, British Medical Bulletin, 53,3 (1996) 522-38.
13. <https://www.analizokulu.net/ders/mitoz-ve-eseysiz-ureme-3>. 23 Haziran 2021.
14. Cabadak, H., H  cre Siklusu ve Kanser, AD   Tıp Fak  ltesi Dergisi, 9,3 (2008) 51-61.
15.   aylak, B., Endometriyum ve over y  ksek dereceli ser  z adenokarsinomlarında p16, p53 ve ki-67 (mib-1) dıřavurumlarının ayırıcı tani ve prognostik parametreler ile iliřkisi, Uzmanlık Tezi, Ankara, 2009.
16. <https://avesis.yildiz.edu.tr/resume/downloadfile/bmansur?key=bab22e4a-0711-43a5-ba49-23f9c7134f7d>. 9 Eyl  l 2021.
17. Aliustaoğlu, M., Temel Kanser Fizyopatolojisi, Klinik Geliřim, (2009) 46-49.

18. Foster, I., Cancer: A cell cycle defect, Radiography, 14 (2008) 144-9.
19. Curran, J. E., Vaughan, T., Lea, R. A., Weinstein, S. R., Morrison, N. A. ve Griffiths, L. R., Association of A vitamin D receptor polymorphism with sporadic breast cancer development, International Journal of Cancer, 83,6 (1999) 723-726.
20. http://evolution.berkeley.edu/evolibrary/glossary/glossary_popup.php?word=mutation. 6 Nisan 2021.
21. Bütüner, B. D. ve Kantarcı, G., Mutasyon, DNA Hasari, Onarım Mekanizmaları ve Kanslerle ilişkisi, Ankara Ecz. Fak. Derg., 35,2 (2006) 149-170.
22. Preston, B. D., Albertson, T. M. ve Herr, J., DNA Replication Fidelity and Cancer, National Institutes of Health, 20,5 (2010) 281–293.
23. https://www.researchgate.net/publication/340075948_DNA_Damage_Response_Pathways_in_Cancer_Predisposition_and_Metastasis. 6 Nisan 2021.
24. https://stringfixer.com/tr/Directed_mutagenesis. 23 Haziran 2021.
25. Dingermann, T., Hänsel, R. ve Zündorf, I., Pharmazeutische Biologie: Molekulare Grundlagen und klinische Anwendung, Springer, (2002) 425.
26. Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K. ve Walter, P., Molecular Biology of the Cell, Fourth Edition, 2002, 527.
27. <https://bioisnotdifficult.blogspot.com/2016/07/mutation.html>. 6 Nisan 2021.
28. <https://bilimfili.com/mutasyon-cesitleri-nelerdir>. 23 Haziran 2021.
29. <https://byjus.com/neet-questions/what-is-an-example-of-a-missense-mutation/>. 6 Nisan 2021.
30. Ringer, D. P. ve Schnipper, L. E., Principles of Cancer Biology. In: Lenhard, R. E., Osteen, R. T. ve Gansler, T. eds., Clinical Oncology Atlanta: American Cancer Society, (2001) 21-35.
31. Berenblum, I., Frontiers of Biology, Carcinogenesis as a Biological Problem. Amsterdam: North-Holland Pub. Co., New York, (1974) 212-224.
32. Hussein, S. P. ve Harris C. C., Molecular epidemiology of human cancer, Recent Results Cancer Res, 154 (1998) 22–36.
33. Bos, J. L., Van Kreijl, C. F., In: Vainio, H., Magee, P. N., McGregor, D. B. ve McMichael, A. J. eds., Genes and Gene Products that Regulate Proliferation and Differentiation: Critical Targets in Carcinogenesis, Mechanisms of Carcinogenesis in Risk Identification, International Agency for Research of Cancer, Lyon, (1992) 57–65.
34. Foulds, L., Neoplastic Development 1, New York: Academic Pres, (1969) 122-134.

35. Sarid, R. ve Gao, S. J., Viruses and Human Cancer: From Detection to Causality, *Cancer Lett.*, 305 (2011), 219-222.
36. <https://www.cancer.org/content/dam/CRC/PDF/Public/6141.00.pdf>. 10 Ekim 2021.
37. İçli, F. ve Akbulut, H. Onkolojiiye Giriş. İn. İliçin, G., Biberoğlu, K., Süleymanlar, G. ve ark: eds. İç Hastalıkları. Güneş Kitapevi, (2005) 2007-2014.
38. Decuadin, D., Geley, S., Hirsch, T., Castedo, M., Marchetti, P., Macho, A., Kofler, R. ve Kroemer, G., Bcl-2 and Bcl-Xl antagonize the mitochondria dysfunction preceding nuclear apoptosis induced by chemotherapeutic agents. *Cancer Res*, 52 (1997) 62-67.
39. Golias, C., Charalabopoulos, A. ve Charalabopoulos, K., Cell proliferation and cell cycle control: a mini review, *The International Journal of Clinical Practice*, 58 (2004) 1134-1141.
40. Carmena, M. ve Earnshaw, W. C., The cellular geography of aurora kinases, *Nature Reviews Molecular Cell Biology*, 4 (2003) 842-854.
41. Akgün, Ş., Küçüksayan, H. ve Akça, H., *Kanser Moleküler Biyolojisi*, 12. Bölüm Tümör Baskılayıcı Genler, Denizli, 2018.
42. Ni, S., Huang, D., Chen, X., Huang, J., Kong, Y., Xu, Y., Du, X. ve Sheng, W., C-kit gene mutation and CD117 expression in human anorectal melanomas, *Human Pathology*, 43 (2012) 801-807.
43. Garnett, M. J. ve Marais, R., Guilty as charged: B-Raf is a human oncogene, *Cancer Cell*, 6,4 (2004) 316-19.
44. https://www.brafv600.com/BRAF_in_melanoma/role_of_BRAF/. 10 Haziran 2021.
45. Shen, H., Yuan, Y., Hu, H. G., Zhang, X., Ye, X., Li, M., Fang, W. ve Zheng, S., Clinical significance of K-ras and BRAF mutations in Chinese colorectal cancer patients, *World Journal of Gastroenterology*, 17 (2011) 809-816.
46. Vasko, V., Ferrand, M., Di Cristofaro, J., Carayon, P., Henry, J. F. ve De Micco, C., Specific pattern of RAS oncogene mutations in follicular thyroids tumors, *Journal of Clinical Endocrinology and Metabolism*, 88,6 (2003) 2745-2752.
47. Theoharis, C., Roman, S. ve Sosa, A. J., The molecular diagnosis and management of thyroid neoplasms, *Current Opinion in Oncology*, 24 (2012) 35-41.
48. <https://debuglies.com/2019/06/06/new-therapy-for-advanced-melanoma-braf-v600e/>. 10 Haziran 2021.
49. Kumar, V., Abbas, A. K. ve Fausto, N., The female genital tract, chapter 22, in Robbins and Cotran Pathologic Basis of Disease, Seventh Edition, Elsevier Saunders, Philadelphia, 1092- 1104, 2005.

50. Singer, G., Oldt III, R., Cohen, Y., Wang, B. G., Sidransky, D., Kurman, R. J. ve Shih, L., Mutations in BRAF and KRAS Characterize the Development of Low-Grade Ovarian Serous Carcinoma, Journal of the National Cancer Institute, 95 (2003) 484-486.
51. Ersöz, C. C., Overin Seröz Adenokarsinomları ve Seröz Borderline Tümörlerinde RAS/RAF/MEK/MAPK Yolağındaki Moleküler Değişikliklerin Araştırılması, Ankara Üniversitesi Tıp Fakültesi Mecmuası, 73,2 (2020) 144-148.
52. Pellegriti, G., Frasca, F., Regalbuto, C., Squatrito, S. ve Vigneri, R., Worldwide increasing incidence of thyroid cancer: update on epidemiology and risk factors, Journal of Cancer Epidemiology, 2013 (2013) 965212.
53. Enewold, L., Zhu, K., Ron, E., Marrogi, A. Z., Stojadinovic, A., Peoples, G. E. ve Devesa, S. S., Rising thyroid cancer incidence in the United States by demographic and tumor characteristics, 1980–2005, Cancer Epidemiol Biomarkers Prev., 18 (2009) 784–91.
54. Dmour, S., Al-Maaitah, A., Eren, T., Ekinci, O. ve Alimoglu, O., Prevalence of BRAF Mutation in Papillary Thyroid Cancer, Dicle Tıp Dergisi/ Dicle Medical Journal, 43,4 (2016) 569-575.
55. Xing, M., Alzahrani, A. S., Carson, K. A., Shong, Y. K., Kim, T. Y., Viola, D., Elisei, R., Bendlova, B., Yip, L., Mian, C., Vianello, F., Tuttle, R. M., Robenshtok, E., Fagin, J. A., Puxeddu, E., Fugazzola, L., Czarniecka, A., Jarzab, B., O’Neil, C. J., Sywak, M. S., Lam, A. K., Riesco-Eizaguirre, G., Santisteban, P., Nakayama, H., Clifton-Bligh, R., Tallini, G., Holt, E. H. ve Sykorova, V., Association between BRAF V600E mutation and recurrence of papillary thyroid cancer, Journal of Clinical Oncology, 33 (2015) 42-50.
56. Schlumberger, M. J., Papillary and follicular thyroid carcinoma, The New England Journal of Medicine, 338 (1998) 297-306.
57. Nakayama, H., Yoshida, A., Nakamura, Y., Hayashi, H., Miyagi, Y., Wada, N., Rino, Y., Masuda, M. ve Imada T., Clinical significance of BRAF (V600E) mutation and Ki-67 labeling index in papillary thyroid carcinomas, Anticancer Research, 27,5B (2007) 3645-3649.
58. Henderson, Y. C., Shellenberger, T. D., Williams, M. D., El-Naggar, A., Fredrick, M. J., Cieply, K. M. ve Clayman, G. L., High rate of BRAF and RET/PTC dual mutations associated with recurrent papillary thyroid carcinoma, Clinical Cancer Research, 15 (2009) 485-91.
59. Xing, M., Recent advances in molecular biology of thyroid cancer and their clinical implications, Otolaryngol Clin North Am., 41 (2008) 1135-1146.
60. Lundgren, C. I., Hall, P., Dickman, P. W. ve Zedenius, J., Influence of surgical and postoperative treatment on survival in differentiated thyroid cancer, British Journal of Surgery, 94 (2007) 571-577.

61. Dahm, R., Friedrich Miescher and the discovery of DNA, *Developmental Biology*, 27, 2 (2005) 274-288.
62. Tufano, R. P., Teixeira, G. V., Bishop, J., Carson, K. A. ve Xing, M., BRAF mutation in papillary thyroid cancer and its value in tailoring initial treatment: a systematic review and meta-analysis, *Baltimore*, 91 (2012) 274-86.
63. Cohen, Y., Xing, M., Mambo E., Guo, Z., Wu, G., Trink, B., Beller, U., Westra, W. H., Ladenson, P. W. ve Sidransky, D., BRAF mutation in papillary thyroid carcinoma, *Journal of the National Cancer Institute*, 95 (2003) 625-627.
64. Şimşek, T. ve Cantürk, N. Z., Tiroid Karsinomlarının Genetik Temeli: Bir Derleme, *Okmeydanı Tıp Dergisi*, 29,1 (2013)1-7.
65. Türkiye Kanser İstatistikleri 2017, T.C. Sağlık Bakanlığı Halk Sağlığı Genel Müdürlüğü, Ankara, 2021.
66. <https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/ovarian-cancer/survival#heading=One>. 3 Şubat 2020.
67. Davies, H., Bignell, G. R., Cox, C., Stephens, P., Edkins, S., Clegg, S., Teague, J., Woffendin, H., Garnett, M. J., Bottomley, W., Davis, N., Dicks, E., Ewing, R., Floyd, Y., Gray, K., Hall, S., Hawes, R., Hughes, J., Kosmidou, V., Menzies, A., Mould, C., Parker, A., Stevens, C., Watt, S., Hooper, S., Wilson, R., Jayatilake, H., Gusterson, B. A., Cooper, C., Shipley, J., Hargrave, D., Pritchard-Jones, K., Maitland, N., Chenevix-Trench, G., Riggins, G. J., Bigner, D. D., Palmieri, G., Cossu, A., Flanagan, A., Nicholson, A., Ho, J. W., Leung, S. Y., Yuen, S. T., Weber, B. L., Seigler, H. F., Darrow, T. L., Paterson, H., Marais, R., Marshall, C. J., Wooster, R., Stratton, M. R. ve Futreal, P. A., Mutations of the BRAF gene in human cancer, *Nature*, 417 (2002) 949-54
68. Kızmaz, M. Z., Paylan, İ. C. Ve Erkan, S., DNA Dizilemenin Tarihsel Gelişimi, *Gaziosmanpaşa Bilimsel Araştırma Dergisi*, 6,2 (2017) 47-53.
69. https://tr.wikipedia.org/wiki/Genetik_kod. 23 Haziran 2021.
70. <https://www.chegg.com/homework-help/questions-and-answers/dna-molecule-question-5-3-caattgataggtcagtcatacat-5-5-g-ttaactatccagtcagttatgta-3-second-q34825567>. 19 Aralık 2021.
71. <https://evrimagaci.org/dna-dizileme-yontemleri-genleri-nasil-diziliyoruz-21>. 9 Ağustos 2021.
72. Madigan, M. T., Martinko, J. M., Bender, K. S., Buckley, D. H. ve Stahl, D. A., *Brock Biology of Microorganisms*, Fourteenth Edition, Pearson Education, Inc., New Jersey, 2015.
73. Anderson, S., Shotgun DNA sequencing using cloned DNase I-generated fragments, *Nucleic acids research*, 9, 13 (1981) 3015-3027.

74. King, J. L., LaRue, B. L., Novroski, N. M., Stoljarova, M., Seo, S. B., Zeng, X., Warshauer, D. H., Davis, C. P., Parson, W. ve Sajantila, A., High-quality and high-throughput massively parallel sequencing of the human mitochondrial genome using the Illumina MiSeq, Forensic Science International: Genetics, 12 (2014) 128-135.
75. <https://www.ncbi.nlm.nih.gov/home/about/mission/>. 26 Temmuz 2021.
76. <http://biyo-informatik.blogspot.com/2016/10/fasta-nedir.html>. 26 Temmuz 2021.
77. <https://zhanggroup.org/FASTA/>. 26 Temmuz 2021.
78. <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE113695>. 26 Temmuz 2021.
79. Üstek, D., Abacı, N., Sırma, S. ve Çakiris, A., Yeni Nesil DNA Dizileme, Deneysel Tıp Araştırma Enstitüsü Dergisi, 1,1 (2011) 11-18.
80. Al-Kuwari, S. M. S. A., Integrated-Key Cryptographic Hash Functions, Doctoral Thesis, University of Bath, Department of Computer Science, UK, 2011.
81. <https://learncryptography.com/hash-functions/hash-collision-attack>. 19 Aralık 2019.
82. <http://repository.essex.ac.uk/22334/1/Gokce%20Caylak%20Kayaturan-repository.pdf>. 23 Aralık 2019.
83. Kayaturan, G. Ç., Representing Shortest Paths in Graphs Using Bloom Filters without False Positives and Applications to Routing in Computer Networks, Doctoral Thesis, University of Essex, Colchester, 2017.
84. https://en.wikipedia.org/wiki/Bloom_filter. 19 Aralık 2019.
85. Schatz, M. C., Trapnell, C., Delcher, A. L. ve Varshney, A., High-throughput sequence alignment using Graphics Processing Units, BMC Bioinformatics, 8 (2007) 474.
86. Chen, S., Huang, T., Wen, T., Li, H., Xu, M. ve Gu, J., MutScan: fast detection and visualization of target mutations by scanning FASTQ data, BMC Bioinformatics, 19 (2018) 16.

ÖZGEÇMİŞ

İlköğretim ve lise eğitimini Trabzon'da tamamladı. 2011 yılında Kocaeli Üniversitesi Bilgisayar Mühendisliği Bölümü'nden mezun oldu. 2018 yılında Karadeniz Teknik Üniversitesi Fen Bilimleri Enstitüsü Bilgisayar Mühendisliği Anabilim Dalı'nda yüksek lisans yapmaya hak kazandı. Kurucu ortağı olduğu Phi Software Yazılım ve Danışmanlık Hiz. Tic. Ltd. Şti'de çalışma hayatını sürdürmektedir.

