

**T.C.
ERCIYES ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ
TARIMSAL BİYOTEKNOLOJİ ANABİLİM DALI**

**ANADOLU YABAN VE ANKARA KEÇİLERİNDE TÜM
GENOM DİZİLENMESİ VE BİYOİNFORMATİK
ANALİZLER**

**Hazırlayan
Saffet TEBER**

**Danışman
Doç. Dr. Osman İBİŞ**

Doktora Tezi

**Mayıs 2024
KAYSERİ**

**T.C.
ERCIYES ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ
TARIMSAL BİYOTEKNOLOJİ ANABİLİM DALI**

**ANADOLU YABAN VE ANKARA KEÇİLERİNDE TÜM
GENOM DİZİLENMESİ VE BİYOBİYOİNFORMATİK
ANALİZLER
(Doktora Tezi)**

**Hazırlayan
Saffet TEBER**

**Danışman
Doç. Dr. Osman İBİŞ**

**Bu çalışma, Erciyes Üniversitesi Bilimsel Araştırma Projeleri Birimi tarafından
FDK-2022-11535 kodlu proje ve TÜBİTAK tarafından 122O844 kodlu proje ile
desteklenmiştir.**

**Mayıs 2024
KAYSERİ**

BİLİMSEL ETİĞE UYGUNLUK

Bu çalışmadaki tüm bilgilerin, akademik ve etik kurallara uygun bir şekilde elde edildiğini beyan ederim. Aynı zamanda bu kural ve davranışların gerektirdiği gibi, bu çalışmanın özünde olmayan tüm materyal ve sonuçları tam olarak aktardığımı ve referans gösterdiğimi belirtirim.

Saffet Teber

İmza

“Anadolu Yaban ve Ankara Keçilerinde Tüm Genom Dizilenmesi ve Biyoinformatik Analizler” adlı Doktora tezi, Erciyes Üniversitesi Lisansüstü Tez Önerisi ve Tez Yazma Yönergesi’ ne uygun olarak hazırlanmıştır.

Hazırlayan

Saffet TEBER

İmza

Danışman

Doç. Dr. Osman İBİŞ

İmza

Tarımsal Biyoteknoloji ABD Başkanı

Prof. Dr. Semih YILMAZ

İmza

TEŞEKKÜR

“Anadolu Yaban ve Ankara Keçilerinde Tüm Genom Dizilenmesi ve Biyoinformatik Analizler” başlıklı tez çalışması, Erciyes Üniversitesi Ziraat Fakültesi Tarımsal Biyoteknoloji Bölümü ve Erciyes Üniversitesi Genom ve Kök Hücre Merkezi (GenKök) Genom Laboratuvarı’nda yürütülmüştür. Tez çalışması kapsamında yapılan biyoinformatik analizler Erciyes Üniversitesi Bilgi İşlem Daire Başkanlığı bünyesindeki sunucular ve TÜBİTAK ULAKBİM, Yüksek Başarım ve Grid Hesaplama Merkezi’nde (TRUBA kaynaklarında) gerçekleştirilmiştir. Tez çalışması Erciyes Üniversitesi Bilimsel Araştırma Projeleri Birimi tarafından ve TÜBİTAK tarafından 1002-B proje türü kapsamında desteklenerek yürütülmüştür. Desteklerinden dolayı TÜBİTAK ve Erciyes Üniversitesi’ne teşekkürlerimi sunarım.

Tez çalışmasını gerçekleştiren Saffet TEBER, 100/2000 YÖK Doktora Bursiyeri’dir. Bu kapsamda desteklerinden dolayı Yükseköğretim Kurulu’na teşekkürlerimi sunarım.

Bilgisini ve desteğini hiçbir zaman esirgemeyen ve her zaman çalışmalarımda bana yardımcı olan danışmanım Doç. Dr. Osman İBİŞ’e sonsuz minnettarım.

Saffet TEBER

Mayıs 2024, KAYSERİ

ANADOLU YABAN VE ANKARA KEÇİLERİNDE TÜM GENOM DİZİLENMESİ VE BİYOİNFORMATİK ANALİZLER

Saffet TEBER

Erciyes Üniversitesi, Fen Bilimleri Enstitüsü
Doktora Tezi, Mayıs 2024
Danışman: Doç. Dr. Osman İBİŞ

ÖZET

Dünya’da yaygın olarak yetiştiriciliği yapılan evcil keçi (*Capra hircus*), en eski evcil hayvanlardan birisidir. Keçiler yaklaşık olarak 10.000 yıl önce Bereketli Hilal’de evcilleştirilmiş ve bunu takiben insan hareketiyle dünya geneline yayılmış, farklı karakterler bakımından yetiştiriciliği yapılan çok sayıda keçi ırkı ortaya çıkmıştır. Bu tez çalışmasında Ankara keçisinin genomu hibrit yöntemle *de novo* olarak birleştirilmiştir ve anotasyonu yapılmıştır. Elde edilen kontig seviyedeki genom birleştirme ~2,6 Gb uzunluğunda ve N50 değeri ~11 Mb olup, toplam 1.738 kontigden oluşmaktadır. Ankara keçisi genomunun skafoldlanması sonucunda 710 skafold, BUSCO değeri %95,7 olarak elde edilmiştir. Ankara keçisi genomunun tekrar içeriği %41,66 olarak belirlenmiş ve homoloji temelli yaklaşımla yapılan gen tahmininde toplam 21.645 gen tahmin edilmiştir. Ülkemizde doğal yayılım gösteren ve evcil keçinin atası olan Anadolu yaban keçisinin (*Capra aegagrus*) tüm genomu ise Illumina kısa okumaları kullanılarak *de novo* birleştirilmiştir. Birleştirilen genom 181.575 kontig içermekte olup, yaklaşık 2,7 Gb boyutundadır. Yaban keçisi genomunun skafoldlanması sonucunda BUSCO değeri %94,7 olan 68.461 skafold elde edilmiştir. Yaban keçisi genomunun tekrar içeriği %42,13 olarak belirlenmiş ve homoloji temelli yaklaşımla yapılan gen tahmininde toplam 21.409 gen tahmin edilmiştir. Kısa okumaların referans keçi genomlarına haritalanması ile Ankara keçisi keçisinde 11.168.617 ve yaban keçisinde 10.935.456 varyant tespit edilmiştir. Ankara keçisi ve yaban keçisi genomları arasında yapılan karşılaştırma ile kromozomal skafoldlar arasında oldukça yüksek sentenik bloklar tespit edilmiştir.

Anahtar Kelimeler: Ankara Keçisi, *Capra hircus*, Yaban keçisi, *Capra aegagrus*, Tüm genom dizileme

WHOLE GENOME SEQUENCING AND BIOINFORMATIC ANALYSIS OF ANATOLIAN WILD AND ANGORA GOATS

Saffet TEBER

Erciyes University, Graduate School of Natural and Applied Sciences

PhD Thesis, May 2024

Supervisor: Assoc. Dr. Osman İBİŞ

ABSTRACT

The domestic goat (*Capra hircus*), which is widely cultivated around the world, is one of the oldest domestic animals. Goats, which serve as many resources for humans, domesticated in the Fertile Crescent approximately 10,000 years ago and subsequently spread around the world with human movement, resulting in a large number of breeds with different characteristics. In this thesis, the genome of Angora goat was *de novo* assembled and annotated by hybrid method. The resulting contig-level genome assembly is ~2.6 Gb in length and has an N50 value of ~11 Mb and consists of a total of 1,738 contigs. As a result of the scaffolding of the Angora goat genome, 710 scaffolds and a BUSCO value of 95.7% were obtained. The repeat content of the Angora goat genome was determined as 41.66% and a total of 21,645 genes were predicted in gene prediction using homology-based approach. The whole genome of Anatolian wild goat (*Capra aegagrus*), another goat distributed in Turkey, was assembled *de novo* using Illumina short reads. The assembled genome contains 181,575 contigs and is ~2.7 Gb in size. Scaffolding of the wild goat genome resulted in 68,461 scaffolds with a BUSCO value of 94.7%. The repeat content of the wild goat genome was determined as 42.13% and a total of 21,409 genes were predicted by homology-based gene prediction. By mapping short reads to reference goat genomes, 11,168,617 and 10,935,456 variants were identified in Angora goat and wild goat, respectively. Comparison between Angora goat and wild goat genomes revealed very high syntenic blocks between chromosomal scaffolds.

Keywords: Angora Goat, *Capra hircus*, Wild goat, *Capra aegagrus*, Whole genome sequencing

İÇİNDEKİLER

ANADOLU YABAN VE ANKARA KEÇİLERİNDE TÜM GENOM DİZİLENMESİ VE BİYOİNFORMATİK ANALİZLER

BİLİMSEL ETİĞE UYGUNLUK.....	ii
YÖNERGEYE UYGUNLUK.....	iii
KABUL VE ONAY	iv
TEŞEKKÜR.....	v
ÖZET	vi
ABSTRACT.....	vii
İÇİNDEKİLER	viii
KISALTMALAR.....	xi
TABLolar LİSTESİ.....	xiii
ŞEKİLLER LİSTESİ	xv
GİRİŞ	1

1. BÖLÜM

GENEL BİLGİLER ve LİTERATÜR ÇALIŞMASI

1.1. Keçiler hakkında genel bilgi.....	4
1.2. Sekanslama sistemleri.....	6
1.3. Genom birleştirme	7
1.4. Genom anotasyonu.....	10
1.5. Mitokondriyal DNA	12
1.6. Genom çalışmalarının önemi	13
1.7. Çiftlik hayvanlarında yapılan genom çalışmaları.....	13
1.8. Keçilerde yapılan genom çalışmaları	17
1.9. Diğer memeli hayvanlarda yapılan genom çalışmaları	19
1.10. Diğer hayvanlarda yapılan genom çalışmaları.....	26
1.11. Genom çalışmalarının tarımsal faydaları	37

2. BÖLÜM

YÖNTEM VE MATERYAL

2.1. Laboratuvar çalışmaları.....	43
2.1.1. Hayvan materyalinin temini	43
2.1.2. Illumina platformu için genomik DNA izolasyonları	44

2.1.3. Ankara keçisinde Oxford Nanopore platformu için yüksek molekül ağırlıklı genomik DNA'nın elde edilmesi	44
2.1.4. Illumina platformunda sekanslama kütüphanelerinin kurulması	45
2.1.5. Ankara keçisi için Oxford Nanopore platformu sekanslama kütüphanesinin hazırlanması	45
2.1.6. Oxford Nanopore MinION platformunun çalıştırılması.....	45
2.2. Biyoinformatik analizler.....	46
2.2.1. Elde edilen Ankara keçisi ve Yaban keçisi kısa okumaların filtrelenmesi .	47
2.2.2. Elde edilen Ankara keçisi uzun okumaların filtrelenmesi	47
2.2.3. Ankara keçisi genomunun hibrit yöntemlerle birleştirilmesi	48
2.2.4. Yaban keçisi genomunun Illumina kısa okumaları ile birleştirilmesi	48
2.2.5. Ankara keçisi ve yaban keçisi genomlarının skafoldlanması.....	48
2.2.6. Ankara keçisi ve yaban keçisi komple mitokondriyal DNA'larının elde edilmesi.....	48
2.2.7. Mitogenom bazlı filogenetik analizler	50
2.2.8. Keçi genomlarının anotasyonu	51
2.2.9. Ankara keçisi ve yaban keçisi genomları arasındaki sinteni	52
2.2.10. Tüm genomlara dayalı filogenetik analizler	52
2.2.11. Keçi genomlarının referans genomlara haritalanması ve varyasyonun belirlenmesi.....	53
2.2.12. Elde edilen ham verilerinin NCBI-SRA veri tabanına yüklenmesi.....	53

3. BÖLÜM

BULGULAR

3.1. Yüksek molekül ağırlıklı genomik DNA izolasyonu	54
3.2. Illumina platformunda dizileme için yapılan genomik DNA izolasyonu.....	54
3.3. Ankara keçisi tüm genomunun Illumina kısa okuma platformunda dizilenmesi	55
3.3. Yaban keçisi tüm genomunun Illumina kısa okuma platformunda dizilenmesi	55
3.4. Yaban keçisi genomunun birleştirilmesi.....	56
3.5. Oxford Nanopore MinION platformu kullanılarak elde edilen okuma sonuçları	57
3.6. Hibrit yöntemle birleştirilen genomun istatistiksel değerlendirilmesi.....	63

3.7. Skafoldlama sonuçları	64
3.8. Genom anotasyon sonuçları	72
3.8.1. Varyant anotasyon sonuçları	72
3.8.2. Tekrar elementleri	75
3.8.3. Gen tahmini	77
3.9. Ankara keçisi ve yaban keçisi genomları arasındaki sinteni	83
3.10. Tüm genomlar kullanılarak gerçekleştirilen filogenetik analizler	87
3.11. Mitogenom birleştirme sonuçları.....	88
3.11.1. Mitogenom organizasyonu	88
3.11.2. Mitogenom tRNA yapıları.....	95
3.11.3.Mitogenomlar temelli filogenetik analizler	97
3.12. Ham okuma verilerinin NCBI-SRA veri tabanına yüklenmesi	100

4. BÖLÜM

TARTIŞMA-SONUÇ ve ÖNERİLER

4.1. Tartışma	101
4.1.1. Genom birleştirme	101
4.1.2. Genom anotasyonu.....	105
4.1.3. Mitogenom organizasyonu	107
4.1.4. Varyant	108
4.1.5. Sinteni.....	108
4.1.6. Filogeni.....	108
4.2. Sonuç ve Öneriler	109
KAYNAKÇA	110
ÖZGEÇMİŞ.....	135

KISALTMALAR

μL	: Mikrolitre
bç	: Baz çifti (Base pair - bp)
BM	: Birleşmiş Milletler
BUSCO	: Evrensel Tek Kopya Ortologların Karşılaştırılması (Benchmarking Universal Single-Copy Orthologs)
circRNA	: Dairesel RNA (Circular RNA)
DNA	: Deoksiribo nükleik asit
FAO	: Gıda ve Tarım Örgütü (Food and Agriculture Organization of the United Nations)
fmol	: femtomol
Gb	: Gigabaz
GB	: Gigabyte
GC	: Guanin-Sitozin
gDNA	: Genomik DNA
GGVD	: Goat genome variation database
GO	: Gen Ontoloji (Gene Ontology)
HMW	: Yüksek Molekül Ağırlıklı (High Molecule Weight)
IUCN	: Dünya Doğayı ve Doğal Kaynakları Koruma Birliği (International Union for Conservation of Nature and Natural Resources)
İnDel	: İnsersiyon-delesyon (insertion-deletion)
Kb	: Kilobaz
kg	: Kilogram
lncRNA	: Uzun kodlamayan RNA (Long non-coding RNA)
Mb	: Megabaz
MHC	: Majör histokompatibilite kompleksi
mL	: Mililitre
MP	: Eş-çift (Mate-Pair)
mtDNA	: mitokondriyal DNA
NCBI	: National Center for Biotechnology Information
ncRNA	: Kodlamayan RNA (non-coding RNA)

NGS	: Yeni Nesil Dizileme (Next Generation Sequencing)
NT	: Neredeyse Tehdit Altında (Near Threatened)
ONT	: Oxford Nanopore Technologies
PacBio	: Pasific Bioscience
PCG	: Protein Kodlayan Gen (Protein Coding Gene)
PE	: Çift uçlu (Paired-End)
QTL	: Kantitatif Özellik Lokusu (Quantitative trait loci)
RNA	: Ribo nükleik asit
rRNA	: Ribozomal RNA
rSNP	: regülatör SNP
SMRT	: Tek Molekül Gerçek Zamanlı (Single Molecule Real Time)
SNP	: Tek Nükleotit Polimorfizmi (Single Nucleotide Polymorphism)
SNV	: Tek nükleotit varyantı (Single nucleotide variant)
SRA	: Sequence Read Archive
TF	: Transkripsiyon faktörü
TFBS	: Spesifik bağlanma bölgelerini hedefleri (Targets specific binding-sites)
tRNA	: Taşıyıcı RNA
UCSC	: University of California Santa Cruz
VU	: Hassas (Vulnerable)

TABLOLAR LİSTESİ

Tablo 1.	Ankara keçisi ve yaban keçisi mitogenomlarının amplifiye edilmesi için Long- Range PCR’da kullanılan primerler	49
Tablo 2.	Filogenetik analizlerde kullanılan keçi mitogenomları	50
Tablo 3.	Ankara keçisi genomunun Illumina platformu ile kısa okumaları	55
Tablo 4.	Yaban keçisi genomunun Illumina kısa okumaları.....	55
Tablo 5.	Yaban keçisinin genom birleştirmesinde elde edilen kontiglerin metrikleri.....	56
Tablo 6.	Yaban keçisi genomunun Illumina kısa okumalar ile birleştirilmesi sonrası BUSCO kullanılarak yapılan genom değerlendirmesinin istatistiksel değerleri	57
Tablo 7.	Ankara keçisi genomunun Oxford Nanopore uzun okuma değerleri	60
Tablo 8.	Guppy ile basecalling sonrası okuma değerleri	62
Tablo 9.	Ankara keçisi genom birleştirmesinde elde edilen kontiglerin metrikleri.	63
Tablo 10.	BUSCO kullanılarak yapılan genom değerlendirmesinin istatistiksel değerlendirmesi.....	64
Tablo 11.	Ankara keçisi genom birleştirmesinde elde edilen skafoldların metrikleri	65
Tablo 12.	Ankara keçisi genomunun skafoldlanma sonrası BUSCO kullanılarak yapılan genom değerlendirmesinin istatistiksel verileri	66
Tablo 13.	Yaban keçisi genom birleştirmesinde elde edilen skafoldların metrikleri.	67
Tablo 14.	Yaban keçisi genomunun skafoldlanma sonrası BUSCO kullanılarak yapılan genom değerlendirmesinin istatistiksel verileri	68
Tablo 15.	Ankara keçisi genom birleştirmesindeki en uzun 31 skafoldun metrikleri	69
Tablo 16.	Ankara keçisi genom birleştirmesinde elde edilen en uzun 31 skafoldun BUSCO verileri.....	70
Tablo 17.	Yaban keçisi genom birleştirmesindeki en uzun 31 skafoldun metrikleri.	70
Tablo 18.	Yaban keçisi genom birleştirmesinde elde edilen en uzun 31 skafoldun BUSCO değerleri.....	71
Tablo 19.	Ankara keçisi varyant anotasyonunun bölgesel dağılım sayıları.....	73
Tablo 20.	Ankara keçisi varyant anotasyonunun nedensel sayıları	74

Tablo 21.	Ankara keçisi genomundaki tekrar bölgelerinin anotasyon sonuçları	75
Tablo 22.	Yaban keçisi genomundaki tekrar bölgelerinin anotasyonu	76
Tablo 23.	Ankara keçisi gen tahmininde BUSCO değerleri	77
Tablo 24.	Ankara keçisi genomunun fonksiyonel anotasyon verileri	79
Tablo 25.	Yaban keçisi gen tahmininde BUSCO verileri	81
Tablo 26.	Yaban keçisi genomunun fonksiyonel anotasyon verileri	82
Tablo 27.	Ankara keçisi mitogenomunun baz kompozisyonu	89
Tablo 28.	Yaban keçisi mitogenomunun baz kompozisyonu	89
Tablo 29.	Ankara keçisi mitogenomun karakterizasyon tablosu	91
Tablo 30.	Yaban keçisi mitogenomun karakterizasyon tablosu	92

ŞEKİLLER LİSTESİ

Şekil 1.	Evcil keçi ve yaban keçisinin taksonomisi	3
Şekil 2.	Bereketli Hilâl bölgesinde evcilleştirilen türlerin tarihi (Zeder, 2008)	4
Şekil 3.	2015 yılında kilometrekareye düşen keçi sayısı (FAO).....	5
Şekil 4.	Yaban keçisinin dünyadaki yayılış alanları (ICUN Red List)	6
Şekil 5.	Çalışmada kullanılan Ankara keçisi	44
Şekil 6.	Ankara keçisi genomunun dizilenmesinde kullanılan iş akışı	46
Şekil 7.	Yaban keçisi genomunun dizilenmesinde kullanılan iş akışı	47
Şekil 8.	HMW DNA izolasyon kiti ile izole edilen Ankara keçisi gDNA'sının agaroz jel görüntüsü.....	54
Şekil 9.	Yaban keçisi genom birleştirmesinin BUSCO sonuçlarının pasta dağılım grafiği.....	57
Şekil 10.	Oxford Nanopore MinION cihazı birinci run sonuçları	58
Şekil 11.	Oxford Nanopore MinION cihazı ikinci run sonuçları.....	58
Şekil 12.	Oxford Nanopore MinION cihazı üçüncü run sonuçları	59
Şekil 13.	Oxford Nanopore MinION platformu 4.run sonuçları.....	60
Şekil 14.	Ankara keçisi kontig BUSCO sonuçları pasta dağılımı.....	64
Şekil 15.	Ankara keçisi genom birleştirmesinde elde edilen skafoldların dairesel görseli.....	65
Şekil 16.	Ankara keçisi skafold BUSCO sonuçları pasta dağılımı	66
Şekil 17.	Yaban keçisi genom birleştirmesinde elde edilen skafoldların dairesel görseli.....	67
Şekil 18.	Yaban keçisi skafold BUSCO sonuçları pasta dağılımı	68
Şekil 19.	Ankara keçisi genom birleştirmesinde elde edilen en uzun 31 skafoldun dairesel gösterimi	69
Şekil 20.	Ankara keçisi genom birleştirmesinde elde edilen en uzun 31 skafoldun BUSCO sonuçları pasta dağılımı	70
Şekil 21.	Yaban keçisi genom birleştirmesinde elde edilen en uzun 31 skafoldun dairesel gösterimi.....	71
Şekil 22.	Yaban keçisi genom birleştirmesinde elde edilen en uzun 31 skafoldun BUSCO sonuçları pasta dağılımı	72
Şekil 23.	Ankara keçisi varyant anotasyonu	72

Şekil 24.	Ankara keçisi varyant anotasyonunun bölgesel dağılımı	73
Şekil 25.	Ankara keçisi varyant anotasyonunun nedensel sonuçları	74
Şekil 26.	Ankara keçisi genomunun tekrar içeriği grafiği	76
Şekil 27.	Yaban keçisi genomunun tekrar içeriği grafiği.....	77
Şekil 28.	Ankara keçisi genom anotasyonunun BUSCO grafiksel gösterimi.....	78
Şekil 29.	Ankara keçisi genomunun fonksiyonel anotasyon sonuçlarının GO dağılımları	80
Şekil 30.	Ankara keçisi genomunun fonksiyonel anotasyon sonuçlarının GO dağılımları	80
Şekil 31.	Yaban keçisi genom anotasyonunun BUSCO grafiksel gösterimi	81
Şekil 32.	Yaban keçisi genomunun fonksiyonel anotasyon sonuçlarının GO dağılımları.....	83
Şekil 33.	Yaban keçisi genomunun fonksiyonel anotasyon sonuçlarının GO dağılımları.....	83
Şekil 34.	Ankara keçisi ve yaban keçisi en uzun 31 skafoldu arasındaki sinteni	84
Şekil 35.	Ankara keçisi ile yaban keçisi en uzun 31 skafoldları arasındaki sintenin dairesel gösterimi	84
Şekil 36.	Yaban keçisi ve referans yaban keçisi arasındaki en uzun 31 skafoldu arasındaki sinteni	85
Şekil 37.	Yaban keçisi ile referans yaban keçisi genomu arasındaki tüm skafoldların sintenisi	85
Şekil 38.	Ankara keçisi ve referans evcil keçi arasındaki en uzun 31 skafoldu arasındaki sinteni	86
Şekil 39.	Ankara keçisinin tüm skafoldları ile referans evcil keçi genomu arasındaki sinteni	86
Şekil 40.	Filogenetik analizlerde kullanılan tüm genomların BUSCO değerlendirmeleri	87
Şekil 41.	Tüm genomlar kullanılarak oluşturulan ML ağacı	88
Şekil 42.	Ankara keçisi ve yaban keçisi mitogenomlarının dairesel formu.....	90
Şekil 43.	Ankara keçisi RSCU-Kodon kullanımı grafiği.....	93
Şekil 44.	Yaban keçisi RSCU-Kodon kullanımı grafiği	94
Şekil 45.	Ankara keçisi mitogenomundaki tRNA'ların tahmini sekonder yapıları ..	95
Şekil 46.	Yaban keçisi mitogenomundaki tRNA'ların tahmini sekonder yapıları ...	96

Şekil 47.	ML yöntemi kullanılarak oluşturulan filogenetik ağaç.....	98
Şekil 48.	ML Filogenetik ağacın büyütülmüş hali	99



GİRİŞ

Birinci nesil sekanslama sistemlerinden olan Sanger sekanslama platformundan sonra gelişen teknolojiye bağılı olarak ikinci ve üçüncü nesil sekanslama platformları olan yeni nesil sekanslama (NGS) teknolojileri geliştirilerek kullanılmaya başlanmış, böylece baz başı maliyet ve iş gücü gereksinimi büyük oranda azalmıştır. Bunu takiben model ve model olmayan organizmaların farklı seviyelerdeki genomlarının dizilenmesinde büyük bir artış olmuştur. Yeni nesil sekanslama platformlarının yaygınlaşması ve çeşitlenmesiyle birlikte de birçok donanımlı araştırma merkezlerindeki laboratuvarlar kendi bünyesinde tüm genom dizileme çalışmalarını gerçekleştirebilmektedir. Yeni nesil sekanslama sistemlerinden 2. nesil dizileme sistemleri kısa ve doğruluğu yüksek okuma oranlarına sahipken, 3. nesil dizileme sistemleri hata oranı nispeten biraz daha yüksek ancak daha uzun okumalar gerçekleştirebilmektedir.

Sekanslama platformlarındaki bu gelişmelerle genomlar, nispeten küçük parçalar halinde dizilerek çeşitli algoritmalar yardımıyla biyoinformatik yöntemlerle birleştirilmektedir. Bu işlemler kısaca kontiglerin elde edilmesi ve sonrasında kromozomların veya kromozom altı seviyedeki skafoldların elde edilmesine dayanmaktadır. Bir türe ait ilk kez veya mevcut referans genom kullanılmadan genom birleştirme yapılıyorsa, bu işlem “*de novo* birleştirme” olarak adlandırılmaktadır.

Genom birleştirmesi yapıldıktan sonra elde edilen birleştirmede genlerin bulunması, tekrar bölgelerinin, ekzonların ve intronların belirlenmesi gibi yapısal işlemler ile yapısal ve fonksiyonel genom anotasyonu basamağı gelmektedir. Bu aşamada elde edilen genom sekanslarında tekrar elementleri bakımından maskeleyerek yapılarak çeşitli yöntem ve algoritmalar ile gen tahmini işlemleri gerçekleştirilmektedir. Böylece türler arasındaki gen içerikleri belirlenerek fonksiyonel anotasyonlar yapılabilmektedir.

Genom dizileme alıřmaları ile gerek tarım bilimlerinde gerekse de saėlık bilimlerinde byk geliřmeler kaydedilmiřtir. Bu sayede nispeten dřk maliyet ve az iř gc ile kalıtsal hastalıkların temelleri, genetik hastalıkların tanısı, tedavisi ve iliřkileri gibi saėlık alanındaki alıřmaları mmkn kılmıřtır. Genom alıřmaları, tarım bilimlerinde ise seleksiyon yntemlerine hız ve hassasiyet kazandırmıř, verim ve kalite gibi parametrelerde byk kazanımlar saėlamıřtır.

Evcil keiler (*Capra hircus*), Dnya apında yaygın olarak yetiřtiriciliėi yapılan bir tr olup insanlar iin nemli bir et, st, lif ve post kaynaėıdırlar. lkemizde kei yetiřtiriciliėi ekonomik olarak nemli bir faaliyet alanı olup, genellikle yerli ırklar ile ekstansif olarak yapılmaktadır. Ankara keisi, tiftik denilen kendine has lifleri ile nem kazanmaktadır. Tiftik dıřında Ankara keilerinden et retimi olarak da faydalanılmaktadır. Diėer yerli kei ırklarımız ile karřılařtırıldıėında Ankara keisi, kk csseli ve nispeten daha ge geliřmekte olup, oėunlukla Ankara ve evresinde yetiřtirilmektedir.

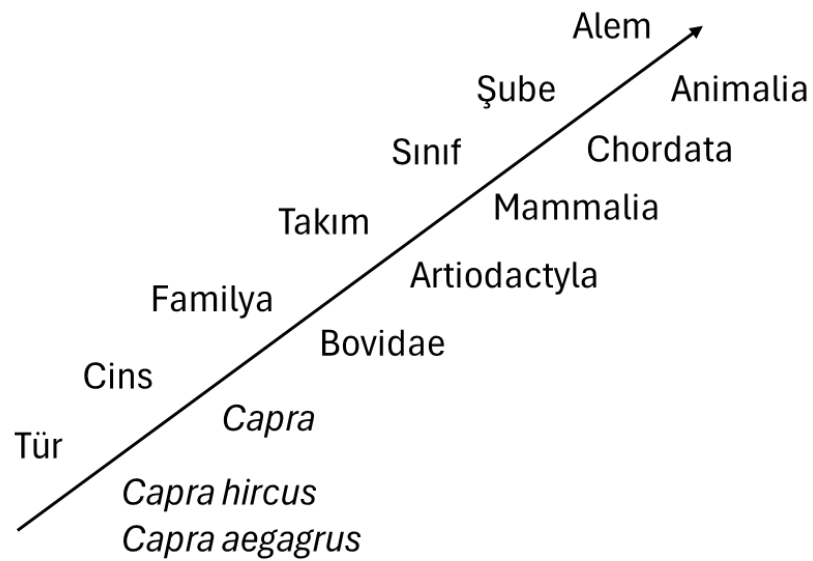
Dnya Doėa Koruma Birliėi (IUCN) tarafından “Tehdit Altında Girebilir (NT)” kategorisinde sınıflandırılan, Dnya genelinde geniř bir alanda ve lkemizde de gney ve doėu blgelerinde yayılıř gsteren bir diėer *Capra* cinsi yesi olan yaban keisi trlerinden Anadolu yaban keisi (*Capra aegagrus*), av turizmi ve genetik eřitlilik bakımından olduka nemlidir.

Bu tez alıřması kapsamında Ankara keisi tm genomunun uzun ve kısa okuma platformları kullanılarak hibrit yntemle, Anadolu yaban keisinin tm genomunun kısa okumalar kullanılarak *de novo* olarak birleřtirilmesi, genom anotasyonlarının gerekleřtirilmesi amalanmıřtır. Ayrıca tez alıřması kapsamında Ankara keisi kısa okumaları referans kei genomuna, Anadolu yaban keisi ise *Capra aegagrus* referans genomuna haritalanarak SNP/InDel’lerin de belirlenmesi, Ankara keisi ile yaban keisi arasındaki sentenik blgelerin arařtırılması ve filogenetik analizlerin gerekleřtirilmesi amalanmıřtır.

1. BÖLÜM

GENEL BİLGİLER VE LİTERATÜR ÇALIŞMASI

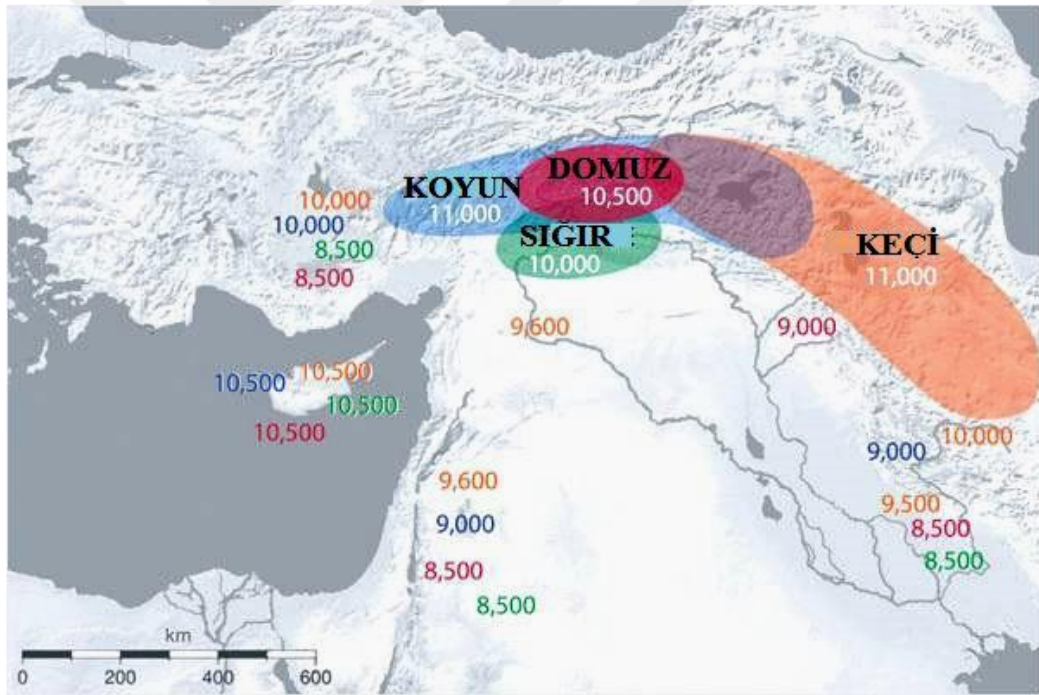
Ülkemiz üç kıtanın birleştiği coğrafi konumu ve iklimi ile oldukça zengin bir genetik çeşitliliğine sahiptir (Kryštufek ve Vohralik, 2001). Anadolu coğrafyasının dağlık topografyası, türleşme ve zoocoğrafik alt bölgelerin oluşumunda önemli bir rol oynamaktadır. Bu topografya, tür içi ve türler arası genetik farklılaşmayı sağlamaktadır (Bilgin, 2011). Türkiye memeli türleri bakımından zengin bir ülkedir. Ülkemizde 6 takım ve 36 familyaya ait yaklaşık 170 memeli türü bulunmaktadır (Kryštufek ve Vohralik, 2009; Karataş vd., 2021; Seyfi vd., 2021). Türkiye'nin doğusunu da kapsayan “Bereketli Hilal” denilen bölge, birçok hayvan için evcilleştirme merkezlerinden birisidir (Şekil 2) ve koyun, keçi, domuz, sığır gibi hayvanların bu bölgede evcilleştirildiği düşünülmektedir (MacHugh ve Bradley, 2001).



Şekil 1. Evcil keçi ve yaban keçisinin taksonomisi

1.1. Keçiler hakkında genel bilgi

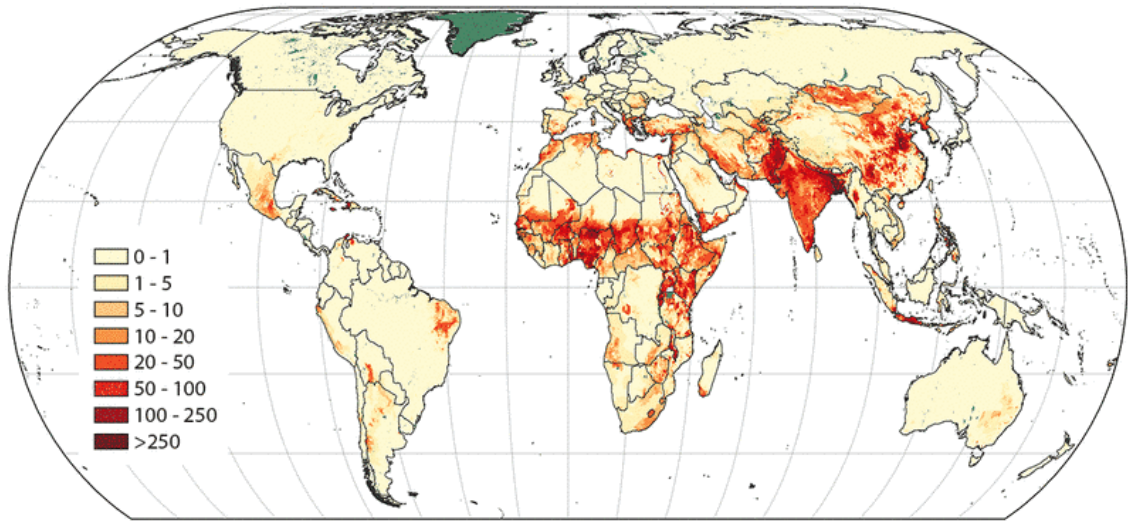
Bovidae familyası *Capra* cinsi içerisinde yer alan *Capra hircus* evcil bir keçi türü olup, en eski evcilleştirilmiş hayvanlardan birisidir (Şekil 1). Dünya genelinde özellikle de Çin, Hindistan ve diğer gelişmekte olan ülkelerde yaygın olarak yetiştiriciliği yapılmaktadır (Zeder vd., 2000) (Şekil 3). Keçiler önemli bir et, süt, lif ve post kaynağı olmalarının yanı sıra, aynı zamanda insan uygarlığında çok eski zamanlardan beri tarımsal, ekonomik, kültürel ve hatta dini rolleri yerine getirmişlerdir (MacHugh vd., 2001). Fosil kanıtlar, keçinin yaklaşık olarak 10.000 yıl önce “Bereketli Hilal” olarak adlandırılan bölgede iki yabani *Capra* türünden (*Capra aegagrus* ve *Capra falconeri*) evcilleştirilmiş olabileceğini ve daha sonra insan göçü ve ticareti ile hızla tüm dünyaya yayılmış olabileceğini göstermektedir (Luikart vd., 2001).



Şekil 2. Bereketli Hilâl bölgesinde evcilleştirilen türlerin tarihi (Zeder, 2008)

Evcil keçi ırkları 10.000 yıllık evcilleştirme geçmişinde, fiziksel görünümünden davranışlara kadar yaban keçilerine kıyasla radikal olarak farklı fenotipik özellikler gösterecek şekilde farklılaşmıştır (Zeder, 2006). Genel olarak bunlar arasında; (a) birkaç boynuzsuz ırkla birlikte daha küçük boynuzlar ve vücut boyutunda bir azalma, (b) artan uysallık, (c) ırklar arasında bir post rengi varyasyonu, (d) insan göçü ve kolonizasyonundan kaynaklanan coğrafi olarak yaygın bir dağılıma sahip olma

sayılabilmektedir (Silanikove, 2000; Zeder, 2006; Cieslak vd., 2011). Bu kazanılmış özelliklerin genetik temelleri günümüzde hala tam anlamıyla açıklığa kavuşturulamamıştır. Keçiler evcil hayvan olarak değerlendirilmesinin yanında, biyomedikal araştırmalar için uygun hayvan modeli olarak, karmaşık özelliklerin genetik temelini araştırılmasında ve peptit ilaçlarının transgenik üretiminde de kullanılabilmektedir (Ebert vd., 1991; Ho Ko vd., 2000).



Şekil 3. 2015 yılında kilometrekareye düşen keçi sayısı (FAO)

Ülkemiz için keçiler ekonomik olarak önemli bir çiftlik hayvanıdır. Günümüzde yetiştiriciliği yapılan yerli keçi ırk ve tipleri Ankara (Tiftik), Norduz, Kilis, Abaza, Gürcü, Malta, Honamlı, Kıl ve Türk Saanen keçisidir (Akçapınar, 1994).

Ankara (Tiftik) keçisi, 13. yüzyılda Orta Asya'dan Türkler tarafından Anadolu'ya getirilmiş ve Ankara civarında günümüzdeki özelliklerini kazanmıştır (Batu, 1951). Ankara keçisi, Türkiye'nin önemli gen kaynaklarından biridir ve yetiştirildiği yerlerde karasal iklim etkilidir. Ankara keçisi yetiştiriciliği, ülkemizde başta Ankara ve İç Anadolu Bölgesi olmak üzere Karadeniz, Doğu ve Güneydoğu Anadolu bölgelerinde de yapılmaktadır. “Angora” keçisi olarak da bilinen Ankara keçisi diğer keçilerden farklı olarak kaşmir keçisi ile benzer özelliklere sahip, tiftik denilen kendine has life sahiptir ve yaklaşık tiftik verimi 3 kg kadardır. Yetiştiriciliğinde özel bakım gerektirmeyen ve küçük cüsseli olan Ankara keçisinin sürü idaresi kolay olup, zorlu çevresel koşullara oldukça dayanıklıdır (TAGEM, 2009).

Ankara keçisinin en önemli verimi tiftiktir. Et, süt, deri ve gübre verimleri ise ikincil verimlerdir. Siyah, kırçıl, bakır sarısı, kahverengi, kurşuni ve bej renklerde de olabilen genellikle beyaz renkte olan tiftik, Ankara keçilerinin ayaklarına kadar tüm vücudunu kaplamaktadır. Tiftik dayanıklı, parlak ve esnek bir yapıya sahiptir. Nemi iyi çekebilen ve ısı yalıtımı sağlayan, ısıya dayanıklı ve kolay boyanan değerli bir tekstil ham maddesidir. Tiftik verimi, ilk kırkımdan sonra artmaya başlar ve genellikle 4-6 yaşına kadar devam eder. Keçinin yaşı ilerledikçe tiftik kalitesi düşmektedir (Erol, 2015).

Bovidae familyası *Capra* cinsi içerisinde yer alan evcil keçinin atası olan Anadolu yaban keçisi (*Capra aegagrus*), Dünya’da Ermenistan, Azerbaycan, Gürcistan, İran, Irak, Pakistan, Rusya, Türkmenistan ve Türkiye’de yayılış göstermektedir (Şekil 4). Ülkemizde Akdeniz, Doğu, Güneydoğu ve Kuzeydoğu kısımlarında yayılış gösteren bu tür, av turizmi, ekolojik turizm ve genetik çeşitlilik bakımından oldukça önemlidir (Turan 1987, Gündoğdu 2006, Paşalı 2014). Günümüzde nesli tükenmekte olan türler arasında “Tehdit Altına Girebilir (NT)” kategorisinde (IUCN, www.iucnredlist.org) yer alan yaban keçisi, geçmiş zamanda avcılık yapan insanlar için önemli bir besin kaynağı olmuştur. Yaşam süreleri doğada 15-20 yıl olup, genellikle sürü halinde bulunurlar (Paşalı, 2014).



Şekil 4. Yaban keçisinin dünyadaki yayılış alanları (ICUN Red List)

1.2. Sekanslama sistemleri

İnsan genom projesi ile başlayan DNA dizileme serüveninde Yeni Nesil Dizileme (NGS) sistemleri, 2000’li yılların başlarında hayatımıza girip verimlerinin yıllar boyunca arttırılması ile model organizmalarının genomlarının dizilenmesinin yanı sıra model olmayan organizmaların da genomlarının dizilenmesine olanak tanımıştır. Sanger tipi

sekanslamaya göre daha hızlı/kısa sürede daha yüksek verimle, daha düşük maliyet ve az iş gücü ile çalışan bu sekanslama platformları ile genom çalışmaları büyük bir ivme kazanmıştır. Teknolojik gelişmeye paralel olarak özellikle son 10 yılda yüzlerce canlının tüm genomu dizilenmiş ve devam eden yüzlerce genom çalışması bulunmaktadır (<https://www.ncbi.nlm.nih.gov/genome/>). Gerek kısa okuma sistemleri gerekse uzun okuma sistemleri veya bu ikisinin beraber kullanılması ile ortaya çıkan genomik verilerin işlenmesi ile hastalıkların genetik temelleri çözülmeye başlanmış, ıslah çalışmalarında kullanılmasıyla verim ve kalitede artışlara sebep olmuş, QTL ve SNP keşfi hızlanmış, karşılaştırmalı genomik ve genomik seleksiyon çalışmalarında genom çapında çalışmaların önünü açmıştır. Günümüzde birbirinden farklı metodolojiye dayalı sekanslama platformları geliştirilmiş olup, en çok kullanılan sistemler kısa okuma için Illumina ve BGI/MGI platformları, uzun okuma için Pasific Biosciences (PacBio) ve Oxford Nanopore Technologies (ONT) platformlarıdır (Amarasinghe vd., 2020).

Illumina gibi 2. nesil dizileme sistemleri, yüksek doğrulukta ve oldukça yüksek kapsamda veri sağlamakta ancak kısa okumalar üretmektedir (Metzker, 2010). Bu platformlarda hata oranı %2'den daha azdır. Bu platformlardan farklı olarak Oxford Nanopore (Jain vd., 2016) sekanslama platformu, binlerce baz uzunluğunda olan daha uzun okumalar üretebilmektedir. Ancak çalışma prensibi gereği bu sistemlerin hata oranı nispeten daha yüksektir ve bu oran %30'a kadar ulaşabilmektedir. Bununla birlikte platformlardaki gelişmelerle her geçen gün hata oranı düşmektedir. Bunun sonucu olarak günümüzde birçok çalışmada sahip olduğu özellikler sebebiyle Oxford Nanopore platformu kullanılmaktadır.

1.3. Genom birleştirme

Sekanslama sistemlerindeki gelişmelerle birlikte DNA, küçük parçalara bölünerek her biri paralel olarak ayrı ayrı dizilenebilmekte (Massively parallel sequencing) ancak bu parçaların sırası korunamamaktadır. Büyük miktarda veri üretebilen (High Throughput Sequencing) bu sistemlerin bu anlamdaki yetersizliğinden dolayı elde edilen ham verilerin analizinde genom birleştirme (Genome assembly) olarak adlandırılan biyoinformatik işlemler kullanılmaktadır. Organizmaların tüm genomları temelde bu yaklaşımla elde edilmektedir. Kısa DNA parçalarının birbirleriyle karşılaştırılarak birleştirilmesi sonucu kontiglerin elde edilmesi ve sonrasında kromozomların veya

kromozom altı seviyedeki skafoldların elde edilmesi üzerine dayanmaktadır (Bayat vd., 2020). Bir türe ait ilk kez veya mevcut referans genom kullanılmadan genom birleştirme işlemi “*de novo* birleştirme” olarak adlandırılmaktadır. Birleştirilen bu genom, aynı veya yakın türlerdeki daha sonraki yapılacak olan genom birleştirme çalışmaları için referans oluşturması açısından oldukça önemlidir. Bu *de novo* genom birleştirme işlemi kompleks ve zorlayıcı bir süreç olup, referans seviyesinde genom elde edilmesi çalışmalarının ilk basamağını oluşturmaktadır (Salzberg vd., 2012; Bradnam vd., 2013).

Referans genomik verisi mevcut olan türlerin tüm genom eldesi çalışmalarında iki genom birleştirme stratejisi bulunmaktadır. Bunlardan ilkinde, okumalar referans genomu karşı haritalanır ve daha sonra alternatif bir konsensüs dizi elde edilir (Vezzi vd., 2011). İkinci yaklaşımda ise, okumalar ilk olarak *de novo* olarak birleştirilir. Daha sonra elde edilen kontigler/skafoldlar kromozomları sıralamak ve yönlendirmek, genom anotasyonları için gen bilgisi almak ve olası yanlış birleştirilmiş kontigleri veya skafoldları tanımlamak için referans genomu karşı hizalanır (Vezzi vd., 2011; Bao vd., 2014). Bazen iki yaklaşımın bir kombinasyonu da uygulanmaktadır (Wang vd., 2014). Bununla birlikte referanslı birleştirme stratejilerinin bazı dezavantajları da vardır. Elde edilen birleştirmeler kullanılan referansa yönelik bazı önyargılar içerebilmektedir. Referans genom ile birleştirilen genom arasında daha fazla farklılaşan bölgeler varsa bunlar yeniden oluşturulamayabilir, eksik olabilir ve bu nedenle hedef birleştirmede çeşitliliğin azalmasına yol açabilmektedir (Schneeberger vd., 2011; Vezzi vd., 2011; Gordon vd., 2016). Ayrıca, referans dizisindeki hatalar ve türler arasındaki kromozomal yeniden düzenlemeler de hatalara yol açabilmektedir (Ekblom vd., 2014).

Genom birleştirme yapılırken verilerin türü ve miktarı süreci etkileyen en önemli faktörlerdendir. Birleştirilen genomun doğruluğunu değerlendirmek için çeşitli yöntemler kullanılmaktadır. Bunlardan en önemlileri birleştirilen genomun kapsamı, bitişikliği/bütünlüğü ve hata oranıdır (Gurevich vd., 2013). Tüm yönleriyle eksiksiz bir genom birleştirme yöntemi bulunmamakta, ancak elde edilen verilere uygun olarak seçilen yöntemler kullanılarak bunlar çözülmeye çalışılmaktadır (Bradnam vd., 2013). Elde edilen birleştirmenin değerlendirilmesinde kullanılan temel standart “N50” değeridir. N50 değeri genom bitişikliğinin belirlenmesinde bir ölçüt olup, toplam birleştirme uzunluğunun yarısındaki en uzun diziler içindeki en kısa kontiğin dizi uzunluğu olarak tanımlanır (Yandell ve Ence, 2012). Genom çalışmalarında kullanılan

bir diğ er  nemli deęer olan kalite skoru (Q skoru, Phred) i in Q10 deęeri %90 doęruluęu, Q20 deęeri %99 doęruluęu ve Q30 ise %99,9 doęruluęu ifade etmektedir (Ewing ve Green, 1998).

Genom birleřtirmede sadece kısa okuma  reten platformlardan elde edilen okumalar kullanılıyorsa birleřtirme doęruluęunu arttırmak i in y ksek genom kapsamı gereklidir. Kısa ve uzun okumaların birlikte kullanıldığı hibrit y ntemlerde kapsama derinlięi, birleřtirme hattı se imi  zerinde daha g  l  bir etkiye sahip olmaktadır. Y ksek kapsamlı uzun okumalar saęlanırsa, birleřtirme yapmak (Koren vd., 2017) ve genom ‘polishing’ (uzun okumalardaki hataların, kısa okumaların bu okumalara haritalanması yoluyla d zeltilmesi) i in kısa okumaları kullanmak da m mk nd r (Walker vd., 2014). Uzun okuma kapsamının d ř k olduęu durumlarda, uzun okuma birleřtirmesinden  nce uzun okumaların kısa okumalar ile d zeltilmesi gerekmektedir (Haghshenas vd., 2016). Uzun okuma verilerinin genom birleřtirme i in yeterli olmadıęı durumlarda ise kısa okumaların birleřtirmesinin bitiřiklięini geliřtirmek i in kullanımı m mk nd r (Bankevich vd., 2012).

Genom birleřtirmede bir dizi kalite kontrol ařamasından sonraki ilk basamak, genomun b y k par aları olan ve kontig adı verilen dizileri, genomdaki sıraları hakkında herhangi bir bilgi olmaksızın elde etmektir. Burada ideal olan her bir kromozom i in kromozomun tamamını kapsayan bir kontig elde etmek olacaktır. Ancak bu durum,  zellikle b y k genomlarla  alıřırken mevcut birleřtirme algoritmaları i in zor bir durumdur. Genom birleřtirmede sonraki adımlar kontigleri sıralayıp, kromozom  zerindeki bořlukları doldurarak ve hataları d zelterek birleřtirmeyi iyileřtirmek olacaktır. Skafold, en az iki veya daha fazla kontięin genomdaki sırasının ve aralarındaki mesafenin tahmin edilerek birleřtirilmesiyle elde edilen kromozom altı dizilere verilen isimdir. Bořluk doldurma, skafoldların kromozom  zerindeki sırasının belirlenerek birleřtirilmesi iřlemidir (Bayat vd., 2020).

B y k geliřmelere raęmen Illumina gibi 2. nesil dizileme sistemleri,  ok sayıda kopya i eren ve GC a ısından zengin b lgelerde bulunan uzun tekrarları a ık bir řekilde   zmede sınırlı olup y ksek oranda par alanmıř genom birleřtirmeleri  retmektedir (Pop ve Salzberg, 2008). Teorik modellemelerle, okuma uzunluklarının 1.000 b ’den 100 b ’ye d řmesiyle genom bitiřiklięinde altı kat veya daha fazla azalmaya neden olabileceęi

gösterilmiştir. Bu kapsamda 150 bç uzunluğunda olan Illumina okumaları genom birleştirmesini zorlaştırmaktadır (Kingsford vd., 2010).

2. nesil dizileme sistemlerindeki sınırlamalarının üstesinden gelmek ve tüm genomlarının birleştirilmesini kolaylaştırmak için daha uzun okumalar üreten dizileme teknolojilerine olan ihtiyaç, üçüncü nesil dizileme teknolojilerinin ortaya çıkmasını sağlamıştır. Okuma uzunluklarının teorik bir üst sınırı olmayan ve >2 milyon nükleotide kadar okuma yapabilen Oxford Nanopore platformları, yakın zamanda geliştirilmiştir (Lu vd., 2016; Payne vd., 2019). Oxford Nanopore dizileme, tüm genomların dizilenmesinde tekrar eden bölgeleri kapsayan uzun okumalar üretebilmekte, böylece daha az parçalanmış ve hatta tam genomların elde edilmesine imkan tanımaktadır.

Genom birleştirmede yüksek kontig sayısının azaltılması ve nispeten düşük olan N50 değerinin artırılması, hem kısa hem de uzun okumaların eşzamanlı kullanımıyla elde edilebilmektedir. Son yıllarda böyle bir hibrit yaklaşımın genom birleştirmelerinin kalitesini iyileştirmede ve gen ailesi genişlemelerinin keşfinde yararlı olduğu gösterilmiştir. En iyi birleştirilmiş genomlardan biri olan *de novo* insan genom dizisi için hibrit bir yaklaşım uygulanmıştır (Mostovoy vd., 2016). Palyaço balığı genomunun yüksek kapsamlı Illumina kısa okumaları ve düşük kapsamlı Nanopore uzun okumaları kullanılarak birleştirilmesinde, yalnızca Illumina kısa verilerinin kullanıldığı birleştirmeye karşılaştırıldığında genom istatistiklerinde önemli iyileşmenin olduğu rapor edilmiştir (Tan vd., 2018).

1.4. Genom anotasyonu

Yeni nesil dizileme sistemleri ile çeşitli organizmalardan kısa süre içerisinde oldukça büyük miktarlarda veriler üretilebilir hale gelmiştir. Bu sistemler ile üretilen verilerin işlenerek hem bireysel hem de popülasyon temelinde çalışılması, elde edilen genom birleştirmenin kalitesine bağlıdır. Başlarda protein kodlayan genlerin belirlenmesini ifade eden genom anotasyonu, günümüzde binlerce bireysel genom üzerindeki tek nükleotidlerin dahi belirlenerek fonksiyonel olarak açıklanmasını içerecek şekilde gelişmiştir.

Organizmadaki mevcut genlerin bulunması ve tanımlanması, genom anotasyonu işleminin büyük bir bölümünü oluşturmaktadır. Burada kapsamlı ve doğru bir yaklaşım

kullanmak oldukça önemli olup, birden fazla bağımsız ve tamamlayıcı algoritma ve yöntemin uygulanması gerekmektedir. Genomların annotate edilmesinde birçok algoritma ve yöntem geliştirilmiş olmasına rağmen birtakım zorluklar hala mevcuttur (Brent, 2005; Li vd., 2019; Mishra vd., 2020).

Genom anotasyonu, yapısal anotasyon ve fonksiyonel anotasyon olarak ikiye ayrılmaktadır. Yapısal anotasyon, DNA'nın özelliklerinin (ekzonlar, intronlar, promotörler, transpozonlar vb.) bulunması iken fonksiyonel anotasyon, biyolojik bilginin yapısal anotasyonla belirlenen gen veya protein dizileriyle ilişkilendirilmesi işlemleridir (Ejigu ve Jung, 2020).

Yapısal anotasyonun ilk basamağı genomdaki tekrar bölgelerinin maskelenmesidir. Bu tekrar bölgeleri prokaryot organizmalarda genomun %0 ila %42'sini oluştururken, ökaryot organizmalarda bu daha fazla olmakla birlikte insan genomun üçte ikisini oluşturmaktadır (Treangen vd., 2009; de Koning vd., 2011). Tekrar dizileri birbirine bitişik ve tipik olarak sentromerde bulunurlar (Barra ve Fachinetti, 2018). Ayrıca bu tekrarlar uzun ve kısa serpiştirilmiş nükleer elementler (LINE'lar ve SINE'ler), DNA transpozonları vb. gibi farklı transpozon element formlarında da genomda bulunabilmektedirler (Bourque vd., 2018). Tekrar maskeleyme araçları, *de novo* tahmin etmeye veya önceden tanımlanmış tekrarların listelerini içeren veri tabanlarına dayanmaktadır (Ejigu ve Jung, 2020).

Protein kodlayan genlerin ve diğer düzenleyici elementlerin belirlenmesi, genom anotasyonunun en temel aşamasıdır. Gen tahmini, özellikle ökaryotik genomlar için karmaşık bir süreçtir (Yandell ve Ence, 2012). Ekzonlar ve alternatif ekleme varyantları arasındaki değişen boyutlardaki intronlar gen yapısı tahminini zorlaştırmaktadır. Birçok gen tahmin yöntemi mevcut olup, bunlar *ab initio* yöntemler, homolojiye dayalı yöntemler ve birleşik yöntemler olarak sınıflandırılabilir (Del Angel vd., 2018).

Ab initio yaklaşımları nükleotid dizisine dayalı gen tahmin metodu olup; promotörleri, kodlayan veya kodlamayan bölgeleri ve genom dizisindeki intron-ekzon bağlantılarını tanımlamak için gizli Markov modeli (HMM) gibi istatistiksel modellere dayanır. Homoloji temelli yaklaşımda EST, cDNA ve/veya protein dizileri benzerlikleri kullanılarak gen tahmini yapılır. Birleşik yöntemlerde ise gen tahmini için *ab initio* ve

homoloji temelli yaklaşımları birleştiren programlardan oluşur (Yu vd., 2006; Ejigu ve Jung, 2020). Günümüzde homoloji temelli yöntemler, mevcut çok miktarda EST ve cDNA dizisi nedeniyle genom anotasyonunda merkezi bir rol oynamaktadır (Sagot vd., 2002).

Fonksiyonel anotasyonda ise BLAST gibi araçlar kullanılarak benzer bir dizi tanımlandıktan sonra işlevsel bir anotasyonun bir genle ilişkilendirilmesi yapılmaktadır. Gen Ontolojisi (GO), gen fonksiyonu için en kapsamlı ve yaygın olarak kullanılan bilgi tabanıdır (Botstein vd., 2000). GO, moleküler fonksiyon (bir gen ürününün moleküler seviyedeki aktivitesi), hücresel bileşen (gen ürününün konumu) ve biyolojik süreç (bir gen fonksiyonunun kullanıldığı biyolojik bir program) bileşenlerini içermektedir (Ejigu ve Jung, 2020). Ayrıca en yaygın kullanılan diğer araçlardan Kyoto Encyclopedia of Gene and Genomes (KEGG) yolak bilgisi (Kanehisa ve Goto, 2000), Reactome Pathway Knowledgebase (Fabregat vd., 2018) insan odaklı moleküler süreçler, Rhea enzim ve metabolik yolaklar (Morgat vd., 2017), ChEBI kimyasal ontoloji (Hastings vd., 2016) ve NCBI's Conserved Domain Database (CDD) korunmuş protein domainleri (Lu vd., 2020) için kullanılmaktadır. Buna ek olarak Database of Genomic Variants (DGV) (MacDonald vd., 2014), dbVar (Lappalainen vd., 2012), Human Gene Mutation Database (HGMD) (Stenson vd., 2012), HGVbase (Fredman vd., 2002) gibi varyasyon bilgisi içeren sistemler de kullanılmaktadır.

1.5. Mitokondriyal DNA

Mitokondriyal genom (mitogenom), ökaryotlarda nüklear genomdan bağımsız, evrimsel analiz, genetik çeşitlilik, filogenetik, popülasyon yapısı, taksonomi vb. gibi çalışmalar için oldukça kullanışlı bilgiler sağlayan dairesel çift zincirli yapıda bir DNA molekülüdür (Delsuc vd., 2003; Hassain vd., 2013; Carpi vd., 2016; Olivieri vd., 2017; Ramos vd., 2018). Buna ek olarak karşılaştırmalı ve evrimsel genomik açısından da önemli bilgiler sağlar (Carpi vd., 2016; Olivieri vd., 2017; Ramos vd., 2018). Mitogenom, memelilerde yaklaşık olarak 15-18 kilobaz uzunluğundadır (Bandelt vd., 2006; Jenuth vd., 1996; Galtier vd., 2009) ve 22 tRNA, 2 rRNA ve 13 mRNA olmak üzere toplam 37 gen bölgesi ile 1 tane kodlama yapmayan kontrol bölgesi (D – loop) taşımaktadır (Ruokonen, 2001; Iborra vd., 2004; Galtier vd., 2009).

1.6. Genom çalışmalarının önemi

Dünyadaki yaşam, başlangıçtaki basitlikten bugün şaşırtıcı derecede bir çeşitliliğe evrilmiştir. Dünya Biyogenom Projesi, ökaryotlar olarak adlandırılan tüm mevcut organizmaların (yaklaşık 2 milyon tür), filogenetik, ekolojik ve yüksek etkili önceliğe sahip olanlardan başlayarak, Dünya üzerinde dijital bir yaşam kitaplığı oluşturmak için yüksek kalitede genom sekanslanması gerektiğini önermiştir. Evrimsel ve ekolojik öneme sahip pek çok sorunun, hayat ağacındaki tüm dallardaki veya doğal ekosistemlerdeki tüm türlerdeki farklılıkları temsil eden tüm genom verileri mevcut olduğunda ele alınabileceğini öne sürülmektedir. Genomik bir yaşam ağacının tüm ekosistemlerde devam eden türleşme, adaptasyon ve organizma bağımlılıkları süreçlerinin anlaşılmasını destekleyeceği düşünülmektedir. Bu keşiflerin filogenetik, evrim, ekoloji, koruma, tarım, biyo-endüstri ve tıp alanlarında uzun süredir devam eden sorunları çözeceği önerilmiştir (Blaxter vd., 2022).

Yaklaşık 2 milyon türün sekanslanması gibi muazzam bir görev, belirli ilgi konularını ele almak için odaklanmış projelerle izlenebilir hale getirilmiştir. Örneğin Omurgalı Genomları Projesi 70.000 omurgalı türünün tümünü sekanslamayı amaçlamaktadır ve her bir taksonomik düzen için referans genomlarla başlamıştır (Rhieet vd., 2021). Darwin Hayat Ağacı Projesi, İngiltere ve İrlanda'daki 70.000 türün tamamını sekanslamayı hedeflemekte olup temsil edilen 4.200 taksonomik ailenin her biri için bir referans genomla başlamıştır (The Darwin Tree of Life Project Consortium). Diğer projeler ise koruma için ikonik, sosyal veya bilimsel açıdan önemli veya benzersiz adaptasyon setleri nedeniyle ilginç olan türlere ve ekosistemlere odaklanmaktadır (Blaxter vd., 2022).

1.7. Çiftlik hayvanlarında yapılan genom çalışmaları

En eski evcil hayvanlardan biri olan domuz (*Sus scrofa*), ~10.000 yıl önce Asya ve Avrupa'da birbirinden bağımsız olarak evcilleştirilmiştir (Larson vd., 2005; Groenen vd., 2012). Domuz önemli bir ticari çiftlik hayvanıdır ve insanlarla fizyolojik ve fenotipik benzerlikleri nedeniyle insan hastalıklarını anlamak için kullanılan uygun bir tıbbi model organizmadır (Sullivan vd., 2001; Meurens vd., 2012). Yang vd. (2019a), Çin yerli ırkı olan Luchuan domuzunun erkek bir bireyinin ultra yüksek kaliteli fazlı kromozom seviyesinde genom birleştirmesini 140X PacBio, ~86X Illumina çift uçlu (paired-end;

PE), ~41X Hi-C ve ~351X BioNano optik haritalama verilerini kombine ederek gerçekleştirmişlerdir. Çalışmada yapılan birincil birleştirme sonucunda 18,03 Mb kontig N50 değeri ve 140,09 Mb skafold N50 değeri ile ~2,58 Gb boyutunda Luchuan domuzunun tüm genomunu elde etmişlerdir. Genom anotasyonunda *ab initio* ve homoloji yaklaşımları ile transkriptom verileri kullanılarak genomda 22.710 protein kodlayan gen (PCG) tahmin edilmiştir. PCG'lere ek olarak 2.835 küçük ncRNA ve 3.066 yeni lncRNA ile 1.019 yeni circRNA gen bölgelerini Luchuan domuzunun genomunda rapor etmişlerdir. Birincil birleştirme ve alternatif haplotip arasındaki karşılaştırma ile alele özgü ekspresyon, epigenetik regülasyon, genom yapısı ve domuzların evrimini incelemek için zengin bir kaynak sağlayan haplotipe özgü çok sayıda allel ortaya çıkarılmıştır. Çalışmada Luchuan ve Duroc domuzları arasındaki ayrılma süresinin yaklaşık olarak 1,7 milyon yıl öncesine dayandığına dair açık moleküler evrimsel kanıtlar sağlanmıştır (Yang vd., 2019a).

Büyüme özellikleri bakımından üstün performans gösteren Duroc domuzu, ticari domuz popülasyonlarının terminal erkek ebeveyni olarak yaygın şekilde kullanılmaktadır (Ding vd., 2018). Du vd. (2021), Duroc domuzunun genomunu birleştirmek için kısa ve uzun okumaların kullanıldığı farklı 9 yazılımla 14 birleştirme kombinasyonunun karşılaştırmalı bir performans değerlendirmesini sunmuşlardır. Çalışmada 100 bç uzunluğunda 230 Gb Illumina çift uçlu okumaları ve üç farklı insört boyutuyla 170 Gb Illumina eş-çift (mate-pair; MP) okumaları ile ortalama uzunluğu 8.510 bç olan toplam 56 Gb PacBio okuması da üretilmiştir. K-mer analizinde Duroc domuzu genom büyüklüğünün 2,5 Gb olduğu tahmin edilmiştir. Entegre bir hibrit *de novo* birleştirme hattı (pipeline) olan HSCG'yi (Hybrid, Scaffold, Chromosomer, Gap-fill) tasarlamışlar ve Duroc domuzu için bir taslak genom elde etmişlerdir. Elde ettikleri yeni genom ile *Sus scrofa* 11.1 arasındaki karşılaştırma sonucunda *S. scrofa* 11.1 genlerinde önemli kırılma noktalarının olduğu gözlenmiştir (Du vd., 2021).

Low vd. (2019), tek molekül dizileme ve kromatin konformasyon yakalama verilerini kullanarak manda (*Bubalus bubalis*) genomunun kromozom düzeyinde bir birleştirmesini gerçekleştirmişlerdir. Ortalama uzunluğu 11,5 kb olan PacBio Sequel okumaları, 2,66 Gb boyutundaki genomun tekrarlayan öğelerin çözülmesine ve sekans bitişikliğinin oluşturulmasına yardımcı olmuş ve sentromerler yayılmış olarak doğru şekilde skafoldlanmıştır. Çalışmada ~75X PacBio Sequel uzun okuma, ~24X Chicago okuma,

~58X Hi-C okuma ve ~82X Illumina çift uçlu okuma verileri kullanılmıştır. Gen anotasyonunun geliştirilmesi için 50 farklı dokudan yaklaşık 15 milyar RNA-seq okuma verisi de kullanılmış olup 20.801 protein kodlayan gen, 8.443 kodlanmayan gen ve 4.465 pseudogen belirlenmiştir. Bu genomun dizi bitişikliğinde insan ve keçi referanslarından daha başarılı olup, majör histo-uyumluluk kompleksi (MHC) gibi bir araya getirilmesi zor gen kümelerinin anotasyonunu kolaylaştırdığını rapor etmişlerdir (Low vd., 2019).

Günümüzde aşırı avlanma ve habitat kaybı nedeniyle tehlike altında olarak sınıflandırılan *Bubalus depressicornis* (Smith 1827), Sulawesi ve Buton Adaları'na özgü yabani bir cüce mandadır ve *Bubalus quarlesi* (Ouwens 1910) ile simpatrik yayılışa sahiptir. *Bubalus* cinsinin (Asya mandaları) yabani türleri için genomik veriler hala eksik olmakla birlikte şu anda veri tabanlarında evcil mandalar için birkaç tüm genom verisi mevcuttur. Bu eksikliği gidermek için Porrelli vd. (2022), *Bubalus depressicornis* genomunun taslak birleştirmesini oluşturmuş ve diğer bufalo genomlarıyla karşılaştırmışlardır. NextSeq 500 Illumina sistemi ile 2x151 bp uzunluğunda üretilen 991.437.058 çift uçlu okumalar, MaSuRCA v.3.3.1 algoritması (Zimin vd., 2013; 2017) kullanılarak birleştirilmiştir. Birleştirilen genomun büyüklüğü 2,56 Gb olup, 103.135 kontigden oluşmaktadır ve en uzun kontig 337.39 kb uzunluğunda elde edilmiştir. N50 ve L50 değerleri sırasıyla 38,73 ve 19,83 kb, ortalama genom kapsamı 44X ve GC içeriği %41,74'tür. Genom bütünlüğünü değerlendirmek için yapılan homoloji temelli yaklaşımda %94,51 tam ve %3,65 kısmi genomik özellik; *de novo* yaklaşımda referansın annotate edilmiş genlerinin %97,14'ünü temsil eden 32.393 gen; memeli ortolog veri tabanına karşı BUSCO ile de %71,1 tam, %11,7 parçalı ve %17,2 kayıp ortolog bölgeler bulunmuştur. Ayrıca tekrar analizi ile genomun %42,12'sinin tekrar bölgesi olduğu tespit edilmiştir (Porrelli vd., 2022).

Sığır uzun süre süt, et, deri ve güç gibi ürünler sağlayarak insan toplumlarında önemli roller oynamaktadır. Şu anda evcilleştirilmiş iki büyük sığır olan taurin (*Bos primigenius taurus*) ve indicine veya zebu (*Bos primigenius indicus*), 250.000 ila 330.000 yıl önce birbirinden farklı iki soya ayrılan soyu tükenmiş vahşi yaban öküzünden (*Bos primigenius*) gelmektedir (Pitt vd., 2019). Taurin sığırları, *Bos primigenius primigenius*'un yaklaşık 10.000 yıl önce Bereketli Hilal'de evcilleştirilmesinden kaynaklanırken, *Bos primigenius nomadicus*'un soyundan gelen indicine sığırlarının yaklaşık 8.000 yıl önce Indus Vadisi'nde evcilleştirildiği düşünülmektedir (MacHugh vd.,

1997; Murray vd., 2010; Porto-Neto vd., 2013; Park vd., 2015; Pitt vd., 2019). Doğal seçim ve ardından insanlar tarafından yönlendirilen evcilleştirme sonrası seçim, sığır genotiplerini modifiye ederek taurin ve indicine sığırları arasında farklı genetik ve fenotip profillerine neden olmuştur. Sığırlarda dizi seviyesindeki genetik varyantlar hakkındaki bilgilerin çoğu *Bos primigenius taurus* popülasyonları için mevcuttur. Fernandes Júnior vd. (2020), 52 Nellore (*Bos primigenius indicus*) boğasının tam bir genomik karakterizasyonu gerçekleştirmiş ve tropikal adaptasyon ile üretken özellikler üzerinde varsayılan etkiye sahip spesifik zebu DNA varyantlarını ortaya çıkarmışlardır. Çalışmada birey başına ortalama 18,50X genom kapsamında veri üretilmiştir. Sığır genomunun uzunluğu 2715,85 Mb (<https://www.ncbi.nlm.nih.gov/genome/82>) ile ortalama olarak kb başına yaklaşık 15 polimorfizm tanımlanmıştır. SNP ve indel mutasyonları, güncel sığır referans genomu (ARS_UCD1.2) kullanılarak tanımlanmış ve Ensembl VEP yazılımı kullanılarak varyantların anotasyonları yapılmıştır. Sonuç olarak toplam 35.753.707 SNP ve 4.492.636 indel tespit edilerek fonksiyonel etkileri annotate edilmiştir. Ayrıca hem bir SNP hem de bir indel içeren, proteinler üzerinde yüksek fonksiyonel etkiye sahip 400 gen belirlenmiştir. Bunlar arasında sığırların enfeksiyonlara karşı bağışıklık tepkisi ve üreme yönleriyle ilişkili *BoLA*; memelilerde ısı düzenleyici koruyucu mekanizmalarla ilgili *HSPA8*, *DNAJC27* ve *DNAJC28* ve memeli türlerinin evriminde önemli genetik faktörler olan birçok koku sinyal yolu ile ilgili genler vurgulanmıştır (Fernandes Júnior vd., 2020).

Evcil koyun (*Ovis aries*) eti, yünü ve sütü için dünya çapında yetiştirilen önemli bir tarımsal türdür. Bu tür için yüksek kaliteli bir referans genom, biyolojik özellikleri etkileyen genetik mekanizmaları keşfetme yeteneğini geliştirmektedir. Ayrıca yüksek kaliteli bir referans genom, gen düzenleyici elemanların kesin fonksiyonel anotasyonuna imkan tanımaktadır. Davenport vd. (2022), yaptıkları bir çalışmada koyun için veri tabanlarından alınan Illumina kısa okuma (55X kapsam), Pacific Biosciences uzun okuma (75X kapsam) ve Hi-C verilerine ek olarak Oxford Nanopore verilerinin ek 50X kapsamıyla kombine edilerek canu v1.9. algoritması ile genom birleştirme yapmışlardır. Birleştirilen kontigler, Salsa v2.2 algoritması ile Hi-C verileri kullanılarak skafoldlanmış, boşluklar PBSuite v15.8.24 algoritması ile doldurulmuş ve Nanopolish v0.12.5 algoritması ile polishing yapılmıştır. PurgeDups v1.0.1 algoritması ile duplike kontigler kaldırılmış ve ardından kromozomlar oryante edilmiştir. Varyantları belirlemek için freebayes v1.3.1 algoritması kullanılmış ve bunları doğrulamak için Merfin ve

BCFtools'tan oluşan bir pipeline ile 2 round polish edilmiştir. ARS-UI_Ramb_v2.0 birleştirmesi 2,63 Gb uzunluğunda olup, Oar_rambouillet_v1.0 ve Oar_v4.0 ile karşılaştırıldığında skafold sayısında 19 ve 38 kat azalma ile iyileştirilmiş bitişikliğe sahiptir (43,18 Mb'lik NG50 değeri ile). ARS-UI_Ramb_v2.0, baz başına daha yüksek doğruluğa ve önceki birleştirmeye göre haritalanmış RNA dizisinden tanımlanan daha az ekleme ve silmeye sahip olarak rapor edilmiştir (Davenport vd., 2022).

Evcil kaz sağladığı besleyici et, yumurta ve yağlı karaciğer ile ekonomik açıdan en önemli tarım hayvanlarından biridir. Fosil kanıtlar, Çin kazlarının 7.000 yıldan daha uzun bir süre önce evcilleştirildiğini göstermektedir (Eda vd., 2022). Üç evcil kazın genom birleştirmeleri yakın zamanda tamamlanmış ancak yalnızca hibrit Tianfu kazının (TFG) kromozom seviyesinde birleştirilmiştir (Li vd., 2020). Ouyang vd. (2022), Çin yerli kazının (Xingguo gri kazı, XGG; *Anser cygnoides*) kromozom düzeyinde bir genom birleştirmesini yapmak için PacBio, Illumina, 10x genomics, BioNano ve Hi-C teknolojilerini içeren hibrit bir *de novo* yaklaşımı kullanmış ve 686,12 Gb sekans verisi (~562,4X genom kapsamı ile) elde etmişlerdir. Ayrıca çalışmada Anatidae'nin evrimi sırasında kaz karaciğerindeki yağ depolama kapasitesinin adaptasyonu analiz edilmiştir. Anatidae'nin evrimi sırasında kazların biyolojik özelliklerini keşfetmek için karşılaştırmalı genom analizi ile birleştirmişlerdir. 994 kazın genomik yeniden sekanslaması, kazların ikili kökenini destekleyen (*Anser cygnoides* ve *Anser anser*) kazların genetik ilişkilerini araştırmak için kullanılmıştır. Çin yerli kazları, Avrupa kazlarından daha yüksek genetik çeşitlilik göstermiştir. Ayrıca çalışmada, Çin evcil kazlarının beyaz tüylerini belirleyen endotelin reseptörü B alt tipi 2'ye (*EDNRB2*) 14-bp'lik bir eklemenin doğal bir mutasyon olduğu bulunmuştur ve “genetik otostopun” bir sonucu olarak bağlantılı alellerin frekansı hızla arttığı ve güçlü yapay seçilim altında beyaz kazların tamamen farklı haplotiplerinin oluşumuna yol açtığı bulunmuştur (Ouyang vd., 2022).

1.8. Keçilerde yapılan genom çalışmaları

Dong vd. (2013), dişi bir Yunnan siyah keçisinin genomunu yaklaşık 2,66 Gb olarak yayınlamış ve bu çalışmada kısa okuma dizileme ve optik haritalama yöntemlerini kullanmışlardır. Genomun dizilenmesinde ~180 bç, 350 bç, 800 bç, 2 kb, 5 kb, 10 kb ve 20 kb insört uzunluklarda 14 çift uçlu kütüphane kurulmuş ve Illumina platformu ile

191,5 Gb yüksek kaliteli okuma (tahmini genom boyutunun 65,6X kapsamında) verisi elde edilmiştir. Genom birleştirmede SOAPdenovo version 1.03 (Li vd., 2010) yazılımı kullanılmış skafold N50 değeri 3,06 Mb, super-skafold N50 değeri 16,3 Mb olarak bulunmuştur. Çalışmada 10 dokudan izole edilen RNA verileriyle birlikte 22.175 protein kodlayan gen annotate edilmiştir. Kaşmir keçisinin birincil ve ikincil foliküllerinin karşılaştırmalı transkriptomik analizi, iki kıl kökü tipi arasında farklı şekilde ifade edilen 51 geni ortaya koymuştur (Dong vd., 2013).

Siyah Bengal keçisi genomu, Illumina HiSeq 2500 platformunda 150 bç uzunlukta okumalarla dizilenmiştir. Üretilen 40 Gb verinin (14X kapsam) *de novo* birleştirmesi ABySS v.2.1.5 assembler (Simpson vd. 2009) algoritması kullanılarak yapılmış, genom 3,04 Gb büyüklüğünde birleştirilmiş ve 26.458 gen tahmin edilmiştir. BUSCO analizi genom birleştirmenin %82,5 bütünlükte olduğunu ortaya koymuştur (Siddiki vd., 2019).

Bickhart vd. (2017), kontig formasyonu için uzun okuma, konsensus doğrulama için kısa okuma, optik ve kromatin etkileşim haritalamayla skafoldlama ile San Clemente keçi ırkı için *de novo* genom birleştirme oluşturmuşlardır. Bu yöntemle sadece 649 boşluk içeren son derece yüksek kromozom uzunluğuna sahip skafoldlar elde etmişler ve genomu skafold N50 değeri 87 Mb ile 2,92 Gb olarak birleştirmişlerdir (Bickhart vd., 2017).

Güncel biyolojik araştırmalarla araştırmacılar evcilleştirilen türler ve onların yabani atalarının karşılaştırılması yoluyla evcilleştirmeyle ilişkili fenotipik farklılıkların altında yatan genleri belirlemeye çalışmaktadırlar (Wiener ve Wilkinson, 2011; Meyer vd., 2013). Dong vd. (2015), keçi evcilleştirmesine ve soy oluşumuna eşlik eden genomik değişiklikleri araştırmak için yaban keçisi (*Capra aegagrus*) referans genomunu oluşturmuşlardır ve yaban keçisi genom dizilemesi Hiseq 2000 cihazı ile 250 bç-5 kb insört uzunluklarda çift uçlu okumalar ve 10kb ile 20kb'lik eş-çift okuma kütüphaneleri (yaklaşık 108,5X çift uç ve 27,7X eş-çift okuma kapsamı, toplamda 381,50 Gb veri) kullanılarak SOAPdenovo (Li vd., 2010) yazılımı ile yapılmıştır. Elde edilen genom CapAeg_1.0'ın kontig N50 boyutu 18.965 bç ve skafold N50 boyutu 2.057.686 bç olup, 23.217 gen annotate edilmiştir. Evcil keçilerde post rengiyle ilişkili olan on üç kopya sayısı varyasyon geni tanımlanmıştır. Bunların arasında *ASIP* gen duplikasyonunun evcil keçilerde açık post rengi fenotipinin oluşumuna katkıda bulunduğu önerilmiştir (Dong vd., 2015).

Ülkemizde yetiştiriciliği yapılan ve yerli ırklarımızdan birisi olan Norduz keçisi üzerine yapılan öncü bir araştırmada 20 keçiden ~120 Gb veri Illumina HiSeq X platformu ile 2 x 150 baz uzunluğunda elde edilmiştir. Birey başına 33,1X genom kapsamında üretilen okumalar DRAGEN pipeline ve GATK kullanılarak işlenmiştir. Okumalar keçi referans genomuna haritalanmış ~665 milyon okuma, referans keçi genomuyla hizalanmıştır. Tüm okumalar başarıyla referans genomuna hizalanmış ve hizalanmış okumaların boyutu ~96,4 Gb'dır. Sıkı filtreleme koşulunda, Norduz keçi genomunda 9.757.980 SNP, 1.536.715 InDel ve 290 CNV tespit edilmiştir. Çalışmada yüksek etkili SNP varyasyonlarının fonksiyonel analizi, klasik tamamlayıcı aktivasyon biyolojik sürecinin keçide önemli ölçüde etkilendiğini göstermiştir. Keçi genomundaki CNV'ler, virüslere karşı savunma, bağışıklık tepkisi ve hücre zarı taşıyıcıları ile ilgili genlerde bulunmuştur. INF ile uyarılan GTPazlar olan *GBP2*, *GBP5* ve memeli ortolog *GBP1*'in keçilerde yüksek kopya sayılarına sahip olduğu gösterilmiştir (Arslan, 2023).

1.9. Diğer memeli hayvanlarda yapılan genom çalışmaları

Köpekler, insanlar tarafından yaklaşık olarak 30.000 yıl önce ilk evcilleştirilen hayvanlardır (Botigue vd., 2017). Basenjiler, Afrika Kongo'daki kabilelerle birlikte yaşayıp avlanan, Orta Afrika kökenli eski bir köpek ırkıdır. Edwards vd. (2021), iki Basenji köpeğinin *de novo* genom birleştirmesini gerçekleştirmişler ve ilk olarak dişi bir bireyin genomu 84,5 Gb Oxford Nanopore PromethION sistemi ile yaklaşık 2,41 Gb genom boyutu için 35X kapsamda birleştirmişlerdir. Ek olarak yaklaşık 47,7X kapsamda 115,1 Gb BGISEQ kısa okumalar ile hata düzeltme yapmışlardır. Hi-C verileriyle kontigler daha uzun skafoldlara birleştirilmiştir. Erkek bir Basenji örneğinin genomu ise Pacific Bioscience Tek Molekül Gerçek Zamanlı dizileme ile yaklaşık 45X kapsamda dizilenmiştir. İki birleştirme için de homoloji temelli gen tahmin programı olan GeMoMa v1.6.2beta algoritması (Keilwagen vd., 2019) ile dokuz referans tür kullanılmış, sırasıyla 27.129 (68.251 transkript) ve 27.783 protein kodlayan gen (65.511 transkript) tahmin edilmiştir. Çalışmada elde edilen genomlar ile daha önceki çalışmalarda elde edilmiş üç Basenji genomu karşılaştırarak aralarındaki yapısal farklılıkları belirlemişlerdir (Edwards vd., 2021).

Karaçalı antilobu (*Hippotragus equinus*), Hippotraginae'nin (Bovidae) ikinci en büyük üyesidir ve Sahra altı mezik ormanlık alanlarda yaygın olarak bulunmaktadır. Gonçalves

vd. (2021), ıslah programından bir erkek karaçalı antilobun tüm genom dizisinin ilk kez *de novo* birleştirmesini ve anotasyonunu bildirmişlerdir. Buna ek olarak tanımlanan altı alt türden beşini temsil eden beş yabancı bireyin yeniden sekanslanması (re-sequencing) yoluyla tek nükleotit varyantlarını (SNV'ler) incelemişlerdir. 10X Genomics Chromium yöntemi kullanılarak elde edilen $1,2 \times 10^9$ çift uçlu okuma Supernova v.2.1.1 (Weisenfeld vd., 2017) yazılımıyla N50 değeri 8,42 Mb olan ve %91,2'lik BUSCO tamlığına sahip 16.880 skafolddan oluşan 2,56 Gb uzunluğunda bir taslak genom (tahmini genom boyutu 3,01 Gb) üretilmiştir. Taslak genom 1,1 Gb (%42,2) tekrarlayan diziler içermekte olup, *de novo* anotasyon ile Augustus v.3.3.2 algoritması (Stanke vd., 2008) kullanılarak 20.518 protein kodlayan gen tanımlanmıştır. Evcil ineğe karşı genom sentenisi, ortalama %92,7'lik bir benzerlik göstermiştir. Beş yabancı bireyin ortalama dizileme derinliği 9,8X olacak şekilde yeniden dizilenmesiyle filtrelenmiş halde 3.4×10^6 biallelik SNV kümesi tanımlanmıştır (Gonçalves vd., 2021).

Cervidler, boynuzların dökülmesi ve yeniden büyümesi ile ayırt edilmektedirler. Kuzey Amerika'da yaygın olarak yayılış gösteren beyaz kuyruklu geyik (*Odocoileus virginianus*), birçok hane için eğlence, ticaret ve gıda kaynağı değerine sahiptir. London vd. (2022), PacBio Sequel II platformunda dizilenen ve Wtdbg2 algoritması kullanılarak birleştirilen tek bir Illinois beyaz kuyruklu geyik DNA'sı kullanılarak üretilen bir genom sunmuşlardır. Çalışmada Omni-C kromatin konformasyon yakalama dizilimi, genom kontiglerini skafoldlamak için kullanılmıştır. 390,6 Gb PacBio, 157,4 Gb Illumina kısa okuma ve 145,1 Gb Omni-C okuma ile yapılan final birleştirme 21,7 Mb'lık bir kontig N50 değeri, 52,4 Mb'lık bir skafold N50 değeri ve %93,1'lik bir BUSCO bütünlüğüne sahip 508 skafolddan oluşan 2,42 Gb boyutunda elde edilmiştir. Otuz altı kromozom psödomolekülü, dizilenmiş tüm genom uzunluğunun %93'ünü kapsamıştır. BRAKER pipeline kullanılarak tahmin edilen toplam 20.651 gen, InterProScan kullanılarak doğrulanmıştır (London vd., 2022).

Equidae'nin evcilleştirilmiş üyelerinden biri olan eşek (*Equus asinus*), yaklaşık 7.000-9.000 yıl önce tropikal ve subtropikal Afrika'da evcilleştirilmiş ve insanlar için ulaşım, gübre ve yiyecek sağlamıştır (Smith ve Pearson, 2005; Han vd., 2014; Wang vd., 2020). Haplotip, hem bireysel hem de popülasyon düzeylerinde genomları anlama konusunda önemli bilgiler sağlamaktadır. Bununla birlikte, pek çok model olmayan organizma üzerindeki araştırmalar, haplotip eksikliğinden dolayı hala bağımsız genetik

varyasyonlara dayanmaktadır. Yapılan bir çalışmada insan uygarlığında hayati bir rol oynayan model olmayan bir organizma olan *Equus asinus* için haplotip birleştirme yapılmıştır. Tek moleküllü gerçek zamanlı dizileme (30X), Illumina kısa okuma dizileme (211X) ve yüksek verimli kromozom konformasyon yakalamasından (56X) elde edilen yüksek derinlikli sekans verilerine dayanarak Dezhou eşeğinin hibrit tek bireysel toplanmış haplotipi tanımlanmıştır. Yüksek derinlikte dizilenmiş Dezhou eşeği için neredeyse tamamlanmış bir haplotip ve eşek popülasyonunun yeniden sekanslama verileri için aşamalı bir kohort oluşturulmuştur. Dahası, Dezhou eşeğinin tam kromozom ölçekli haplotipini %99,7'den fazla faz oranıyla tanımlanmıştır. 39 milyondan fazla genetik varyasyona sahip bir eşek haplotipi veri seti oluşturmak için 156 eşekten oluşan bir kohort da fazlı hale getirilmiştir (Miao vd., 2022).

Günümüzde 1.400'den fazla türü tanımlanan yarasalar, mevcut tüm memeli türlerinin yaklaşık %20'sini oluşturmaktadır. Bu kapsamda Jebb vd. (2020), altı yarasa türünün (*Rhinolophus ferrumequinum*, *Rousettus aegyptiacus*, *Phyllostomus discolor*, *Myotis myotis*, *Pipistrellus kuhlii* ve *Molossus molossus*) ilk referans kalitesinde genomlarını oluşturmak için uzun okuma dizileme ve “state of-the-art scaffolding” protokolleri (Teeling vd., 2018) kullanmışlardır. Çalışmada 100X PacBio okuması, 10X Illumina okuması ve Hi-C konfirmasyon yakalama yöntemleri kullanılmıştır. RNA-seq verileri ise PacBio ile çeşitli dokulardan elde edilmiştir. İlgili çalışmada altı yarasa türünün hepsinde yüksek bitişikliğe sahip genomlar elde edilmiştir. Kromozom seviyesindeki skafoldlarda N50 değeri 92–171,1 Mb'dır. Yüksek oranda eksiksiz gen anotasyonlarını elde etmek için *de novo* ve homoloji gen tahminlerinin yanı sıra kısa ve uzun okuma transkriptomiklerle ‘Tool to infer Orthologs from Genome Alignments (TOGA)’ yazılımından gen projeksiyonları entegre edilmiştir. Sonuç olarak altı yarasa türü için 19.122 ila 21.303 protein kodlayan gen annotate edilmiştir. Laurasiatheria'daki yarasaların filogenetik pozisyonunu çözmek için, genomun ortolog protein kodlayan ve kodlama yapmayan bölgelerin kapsamlı setlerine çeşitli filogenetik yöntemler uygulanmış ve Scrotifera içinde yarasalar için bir bazal orjin belirlenmiştir (Jebb vd., 2020).

Koalalar, IUCN kırmızı listesinde VU (Vulnerable, Hassas) kategorisinde sınıflandırılan ikonik bir keseli memeli türüdür. Blanchard vd. (2023), koala popülasyonlarının karşılaştırmalı genomiklerinde ve retroviral entegrasyon çalışmalarında kullanılmak üzere Güney Avustralya koalasına olan "Wilpena"nın tüm genomunu elde etmişlerdir. 58

GB ONT verisi (ortalama okuma uzunluğu 24 kb ve ortalama Q skoru 13,7 olan 2.572.260 okuma) ve 1,289 milyon (2x150bp >Q30) Illumina okuması ile Flye algoritması kullanarak genomu birleştirmişlerdir. Birleştirilen genom uzunluğu 3.233.824.327 bp olup, ilk önce Medaka algoritması kullanılarak uzun okuma polishing ve sonrasında Illumina verileri kullanılarak Polca algoritması ile son polishing yapılmıştır. Sonuçta 1.265 kontig ve toplam genom boyutu 3.234.982.288 bp'lık bir uzunlukta final genom elde edilmiş ve birleştirilen genomun N50 değeri 48,8 Mb olmuştur. Genom, RED algoritması kullanılarak “soft” parametresi ile maskelenmiş ve veri tabanında mevcut olan birden fazla biyolojik dokuya ait koala RNAseq verileriyle birlikte braker2 algoritması ile analizi sonucunda toplam 52.384 gen tahmin edilmiştir. EggNOG mapper algoritması kullanılarak yapılan fonksiyonel anotasyon işlemiyle transkripsiyon desteğine sahip toplam 27.669 gen tanımlanmıştır (Blanchard vd., 2022).

Halka kuyruklu lemur (*Lemur catta*), Madagaskar'a özgü karizmatik bir primattır. Bu hayvanlar oldukça ilginç ve popüler olup, sayıları vahşi doğada 2.500 bireyin altına düşmüş ve IUCN tarafından nesli tükenmekte olan bir tür olarak sınıflandırılmıştır. Palmada-Flores vd. (2022) yaptıkları çalışmada *L. catta* (mLemCat1) için Omurgalı Genomları Projesi standartlarına uygun yeni bir yüksek kaliteli referans genom birleştirmesi oluşturmuştur. Bu yeni uzun okuma birleştirmesinde tahmin edilen genom boyutu 3,21 Gb olarak sunulmuş ve mLemCat1 genomunun birleştirilmesi için 86,43X kapsama sahip 10X bağlantılı okuma (linked-reads) verisi, 66,68X kapsama sahip Arima verisi, 154,57X kapsama sahip Bionano verisi ve 62,88X kapsama sahip PacBio verisi kullanılmıştır. Elde edilen genom birleştirmenin bitişikliği ve bütünlüğü 90,982 Mb skafold N50 değeri ve 10,570 Mb kontig N50 değerleri ile son derece yüksektir (Palmada-Flores vd., 2022).

Boz ayı (*Ursus arctos*), dünyadaki ikinci en büyük ve en yaygın karasal etçil bir türdür ve son zamanlarda insan metabolik hastalıkları için tıbbi bir model olarak önem kazanmıştır. Yapılan bir çalışmada, bir erkek Kuzey Amerika boz ayısının tamamen fazlı bir kromozom düzeyinde birleştirmesi için elde edilen Pacific Biosciences (PacBio) Hi-Fi (32X kapsam) verileri ile veri tabanlarındaki mevcut Hi-C verileri birlikte değerlendirilerek Hifiasm (Cheng vd., 2021) ve yak (<https://github.com/lh3/yak/releases/tag/v0.1>) algoritmaları kullanılmıştır. Final genom boyutu 70,08 Mb skafold N50 uzunluğu ve 43,94 Mb kontig N50 uzunluğu ile 2,47

Gb'dır. BUSCO analizi, Mammalia veri tabanındaki tek kopya ortologların %94,5'inin bu genomda bulunduğunu ortaya çıkarmıştır. Ayrıca tekrarlayan elementler genomun %44,48'ini oluşturmuş ve toplam 20.480 protein kodlayan gen belirlenmiştir. Çalışmada tüm genom hizalamasına dayanarak kutup ayısıyla boz ayının yüksek oranda sentenik olduğu bulunmuş, ayrıca 7.246 tek kopya ortolog verinin kullanılmasıyla gerçekleştirilen filogenetik analizde ise Ursidae için şu anda önerilen tür ağacını desteklediği bildirilmiştir (Armstrong vd., 2022).

Amerikan vizonu (*Neogale vison*), Kuzey Amerika'ya özgü etçil bir memelidir ve Mustelidae ailesinin bir üyesidir. Karimi vd. (2022), yaptıkları bir çalışmada PacBio (21X), Hi-C (57X) ve kısa okumalar (BGI, 38X) kullanarak Amerikan vizonu için yüksek kaliteli bir kromozom düzeyinde genom birleştirmesi elde etmişlerdir. Birleştirmenin ilk taslağı 2.884.047 PacBio uzun okuması kullanılarak oluşturmuştur. Bu çalışmada WTDBG2, Flye, Hifiasm, Hicanu ve IPA birleştirme yazılımları denenmiş ve en iyi birleştirme Hifiasm ile elde edilmiştir. Daha sonra BGI kısa okumaları ile Polca algoritması kullanılarak polishing yapılmıştır. Hi-C verilerinin ilk taslak birleştirmeye entegrasyonu ile 220 Mb'lık N50 değerine sahip 183 skafold içeren birleştirme elde edilmiştir. Amerikan vizonunun bu boşluksuz genom birleştirmesinin (ASM_NN_V1), tüm genomun yaklaşık %98,6'sının 15 kromozom tarafından kapsandığı toplam 2,68 Gb uzunluğunda olduğunu rapor etmişlerdir. NCBI Ökaryotik Genom Anotasyon hattı kullanılarak Amerikan vizon genomunda toplamda 25.377 gen tahmin edilmiştir (Karimi vd., 2022).

Sibirya sincabı *Tamias sibiricus* (Laxmann, 1769), Rodentia takımının Sciuridae familyasındaki Xerinae alt familyasına ait bir türdür. Li vd. (2022) Illumina dizileme, PacBio dizileme ve kromozom yapı yakalama tekniğini kullanarak *Tamias sibiricus*'un kromozom ölçeğinde genomunu birleştirmiştir. Çalışmada 132,39 Gb Illumina, 111,63 Gb PacBio, 217,38 Gb Hi-C, 3,89 Gb RNA-seq verisi kullanılmıştır. PacBio okumalarının birleştirilmesinde Canu v1.8 algoritması kullanılmış ve Illumina okumaları Pilon v1.23 algoritması ile polish edilmiştir. 3D-DNA ve Juicebox v1.11.08 algoritmaları ile Hi-C okumaları kullanılarak kromozomlara atama yapılmıştır. Birleştirilen genom 172,61 Mb skafold N50 uzunluğu ile 2,64 Gb boyutundadır. Toplam 2,59 Gb genom verisi, %98,03'e varan bir oranla 19 kromozoma (28.70 ila 222,90 Mb aralığında) atanmış ve oryante edilmiştir. Ek olarak 25.311 protein kodlayan genin ortalama gen uzunluğu

32.936 bç olarak tahmin edilmiş ve bu genlerin %94,73'ü fonksiyonel olarak annotate edilmiştir. Genomun %38,87'sini oluşturan ve toplam boyutu 1,03 Gb olan tekrarlayan diziler tespit edilmiştir (Li vd., 2022).

Sarı gerdanlı sansar (*Martes flavigula*), Mustelidae familyasının *Martes* cinsine aittir ve adını göğsünde ve boğazında göze çarpan sarı tüylerinden almıştır (Grassman vd., 2005). Sarı gerdanlı sansar, Asya'nın büyük bir kısmında yayılış gösteren ve çok çeşitli habitatları işgal eden orta büyüklükte bir etoburdur. Mei vd. (2023) yaptıkları bir çalışmada Oxford Nanopore ve Hi-C teknolojileri kullanılarak bu organizmanın yüksek kaliteli bir genom birleştirmesini bildirmişlerdir. 150 bç çift uçlu okumalar oluşturmak için Illumina HiSeq X Ten platformu kullanılmış ve bunlar, genom birleştirmesinin 96,70 katını kapsayan 1,58 milyar okuma (236,83 Gb) ham dizi verisi sağlamıştır. Nanopore okumaları PromethION platformu ile elde edilmiş ve toplamda 27,76 milyon okuma (264,40 Gb) ham dizi verisi elde edilmiş, bu da genom birleştirmenin 107,96 katı kapsamındadır. Ortalama okuma uzunluğu ve N50 uzunluğu 9,53 kb ve 17,43 kb olup, en uzun okuma kapsamı 204,65 kb'dır. Hi-C kütüphaneleri, Illumina NovaSeq6000 platformunda 150 bp PE modunda dizilenmiş ve genom birleştirmenin 105,07 katını kapsayan 257,31 Gb Hi-C okuması elde edilmiştir. Ek olarak yedi dokudan RNA izole edilmiş ve toplam 60,43 Gb ham okuma sağlayan Illumina Novaseq6000 platformunda transkriptom dizileme gerçekleştirilmiştir. Genom birleştirmede NextDenovo v2.0-beta.1 algoritması (<https://github.com/Nextomics/NextDenovo>) ve genomun polish edilmesinde Nextpolish v1.0.5 algoritması kullanılmıştır. Ham Hi-C verileri öncelikle Hi-C-Pro v2.8.0 algoritması kullanılarak filtrelenmiştir. Ardından, Fastp ile kalite kontrolü sonrası temiz Hi-C verileri, Bowtie2 v2.3.2 algoritması kullanılarak sansar genom birleştirmesine haritalanmıştır. Hi-C verileriyle birleştirildiğinde, kromozom düzeyinde bir genom üretmek için LACHESIS yazılım uygulanmıştır. Ek olarak, kontigler arasındaki etkileşim gücüne ve Juicebox algoritması kullanarak bir bağlantı haritasına dayalı olarak yanlış birleştirilmiş kontigler manuel olarak düzeltilmiştir. Çalışmada elde edilen final genomun toplam boyutu 2.449,15 Mb olup, N50 uzunluğu 68,60 Mb olan 215 kontigden oluşmaktadır. Hi-C analizi kullanılarak, birleştirilen dizilerin 2.419,20 Mb'ı (%98,78) 21 bağlantı grubuna atanmıştır. Merqury değerlendirmesi, genomun 43,75 QV değeriyle %94,95 oranında tamamlandığını ortaya koymuştur. Ek olarak, genomun yaklaşık %39,74 tekrar dizisinden oluştuğu bulunmuştur. Protein kodlayan genlerin tespitinde *ab*

initio tahmin, homoloji temelli tahmin ve transkriptom-temelli yöntemler kullanılmış, bunun sonucunda da 20.464 protein kodlayan genden 20.322 (%99,31) gen için tahmin ve fonksiyonel anotasyon başarıyla gerçekleştirilmiştir (Mei vd., 2023).

Roborovski cüce hamsteri *Phodopus roborovskii*, Cricetinae alt familyasındaki yedi hamsterden biri olan *Phodopus* cinsine aittir. Fareler, sıçanlar veya gelincikler gibi diğer kemirgenlere benzer olarak, hamsterler de bir dizi hastalık için önemli hayvan modelleri olarak kullanılmaktadır. Andreotti vd. (2022) yaptıkları bir çalışmada *Phodopus roborovskii* genomunu dizilemişlerdir. Çalışmada elde edilen final birleştirmenin toplam uzunluğu 2,38 Gb olup, N50 değeri 25,78 Mb ve L50 değeri 30 olan toplam 2.078 (2.055 kontig > 50 kb) kontigden oluşmaktadır. QUAST' ile yapılan genom birleştirme değerlendirmesine göre 676,47 milyon paired-end kısa okumaların %99,75'i 80X kapsamda ve 4,13 milyon uzun okumaların %99,74'ü 34X kapsamda haritalanmıştır. Canu algoritmasının ürettiği ham birleştirme %79,3 bütünlüğe sahipken, bu değer Racon algoritması kullanılarak yapılan polishing ile %85,9'a, Medaka algoritması kullanılarak yapılan polishing ile %88,8'e ve son olarak POLCA algoritması ile kısa okumaların kullanıldığı polishing ile %92,0'ye ulaşmıştır. Ayrıca çalışmada 4 adet RNA-seq örneği ile GeMoMa algoritması kullanılarak 21.836 transkript tespit edilmiş ve 17.844 gen lokusu annotate edilmiştir (Andreotti vd., 2022).

Hawaii Adaları birçok endemik flora ve faunanın yok olmaya sürüklendiği veya insan faaliyetinin bir sonucu olarak giderek artan bir şekilde tehdit edildiği bir biyolojik çeşitlilik sıcak noktasıdır. Adalara özgü sadece iki memeli türü vardır. Bunlar Hawaii Hoary yarasası ve Hawaii keşiş fokudur. Hawaii foku (HMS), *Neomonachus* cinsinin tropikal kulaksız foklarının günümüze ulaşan tek türüdür. Hawaii fokunun darboğaz ve müteakip akraba çiftleşmesi nedeniyle genetik heterojenliğinin azaldığı bildirilmiştir. Mohr vd. (2022) yaptıkları bir çalışmada 10X Genomics linked reads (61X kapsam), Nanopore sekanslama (7,3X kapsam), Bionano Genomics Optik Genom Haritalama ve Hi-C (60,8X) yöntemlerini kullanarak erkek bir Hawaii fokunun kromozomal referans birleştirmesini 2.376.537.768 bç genom büyüklüğü ile rapor etmişlerdir. Çalışmada yapılan final birleştirme 27.736 gen içermekte olup 16 otozom, bir X ve Y kromozomlarının kısımlarından oluşmaktadır. Bu hayvandaki varyantlar diğer HMS'lerle ve dizilenen bir insan örneğiyle karşılaştırılmış ve insanda görülen varyasyonun yaklaşık

%12'sini doğrulamıştır. Referans hayvanın HMS'yi temsil ettiğini doğrulamak için diğer 10 bireyle karşılaştırılmış ve hepsinde benzer şekilde düşük varyasyon görülmüştür.

1.10. Diğer hayvanlarda yapılan genom çalışmaları

Bal arısı (*Apis mellifera*)'nın yüksek oranda arı sütü üreten bir ırkı olan HRJHB, Çin'de İtalyan bal arılarının (*Apis mellifera ligustica*) art arda yapay seçilimi yoluyla elde edilmiştir. Cao vd. (2021) 33,37 Gb PacBio uzun okumaları, 10,77 Gb Illumina kısa okumaları ve 32,79 Gb Hi-C okumaları kullanarak HRJHB ırkı için kromozom ölçeğinde bir genom dizisi elde etmişlerdir. Elde edilen genom, skafold N50 değeri 13,6 Mb olan 222 Mb uzunluğunda 16 pseudo-kromozomdan oluşmaktadır. BUSCO analizi, bu genom için %99,3 bütünlük değeri ortaya koymuştur. Genomun %8,11'i tekrarlayan element olarak karakterize edilmiş olup *ab initio* yaklaşımı, 18,05 Gb transkriptom verisi ve homoloji yöntemleri kullanılarak MAKER yazılımı ile 12.288 protein kodlayan gen tahmin etmişlerdir. HRJHB ırkı ve yakından ilişkili İtalyan bal arıları arasında tüm genom hizalama analizi yoluyla bir kromozom inversiyonu tanımlanmıştır (Cao vd., 2021).

Küçük Çayır Tavuğu (*Tympanuchus pallidicinctus*; LEPC), süslü ve üreme mevsimindeki muhteşem gösterileriyle tanınan ikonik bir Kuzey Amerika çayır tavuğudur. NCBI'da referans genomu bulunan diğer *Tympanuchus* türünde 2.186 skafold, 12 Mb'lık bir N50 değerine sahip olup, annotate edilmemiş olan (GCA_001870855.1) Greater Prairie-Chicken'dir. Black vd. (2023) yaptıkları bir çalışmada LEPC referans genomunun eldesi için PacBio Circular Consensus Reads (CCS) ile ortalama 14 kb okuma uzunluğuyla 4 milyon okuma (65,8 Gb veri) elde etmişler ve 60-mer analizi ile 0,987 Gb'lik bir haploid genomu tahmin etmişlerdir. PacBio CCR okumaları, Nextdenovo algoritması kullanılarak birleştirilmiş ve genom 45 Mb'lık bir N50 değeri ile 206 skafolddan oluşmuştur. Ayrıca 15.563 tahmin edilen protein kodlayan gen belirlenmiştir. Ek olarak 40.823 mRNA ve 9.527 kodlamayan mRNA'lar, LEPC skafold seviyesindeki birleştirmesinde annotate edilmiştir (Black vd., 2023).

Tang vd. (2021), en büyük yengeç olan ve ticari olarak değerlendirilen mavi kral yengeci (*Paralithodes platypus*) genomunu kromozom seviyesinde birleştirmiş ve N50 değeri 51,15 Mb olan 104 kromozom elde etmişlerdir. Çalışmada yaklaşık 447,62 Gb BGI kısa insört okuma, ~229,39 Gb PacBio uzun okuma, ~321,74 Gb Hi-C dizileme ve ~18,39 Gb

transkriptom verisi üretilmiştir. İlk olarak okumalar filtreleme işlemlerinden sonra wtdbg, racon, pilon ve 3D *de novo* assembly yazılımları ile kombine edilmiş ve okumalar 4,77 Gb genomla birleştirilmiştir. Genom anotasyonunda *de novo*, homoloji tabanlı yöntemler ile transkriptom verileri kullanılmış ve 28.287 protein kodlayan gen tanımlanmıştır. Çeşitli Crustacea türleri ile diğer birkaç temsili türün yüksek kaliteli genom birleştirmeleri kullanılarak gerçekleştirilen filogenetik analiz, mavi kral yengeci ile Çin mitten yengeci ve yüzen yengeç de dahil olmak üzere gerçek yengeçlerle yakın bir ilişkiye sahip olduğunu ve ~ 272,5 milyon yıl önce ayrıldığını ortaya koymuştur (Tang vd., 2021).

Hawaii'ye 1900'lerin başında getirilen Japon beyaz gözü veya ötücü beyaz gözü (*Zosterops japonicus*), günümüzde takımadalarda en çok bulunan karasal kuştur. Zosteropidae ailesi özellikle de *Zosterops* cinsi, kuşlardaki ve bir bütün olarak omurgalılarındaki en hızlı farklılaşma olaylarından birini temsil etmektedirler (Moyle vd., 2009). Venkatraman vd. (2021) yapmış oldukları çalışmada istilacı bir tür olan *Z. japonicus*'un ilk genomunu sunmuşlardır. Hiseq X lane ile 122,6 Gb (85X kapsam) veri üretmişler ve genom boyutunun 1,1 Gb olduğunu tahmin etmişlerdir. Supernova 2.0 algoritması (Weisenfeld vd., 2017) kullanarak yapılan final birleştirmede N50 değeri 9,69 Mb ile 29.443 skafold elde edilmiştir. Genomun annotate edilmesi için AUGUSTUS yazılımı ile tek bir örnekte 24.213 gen ve çok örnekli veri setinde ise 21.571 gen belirlemişler, BRAKER yazılımı ile bu sayı 74.311 gen ile 69.835 gen olmuştur. Ayrıca çalışmada 47 örnekten elde edilen RNA-seq verileri ile tek bir örnekten alınan verilerin anotasyon üzerine etkisi karşılaştırılmıştır. Daha fazla verinin AUGUSTUS kullanıldığında annotate genlerin sayısında önemli ölçüde değişme olmadığını, ancak BRAKER2 yazılımı kullanıldığında sayının önemli ölçüde arttığını vurgulamışlardır (Venkatraman vd., 2021).

Anatidae ördekler, kazlar ve kuğular gibi Anseriformes takımına ait kuşların çoğu dahil olmak üzere 43 cins ve 174 tür içeren Anseriformes takımına sahip antik Aves sınıfının bir familyasıdır ve yüzen kuşların en önde gelen familyasıdır (Hoyo vd., 1992). Anatidae, büyük ekonomik değeri olan çok sayıda su kuşu türü içermesine rağmen genetik çeşitlilik temeli yeterince araştırılmamıştır. Zhao vd. (2022), Güney Çin'de bir tür olan Aslan başlı kazın (*Anser cygnoides*) kromozom düzeyinde bir genom birleştirmesi sunmuşlardır. Illumina sekanslama (111,25 Gb), PacBio SMRT sekanslama (185,37 Gb), BioNano optik haritalama (468,4 Gb) ve Hi-C (230 Gb) ile üretilen okumaların birleştirilmesinde

Canu algoritması kullanılmış ve Pilon algoritması ile polish edilmiştir. RNAseq ile 9 farklı dokudan 760 Gb ham okuma elde edilmiştir. Genom bütünlüğü BUSCO ile değerlendirilmiş ve genom anotasyonunda MAKER algoritması kullanılmıştır. Genom çapında ilişkilendirme için varyantların tespitinde BWA, Sam Tools ve GATK4 kullanılmıştır. Birleştirilen genom 1,19 Gb boyutunda olup, 20,59 Mb N50 uzunluğuna sahip 1.859 kontigden oluşmaktadır ve BUSCO bütünlüğü %96,92'dir. 40 psödokromozomda birleştirilmiş genom, 21.208 protein kodlayan geni içermektedir. Yapılan karşılaştırmalı genomik analiz, kazların ve ördeklerin yaklaşık 28,42 milyon yıl önce ayrıldığını ve kazların büyük gen ailesi genişlemesi ve daralması geçirdiğini ortaya koymuştur. Ayrıca bu çalışmada Wuzong kazı, Huangzong kazı, Magang kazı ve Aslan başlı kaz dahil olmak üzere farklı kaz ırklarında vücut ağırlığı ile ilişkili genetik markırları belirlemek için 44.858 tek nükleotid polimorfizmi (SNP'ler) saptayan ortalama 1.520,6 Mb ham veri sağlayan genom çapında bir ilişkilendirme çalışması da gerçekleştirilmiştir. Böylece 6 SNP'nin vücut ağırlığı ile önemli ölçüde ilişkili olduğunu ve 25'inin potansiyel olarak ilişkili olduğunu gösterilmiştir (Zhao vd., 2022).

Göçmen bir su kuşu olan yaban kuğu kazı, evcil kaz ırklarına göre çok farklı özelliklere sahip olup en dikkate değer olanı yaban kuğu kazının uzun mesafe uçuş yeteneğine sahip olmasıdır. Kuğu kazı, IUCN Tehdit Altındaki Türlerin Kırmızı Listesi'nde savunmasız olarak listelenmiştir ve kuğu kazının küresel nüfusunun yaklaşık 60.000–90.000 olduğu tahmin edilmektedir (<http://www.iucnredlist.org/>). Ni vd. (2023), yalnızca hayvan koruma alanında çalışan araştırmacılar için değil, aynı zamanda kaz genomik ıslahı ve daha ileri genomik araştırmalara odaklananlar için de değerli bir kaynak sağlamak üzere 10 yaban kuğu kazı için tüm genom dizileme yapmışlardır. Dizileme 350 bç'lik bir insört boyutu ile çift uçlu kütüphaneler kurularak Illumina HiSeq platformunda yapılmıştır. Ham veriler üzerinde kalite kontrolü gerçekleştirmek için FASTP (Chen vd., 2018) yazılımı kullanılmıştır. Temiz okumalar varsayılan parametrelerle BWA-MEM algoritması (Li ve Durbin, 2009) kullanılarak kuğu kaz genomuna (PRJNA826973) hizalanmıştır. SNP'ler, GATK yazılımı (McKenna vd., 2010) kullanılarak calling yapılmıştır. SNP'ler, VCFtools (Danecek vd., 2011) kullanılarak minör alel frekansı <0,05 ve kayıp oranı >0,10 olacak şekilde filtrelenmiştir. SNP'lerin kromozomlar üzerindeki dağılımı, R “CMplot” paketi kullanılarak çizilmiştir. Çalışmada toplamda 162,6 Gb veri üretilmiştir. Bir birey, elde edilen 39,6 Gbp veri ile derinlemesine

dizilenmiş, diğer örnekler ise yaklaşık 13,7 Gbp veri ile dizilenmiştir. Çalışmadaki ortalama haritalanma oranı 0,90'dır (eşleme kalite değeri ≥ 20). Son olarak toplam 10.727.005 SNP (minör allel frekansı ≥ 0.05) elde edilmiş olup, toplam 7.646.295 (%71,28) transisyon (Ts) ve 3.080.710 (%28,72) transversiyon (Tv) gözlemlenmiştir (Ni vd., 2023).

Balıklar, çeşitli cinsiyet belirleme sistemleri ve üreme modelleri sergilerler (Bachtrog vd., 2014; Nagahama vd., 2021). Gonokorizm ve hermafroditizm, eşeyli üreyen organizmalardaki en önemli üreme modelleridir ve genellikle teleost balıklarında görülür (Bachtrog vd., 2014; Jarne ve Auld, 2006; Nagahama vd., 2021). 1.500'den fazla türde hermafroditizm belgelenmiştir (De Mitcheson ve Liu, 2008). Gonokorizm ve hermafroditizmin evrimsel yönü, çözülmesi gereken ilgi çekici bir gizemdir. Özel geçici hermafrodit aşaması, *Larimichthys polyactis* hermafrodit oluşumunu incelemek için çekici bir model yapmaktadır. Bununla birlikte Doğu Asya'daki en ünlü ticari balık türü olan *L. polyactis* ile *L. crocea* arasındaki köken ve evrimsel ilişki belirsizliğini korumaktadır. Xie vd. (2023) yaptıkları bir çalışmada, PacBio (48 Gb ham veri; ~69X kapsam), Illumina (128 Gb ham veri; ~183X kapsam), 10X Genomics (137,28 Gb ham veri; 196X kapsam) ve BioNano (117 Gb) platformlarını kullanarak Overlap-Layout-Consensus; DBG2OLC ve falcon (version 1.2.4) algoritmaları ile ~706 Mb uzunluğunda (kontig N50 değeri 1,21 Mb ve skafold N50 değeri 4,52 Mb) *L. polyactis* için genom birleştirme yapmışlardır. 4.584 BUSCO grubundan 4.305'inin (~%94,0) genomda tamamen bulunduğunu ortaya koyan BUSCO analizi yapılmış, bu da *de novo* birleştirmenin yüksek düzeyde eksiksiz olduğunu göstermiştir. Toplam 157 Mb (~%22) genom dizisinin, 38 Mb (%5,34) uzun-terminal tekrar (LTR) retrotranspozonları ve 67 Mb (%9,57) DNA transpozonları dahil olmak üzere tekrarlayan öğeler içerdiği tahmin edilmiştir. Homoloji tabanlı, *de novo*, transkriptom verileri ile (her iki cinsiyet için 13 ana organdan ve 18 gonadal gelişim zaman noktasından elde edilen) 25.233 protein kodlayan geni rapor etmişlerdir. Çalışmadaki filogenomik analiz, *L. polyactis*'in yaklaşık 25,4 milyon yıl önce ortak ata olan *L. crocea*'dan ayrıldığını öne sürmüştür. Çalışmada elde edilen yüksek kaliteli genom birleştirmesi, iki tür arasında büyük kromozom bölünmesi veya füzyon olayları olmadan birkaç kromozom içi yeniden düzenlemeyi ve yer değiştirmeyi ortaya çıkaran karşılaştırmalı genomik analizi mümkün kılmış; *Dmrt1* geni, *L. polyactis*'te erkeğe özgü gen olarak tanımlamıştır. Transkriptom analizi, *Dmrt1* ve

bunun yukarı upstream regülatör geninin (*rnf183*) ifadesinin her ikisinin de cinsel olarak dimorfik olduğunu göstermiştir. *Rnf183*, *rnf223* ve *rnf225*'in iki paralogundan farklı olarak, yalnızca *Larimichthys* ve *Lates*'te mevcut olup, diğer teleost türlerinde olmayışı, onun soya özgü çoğaltmadan kaynaklandığını veya diğer teleostlarda kaybolduğunu öne sürmüştür. Son olarak çalışmada yapılan filogenetik analizin, erkek *L. polyactis*'teki hermafrodit aşamasının *dmrt1*'in dizi evrimi ile açıklanabileceğini göstermiştir (Xie vd., 2023).

Çipura levreği (*Micropterus salmoides*), av hayvanı veya evcilleştirilmiş balık olarak yaygın bir şekilde yetiştiriciliği yapılan kozmopolit bir türdür. He vd. (2022), Illumina çift uçlu dizileme (58,51 Gb), PacBio tek molekül dizileme (94,69 Gb) ve Hi-C (77,53 Gb) teknolojilerinin bir kombinasyonu ile *M. salmoides*'in yüksek kaliteli bir kromozom düzeyinde referans genomunu elde etmişlerdir. Canu ve SMARTdevono algoritmaları ile birleştirilen genom Pilon algoritması ile polish edilmiş ve sonuç olarak genom 15,68 Mb'lık bir kontig N50 değeri ve 35,77 Mb'lık skafold N50 değeri ile 844,88 Mb uzunlukta birleştirilmiştir. Birleştirilen genom dizilerinin yaklaşık %99,9'u (844,00 Mb) 23 kromozoma atanmıştır. Birleştirilen genomun %38,19'u tekrar dizisi olarak karakterize edilmiştir. *Ab initio* temelli, homoloji temelli, unigen temelli yaklaşımlarla ve 7 dokudan RNA seq yapılarak elde edilen 20 Gb temiz veri ile 2.693 kodlama yapmayan RNA belirlenmiştir. 3.415 gen ailesinden toplam 26.370 protein kodlayan gen tahmin edilmiş ve bunların %97,69'u işlevsel olarak annotate edilmiştir (He vd., 2022).

Japon yılan balıkları (*Anguilla japonica*), ticari olarak önemli bir tür olup yoğun bir şekilde yetiştiriciliği yapılmaktadır. Günümüzde bu ve ilgili türlerin (Amerikan ve Avrupa yılan balıkları) ticari olarak üremesi zor olduğundan yabani stok su ürünleri yetiştiriciliği için kullanılmaktadır. Ayrıca iklim değişikliği, habitat kaybı, su kirliliği ve değişen okyanus akıntıları yılan balığı popülasyonlarını olumsuz etkilemektedir. Wang vd. (2022), yüksek kaliteli bir Japon dişi yılan balığının referans genomunu ONT uzun okumalarının (234X kapsam; 239,64 Gb), PacBio CLR okumalarının (261X kapsam; 267 Gb), 10X Chromium bağlantılı okumaların (313X kapsam; 319,7 Gb), Hi-C verilerinin (48X kapsam; 48,99 Gb), Illumina kısa okumalarının (148X kapsam; 151,89 Gb) ve mate-pair okumaların (127 kapsam; 130,5 Gb) entegrasyonu yoluyla elde etmişlerdir. Çalışmada Canu v2, Wtdbg2 v2.5 ve Flye v2.71 algoritmaları kullanılarak ONT uzun okumaları ayrı ayrı birleştirilmiş ve kontigler Quickmerge kullanılarak omurgalı

türlerinde kontig N50 ile tam genlerin yüzdesi arasında bir denge elde etmek için birleştirilmiştir. ONT okumalarını kullanarak birleştirme hatalarını kendi kendine düzeltmek için 2 tur Racon v1.4.16 algoritması ve 1 tur Medaka v1.6.1 algoritması kullanılmıştır. PacBio CLR okumaları daha sonra 2 tur Racon kullanılarak hata düzeltme için dahil edilmiştir. Son adım olarak 2 tur Pilon v1.23 algoritması kullanarak Illumina kısa okumaları ve mate-pair kütüphaneleri entegre edilerek birleştirme geliştirilmiştir. Tigrint v1.1.2 ve ARKS v1.0.3 algoritmaları, 10x bağlantılı okumalardan alınan paylaşılan barkodlara göre yanlış birleştirilmiş kontigleri düzeltmek ve kontigleri skafoldlara bağlamak için kullanılmıştır. Skafoldlar için Bionano optik haritalamada kullanılan 3 enzimi entegre etmek için OMGS kullanılmıştır. Dovetail Omni-C kütüphanesinden Hi-C verilerine dayalı olarak 3-dimensional DNA kullanarak skafoldlar daha da genişletilmiş ve skafoldlar JuiceBox v1.11.08 tarafından manuel olarak iyileştirilmiştir. Sonuç olarak genom 1,028 Gb boyutunda, kontig N50 değeri 21,48 Mb ve skafold N50 değeri 58,7 Mb ile elde edilmiştir. BUSCO veri tabanındaki *actinopterygii_odb10*'a göre ışıın yüzgeçli balıklardaki tek kopyalı doğrudan homologların %94'ü Japon yılan balıklarında toplanmıştır. Tekrarlanan elementler, tüm genomun %30,49'unu oluşturduğu belirlenmiştir. *De novo*, homoloji tabanlı ve transkriptom tabanlı anotasyonlar dahil olmak üzere genomdaki protein kodlayan genleri annotate etmek için üç yöntem kullanılmıştır. 29.982 kodlayan gen belirlenmiş ve %97,44'ü (29.219) fonksiyonel olarak annotate edilmiştir. Çeşitli kromozomal füzyonlar ve bölünmeler, *Anguilla* soyunda atalara ait protokromozomal sayı olan 25'i 19'a indirmiştir. Üçüncü tur tüm genom kopyaları (3R-WRD'ler) olaylarında büyük kromozom yeniden düzenlemelerinin meydana geldiğini göstermişlerdir. Genişletilmiş gen ailelerinin filogenetik analizi, koku alma reseptörlerinin ve voltaj kapılı Ca^{2+} kanallarının önemli ölçüde genişlediğini göstermiştir (Wang vd., 2022).

Smith vd. (2022), göl balığının (*Salvelinus namaycush*) kromozom-tutturulmuş genom birleştirmesini ve anotasyonunu yapmışlardır. Dizileme için PacBio uzun okuma, Illumina çift uçlu okuma, Hi-C dizileme ve bağlantı haritası kullanılarak yüksek oranda bitişik 4.120 skafolddan (7.378 kontig) oluşan genom elde edilmiştir (kontig N50 değeri 1,8 Mb ve skafold N50 değeri 44,975 Mb). Genomun %84,7'si 42 kromozom seviyesinde skafolda atanmış bu birleştirmenin şu anda mevcut olan en iyi salmonid genomları ile aynı seviyede olduğu bildirilmiştir (Smith vd., 2022).

'Red drum' olarak da bilinen *Sciaenops ocellatus*, Çin'de yaygın olarak yetiştirilen ticari değeri yüksek bir balıktır. Son yıllarda *S. ocellatus*'da genetik araştırmalar ve moleküler ıslah yoğun olarak çalışılmasına rağmen tüm genom verileri yetersizdir. Xu vd. (2021), genom anotasyonu ile birlikte *S. ocellatus* türünün yüksek kaliteli genomunu yayınlamışlardır. Bu çalışmada 39,31 Gb temiz sekans verisiyle 17mer frekans dağılımında genom boyutu tahmini yapılmış ve genom boyutunun 711,58 Mb olduğu tahmin edilmiştir. Daha sonra 10X verileri Supernova algoritması ile birleştirilmiştir. Birleştirme kalitesini arttırmak için krskgf ve Gapcloser algoritmaları kullanılarak genomdaki boşluklar doldurulmuştur. Genom birleştirmesini iyileştirmek için Hi-C verisi kullanılmış, ilk olarak veriler HiC-Pro algoritması ile filtrelenmiştir. Bowtie2 algoritması ile ham okumalar taslak genoma haritalanmıştır. Jucier ve 3D *de novo* algoritmaları kullanılarak skafoldlar 24 pseudo-kromozoma birleştirilmiştir. Final genom boyutu 686,62 Mb olup kontig N50 değeri 99,71 kb ve skafold N50 değeri 25,62 Mb olarak bildirilmiştir. Hi-C teknolojisi kullanılarak toplam skafoldların %87,82'sini içeren 24 pseudo-kromozom elde edilmiştir. Farklı veri tabanlarındaki proteinlere %99,93 homolog olan 20.053 protein kodlayan gen annotate edilmiştir. Ek olarak toplamda birleştirilen genomun %19,45'ine denk gelen 133,54 Mb tekrar dizisi tespit edilmiştir (Xu vd., 2021).

Cobitoidea, Cypriniformes'teki iki süper aileden biri olmasına rağmen Cobitoidea balıkları için sadece birkaç genom dizilenmiştir. Yang vd. (2019b), 252,90 Gb (256X kapsam) kısa Illumina okumaları ve 31,60 Gb (50X kapsam) uzun PacBio Sequel okumaları ile Platanus (Kajitani vd., 2014) ve DBG2OLC (Ye vd., 2016) algoritmalarını kullanarak genom birleştirme yapmışlardır. Birleştirilen genom 2,87 Mb N50 değeri ile 583,47 Mb uzunluğunda olup, tahmini genomun (638,07 Mb) %91,44'ünü kapsamaktadır. Hi-C tabanlı kromatin temas haritaları kullanılarak genom birleştirmesinin %99,31'i 25 kromozoma dağıtılmış ve N50 değeri 22,3 Mb olarak elde edilmiştir. Gen anotasyon bütünlüğü BUSCO analizi ile değerlendirilmiş ve 2.586 korunmuş genden 2.470'i (%95,5) birleştirilen genomda bulunmuştur. Tekrarlayan elementlerin tüm genomun %33,08'ine ulaştığı hesaplanmış ve 25.406 protein kodlayan gen belirlenerek bunların %92,59'u işlevsel olarak açıklanmıştır (Yang vd., 2019b).

Çevresel tahribat ve aşırı avlanma nedeniyle bu türün popülasyonu tehdit altında olan kaba deri heykeltıraş (*Trachidermus fasciatus*), Çin, Kore ve Japonya'nın kıyı bölgelerinde bulunan küçük, etobur ve katadrom bir balıktır (Islam vd., 2007; Wang ve

Chen, 2010). Nesli tükenmekte olan bu türü korumak için genomun dizilenmesi önemlidir ve bu tür için bir referans genomun olmayışı, araştırmaları büyük ölçüde engellemektedir. Bu eksikliği gidermek için Xie vd. (2021) yaptıkları bir çalışmada *Trachidermus fasciatus* genomunu kapsamlı bir şekilde birleştirmek için NextDenovo algoritması ile Nanopore uzun okuma (4 milyon okuma; 161X kapsam), Illumina kısa okuma ve Hi-C yöntemlerini (toplamda 359 milyon okuma; 100X kapsam) kullanmışlardır. Sonuç olarak tahmin edilen genom boyutu 542,6 Mb ($2n = 40$) ve 24,9 Mb'lık bir skafold N50 değeri ile kromozom düzeyinde yüksek kaliteli bir genom birleştirmesi elde edilmiştir. Birleştirilen genomun BUSCO değeri %96'dan yüksektir. Homoloji temelli yaklaşım ve sekiz farklı dokudan izole edilen RNA'lar (434 milyon okuma) ile 35.093 mRNA transkripti ile toplam 19.147 protein kodlayan gen belirlenmiştir. Ayrıca çalışmada RNR adlı yeni bir gen bulma stratejisi gösterilmiş ve toplamda 51 yeni gen tanımlanmıştır (Xie vd., 2021).

Kingan vd. (2019), tek bir *Anopheles coluzzii* sivrisineğinden elde edilen yüksek kaliteli bir *de novo* genom birleştirmesini yapmışlardır. Çalışmada sadece 100 ng başlangıç genomik DNA'dan bir SMRTbell kütüphanesi oluşturmak için DNA kesme ve boyut seçimi olmadan değiştirilmiş bir SMRTbell kütüphane oluşturma protokolü kullanılmıştır. Örnek, chemistry 3.0 ve software v6.0 ile Sequel sisteminde dizilenmiş, ortalama olarak 20 saatlik filmlerle SMRT Hücresi başına 25 Gb dizi üretilmiştir. Ardından FALCON-Unzip algoritması ile diploid *de novo* genom birleştirmesi yapılmıştır. Elde edilen birleştirme, yüksek bitişikliğe (kontig N50 değeri 3,5 Mb) ve bütünlüğe (korunmuş genlerin %98'inden fazlası mevcut ve tam uzunluktadır) sahiptir. Ayrıca bu tek bir böcek örneğine ait genom birleştirmesi ile önceden yerleştirilememiş 667 gen, AgamP4 PEST referansındaki uygun kromozomal bölgelere yerleştirmiştir. Ek olarak genomun 1/3'ünden fazlası için maternal (anaya ait) ve paternal (babaya ait) haplotipleri çözmeyi başarmışlardır. Tek bir diploid bireyden genomun dizilenmesi ve birleştirilmesiyle, yalnızca iki haplotip elde etmişler ve birden fazla bireyin birlikte kullanıldığı çalışmalara kıyaslandığında bunun birleştirme sürecini basitleştirdiği bildirilmiştir. Bu çalışmada sunulan yöntemin 1 Gb genom boyutu başına 100 ng kadar düşük başlangıç DNA miktarlarına sahip örneklerle uygulanabileceği vurgulanmıştır (Kingan vd., 2019).

Tropikal mercan resifleri, yüz binlerce makroskopik ve bilinmeyen sayıda mikroskobik tür barındıran yüksek biyolojik çeşitlilik seviyeleri ile bilinmektedirler (Fisher vd., 2015).

Bongaerts vd. (2021), Hint-Pasifik'teki en yaygın mercanlardan biri olan *Pachyseris speciosa* için anotasyonu ile beraber bir genom çalışması gerçekleştirmişlerdir. Bu çalışmada PacBio tek molekül uzun okumalar ile yüksek bitişikliğe sahip olarak genomu 984 Mb uzunluğunda elde etmişlerdir. N50 değeri 766,6 kb olan 2.368 kontigde toplamda 39.160 protein kodlayan gen tahmin edilmiştir. Tüm genom yeniden dizilemeye dayalı demografik modelleme çalışması ile morfolojik kamuflajın (mikro ve makromorfolojik özellikler çerçevesinde), yakın zamandaki farklılaşmadan ziyade eski morfolojik stasisden (türlerin evrimsel sürecinde evrimsel değişikliğin hiç olmadığı veya çok az olduğu süreç) kaynaklandığını vurgulamışlardır (Bongaerts vd., 2021).

Sert kabuklu midye (*Mytilus coruscus*), yetiştiriciliği yapılan ekonomik açıdan önemli bir kabuklu deniz hayvanıdır. Yoğun avlanma nedeniyle midye stoklarının çoğu önemli ölçüde azalmıştır. Bu türün doğal dağılımını, genetiğini, üremesini ve yetiştirilmesini inceleme çabaları, yüksek kaliteli bir referans genomun olmaması nedeniyle engellenmiştir. Bu eksikliği gidermek için Li vd. (2020), *M. coruscus*'un hibrit yüksek kaliteli bir referans genomunu Oxford Nanopore PromethION platformu kullanarak 11,31 milyon (161,04 Gb) okuma ve sekans hatalarını doğru bir şekilde düzeltmek için de Illumina HiSeq X Ten platformunda 288,22 milyon (86,47 Gb) paired-end okuma ile birleştirmişlerdir. Uzun okumalar filtrelemeden sonra wtdbg2 kullanılarak birleştirilmiştir. Uzun okuma polishing işlemi Minimap2 ve Racon kullanılarak yapılmıştır. Son polishing Illumina kısa okumaları ile Pilon kullanılarak yapılmıştır. Elde edilen genomun toplam uzunluğu 1,90 Gb olup 898 kb'lik bir skafold N50 değeri ile 10.484 skafolddan oluşmaktadır. *M. coruscus* genom birleştirmesinin *ab initio* anotasyonunda Braker algoritması kullanılarak toplam 42.684 gen tanımlamıştır. Ayrıca 1,01 Gb (birleştirmenin %52,83'ü) tekrar elementi, 278 rRNA ve %1,64'lük yüksek bir heterozigotluk belirlenmiştir. BUSCO değerlendirmesi %91,09'u bütünlük ortaya koymuştur.

Organizmaların çevresel adaptasyonu evrimsel biyolojide yoğun çalışılan bir konu olmasına rağmen, yüksek rakım ortamına genetik adaptasyon ektotermik hayvanlarda zayıf karakterize edilmiştir. Skuamatlar muazzam ekolojik plastisite ve karyotip çeşitliliği ile en çeşitli karasal omurgalılar arasındadır ve adaptasyonun genetik izlerini araştırmak için benzersiz bir model sistemlerdir. Li vd. (2023), yaptıkları çalışmada tek moleküllü gerçek zamanlı, bağlantılı okuma, ikinci nesil dizileme ve Hi-C skafoldlama dahil olmak

üzere yeni nesil dizileme ve skafoldlama teknolojilerinin bir kombinasyonunu kullanılarak Moğol racerunner'ın (*Eremias argus*) ilk kromozom seviyesinde genom birleştirmesini oluşturmuşlardır. Elde edilen birleştirme 19 kromozom ile kontig N50 değeri 3,07 Mb, skafold N50 değeri 96,73 Mb ve genom boyutu 1,63 Gb olarak rapor edilmiş ve karşılaştırmalı genomik analizleri ile çoklu kromozom fizyon/füzyon olaylarının kertenkelelere özgü olduğu bulunmuştur. Ayrıca deniz seviyesinden ~80 ila ~2600 m yüksekliklerden toplanan 61 Moğol racerunner bireyinin genomları da çalışmada dizilenmiştir. Popülasyon genomik analizleri, yüksek rakımlara endemik ve yüksek rakımdaki popülasyonlarında güçlü seçici taramalar altında birçok yeni genomik bölgeyi ortaya çıkarmıştır. Bu genomik bölgelerdeki genlerin esas olarak enerji metabolizması ve DNA hasarı onarım yolları ile ilişkili olduğu belirtilmiştir. Son olarak kertenkelelerin yüksek rakımlarda hipoksiye karşı toleransını artırabilecek iki *PHF14* ikamesini belirlenmiş ve doğrulanmıştır (Liv d., 2023).

Doğu bahçe kertenkelesi *Calotes versicolor* (Daudin, 1802), *Calotes* cinsi, *Draconinae* alt familyası ve *Agamidae* familyasında yer alan (<http://www.reptile-database.org/>) orta ila büyük boyutlu bir kertenkeledir. *Calotes versicolor*'un böcek popülasyonlarını kontrol etmede rol oynadığı bilinmekte olup, aynı zamanda yılan ve kuşlara av ögesi olarak hizmet etmektedir (Matyot, 2004; Sudasinghe ve Somaweera, 2015). Wang vd. (2023), yaptıkları bir çalışmada Oxford Nanopore Teknolojileri (ONT) ve DNBSEQ sekanslama teknolojilerini kullanarak *Calotes versicolor*'un ilk yüksek kaliteli genomunu elde etmişlerdir. Bu birleştirilen genom, *Draconinae* alt familyasının ilk ve tek yüksek kaliteli genomudur. Genom birleştirmesi için ortalama uzunluğu 15.716 bp ve uzun okumaların N50'si 27.282 bp olan toplamda yaklaşık 103,75 gigabaz (63X kapsamda) ONT uzun okumaları üretilmiştir. Ayrıca ~222,61 GB temiz tüm genom dizileme verisi (138X kapsamda) ve gen anotasyonu için ~30,56 GB RNA-seq verisi üretilmiştir. Çalışmada *C. versicolor* genomunun 1,70 GB boyutunda olduğu tahmin edilmiş ve NextDenovo v 2.5.0, (<https://github.com/Nextomics/NextDenovo>) kullanılarak N50 değeri 91,60 Mb olan 104 kontigde toplam 1,61 GB uzunluğunda bir taslak genom birleştirmesi yapılmıştır. BUSCO analizi ile 3.354 BUSCO geninin %98'i (vertebrata_odb10) belirlenmiştir. Genomun %40,29'unu oluşturan 650,26 Mb tekrar elementi belirlenmiştir. Homoloji tabanlı, *de novo* ve RNA-seq destekli genleri birleştirilerek finalde 17.547

protein kodlayan gen tahmin edilmiştir. Ek olarak *C. versicolor* genomunda 195 miRNA, 744 tRNA, 404 rRNA ve 384 snRNA tahmin edilmiştir.

Anolis kertenkeleleri Güney, Orta ve Kuzey Amerika'nın kıtasal neotropik bölgelerinde ve Batı Hint Adaları ile Doğu Pasifik Okyanusu'ndaki adalarda dağılmış 400'den fazla küçük ila orta boy kertenkele türünden oluşmaktadır (Losos, 2009). *Anolis sagrei* en yaygın olarak yerde bulunan ve ağaçların gövdelerinin altına tünemiş orta büyüklükte böcekçil bir kertenkeledir (Williams, 1969). Geneva vd. (2022), çeşitli disiplinlerde yaygın olarak incelenen ve yüksek kaliteli bir referans genomunun eksik olduğu bir kertenkele türü olan kahverengi bir Anolis kertenkeleleri olan *Anolis sagrei* için oldukça eksiksiz bir kromozom ölçekli genom birleştirmesi sunmuşlardır. meraculous v2.2.2.5 algoritması kullanılarak yapılan genom birleştirme, HiRise v2.1.6-072ca03871cc algoritması kullanılarak Hi-C ve Chicago okumaları ile skafoldlanmıştır. Elde edilen yüksek bitişikliğe sahip genomun (N50 değeri 253,6 Mb) anotasyon bütünlüğü açısından mevcut sürüngen ve yılan genomlarının büyük çoğunluğundan iyi olduğu raporlanmıştır. BUSCO değerlendirmesi tek kopya genler için %95,3 olup *Anolis sagrei* genomunun %51,7'si tekrar bölgesi olarak tahmin edilmiştir. Braker v2.0.5 algoritması kullanılarak yapılan anotasyonda 20.033 gen tahmin edilmiştir. Bu birleştirilen genom ve popülasyon yeniden sekanslama verilerinin analizi yoluyla tekrarlayan element birikiminin tarihi incelenmiş, X kromozomu tanımlanmış ve otozomlar ile *A. sagrei*'nin cinsiyet kromozomlarına yol açan füzyonların evrimsel tarihi için bir hipotez önerilmiştir (Geneva vd., 2022).

Quasipaa spinosa, yetişkin erkeklerde bulunan dikenli göğsüyle dikkat çeken bir Asya Dicroglossidae türüdür. Hu vd. (2022), PacBio uzun okuma (276,17 Gb) ve Hi-C kullanarak ilk kromozomal seviyedeki *Q. spinosa* genomu bildirmişlerdir. Çalışmada genom boyutu 2,70 Gb olarak tahmin edilmiştir. Ayrıca 234,75 Gb BGI kısa okuması hata düzeltme için üretilmiştir. Smartdenovo (<https://github.com/ruanjue/smartdenovo>) kullanılarak birleştirilmiş genomun toplam uzunluğu 2.839.292.578 bp olup kontig N50 değeri 3,79 Mb, skafold N50 değeri 327,44 Mb'dır. Birleştirilen genom dizilerinin uzunluğunun yaklaşık %99,30'u Hi-C okumalarının yardımıyla 13 kromozoma atanmıştır. Homoloji, transkriptom verisi (15,37 Gb) ve *de novo* yöntemler ile toplamda 26.173 protein kodlayan gen tahmin edilmiş ve genlerin %95,98'i fonksiyonel olarak annotate edilmiştir. Anotasyonu yapılan genler BUSCO değerlendirmesine göre tam

omurgalı çekirdek gen setinin toplam %92,10'unu kapsamıştır. Çalışmada elde edilen bulgulara göre *Q. spinosa* dikroglossid kardeş takson *Nanorana parkeri*'den yaklaşık 41 milyon yıl önce ayrılmaya başlamıştır. Elde edilen *Q. spinosa* genomu, kurbağalarda belirli bir kromozom evrimsel tarihini temsil eden *Xenopus tropicalis* ile karşılaştırıldığında belirgin kromozomal bölünmeler ortaya çıkarmıştır. Popülasyon analizi, Çin'deki *Q. spinosa*'nın doğu ve batı genetik kümelerine ayrılabilceğini ve batı popülasyonunun doğu popülasyonundan daha fazla çeşitlilik gösterdiğini ortaya koymuştur (Hu vd., 2022).

Hymenolepis diminuta, omurgalı hayvanların ve insanların önemli patojenleri olduğu bilinen üyeleri içeren büyük bir asalak yassı kurt grubu olan Cestoda sınıfının iyi tanımlanmış bir üyesidir. Deneysel parazitolojide model organizma olarak kullanılmasına rağmen tüm genomu henüz yayınlanmamıştır. Nowak vd. (2019) Illumina paired-end, Illumina mate-pair ve MinION Oxford Nanopore uzun okumanın (5,1 GB veri) kombinasyonu ile hibrit *de novo* genom birleştirmesi yapmış ve *H. diminuta* genom birleştirmesini büyük ölçüde geliştirmişlerdir. Sonuçlar hibrit dizileme yaklaşımının yüksek kaliteli veri elde etmek için tercih edilen yöntem olduğunu göstermiştir. Final genom birleştirmesi 75 kb'lık kontig N50 boyutu ve 2,3 Mb'lık bir skafold N50 boyutu ile 177 Mb uzunluğundadır. Çalışmada en eksiksiz cestod genom birleştirmelerinden biri elde edilmiş ve 15.169 potansiyel protein kodlayan geni annotate etmişlerdir (Nowak vd., 2019).

1.11. Genom çalışmalarının tarımsal faydaları

İnsanlardaki karmaşık özelliklerle bağlantılı dizi varyasyonunun keşfinde genom çapında ilişkilendirme çalışmalarının başarısı, çiftlik hayvan türlerinde yüksek verimli SNP genotipleme deneylerine olan ilgiyi artırmış ve temel hedef QTL tespiti ile genomik seleksiyon olmuştur. Tosser-Klopp vd. (2014), keçiler için 50.000-60.000 SNP'lik bir çipi tasarlamak için yaptıkları çalışmada Uluslararası Keçi Genomu Konsorsiyumu olarak birleştirilen üç SNP keşif projesinin federasyonu aracılığıyla biyolojik ve teknik özellikleriyle birlikte bir veri tabanında depolanan keçi genomunda yaklaşık on iki milyon yüksek kaliteli SNP varyantı (Alpine, Boer, Creole, Katjang, Saanen ve Savanna ırkları için toplam 97 hayvandan) belirlemişlerdir. SNP'leri doğrulamak için on ırk ilave edilmiş ve 52.295 lokus başarılı bir şekilde genotiplendirilmiştir. Çalışmada Illumina

tasarım araçlarıyla tamamlanan bir kontig genom birleştirmesi üzerinde esas olarak tüm genom dizileme ve haritalama stratejisinin bu GoatSNP50 çipi üretmede etkili olduğu bildirilmiştir (Tosser-Klopp vd., 2014).

Tibet koyunu, Qinghai plato bölgesinde yaklaşık 2.500 yıl önce ortaya çıkmış en eski Çin koyun ırklarından biridir. Farklı coğrafi koşullara göre Tibet koyunları farklı ırklara evrilmiştir. Li vd. (2021), 5 temsili Tibet koyun ırkının tüm genomunun yeniden sekanslamasını gerçekleştirmişlerdir. Toplamda 2,70 milyar okuma (404,56 Gb) üretilmiş ve 31,36 milyon SNP belirlenmiştir. Yapılan karşılaştırmalı genomik analiz, yakın bir genetik ilişki ile bu ırkların farklı kladlara ayrılacaklarını göstermiştir. Ortak selektif bölgelere sahip bazı genler (*GHR*, *BMP15* ve *CPLANE1*) daha yüksek rakımda yaşayan farklı ırklarda hipoksik adaptasyon için zenginleştirilmiştir. Ek olarak *BSND*, *USP24*, *NCAPG* ve *LCORL* gibi genlerde fiziksel özellikler ile ilgili ırk spesifik selektif bölgeler bulunmuştur (Li vd., 2021).

İstenen fenotiplerin altında yatan genetik varyasyonların belirlenmesi, mevcut çiftlik hayvanları genetik araştırmalarının ana zorluklarından biridir. Yüksek verimli transkriptom dizileme, çeşitli türlerdeki zengin genetik varyantları çözmenin etkili bir yolu olarak kabul edilmektedir. Ma vd. (2018), Çin'de nesli tükenmekte olan bir koyun ırkı olan yağlı kuyruklu Lanzhou koyununun bu özelliğinin altında yatan genetik mekanizmayı araştırmak için yağlı ve ince kuyruklu üç koyun ırkından alınan doku örneklerinde (longissimus dorsi kası, perinefrik yağ ve kuyruk yağı) RNA-seq kullanmışlardır. Lanzhou koyun ırkındaki yağlı kuyruktan sorumlu olan genleri ve genetik ağları tanımlamak için dokuz ayrı koyunda toplamda 7.122.920 SNP ve 901.518 indel tespit edilmiş, bunlardan 606.952 SNP (%8,5) ve 77.633 indel (%8,6) koyunlardaki yağ özellikleriyle ilişkili QTL ile örtüşmüştür. Lanzhou koyununa spesifik 26.613 SNP bulunmuş ve yağlı kuyruklu başka bir koyun ırkında bildirilen aynı genomik pozisyonda bulunan 44 SNP tespit edilmiştir. 33 SNP'nin, koyun kuyruklarında yağ birikimi için güçlü bir aday genomik bölge olduğu bildirilen kromozom 3 bölgesine (39,58-40,91 Mb) seçici olarak dağıldığı bildirilmiş ve ayrıca üç genin (*CREB1*, *WDR92* ve *ETAA1*) yağlı kuyruk gelişimi ile ilişkili olabileceğini ileri sürülmüştür (Ma vd., 2018).

Evrimsel tarihi ve uyum sürecini anlamak hem eski hem de modern genomik verilerden elde edebileceğimiz bilgiye bağlıdır. Antik ve modern keçi örneklerinden elde edilen tüm

genom dizileme verilerinin mevcut olmasıyla, bu önemli kaynakların verimli bir şekilde yeniden kullanılmasını sağlayan kullanıcı dostu bir veri tabanına ihtiyaç duyulmuştur. Fu vd. (2021), keçi genom varyasyonlarını içeren kapsamlı bir veri tabanı olan GGVD'yi (Goat Genome Variation Database) (<http://animal.nwsuaf.edu.cn/GoatVar>) oluşturmak için 208 modern evcil keçi, 24 yaban keçisi (bezoar), 46 yabani dağ keçisi ve 82 antik keçinin genom verilerini kullanmışlardır. GGVD, toplam ~41,44 milyon SNP, ~5,14 milyon indel, 6.193 seçilmiş lokus ve 112 introgresyon bölgesini barındırmaktadır. Görselleştirme için oluşturulan araçlara ek olarak sekans özelliklerine erişimi kolaylaştırmak için UCSC Genom Tarayıcısı, BLAT, BLAST, LiftOver ve pcadapt araçlarını da GGVD'ye entegre etmişlerdir (Fu vd., 2021).

Sığırlar, günümüzden yaklaşık 10.000 yıl önce evcilleştirilmiş olup insanlığa et, süt, deri ve iş gücü sağlamıştır. Et-süt kombine amaçlı bir ırk olan Simental sığırı, dünyada en yaygın yetiştiriciliği yapılan sığır ırklarından biridir. Tek nükleotid polimorfizmleri (SNP'ler) ve kopya sayısı varyasyonları (CNV'ler), evrimsel ve fenotipik çeşitlilikte çok önemli roller oynayan iki ana genomik varyanttır. CNV'ler, 50 bp ile 5 Mb arasında değişmekte olup SNP'lerle karşılaştırıldığında, CNV'nin gen dozajı, kodlama dizisi ve uzun menzilli genlerin düzenlenmesi yoluyla fonksiyon, fenotip ve evrim üzerinde daha büyük bir etkisi vardır (Beckmann vd., 2007; Henrichsen vd., 2009; Mills vd., 2011). Sun vd. (2023) yaptıkları bir çalışmada 30 Simental sığırını ~25X ortalama genom kapsamı ile ~2,17 milyar çift uçlu okuma üreterek sekanslamış ve tek moleküllü uzun okuma verilerini kullanarak yüksek sperm hareketliliği (HSM) ve zayıf sperm hareketliliği (PSM) olan Simental boğaların genetik varyasyonlarını keşfetmek için kapsamlı bir analiz gerçekleştirmişlerdir. Çalışmada Simental boğalarda toplamda ~15 milyon SNP ve 2.944 CNV bölgesi (CNVR) tespit edilmiştir. Bir dizi pozitif seçilmiş gen (PSG'ler) ve CNVR'lerin bağışıklık, kas gelişimi, üreme vb. içeren kantitatif özellik lokusları (QTL'ler) ile örtüştüğü bulunmuştur. Ek olarak önemli ekonomik özellikleri iyileştirmek için yapay üreme ile ilgili olabilecek iki yeni varyant tespit edilmiştir. Ayrıca, erkek fertilitesi ile fonksiyonel olarak ilişkili bir dizi gen ve yolak tanımlanmıştır. Dikkat çekici bir şekilde SPAG16'daki bir CNV'nin tüm zayıf sperm motiliteli boğalarda ve yüksek sperm motiliteli boğaların yarısında tamamen silindiği belirtilmiş ve bu da boğa doğurganlığında çok önemli bir rol oynayabileceği vurgulanmıştır (Sun vd., 2023).

İslah veya genetik iyileştirme, evcilleştirmeyi izleyen yapay seçilim sürecini ifade eder ve bu haliyle modern tarım ve hayvansal üretim üzerinde büyük bir etkisi olmuştur. İyileştirme genellikle ekonomik performansı büyük ölçüde etkileyen özelliklere odaklanır. Bu nedenle iyileştirmenin altında yatan genetik temeli anlamak, ekonomik özellikleri kontrol eden genlerin tanımlanmasına katkıda bulunacak ve gelecekteki mahsul ve hayvan yetiştiriciliğini kolaylaştıracaktır. İpekböceği (*Bombyx mori*), ekonomik açıdan önemli bir böcektir. İpek gelişmiş bir tekstil malzemesidir. İpekböcekçiliği başta Çin ve Japonya olmak üzere birçok ülkenin ekonomik ve kültürel gelişimine önemli katkılar sağlamıştır (Fang vd., 2015). İpekböceği iki aşamalı, ardışık bir evrimden, evcilleştirmeden ve ardından yapay seçilim yoluyla daha fazla genetik iyileştirmeden geçmiştir. İpekböceği benzersiz, tamamen evcilleştirilmiş, ekonomik açıdan önemli bir omurgasızdır. Genetik iyileştirme, ipeğe bağlı özellikler açısından ipekböceği üzerinde büyük bir etkiye sahip olmuştur. Li vd. (2021), ipekböceği ıslahında/iyileştirmesinde güçlü seçim altındaki genomik bölgeleri belirlemek için yerel ve genetik olarak geliştirilmiş ipekböceği hatları üzerinde tüm genomik dizileme yapmışlardır. Toplamda 27,13 Gb veri üretilmiş ve haritalama sonuçlarına dayanarak varyantlar iki havuz olacak şekilde toplanmıştır. Yerel ve geliştirilmiş havuz verilerinde 5.482.279 SNP belirlenmiş, bunların 4.962.777'si yerel havuzlarda ayrışırken, geliştirilmiş havuz 4.631.701 SNP içermiştir. Bu SNP'ler arasında her iki havuzda da 4.112.199 SNP tanımlanmıştır. Benzer şekilde 1.601.653 SNP, iki havuzda birlikte ayrılmış olan 1.626.197 indel tespit edilmiştir. SNP'lerin anotasyonu, %71,16'sının genler arası bölgede bulunduğunu, sadece %0,62'sinin gen kodlama bölgesinde bulunduğunu göstermiş, çoğu SNP'nin gen işlevi üzerinde hiçbir etkisi olmadığı rapor edilmiştir. Tarama tanımlaması için kapsama derinliği 5'e eşit veya daha büyük olan SNP'ler filtrelenmiş ve toplamda 3.638.723 SNP elde edilmiştir. “Genomic-wide selective sweeping” analizi ile güçlü seçim sinyallerine sahip 24 genomik bölge belirlenmiştir. Bunlardan sekizi ipekböceği yetiştiriciliğinin altında yatan 13 aday geni içermektedir. Çalışmada bu genlerden altısı, nöral sinyal yanıtıyla ilgili fonksiyonlarla annotate edilmiştir. Çalışmada elde edilen sonuçların iyileşmenin ipekböceğinin sinir sistemini etkilemiş olabileceği öne sürülmüştür (Li vd., 2021).

Regülatör SNP'ler (rSNPs), promotör bölgelerde yer alan ve bağlanma bölgelerine transkripsiyon faktörlerinin bağlanma afinitesini değiştirerek gen ekspresyonunu

değiřtirmesi bakımından yüksek potansiyele sahip SNP'lerdir. Bu rSNP'ler spesifik özellik ve hastalıklardaki nedenselliklerinden dolayı yaşam bilimlerinde oldukça önemlidirler. Transkripsiyon faktörleri (TFs), promotörler gibi düzenleyici bölgelerde transkripsiyon faktörü bağlanma bölgesi (TFBSs) olarak bilinen kısa DNA motiflerine spesifik olarak bağlanarak transkripsiyonel gen regülasyonunu yönetirler. Kless vd. (2021), sığır, domuz, tavuk, koyun, at, keçi ve köpek gibi tarımsal öneme sahip yedi evcil hayvan türünün ilk rSNP verilerini içeren agReg-SNPdb veri tabanını oluşturmuş ve web arayüzü ile kullanışlı hale getirmişlerdir (<https://azifi.tz.agrar.uni-goettingen.de/agreg-snpdb/>). Çalışmada rSNP'leri belirlemek için biyoinformatik bir hat kurulmuş ve varsayılan TF'lerin bağlanma afinitesini etkileyen ve TFBS'lerde lokalize olan toplamda 10.623.512 rSNP belirlenmiştir (Kless vd., 2021).

En önemli tarım hayvanlarından biri olan koyun yün, deri ve et sağlayan ekonomik değeri yüksek bir hayvandır. Koyunlar eti için yetiştirildiğinde kaburga sayısı, koyun endüstrisinde önemli bir ekonomik özelliktir. Ancak koyunlarda kaburga sayısını düzenleyen genetik mekanizmalar tam olarak anlaşılamamıştır. Çin'de prolifik bir yerli ırk olan Hu koyunu erken büyüme ve gelişme, erken cinsel olgunluk ve yüksek üreme oranı gibi özelliklerine sahiptir. Zhang vd. (2021) yaptıkları bir çalışmada koyunlarda kaburga sayısındaki artışı etkileyen önemli aday genlerin belirlenmesini amaçlamışlardır. Çalışmada artmış kaburga sayısına (R14) sahip 36 Hu koyunun ve normal kaburga sayısına (R13) sahip 36 koyunun tüm genom yeniden sekanslamasını gerçekleştirilmiştir. Hayvan başına ortalama kapsam 8,63X olmak üzere toplam 1825,83 Gb ham veri üretmişlerdir. Üç yöntem (fiksasyon indeksi (F_{ST}), Fisher'in kesin testi ve Ki-kare testi) kullanılarak yapılan analiz, 206 geni temsil eden üç yöntemin sonuçları arasında 219 tek nükleotid polimorfizm bölgesinin örtüştüğünü göstermiştir. GO ve KEGG analizleri, genlerin temel olarak gelişim sürecinin düzenlenmesi, inorganik anyon taşınması, hücresel biyosentetik süreç, sıkı bağlantı, oksitosin sinyal yolu ve aritmojenik sağ ventriküler kardiyomiyopati ile ilişkili olduğunu göstermiştir. Genotipleme ve ilişkilendirme analizi için önemli yollara göre dört mutasyon seçilmiş ve üç sinonim mutasyonun kaburga sayısı ile önemli ölçüde ilişkili olduğu bulunmuştur. Ayrıca *CPOX*, *KCNH1* ve *CPQ* genlerinin Hu koyunlarında kaburga sayısı üzerinde birleşik bir etkiye sahip olduğu ortaya çıkarılmıştır (Zhang vd., 2021).

Sütçü keçiler, özellikle gelişmekte olan ülkeler için ekonomik kalkınmada önemli rol oynayan en eski süt hayvanlarından biridir. Bununla birlikte süt keçilerinin spesifik özellikleri üzerine çok az çalışma yapılmıştır. Xiong vd. (2022), 48 keçinin genomunu dizileyerek ve veri tabanlarındaki 41 keçi genom verisini kullanarak 89 keçinin tüm genom verilerini toplamışlardır. Çalışmada beş sütçü keçi ırkı (Saanen, Nubian, Alpine, Toggenburg ve Guanzhong; n = 24, 15, 11, 6, 6) ve üç sütçü olmayan keçi ırkı (Guishan keçisi, Longlin keçisi, Yunshang kara keçisi; n = 6, 15, 6) kullanılmıştır. Sütçü keçiler ve sütçü olmayan keçilerin genomlarının karşılaştırılması yoluyla sütçü keçilerin genetik çeşitliliği ve seleksiyon özellikleri analiz edilmiştir. Sonuçlar sekiz keçinin Avrupa, Afrika ve Çin yerli keçi popülasyonlarının üç alt grubuna ayrılabilceğini ve ayrıca Avustralya Nubian, Toggenburg ve Avustralya Alplerinin en yüksek bağlantı dengesizliğine, en düşük nükleotit çeşitliliğine ve daha yüksek bir akrabalık katsayısına sahip olduğunu göstermiştir. Çalışmada bu sonuçların güçlü bir şekilde yapay olarak seleksiyona maruz kaldıklarını işaret ettiği vurgulanmıştır. Ek olarak sütçü keçilerin özgülüğü ile ilgili birkaç aday gen (özellikle süt üretim özellikleri (*GHR*, *DGAT2*, *ELF5*, *GLYCAM1*, *ACSBG2*, *ACSS2*), üreme özellikleri (*TSHR*, *TSHB*, *PTGS2*, *ESR2*), bağışıklık ile ilişkili genler (*JAK1*, *POU2F2*, *LRRC66*)) belirlenmiştir (Xiong vd., 2023).

2. BÖLÜM

YÖNTEM VE MATERYAL

2.1. Laboratuvar çalışmaları

Tez çalışması kapsamında laboratuvar çalışmaları Erciyes Üniversitesi Ziraat Fakültesi Tarımsal Biyoteknoloji Bölümü'nde ve Erciyes Üniversitesi Genom ve Kök Hücre Merkezi'nde (GenKök) yürütülmüştür.

2.1.1. Hayvan materyalinin temini

Çalışmada tüm genomunun dizilenmesi için kullanılan Ankara keçisine (Şekil 5) ait kas dokusu mezbaha materyali olarak temin edilmiştir. Çalışmaya dahil olan bir diğer hayvan olan yaban keçinin kas dokusu ise doğada ölü olarak bulunan hayvandan alınmıştır. Elde edilen dokular için 01/12/2021 tarihli ve 21/254 sayılı Erciyes Üniversitesi Hayvan Deneyleri Yerel Etik Kurulu (EÜHADYЕК) raporunda HADYЕК yönergesinin 8. maddesi K fıkrasının 2. bendinde belirtildiği şekilde “Ölü hayvan veya dokusu, mezbaha materyalleri, atık fetuslar ile yapılan prosedürler” kapsamında değerlendirilmiş ve HADYЕК iznine tabi olmadığı yönündeki oybirliği kararı esas alınmıştır.



Şekil 5. Çalışmada kullanılan Ankara keçisi

2.1.2. Illumina platformu için genomik DNA izolasyonları

Elde edilen örneklerden Illumina kısa okuma platformu sekanslama kütüphanesinin hazırlanması için genomik DNA (gDNA) izolasyonu QIAGEN DNeasy Blood & Tissue Kit (Cat. No. 69504) kullanılarak, üreticinin protokolü doğrultusunda gerçekleştirilmiştir. gDNA'lar %0,8'lik agaroz jelde koşturulmuş ve hem spektrofotometrik hem de Qubit Florometrik yöntem kullanılarak saflık, kalite ve miktar yönünden değerlendirilmiştir. Qubit Florometrik yöntem için Invitrogen BR assay kit (Cat. No. Q33265) kullanılarak üreticinin protokolü takip edilerek gerçekleştirilmiştir.

2.1.3. Ankara keçisinde Oxford Nanopore platformu için yüksek molekül ağırlıklı genomik DNA'nın elde edilmesi

Ankara keçisi kas dokusu örneğinden yüksek molekül ağırlıklı genomik DNA (HMW-gDNA), Monarch HMW DNA izolasyon kiti (Cat. No. T3060S) kullanılarak üreticinin talimatları doğrultusunda izole edilmiştir. HMW gDNA izolasyonu, her bir sekanslama kütüphanesinin hazırlanmasında yeniden tekrarlanmıştır. İzole edilen HMW gDNA'lar %0,8'lik agaroz jelde koşturulmuş ve ek olarak hem spektrofotometrik hem de Qubit Florometrik yöntem kullanılarak saflık, kalite ve miktar yönünden değerlendirilmiştir.

2.1.4. Illumina platformunda sekanslama kütüphanelerinin kurulması

Qiagen DNeasy Blood & Tissue Kit ile elde edilen Ankara keçisi ile yaban keçisi gDNA'ları kullanılarak ve Illumina DNA PCR-Free Library Prep Standard Input protokolü takip edilerek kısa okuma sekanslama kütüphaneleri kurulmuştur. Sekanslama kütüphanesinin kurulmasında kullanılan protokol; kısaca Genomik DNA'nın tagmentasyonu, tagmentasyon sonrası temizlik, indeks ligasyonu, kütüphanelerin temizlenmesi, kuantifikasyon ve kütüphanelerin havuzlanması ve normalizasyon şeklinde olup son olarak cihaza yüklenmesidir.

2.1.5. Ankara keçisi için Oxford Nanopore platformu sekanslama kütüphanesinin hazırlanması

Elde edilen gDNA'lar, Oxford Nanopore MinION uzun okuma platformunda dizilenmesi için 100-300 fmol olacak şekilde normalize edilmiştir. Daha sonra konsantrasyonu ayarlanan gDNA'lar ile Oxford Nanopore tarafından geliştirilen SQK-LSK 110 ligasyona dayalı genomik DNA dizileme protokolü takip edilerek uzun okuma sekanslama kütüphanesi kurulmuştur. Uzun okuma sekanslama kütüphanelerinin hazırlanma basamakları; kısaca gDNA'nın end-prep enzimatik reaksiyonu, kütüphanenin AMPure XP beads ile saflaştırılması, Qubit florometre kullanılarak kütüphane konsantrasyonunun belirlenmesi, adaptör ligasyonu, temizlik aşamaları ile birlikte AMPure XP beads kullanılarak tekrardan saflaştırılması ve Qubit ile konsantrasyonunun belirlenmesi şeklindedir.

Ligasyon aşamasında Long fragment buffer (LFB) kullanılmış ve >3kb DNA parçaları ile kütüphane kurulmuştur. Son olarak elde edilen kütüphane 5-50 fmol konsantrasyon olarak cihaza yüklenmiştir. Sekans kütüphanelerinin hazırlanma aşaması her bir yükleme için tekrarlanmış ve toplamda 4 kütüphane kurularak okumalar üretilmiştir.

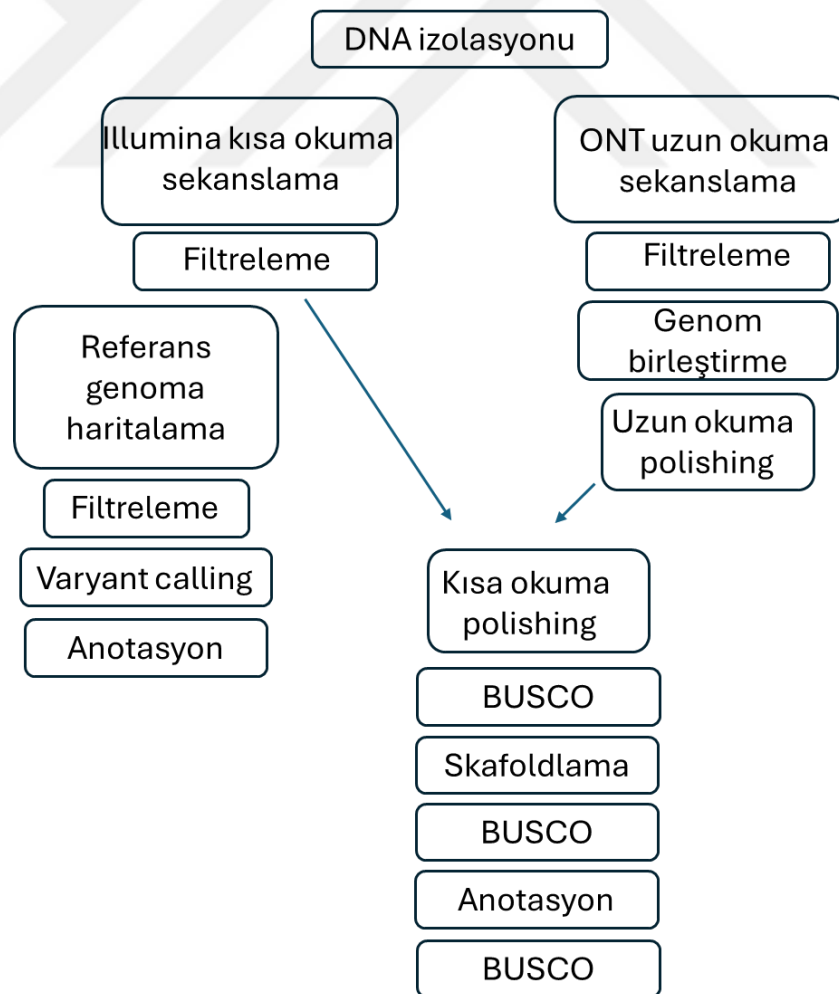
2.1.6. Oxford Nanopore MinION platformunun çalıştırılması

Kütüphane hazırlığı tamamlanan her bir örnek, SQK-LSK 110 protokolü takip edilerek üreticinin talimatları doğrultusunda yükleme bufferları eklenerek hazır hale getirilmiş ve sonra sekanslama kütüphanesi Oxford Nanopore MinION R9.4.1 flowcell (FLO-MIN106)'e 75 µL hacimde yüklenerek 72 saate kadar çalıştırılmıştır. MinION cihazı

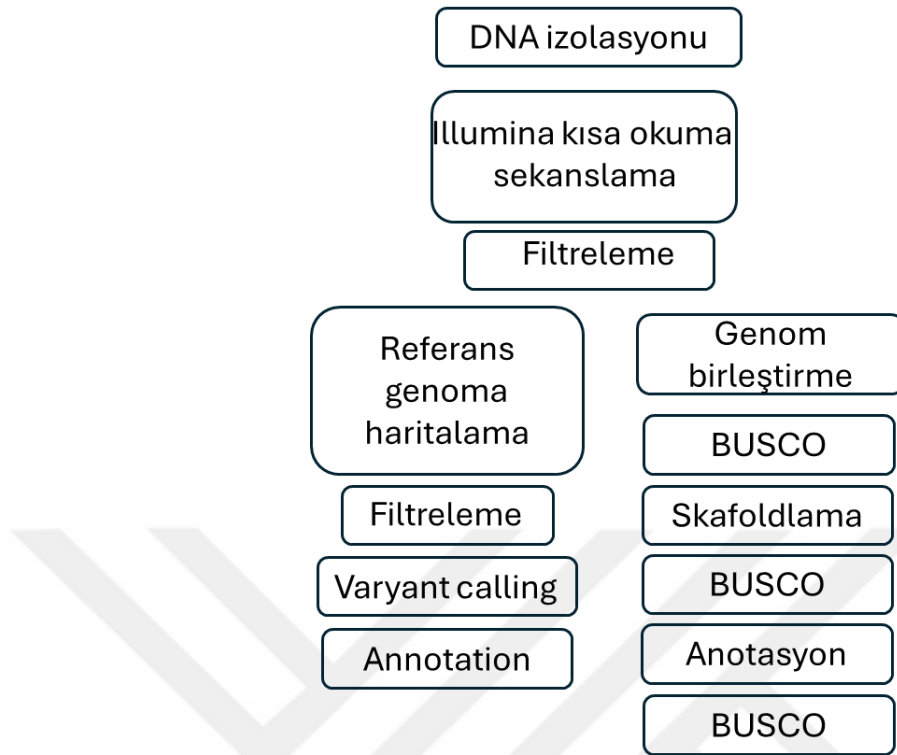
toplamda 4 kere, 4 adet flowcell kullanılarak çalıştırılmıştır. MinION ile üretilen fast5 verileri ilk olarak MinKNOW programı (<https://community.nanoporetech.com/downloads>) ile “fast basecalling” modunda basecalling yapılmış ve MinKNOW programında yapılan basecalling işleminde Q skoru 8’in altında olan okumalar, hatalı okuma olarak sınıflandırılmıştır.

2.2. Biyoinformatik analizler

Tez çalışması kapsamında yapılan biyoinformatik analizler Erciyes Üniversitesi Bilgi İşlem Daire Başkanlığı bünyesindeki sunucular ve TÜBİTAK ULAKBİM, Yüksek Başarım ve Grid Hesaplama Merkezi’nde (TRUBA kaynaklarında) gerçekleştirilmiştir. Ankara keçisi ve yaban keçisi iş akışları Şekil 6 ve 7’de verilmiştir.



Şekil 6. Ankara keçisi genomunun dizilenmesinde kullanılan iş akışı



Şekil 7. Yaban keçi genomunun dizilenmesinde kullanılan iş akışı

2.2.1. Elde edilen Ankara keçi ve Yaban keçi kısa okumaların filtrelenmesi

Ankara keçisine ait ~300 Gb (100X genom kapsamı) Illumina NovaSeq çift uçlu kısa okumaları ile yaban keçisine ait ~210 GB (70X genom kapsamı) Illumina kısa okumaları Geneious Prime programı içerisindeki trimleme aracı olan BBduk (<https://www.geneious.com>) kullanılarak (minimum okuma >50 bp; kalite skoru >30) filtrelenmiştir.

2.2.2. Elde edilen Ankara keçi uzun okumaların filtrelenmesi

Oxford Nanopore MinION platformu kullanılarak elde edilen yaklaşık 314,49 GB fast5 dosyaları Guppy v6.4.6+ae70e8f (<https://community.nanoporetech.com/downloads>) yazılımı ile “high accuracy” ve CPU modları ile Erciyes Üniversitesi Bilgi İşlem Daire Başkanlığı altyapısındaki sunucular kullanılarak “basecalling” yapılmıştır. Basecalling işleminden sonra elde edilen yaklaşık 61,8 GB fastq (~32 GB fasta) veri dosyalarında adaptörlerin kaldırılması için Porechop v0.2.4 (<https://github.com/rrwick/Porechop>), 1.000 bp’den kısa okumaların atılması ve <Q10 düşük kalite skoruna sahip okumaların çıkarılması için de NanoFilt v2.6.0 (De Coster vd., 2018) yazılımları kullanılmıştır.

2.2.3. Ankara keçisi genomunun hibrit yöntemlerle birleştirilmesi

Filtrelenmiş Oxford Nanopore uzun okumaları TÜBİTAK TRUBA HPC platformunda Flye v2.8.1 (Kolmogorov vd., 2019) algoritması kullanılarak –nano-raw modunda *de novo* olarak birleştirilmiştir. Elde edilen taslak genom ilk olarak Medaka v1.6.1 (<https://github.com/nanoporetech/medaka>) ile Oxford Nanopore uzun okumalarının kullanılmasıyla “polish” edilmiş ve daha sonra Polca (MaSuRCA v4.0.9) (Zimin ve Salzberg, 2020) kullanılarak Illumina okumaları ile “polish” edilmiştir. Genom bütünlüğü BUSCO (Manni vd., 2021) ile değerlendirilmiştir.

2.2.4. Yaban keçisi genomunun Illumina kısa okumaları ile birleştirilmesi

Yaban keçisi kısa okumalarının birleştirilmesi MaSuRCA v4.0.9 (Zimin ve Salzberg, 2020) algoritması kullanılarak TÜBİTAK TRUBA HPC platformunda *de novo* olarak yapılmıştır. Genom bütünlüğü BUSCO (Manni vd., 2021) ile değerlendirilmiştir.

2.2.5. Ankara keçisi ve yaban keçisi genomlarının skafoldlanması

Kontig seviyesinde elde edilen Ankara keçisi ve yaban keçisi genomları RagTag (Alonge vd., 2022) algoritması kullanılarak referans genom destekli olarak skafoldlanmıştır. Skafoldlama sonrası genomlar BUSCO (Manni vd., 2021) ile tekrar değerlendirilmiştir. Elde edilen genom birleştirmelerinin metrikleri QUAST (Mikheenko vd., 2018) ile değerlendirilmiştir. Ankara keçisi ve yaban keçisi genomlarının görselleştirilmesi için blobtools (Challis vd., 2020) kullanılmıştır.

2.2.6. Ankara keçisi ve yaban keçisi komple mitokondriyal DNA’larının elde edilmesi

Keçi mitogenomlarının elde edilmesinde izole edilen gDNA’lar kullanılarak çakışan iki fragment şeklinde AL1_2024L - BH1_13002H (yaklaşık 11.000 baz ampikon uzunluğu) ve LuLu_12690L - LuLu_2503H (yaklaşık 6.500 baz ampikon uzunluğu) primerleri ile Long-Range PCR metodu kullanılarak elde edilmiştir. PCR’da “NEB LongAmp® Taq 2X Master Mix (M0287S, NEB)” enzimi kullanılmış olup, PCR bileşenleri ve döngü şartları sırasıyla, 1X LongAmp Taq 2X Master Mix (25 µL), her bir primerden 0.5 µM, ~30 ng gDNA ve bir döngü 94 °C’de 1 dakika ilk denatürasyon aşamasını takip eden 30 döngü 94 °C’de 30 saniye denatürasyon, I. yön primer çifti için 58 °C’de 30 saniye primer

bağlanma, II. yön primer çifti için 55 °C’de 30 saniye primer bağlanma, I. yön primer çifti için 65 °C’de 11 dakika uzama, II. yön primer çifti için 7 dakika uzama aşamalarını takip eden bir döngü 65 °C’de 10 dakika son uzama aşamaları ile gerçekleştirilmiştir. Daha sonra elde edilen amplikonlar “Quibt dsDNA HS Kit” kullanılarak miktarları belirlenmiş ve seyreltilerek sekanslama kütüphaneleri kurulmuş ve Illumina MiSeq platformunda MiSeq Reagent Kit v3 kullanılarak dizilenmiştir.

Tablo 1. Ankara keçisi ve yaban keçisi mitogenomlarının amplifiye edilmesi için Long- Range PCR’da kullanılan primerler

Primer Seti	Primer Adı	Primer Sekansı	Tm derecesi (C°)	Uzama Zamanı (dakika)	Amplikon boyutu (bp)
1	AL1_2024L	5'-GACCGTGCAAAGGTAGCATAATC-3'	58	11	~11.000
	BH1_13002H	5'-TAATTAAGGGCTCAGGCGTTG-3'			
2	LuLu_12690L	5'-TCGGACTAATAATTGTAACCATCG-3'	55	7	~6.500
	LuLu_2503H	5'-CTCAGATCACGTAGGACTTTAATC-3'			

MiSeq platformu kullanılarak elde edilen ham okumalar (Ankara keçisi için 481.398; yaban keçisi için 471.048 okuma), kalite kontrol, filtreleme ve daha sonraki analizler için Geneious Prime programına (<https://www.geneious.com>) yüklenmiştir. İlk olarak okumalar eşleştirilmiş ve okumalardan adaptörlerin çıkarılması, düşük kaliteli (kalite skoru <25) ve kısa okumaların (<50 bp) atılması için Geneious Prime programı içerisindeki trimleme aracı BBDuk kullanılmıştır.

Filtrelenen Ankara keçisi kısa okumaları ve filtrelenmiş yaban keçisi kısa okumaları, NCBI GenBank’den indirilen domestik ve yaban keçisi referans mitogenomlarına (*Capra hircus* erişim no: NC_005044, *Capra aegagrus* erişim no: NC_028161) Geneious Prime programındaki geneious mapping modülü kullanılarak haritalanması yoluyla elde edilmiştir. Daha sonra elde edilen mitogenomlar GetOrganelle (Jin vd., 2020) ile doğrulanmış ve hem referans mitogenomlar hem de MITOS2 webserver (Donath vd., 2019) ile anotasyonları yapılmıştır. tRNA’lar VARNA (Darty vd., 2009) ve dairesel genomlar CGview (Grant vd., 2023) ile çizilmiştir.

PCR ile çoğaltılarak elde edilen mitogenomların doğrulanması ve kontrol edilmesi için keçilerin tüm genom verileri ile referans mitogenomlar kullanılarak tekrar birleştirme

yapılmış ve birbirleri ile Geneious Prime programı içerisindeki MAFFT (Katoh ve Standley, 2016) hizalama aracı ile karşılaştırılmıştır.

2.2.7. Mitogenom bazlı filogenetik analizler

Komple mtDNA bazlı filogenetik ağacın oluşturulması için NCBI GenBank'dan *Capra* cinsi içerisindeki mitogenomlar ve dış grup olarak *Bos taurus* (Erişim no: NC_006853) ve *Bos primigenius* (Erişim no: NC_013996) verileri indirilmiştir. Ayrıca henüz yayınlanmamış mevcut keçi mitogenom verileri de (Al-Abbasi, 2019) veri setine dahil edilmiştir (Tablo 2). D-loop bölgeleri çıkarılarak oluşturulan veri seti Geneious Prime'daki MAFFT (Katoh ve Standley, 2016) kullanılarak hizalanmıştır. Düşük hizalanan bölgeler Gblocks server (Castresana, 2000) kullanılarak hizalamadan çıkarılmıştır. Filogenetik ağaç, jModelTest2 v.2.1.10 (Posada, 2008) ile belirlenen optimal nükleotid ikame modeli (1.000 bootstrap ve GTR +I +G modeli) ile MEGA 11'deki (Tamura vd., 2021) maksimum likelihood (ML) yöntemi kullanılarak oluşturulmuştur.

Tablo 2. Filogenetik analizlerde kullanılan keçi mitogenomları

Koleksiyon No	Tür / Irk	Cinsiyet	Lokasyon	Mitogenom Uzunluğu (bp)	Referans
23CH1760-TR	<i>Capra hircus</i> (Türk Saanen keçisi)	?	Lüleburgaz, Kırklareli, Türkiye	16.640	İbiş, Yayınlanmamış veri
24CH1Ad-TR	<i>Capra hircus</i> (Kıl keçisi)	?	Adana, Türkiye	16.643	İbiş, Yayınlanmamış veri
25CH3Ad-TR	<i>Capra hircus</i> (Kıl keçisi)	?	Adana, Türkiye	16.640	İbiş, Yayınlanmamış veri
26CH5Ad-TR	<i>Capra hircus</i> (Kıl keçisi)	?	Adana, Türkiye	16.641	İbiş, Yayınlanmamış veri
28CHSn2-TR	<i>Capra hircus</i> (Türk Saanen keçisi)	?	Yozgat, Türkiye	16.640	İbiş, Yayınlanmamış veri
1_Angora-TR	<i>Capra hircus</i> (Ankara keçisi)	E	Ankara, Türkiye	16.644	Bu çalışma
3_Hair-TR	<i>Capra hircus</i> (Kıl keçisi)	E	Kayseri, Türkiye	16.651	İbiş, Yayınlanmamış veri
2_Kilis-TR	<i>Capra hircus</i> (Kilis goat)	E	Kayseri, Türkiye	16.649	İbiş, Yayınlanmamış veri
29_IRQ	<i>Capra hircus</i> (Lokal keçi)	D	Samarra, AL-Mamel, Irak	16.642	Al-Abbasi, 2019

30_IRQ	<i>Capra hircus</i> (Belirsiz)	E	Hillah, Babylon, Irak	16.641	Al-Abbasi, 2019
31_IRQ	<i>Capra hircus</i> (Lokal keçi)	D	Hit, AL- Anbar, Irak	16.640	Al-Abbasi, 2019
32_IRQ	<i>Capra hircus</i> (Kıbrıs keçisi)	D	Hit, AL- Anbar, Irak	16.640	Al-Abbasi, 2019
33_IRQ	<i>Capra hircus</i> (Shami keçisi)	D	Hit, AL- Anbar, Irak	16.640	Al-Abbasi, 2019
34_IRQ	<i>Capra hircus</i> (Kıbrıs keçisi)	E	Hit, AL- Anbar, Irak	16.640	Al-Abbasi, 2019
35_IRQ	<i>Capra hircus</i> (Shami keçisi)	D	Hit, AL- Anbar, Irak	16.639	Al-Abbasi, 2019
37_IRQ	<i>Capra hircus</i> (Lokal keçi)	D	Jamiah, Samarra, Irak	16.641	Al-Abbasi, 2019
38_IRQ	<i>Capra hircus</i> (Lokal keçi)	D	Jamiah, Samarra, Irak	16.640	Al-Abbasi, 2019
39_IRQ	<i>Capra hircus</i> (Lokal keçi)	E	Hit, AL- Anbar, Irak	16.641	Al-Abbasi, 2019
91_IRQ	<i>Capra hircus</i> (Belirsiz)	E	Hillah, Babylon, Irak	16.640	Al-Abbasi, 2019
11_IRQ	<i>Capra hircus</i> (Lokal keçi)	D	Hillah, Babylon, Irak	16.640	Al-Abbasi, 2019
30_2_IRQ	<i>Capra hircus</i> (Kıbrıs keçisi)	E	Hadi, Samarra, Irak	16.640	Al-Abbasi, 2019
50_IRQ	<i>Capra hircus</i> (Belirsiz)	E	Hillah, Babylon, Irak	16.639	Al-Abbasi, 2019
71_IRQ	<i>Capra hircus</i> (Belirsiz)	D	Hillah, Babylon, Irak	16.641	Al-Abbasi, 2019
51_IRQ	<i>Capra hircus</i> (Belirsiz)	E	Hillah, Babylon, Irak	16.640	Al-Abbasi, 2019
70_IRQ	<i>Capra hircus</i> (Belirsiz)	D	Hillah, Babylon, Irak	16.639	Al-Abbasi, 2019
90_IRQ	<i>Capra hircus</i> (Belirsiz)	D	Hillah, Babylon, Irak	16.642	Al-Abbasi, 2019
1573_CaCa_TR	<i>Capra aegagrus</i> (Yaban keçisi)	E	Akseki, Antalya, Türkiye	16.623	İbiş, Yayınlanmamış veri
1732_CaCa_TR	<i>Capra aegagrus</i> (Yaban keçisi)	?	Derebucak, Konya, Türkiye	16.622	Bu çalışma
1763_CaAe_TR	<i>Capra aegagrus</i> (Yaban keçisi)	?	Yusufeli, Artvin, Türkiye	16.639	İbiş, Yayınlanmamış veri
1906_RubCa_TR	<i>Rupicapra rupicapra</i> (Kuzey dağ keçisi)	?	Artvin, Türkiye	16.434	İbiş, Yayınlanmamış veri

2.2.8. Keçi genomlarının anotasyonu

Keçi genomlarının annotate edilmesinde ilk olarak genomlarda RepeatMasker (<http://www.repeatmasker.org>) kullanılarak tekrar bölgeleri belirlenmiş ve maskelenmiştir. Daha sonra gen anotasyonunun gerçekleştirilmesi için homoloji tabanlı

yöntem kullanılmıştır. Genomlardaki genlerin tahmini GeMoMa (Keilwagen vd., 2018) kullanılarak yapılmıştır. Homoloji temelli yaklaşımda evcil keçi (*Capra hircus*) referans genomu ve anotasyonu kullanılmıştır. tRNA'ların tespitinde tRNAscan-SE 2 (Chan vd., 2021) ve rRNA'ların tespitinde ise Barrnap (<https://github.com/tseemann/barrnap>) yazılımları kullanılmıştır. Fonksiyonel anotasyon eggNOG-mapper (Carlos vd., 2021) ile yapılmıştır.

2.2.9. Ankara keçisi ve yaban keçisi genomları arasındaki sinteni

Keçi genomları arasındaki sintenin belirlenmesinde ilk olarak skafold seviyedeki genomlar minimap2 (Li, 2018) algoritması kullanılarak hizalanmıştır. Elde edilen hizalama sonrası sintenik ilişkiler NGenomeSyn yazılımı (He vd., 2023) ile görselleştirilmiştir. Sintenik bölgelerin görselleştirilmesi için 10 milyondan büyük skafoldlar seçilmiştir. Sintenik bölgelerin görselleştirilmesinde ayrıca JupiterPlot (<https://github.com/JustinChu/JupiterPlot>) yazılımı da kullanılmış ve hem tüm skafoldlar hem de kromozomal skafoldlar kullanılarak dairesel görselleştirmeler yapılmıştır.

2.2.10. Tüm genomlara dayalı filogenetik analizler

Tüm genom temelli filogenetik analizlerde bu tez çalışmasında elde edilen Ankara keçisi ve yaban keçisi genomlarına ek olarak NCBI Genbak'dan *Ovis ammon polii* x *Ovis aries* (GCA_023701655), *Capricornis sumatraensis* (GCA_032405125), *Ovis aries* (GCA_016772045), *Capra hircus* (GCA_001704415), *Ovis ammon polii* (GCA_028583565), *Capra aegagrus* (GCA_000978405), *Ovis orientalis* (GCA_014523465), *Budorcas taxicolor* (GCA_023091745), *Ovis canadensis* (GCA_004026945), *Ovibos moschatus* (GCA_021462335), *Pseudois nayaur* (GCA_003182575), *Capra sibirica* (GCA_003182615), *Ovis nivicola lydekkeri* (GCA_903231385), *Capra ibex* (GCA_006410555), *Hemitragus hylocrius* (GCA_004026825), *Oreamnos americanus* (GCA_009758055), *Capra hircus* (GCA_015443085), *Capra hircus* (GCA_026652205), *Capra hircus* (GCA_000317765), *Capra hircus* (GCA_004361675) ve dış grup olarak *Bos taurus* (GCA_002263795) genomları indirilmiştir. Filogenetik analizlerde ilk olarak her bir genom ayrı ayrı BUSCO ile vertebrata veri tabanına karşı (3.354 gen) değerlendirilmiştir. Daha sonra elde edilen tek kopya genler MUSCLE kullanılarak hizalanmıştır. Hizalanan diziler trimAI

(<https://github.com/inab/trimal>) kullanılarak trimlenmiştir. Trimlenen diziler supermatrix hizalaması için birleştirilmiş ve ModelFinder (Kalyaanamoorthy vd., 2017) yazılımı kullanılarak en uygun baz değişim modeli belirlenmiştir. Son olarak PhyML yazılımı (Guindon vd., 2010) ile 1.000 bootstrap desteğinde ML yöntemiyle filogenetik ağaç elde edilmiştir.

2.2.11. Keçi genomlarının referans genomlara haritalanması ve varyasyonun belirlenmesi

Keçi genomlarındaki varyasyonun belirlenmesi için evcil keçi referans genomu (ARS1.2, GCF_001704415.2) ve yaban keçisi referans genomu (CapAeg_1.0, GCA_000978405.1) NCBI Genome (<https://www.ncbi.nlm.nih.gov/genome/>) veri tabanından fasta formatında indirilmiştir. İlk olarak çalışma kapsamında elde edilen filtrelenmiş Illumina kısa okumaları referans genomlara BWA-MEM (Li, 2013) algoritması ile haritalanmış ve sam dosyaları elde edilmiştir. Daha sonra elde edilen sam dosyaları samtools (Danecek vd., 2021) kullanılarak bam formatına dönüştürülmüştür. Picard (<http://broadinstitute.github.io/picard/>) kullanılarak bam dosyalarında düzenleme ve duplike okumaların çıkarılması işlemleri yapılmıştır. Daha sonra freebayes (Garrison ve Marth, 2012) kullanılarak SNP'ler ve inDeller belirlenmiştir. Elde edilen SNP/İnDellerin anotasyonunda Ensembl Variant Effect Predictor (VEP) annotation tool algoritması (McLaren vd., 2016) kullanılmıştır.

2.2.12. Elde edilen ham verilerinin NCBI-SRA veri tabanına yüklenmesi

Çalışma kapsamında elde edilen Guppy kullanılarak basecalling yapılmış fastq dosya formatındaki Oxford Nanopore MinION uzun okumaları, SRA Submission Portal Wizard (<https://submit.ncbi.nlm.nih.gov/subs/sra/>) kullanılarak NCBI-SRA veri tabanına (<https://www.ncbi.nlm.nih.gov/sra>) kaydedilmiştir.

3. BÖLÜM

BULGULAR

3.1. Yüksek molekül ağırlıklı genomik DNA izolasyonu

Tez çalışması kapsamında Ankara keçisi kas dokusundan Monarch HMW DNA izolasyon kiti kullanılarak izole edilen yüksek molekül ağırlıklı gDNA'nın miktarı, Qubit ile ortalama olarak 48 ng/ μ L olarak belirlenmiştir. İzole edilen gDNA 0,8'lik agaroz jelde koşturulmuş ve yüksek kalitede ve miktarda olduğu gözlenmiştir (Şekil 8). Ayrıca gDNA, Nanodrop ile spektrofotometrik olarak ölçülmüş ve değeri 47,18 ng/ μ L olarak bulunmuştur. OD 260/280 ve OD 260/230 saflık değerleri de sırasıyla 1,91 ve 1,81 olarak elde edilmiştir.



Şekil 8. HMW DNA izolasyon kiti ile izole edilen Ankara keçisi gDNA'sının agaroz jel görüntüsü

3.2. Illumina platformunda dizileme için yapılan genomik DNA izolasyonu

Ankara keçisi kas dokusu örneğinden QIAGEN DNeasy Blood & Tissue Kit kullanılarak izole edilen gDNA'nın miktarı Qubit ile 94,5 ng/ μ L, Nanodrop ile 112,77 ng/ μ L olarak

elde edilmiştir. Nanodrop ölçümünde OD 260/280 değeri 1,91; OD 260/230 değeri ise 2,38'dir. Yaban keçisi kas dokusu kullanılarak izole edilen gDNA'nın miktarı Qubit ile 25,3 ng/uL olarak elde edilmiştir. Yaban keçisi örneğinin gDNA miktarı Nanodrop ile 21,64 ng/uL olarak belirlenmiş ve OD 260/280 değeri 1,87 ve OD 260/230 değeri 2,26 olarak tespit edilmiştir.

3.3. Ankara keçisi tüm genomunun Illumina kısa okuma platformunda dizilenmesi

Illumina Novaseq platformunda Ankara keçisi gDNA'sının dizilenmesi sonucunda 1.110.684.815 X 2 okuma ile toplamda ~300 GB okuma verisi elde edilmiştir (Tablo 3). Filtreleme sonrası read1 için 1.006.004.151 ve read2 için 969.082.666 okuma elde edilmiştir.

Tablo 3. Ankara keçisi genomunun Illumina platformu ile kısa okumaları

Ham okumalar				
Okuma	Dizi	Minimum	Maximum	Ortalama
Ank-R1	1.110.684.815	35	151	148,4
Ank-R2	1.110.684.815	151	151	151
Filtreleme sonrası okumalar				
Okuma	Dizi	Minimum	Maximum	Ortalama
Ank2-R1	1.006.004.151	50	151	139
Ank2-R2	969.082.666	50	134	124,6

3.3. Yaban keçisi tüm genomunun Illumina kısa okuma platformunda dizilenmesi

Illumina Novaseq platformunda yaban keçisi gDNA'sının dizilenmesi sonucunda 728.931.687 X 2 okuma ile toplamda ~210 GB okuma verisi elde edilmiştir (Tablo 4). Filtreleme sonrası read1 için 693.911.410 ve read2 için 672.771.999 okuma elde edilmiştir.

Tablo 4. Yaban keçisi genomunun Illumina kısa okumaları

Ham okumalar				
Okuma	Dizi	Minimum	Maximum	Ortalama
Aeg-R1	728.931.687	35	151	148,4
Aeg-R2	728.931.687	151	151	151
Filtreleme sonrası				
Okuma	Dizi	Minimum	Maximum	Ortalama
Aeg2-R1	693.911.410	50	151	143,3
Aeg2-R2	672.771.999	50	151	141,6

3.4. Yaban keçisi genomunun birleştirilmesi

Yaban keçisinin genom birleştirilmesi için MaSuRCA algoritması kullanılmış ve toplam 181.575 kontigden oluşan bir birleştirme elde edilmiştir. Elde edilen taslak genomun en uzun kontigi 332.737 bç iken en kısa olan 100 bç'dir. Ortalama uzunluk ise 14.827 bç olup, genom boyutu ~2,7 Gb olarak birleştirilmiştir. Yapılan genom birleştirmede 1.000 bazdan uzun 125.551 kontig, 5.000 bazdan uzun 98.389 kontig, 10.000 bazdan uzun 76.305 kontig, 25.000 bazdan uzun 36.961 kontig ve 50.000 bazdan uzun 12.338 kontig bulunmaktadır. Genom birleştirmenin GC içeriği %42,1 olup N50 değeri 36.402 bazdır (Tablo 5).

Tablo 5. Yaban keçisinin genom birleştirmesinde elde edilen kontiglerin metrikleri

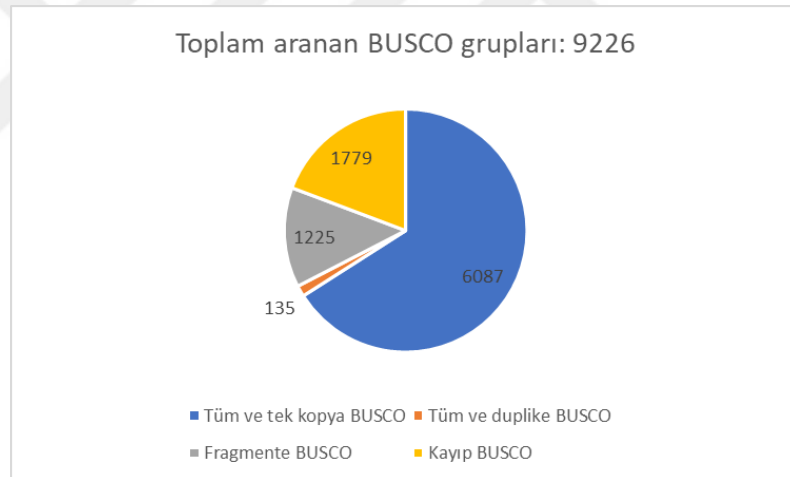
# Kontig (≥ 0 bp)	181.575
# Kontig (≥ 1000 bp)	125.551
# Kontig (≥ 5000 bp)	98.389
# Kontig (≥ 10000 bp)	76.305
# Kontig (≥ 25000 bp)	36.961
# Kontig (≥ 50000 bp)	12.338
En uzun kontig	332.737
Toplam uzunluk	2.683.373.112
Toplam uzunluk (≥ 0 bp)	2.692.141.911
Toplam uzunluk (≥ 1000 bp)	2.674.838.042
Toplam uzunluk (≥ 5000 bp)	2.601.458.755
Toplam uzunluk (≥ 10000 bp)	2.438.396.582
Toplam uzunluk (≥ 25000 bp)	1.789.201.245
Toplam uzunluk (≥ 50000 bp)	925.724.518
N50	36.402
N90	10.609
auN	45.497
L50	22.136
L90	74.038
GC (%)	42,1

Elde edilen yaban keçisi genom birleştirmesinin bütünlüğü BUSCO ile “mammalia_odb10” veri setine karşı değerlendirilmiş ve %67,5’lik bir bütünlük değeri elde edilmiştir. Yapılan BUSCO değerlendirmesinde toplam 9.226 çekirdek genden 6.087’si tek kopya, 135 gen duplike olarak genomda bulunurken; 1.225 gen fragmente ve

1.779 gen kayıp olarak belirlenmiştir. Bütünlük değerlerinin özeti Tablo 6 ve Şekil 9’da verilmiştir

Tablo 6. Yaban keçisi genomunun Illumina kısa okumalar ile birleştirilmesi sonrası BUSCO kullanılarak yapılan genom değerlendirmesinin istatistiksel değerleri

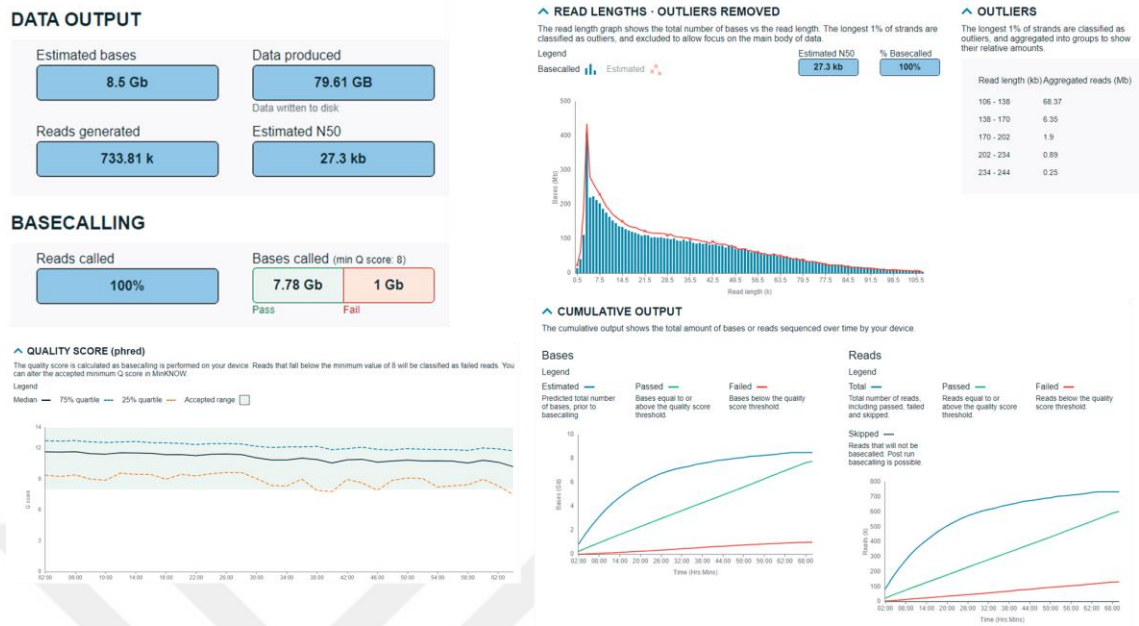
Veri seti: mamalia_odb10	
Toplam bütünlük	%67,5
Tüm BUSCO	6.222
Tüm ve tek kopya BUSCO	6.087
Tüm ve duplike BUSCO	135
Fragmente BUSCO	1.225
Kayıp BUSCO	1.779
Toplam aranan BUSCO grupları	9.226



Şekil 9. Yaban keçisi genom birleştirmesinin BUSCO sonuçlarının pasta dağılım grafiği

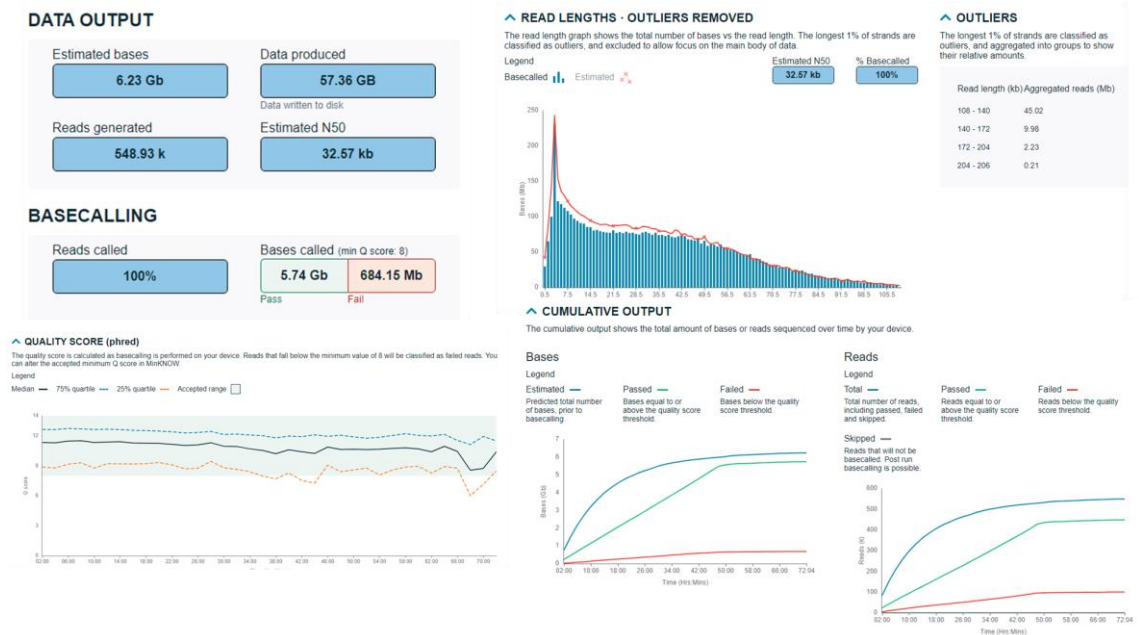
3.5. Oxford Nanopore MinION platformu kullanılarak elde edilen okuma sonuçları

Oxford Nanopore MinION cihazı toplamda 4 kez ve her birinde 72 saat süreyle çalıştırılmış, toplamda 314,49 GB fast5 dosyası üretilmiştir. MinKNOW programı ile fast basecalling yapılmış ve yaklaşık 32 Gb (~10X genom kapsamı) fasta dosyası ortalama Q skoru 12 olarak elde edilmiştir (Tablo 7, Şekil 10-13).



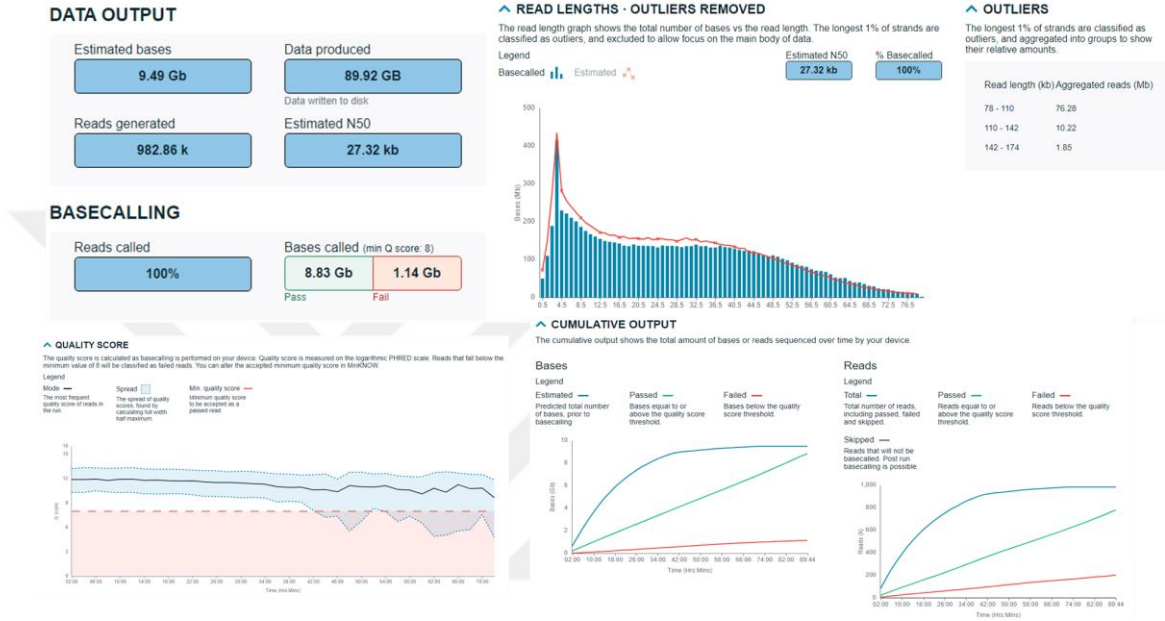
Şekil 10. Oxford Nanopore MinION cihazı birinci run sonuçları

Ankara keçisinin Oxford Nanopore MinION uzun okuma platformunda birinci run'ında toplam 79,61 GB veri üretilmiştir. Üretilen verilerin MinKNOW programında basecalling yapılması sonucunda N50 değeri 27,3 kb olan yaklaşık 733,81 bin (8,5 Gb) okuma elde edilmiştir. Üretilen verilerin yaklaşık 7,78 Gb'ı basecalled olurken 1 Gb veri, Q skoru 8'in altında kaldığı için hatalı olarak değerlendirilmiştir.



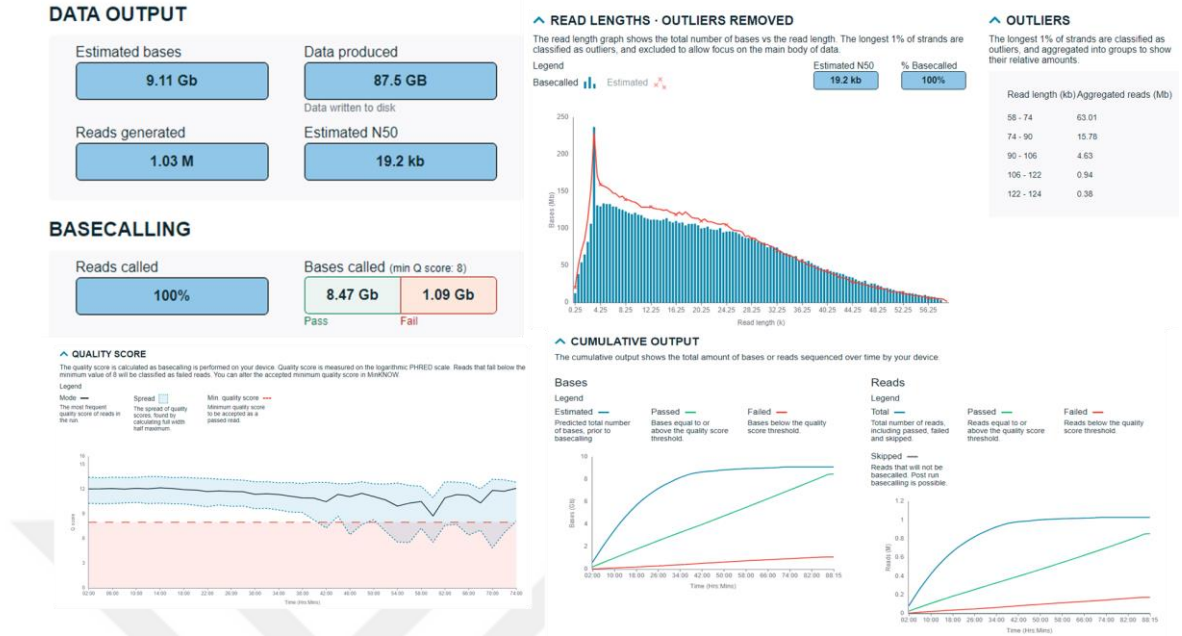
Şekil 11. Oxford Nanopore MinION cihazı ikinci run sonuçları

Ankara keisinin Oxford Nanopore MinION uzun okuma platformunda ikinci run’ında toplam 57,36 GB veri retilmiřtir. retilen verilerin MinKNOW programında basecalling yapılması sonucunda N50 deėeri 32,57 kb olan yaklaşık 548,93 bin (6,23 Gb) okuma elde edilmiřtir. retilen verilerin yaklaşık 5,74 Gb’ı basecalled olurken 684,15 Mb veri, Q skoru 8’in altında kaldığı iin hatalı olarak deėerlendirilmiřtir.



řekil 12. Oxford Nanopore MinION cihazı nc run sonuları

Ankara keisinin Oxford Nanopore MinION uzun okuma platformunda nc run’ında toplam 89,92 GB veri retilmiřtir. retilen verilerin MinKNOW programında basecalling yapılması sonucunda N50 deėeri 27,32 kb olan yaklaşık 982,86 bin (9,49 Gb) okuma elde edilmiřtir. retilen verilerin yaklaşık 8,83 Gb’ı basecalled olurken 1,14 Gb veri, Q skoru 8’in altında kaldığı iin hatalı olarak deėerlendirilmiřtir.



Şekil 13. Oxford Nanopore MinION platformu 4.run sonuçları

Ankara keşisinin Oxford Nanopore MinION uzun okuma platformunda dördüncü run'ında toplam 87,5 GB veri üretilmiştir. Üretilen verilerin MinKNOW programında basecalling yapılması sonucunda N50 değeri 19,2 kb olan yaklaşık 1,03 milyon (9,11 Gb) okuma elde edilmiştir. Üretilen verilerin yaklaşık 8,47 Gb'ı basecalled olurken 1,09 Gb veri, Q skoru 8'in altında kaldığı için hatalı olarak değerlendirilmiştir.

Tablo 7. Ankara keşisi genomunun Oxford Nanopore uzun okuma değerleri

Run	Elde edilen baz miktarı	Üretilen veri	Üretilen okuma	Tahmini N50	Basecalling okuma	Hatalı okuma	Süre
1	8,5 Gb	79,61 GB	733,81 bin	27,3 kb	7,78 Gb	1 Gb	72 saat
2	6,23 Gb	57,36 GB	548,93 bin	32,57 kb	5,74 Gb	684,15 Mb	72 saat
3	9,49 Gb	89,92 GB	982,86 bin	27,32 kb	8,83 Gb	1,14 Gb	72 saat
4	9,11 Gb	87,6 GB	1,03 milyon	19,2 kb	8,47 Gb	1,09 Gb	72 saat

Oxford Nanopore MinION cihazı kullanılarak üretilen toplam 314,49 GB fast5 dosyası, baz doğruluğunu iyileştirmek için Guppy ile tekrardan basecalling yapılmıştır. MinKNOW programında “fast” modda yapılan bu basecalling işlemi Guppy ile “high accuracy” modda yeniden yapılmıştır. Bu sayede elde edilen okumaların N50 değerleri, baz sayıları vb. diğer değerlerde artışlar elde edilmiştir. Birinci run'ın Guppy yazılımı ile yeniden basecalling işlemi sonrasında ortalama okuma uzunluğu 13.614 baz, ortalama okuma kalite skoru 12,9, N50 değeri 27.904 ve toplam baz miktarı 7.798.153.917

olmuştur. Ayrıca okumaların %94,4'lük kısmına denk gelen 7.414,1 Mb dizinin kalite skoru 10'un üzerindedir (Tablo 8). İkinci run'ın Guppy yazılımı ile yeniden basecalling işlemi sonrasında ortalama okuma uzunluğu 14.643 baz, ortalama okuma kalite skoru 12,7, N50 değeri 33.313 ve toplam baz miktarı 5.735.723.117 olmuştur. Ayrıca okumaların %92,7'lik kısmına denk gelen 5.377,1 Mb dizinin kalite skoru 10'un üzerindedir (Tablo 8). Üçüncü run'ın Guppy yazılımı ile yeniden basecalling işlemi sonrasında ortalama okuma uzunluğu 12.902 baz, ortalama okuma kalite skoru 12,7, N50 değeri 28.064 ve toplam baz miktarı 8.833.663.096 olmuştur. Ayrıca okumaların %94,5'lik kısmına denk gelen 8.367,2 Mb dizinin kalite skoru 10'un üzerindedir (Tablo 8). Dördüncü run'ın Guppy yazılımı ile yeniden basecalling işlemi sonrasında ortalama okuma uzunluğu 10.071 baz, ortalama okuma kalite skoru 12,6, N50 değeri 19.513 ve toplam baz miktarı 8.520.272.710 olmuştur. Ayrıca okumaların %93,2'lik kısmına denk gelen 7.983,3 Mb dizinin kalite skoru 10'un üzerindedir (Tablo 8).

Tablo 8. Guppy ile basecalling sonrası okuma değerleri

Genel özet	1.run	2.run	3.run	4.run
Aktif kanallar	494	489	490	474
Ortalama okuma uzunluğu	13.614,10	14.643,30	12.902,50	10.071,30
Ortalama okuma kalitesi	12,9	12,7	12,7	12,6
Medyan okuma uzunluğu	6.698,00	6.476,00	6.129,00	5.737,00
Medyan okuma kalitesi	14,2	14,1	13,9	13,9
Okuma sayısı	572.801,00	391.696,00	684.647,00	845.995,00
Okuma N50 uzunluğu	27.904,00	33.313,00	28.064,00	19.513,00
STDEV okuma uzunluğu	16.645,30	18.192,60	14.980,10	10.829,60
Toplam baz	7.798.153.917	5.735.723.117	8.833.663.096	8.520.272.710
Kalite sınırlarının üzerindeki okumaların sayısı, yüzdesi ve megabazlar				
>Q5	572.801 (%100) 7798,2Mb	391.696 (%100) 5735,7Mb	684.646 (%100) 8833,7Mb	845.994 (%100) 8520,3Mb
>Q7	572.800 (%100) 7798,2Mb	391.693 (%100) 5735,7Mb	684.634 (%100) 8833,6Mb	845.989 (%100) 8520,3Mb
>Q10	540.825 (%94,4) 7414,1Mb	362.996 (%92,7) 5377,1Mb	647.251 (%94,5) 8367,2Mb	788.809 (%93,2) 7983,3Mb
>Q12	465.140 (%81,2) 6448,8Mb	304.179 (%77,7) 4538,0Mb	539.782 (%78,8) 7051,4Mb	653.708 (%77,3) 6734,1Mb
>Q15	193.877 (%33,8) 2649,7Mb	134.839 (%34,4) 2003,0Mb	187.119 (%27,3) 2375,4Mb	245.944 (%29,1) 2542,8Mb
İlk 5 en yüksek ortalama basecall kalite puanı ve okuma uzunlukları				
1	90,0 (881)	22,7 (1.459)	65,8 (2.162)	90,0 (338)
2	54,8 (2.510)	22,0 (2.509)	60,5 (811)	90,0 (215)
3	49,8 (1.546)	21,5 (1.434)	52,0 (2.371)	90,0 (526)
4	43,6 (1.167)	21,4 (4.707)	51,2 (3.049)	90,0 (199)
5	24,1 (2.046)	21,4 (3.450)	48,5 (1.239)	90,0 (188)
En uzun okunan ilk 5 ve ortalama basecall kalite skoru				
1	253.092 (12,2)	212.495 (14,7)	177.550 (10,8)	130.567 (11,0)
2	235.600 (11,7)	205.577 (12,6)	171.293 (10,3)	128.142 (13,4)
3	232.490 (12,7)	202.077 (12,6)	158.799 (12,3)	128.066 (11,2)
4	219.215 (11,9)	194.879 (13,0)	157.323 (13,7)	126.553 (10,5)
5	214.157 (12,8)	191.478 (11,1)	153.078 (13,0)	126.012 (11,8)

3.6. Hibrit yöntemle birleştirilen genomun istatistiksel değerlendirilmesi

Hibrit genom birleştirme yöntemi ile elde edilen genomun uzunluğu 2.612.318.391 bç, toplam kontig sayısı 1.738, N50 değeri 10.940.225, en uzun fragman 57.345.006 ve genom kapsamı 9X olarak elde edilmiştir. Yapılan genom birleştirmede 1.000 bazdan uzun 1.681 kontig, 5.000 bazdan uzun 1.458 kontig, 10.000 bazdan uzun 1.270 kontig, 25.000 bazdan uzun 1.074 kontig ve 50.000 bazdan uzun 925 kontig bulunmaktadır. Genom birleştirmenin GC içeriği %42,98'dir (Tablo 9).

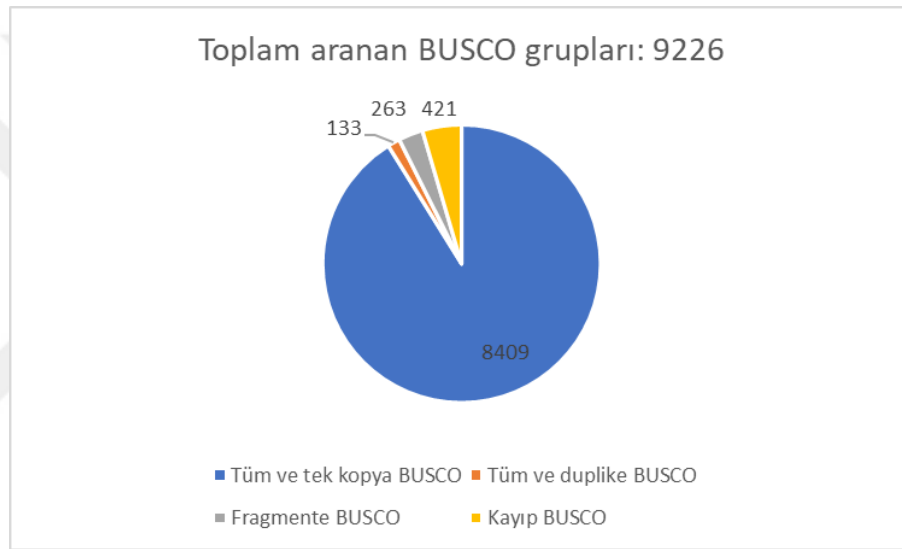
Tablo 9. Ankara keçisi genom birleştirmesinde elde edilen kontiglerin metrikleri

# Kontig (>= 0 bp)	1.738
# Kontig (>= 1000 bp)	1.681
# Kontig (>= 5000 bp)	1.458
# Kontig (>= 10000 bp)	1.270
# Kontig (>= 25000 bp)	1.074
# Kontig (>= 50000 bp)	925
En uzun kontig	573.96.725
Toplam uzunluk	2.614.494.707
Toplam uzunluk (>= 0 bp)	2.614.495.581
Toplam uzunluk (>= 1000 bp)	2.614.456.560
Toplam uzunluk (>= 5000 bp)	2.613.813.018
Toplam uzunluk (>= 10000 bp)	2.612.466.152
Toplam uzunluk (>= 25000 bp)	2.609.444.531
Toplam uzunluk (>= 50000 bp)	2.604.035.824
N50	10.952.280
N90	1.632.458
auN	13.145.023
L50	72
L90	282
GC (%)	41,98

BUSCO ile genom birleştirmesinin bütünlüğü 'mammalian database'e karşı değerlendirilmiş ve tek kopya %91,1, duplike %1,4, fragmente %2,9, kayıp %4,6 olmak üzere elde edilen Ankara keçisi genomu %92,5 bütünlük göstermiştir. Yapılan BUSCO değerlendirmesinde toplam 9.226 çekirdek genden 8.409'u tek kopya, 133 gen duplike olarak genomda bulunurken; 263 gen fragmente ve 421 gen kayıp olarak belirlenmiştir (Tablo 10, Şekil 14).

Tablo 10. BUSCO kullanılarak yapılan genom değerlendirmesinin istatistiksel değerlendirmesi

Veri seti: mamalia_odb10	
Toplam bütünlük	%92,5
Tüm BUSCO	8.542
Tüm ve tek kopya BUSCO	8.409
Tüm ve duplike BUSCO	133
Fragmente BUSCO	263
Kayıp BUSCO	421
Toplam aranan BUSCO grupları	9.226



Şekil 14. Ankara keçisi kontig BUSCO sonuçları pasta dağılımı

3.7. Skafoldlama sonuçları

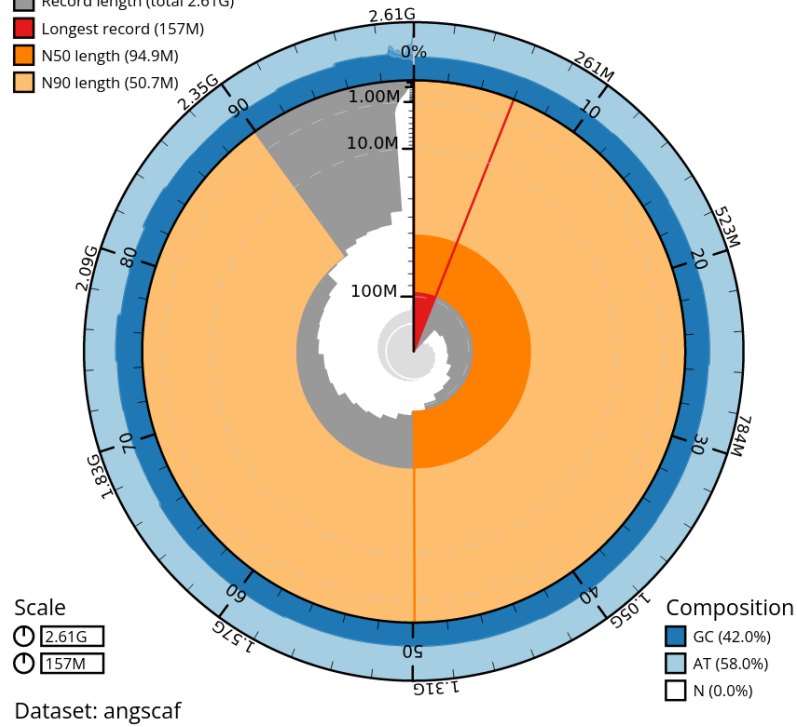
Ankara keçisi genomunun skafoldlanması sonucunda 2.614.597.507 bp uzunluğunda 710 skafold elde edilmiştir. Ek olarak skafoldlanan genom, 102.800 bp gap içermektedir. Yapılan skafold seviyedeki genom birleştirmede 1.000 bazdan uzun 659 kontig, 5.000 bazdan uzun 487 kontig, 10.000 bazdan uzun 334 kontig, 25.000 bazdan uzun 209 kontig ve 50.000 bazdan uzun 130 kontig bulunmaktadır. Genom birleştirmenin GC içeriği %41,98 olup, N50 değeri 94.917.083'dür (Tablo 11, Şekil 15).

Tablo 11. Ankara keçisi genom birleştirmesinde elde edilen skafoldların metrikleri

# Kontig (≥ 0 bp)	710
# Kontig (≥ 1000 bp)	659
# Kontig (≥ 5000 bp)	487
# Kontig (≥ 10000 bp)	334
# Kontig (≥ 25000 bp)	209
# Kontig (≥ 50000 bp)	130
En uzun kontig	157.094.010
Toplam uzunluk	2.614.597.507
Toplam uzunluk (≥ 0 bp)	2.614.598.381
Toplam uzunluk (≥ 1000 bp)	2.614.563.601
Toplam uzunluk (≥ 5000 bp)	2.614.071.830
Toplam uzunluk (≥ 10000 bp)	2.612.977.768
Toplam uzunluk (≥ 25000 bp)	2.611.021.176
Toplam uzunluk (≥ 50000 bp)	2.608.241.977
N50	94.917.083
N90	50.682.291
auN	94.069.613
L50	11
L90	26
GC (%)	41,98

Record statistics

- Log10 record count (total 710)
- Record length (total 2.61G)
- Longest record (157M)
- N50 length (94.9M)
- N90 length (50.7M)

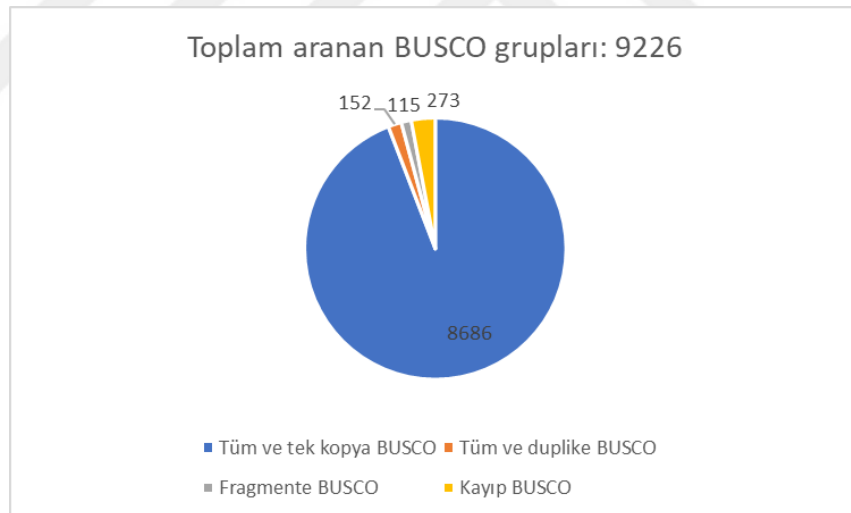


Şekil 15. Ankara keçisi genom birleştirmesinde elde edilen skafoldların dairesel görseli

Skafoldlanan genom BUSCO ile değerlendirilmiş ve %95,7 bütünlük değeri (S: %94,1, D: %1,6, F: %1,2, M: %3,1) elde edilmiştir. Yapılan BUSCO değerlendirmesinde toplam 9.226 çekirdek genden 8.686'sı tek kopya, 152 gen duplike olarak genomda bulunurken; 115 gen fragmente ve 273 gen kayıp olarak belirlenmiştir (Tablo 12, Şekil 16).

Tablo 12. Ankara keçisi genomunun skafoldlanma sonrası BUSCO kullanılarak yapılan genom değerlendirmesinin istatistiksel verileri

Veri seti: mamalia_odb10	
Toplam bütünlük	%95,7
Tüm BUSCO	8.838
Tüm ve tek kopya BUSCO	8.686
Tüm ve duplike BUSCO	152
Fragmente BUSCO	115
Kayıp BUSCO	273
Toplam aranan BUSCO grupları	9.226



Şekil 16. Ankara keçisi skafold BUSCO sonuçları pasta dağılımı

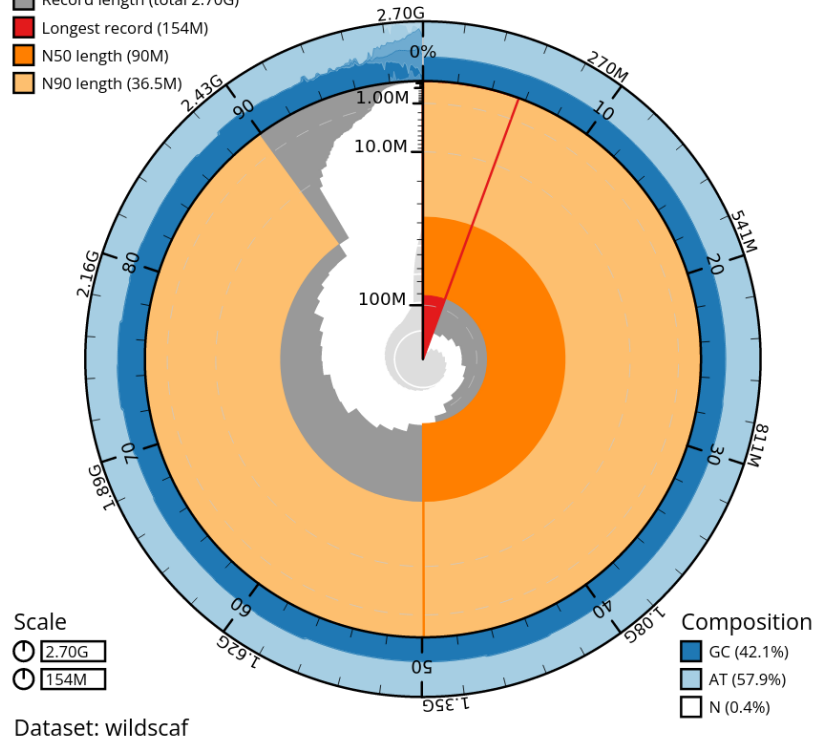
Yaban keçisi genomunun skafoldlanması sonucunda 2.694.791.249 bç uzunluğunda 68.461 skafold elde edilmiştir. Ek olarak skafoldlanan genom, 11.311.400 bç gap içermektedir. Yapılan skafold seviyedeki genom birleştirmede 1.000 bazdan uzun 13.532 kontig, 5.000 bazdan uzun 2.001 kontig, 10.000 bazdan uzun 1.026 kontig, 25.000 bazdan uzun 391 kontig ve 50.000 bazdan uzun 214 kontig bulunmaktadır. Genom birleştirmenin GC içeriği %42,1 ve N50 değeri 90.047.070'dir (Tablo 16, Şekil 17).

Tablo 13. Yaban keçisi genom birleştirmesinde elde edilen skafoldların metrikleri

# Kontig (≥ 0 bp)	68.461
# Kontig (≥ 1000 bp)	13.532
# Kontig (≥ 5000 bp)	2.001
# Kontig (≥ 10000 bp)	1.026
# Kontig (≥ 25000 bp)	391
# Kontig (≥ 50000 bp)	214
En uzun kontig	153.816.560
Toplam uzunluk	2.694.791.249
Toplam uzunluk (≥ 0 bp)	2.703.453.311
Toplam uzunluk (≥ 1000 bp)	2.686.840.392
Toplam uzunluk (≥ 5000 bp)	2.664.456.372
Toplam uzunluk (≥ 10000 bp)	2.657.746.290
Toplam uzunluk (≥ 25000 bp)	2.647.977.526
Toplam uzunluk (≥ 50000 bp)	2.641.896.411
N50	90.047.070
N90	44.312.406
auN	84.890.275
L50	12
L90	29
GC (%)	42,1

Record statistics

- Log10 record count (total 68.5k)
- Record length (total 2.70G)
- Longest record (154M)
- N50 length (90M)
- N90 length (36.5M)

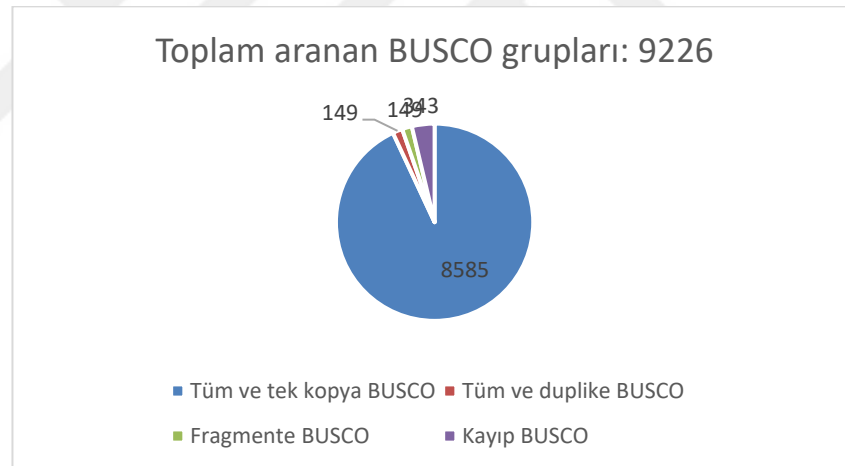


Şekil 17. Yaban keçisi genom birleştirmesinde elde edilen skafoldların dairesel görseli

Skafoldlanan genom BUSCO ile değerlendirilmiş ve %94,7 bütünlük değeri (S: %93,1, D: %1,6, F: %1,6, M: %3,7) elde edilmiştir. Yapılan BUSCO değerlendirmesinde toplam 9.226 çekirdek genden 8.585'i tek kopya, 149 gen duplike olarak genomda bulunurken; 149 gen fragmente ve 343 gen kayıp olarak belirlenmiştir (Tablo 14, Şekil 18).

Tablo 14. Yaban keçisi genomunun skafoldlanma sonrası BUSCO kullanılarak yapılan genom değerlendirmesinin istatistiksel verileri

Veri seti: mamalia_odb10	
Toplam bütünlük	%94,7
Tüm BUSCO	8.734
Tüm ve tek kopya BUSCO	8.585
Tüm ve duplike BUSCO	149
Fragmente BUSCO	149
Kayıp BUSCO	343
Toplam aranan BUSCO grupları	9.226



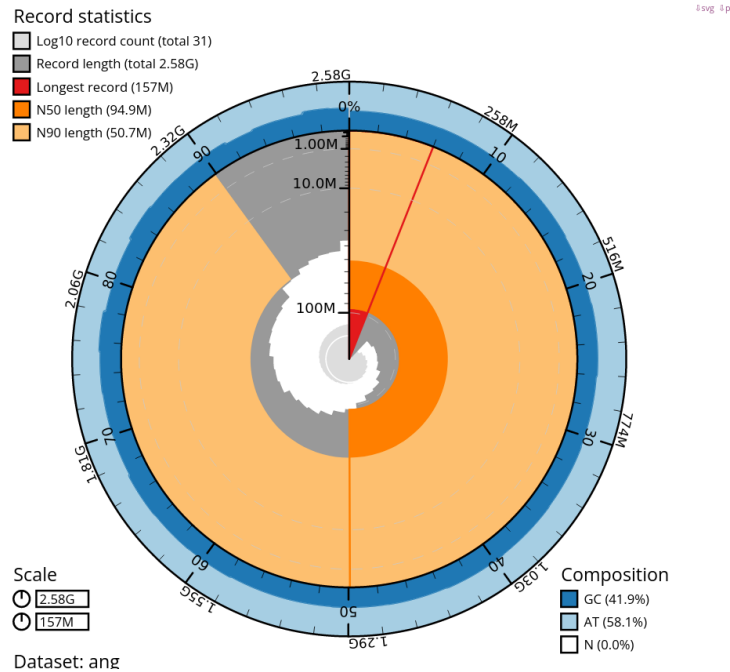
Şekil 18. Yaban keçisi skafold BUSCO sonuçları pasta dağılımı

Ankara keçisi ve Yaban keçisi genom birleştirmelerinde en uzun (kromozomal seviyedeki) 31 skafold çıkarılmış ve Tablo 15 ve 17'de metrikleri verilmiştir (Şekil 19 ve 21). Bu kromozomal seviyedeki skafoldlar BUSCO ile mamalia_odb10 veri tabanına karşı değerlendirilmiş 9.226 çekirdek genden tek kopya olan 8.605'si Ankara keçisi 31 kromozomal skafoldunda ve 8.227'si de yaban keçisi 31 kromozomal skafoldunda bulunmuştur. Ankara keçisi 31 kromozomal skafoldunun BUSCO bütünlük değeri %94,9, yaban keçisi 31 kromozomal skafoldunun BUSCO bütünlük değeri ise %90,7

olarak elde edilmiştir. Ankara keçisi ile yaban keçisi kromozomal skafoldlarının BUSCO değerlendirme sonuçları Tablo 16 ve 18 ile Şekil 20 ve 22’de verilmiştir.

Tablo 15. Ankara keçisi genom birleştirmesindeki en uzun 31 skafoldun metrikleri

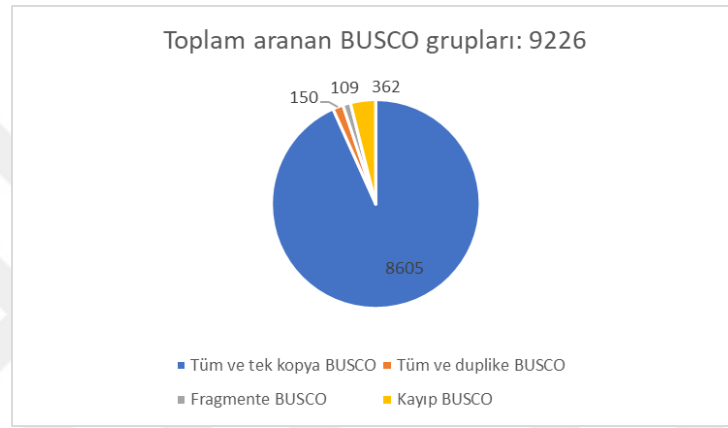
# Kontig (≥ 0 bp)	31
# Kontig (≥ 1000 bp)	31
# Kontig (≥ 5000 bp)	31
# Kontig (≥ 10000 bp)	31
# Kontig (≥ 25000 bp)	31
# Kontig (≥ 50000 bp)	31
En uzun kontig	157.094.010
Toplam uzunluk	2.578.578.506
Toplam uzunluk (≥ 0 bp)	2.578.578.506
Toplam uzunluk (≥ 1000 bp)	2.578.578.506
Toplam uzunluk (≥ 5000 bp)	2.578.578.506
Toplam uzunluk (≥ 10000 bp)	2.578.578.506
Toplam uzunluk (≥ 25000 bp)	2.578.578.506
Toplam uzunluk (≥ 50000 bp)	2.578.578.506
N50	94.917.083
N90	50.682.291
auN	95.374.487
L50	11
L90	26
GC (%)	41,94



Şekil 19. Ankara keçisi genom birleştirmesinde elde edilen en uzun 31 skafoldun dairesel gösterimi

Tablo 16. Ankara keçisi genom birleştirmesinde elde edilen en uzun 31 skafoldun BUSCO verileri

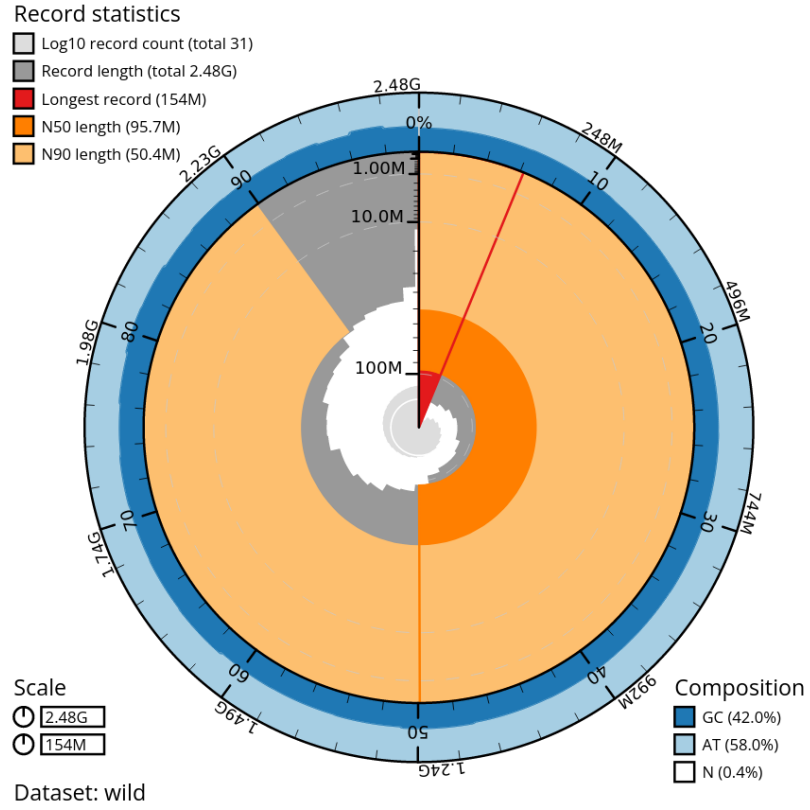
Veri seti: mamalia_odb10	
Toplam bütünlük	94.9%
Tüm BUSCO	8755
Tüm ve tek kopya BUSCO	8605
Tüm ve duplike BUSCO	150
Fragmente BUSCO	109
Kayıp BUSCO	362
Toplam aranan BUSCO grupları	9226



Şekil 20. Ankara keçisi genom birleştirmesinde elde edilen en uzun 31 skafoldun BUSCO sonuçları pasta dağılımı

Tablo 17. Yaban keçisi genom birleştirmesindeki en uzun 31 skafoldun metrikleri

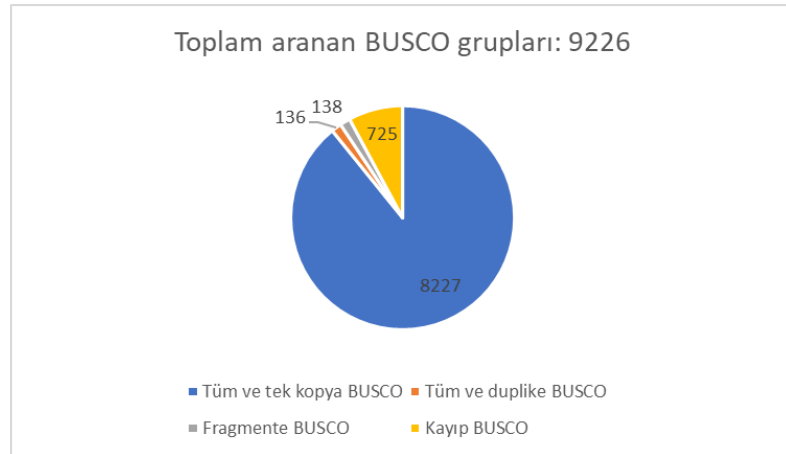
# Kontig (≥ 0 bp)	31
# Kontig (≥ 1000 bp)	31
# Kontig (≥ 5000 bp)	31
# Kontig (≥ 10000 bp)	31
# Kontig (≥ 25000 bp)	31
# Kontig (≥ 50000 bp)	31
En uzun kontig	153.816.560
Toplam uzunluk	2.478.849.545
Toplam uzunluk (≥ 0 bp)	2.478.849.545
Toplam uzunluk (≥ 1000 bp)	2.478.849.545
Toplam uzunluk (≥ 5000 bp)	2.478.849.545
Toplam uzunluk (≥ 10000 bp)	2.478.849.545
Toplam uzunluk (≥ 25000 bp)	2.478.849.545
Toplam uzunluk (≥ 50000 bp)	2.478.849.545
N50	95.698.409
N90	50.367.326
auN	92.052.797
L50	11
L90	25
GC (%)	42,01



Şekil 21. Yaban keçisi genom birleştirmesinde elde edilen en uzun 31 skafoldun dairesel gösterimi

Tablo 18. Yaban keçisi genom birleştirmesinde elde edilen en uzun 31 skafoldun BUSCO değerleri

Veri seti: mamalia_odb10	
Toplam bütünlük	90.7%
Tüm BUSCO	8363
Tüm ve tek kopya BUSCO	8227
Tüm ve duplike BUSCO	136
Fragmente BUSCO	138
Kayıp BUSCO	725
Toplam aranan BUSCO grupları	9226



Şekil 22. Yaban keçisi genom birleştirmesinde elde edilen en uzun 31 skafoldun BUSCO sonuçları pasta dağılımı

3.8. Genom anotasyon sonuçları

3.8.1. Varyant anotasyon sonuçları

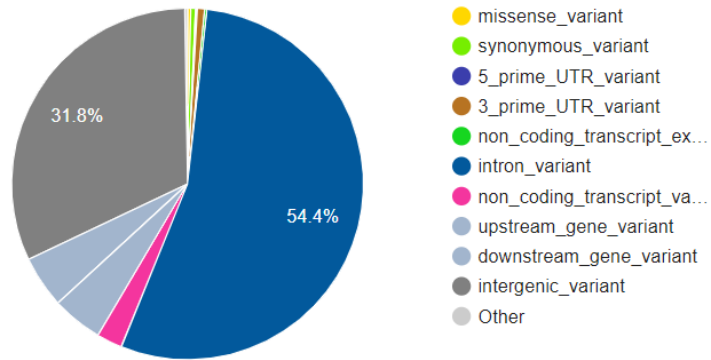
Yaban keçisi Illumina kısa okumalarının referans yaban keçisi genomuna haritalanması sonucu toplamda 10.935.456 varyant tespit edilmiştir. Bu varyantlardan 9.256.756'sı SNP, 575.054'ü insersiyon ve 559.390'ı delesyon olarak annotate edilmiştir.

Ankara keçisi Illumina kısa okumalarının referans evcil keçi genomuna haritalanması sonucu toplamda 11.168.617 SNP tespit edilmiş ve bunlardan 10.775.084'ünün anotasyonu yapılmıştır. Bunlardan 231.463'ü indel, 431.767'ü substitution, 522.054'ü insersiyon, 543.088'i delesyon ve 9.046.712'si SNV olarak tespit edilmiştir (Şekil 23).



Şekil 23. Ankara keçisi varyant anotasyonu

Anotasyonu yapılan SNP'lerin sonuçları Tablo 19'de sunulmuş olup, en çok bulunan 10.494.972 ile intron varyantlarıdır (Şekil 24).

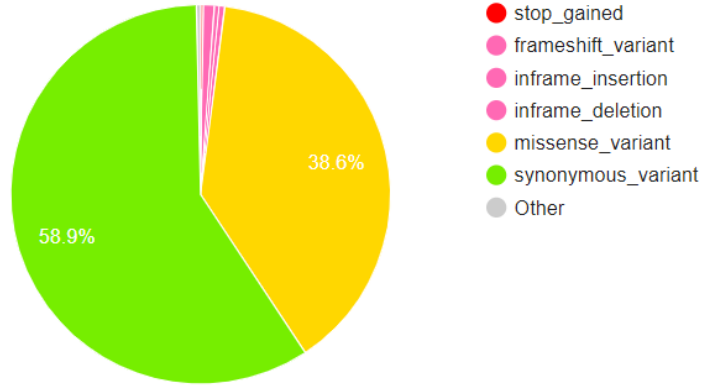


Şekil 24. Ankara keçisi varyant anotasyonunun bölgesel dağılımı

Tablo 19. Ankara keçisi varyant anotasyonunun bölgesel dağılım sayıları

Bölge	Miktar
intron_variant	10.494.972
intergenic_variant	6.143.104
downstream_gene_variant	921.549
upstream_gene_variant	901.135
non_coding_transcript_variant	453.813
3_prime_UTR_variant	131.481
synonymous_variant	85.095
missense_variant	55.785
non_coding_transcript_exon_variant	42.218
5_prime_UTR_variant	30.283
splice_polypyrimidine_tract_variant	23.725
splice_region_variant	14.982
splice_donor_region_variant	2.759
frameshift_variant	1.328
inframe_deletion	697
splice_donor_5th_base_variant	646
inframe_insertion	585
stop_gained	370
splice_acceptor_variant	360
splice_donor_variant	248
coding_sequence_variant	161
stop_lost	133
start_lost	87
protein_altering_variant	80
stop_retained_variant	39
start_retained_variant	6
incomplete_terminal_codon_variant	1

Ayrıca Ankara keçisinde gerçekleştirilen varyant analizi sonucunda büyük bir çoğunluğu sinonim varyant olan (%58,9) küçük varyantlar belirlenmiştir. Bunları takip eden ikinci en çok bulunan varyant tipi ise missense varyantlarıdır (%38,6). Elde edilen sonuçlar Şekil 25 ve Tablo 20’de gösterilmiştir.



Şekil 25. Ankara keçisi varyant anotasyonunun nedensel sonuçları

Tablo 20. Ankara keçisi varyant anotasyonunun nedensel sayıları

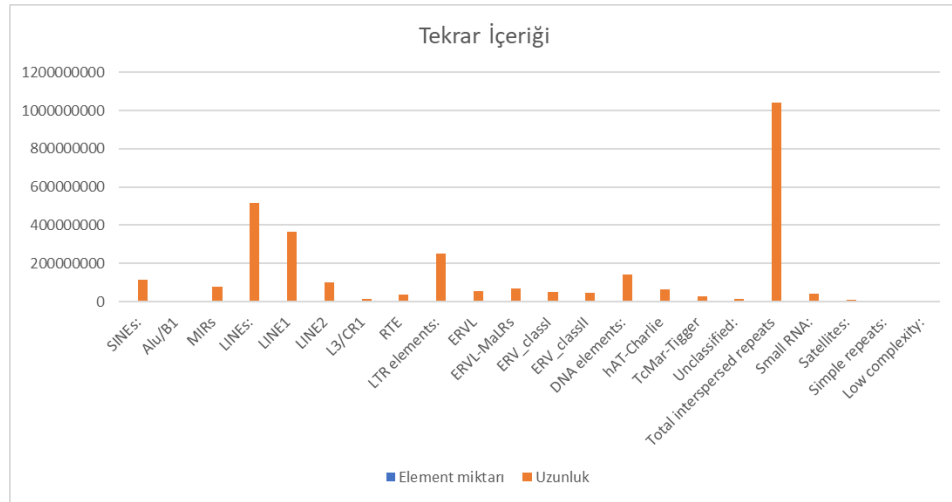
Nedensel Tip	Miktar
synonymous_variant	85.095
missense_variant	55.785
frameshift_variant	1.328
inframe_deletion	697
inframe_insertion	585
stop_gained	370
coding_sequence_variant	161
stop_lost	133
start_lost	87
protein_altering_variant	80
stop_retained_variant	39
start_retained_variant	6
incomplete_terminal_codon_variant	1

3.8.2. Tekrar elementleri

Tekrar bölgesi analizi ile Ankara keçisi genomunun tekrar içeriği %41,66 olarak belirlenmiştir (Tablo 21). GC içeriği %41,98 olan 2.614.598.381 bç uzunluğundaki genomda en çok %39,78 oranla interspersed tekrarlar tespit edilmiştir (Şekil 26).

Tablo 21. Ankara keçisi genomundaki tekrar bölgelerinin anotasyon sonuçları

Toplam uzunluk: 2.614.598.381 bç			
GC içeriği: %41,98			
Maskelenen baz: 1.089.319.018 bç (%41,66)			
	Element miktarı	Uzunluk	Yüzde (%)
SINE'ler:	1.208.564	115.783.198 bç	%4,43
Alu/B1	30.764	1.596.331 bç	%0,06
MIRs	600.847	77.323.316 bç	%2,96
LINE'ler:	1.389.518	516.400.092 bç	%19,75
LINE1	817.848	366.847.504 bç	%14,03
LINE2	398.171	99.161.033 bç	%3,79
L3/CR1	59.102	12.168.706 bç	%0,47
RTE	111.434	37.803.782 bç	%1,45
LTR elementler:	1.911.447	249.355.250 bç	%9,54
ERV1	326.560	54.657.793 bç	%2,09
ERV1-MaLRs	342.392	69.852.359 bç	%2,67
ERV_classI	397.769	52.250.336 bç	%2,00
ERV_classII	619.610	46.271.869 bç	%1,77
DNA elementler:	1.333.836	144.257.706 bç	%5,52
hAT-Charlie	611.756	66.559.821 bç	%2,55
TcMar-Tigger	191.085	27.808.039 bç	%1,06
Sınıflandırılmamış:	232.878	14.245.069 bç	%0,54
Toplam dağılmış tekrarlar		1.040.041.315 bç	%39,78
Small RNA:	812.661	39.971.767 bç	%1,53
Satellitler:	160.439	8.925.057 bç	%0,34
Basit tekrarlar:	0	0 bç	%0,00
Düşük kompleksler:	0	0 bç	%0,00



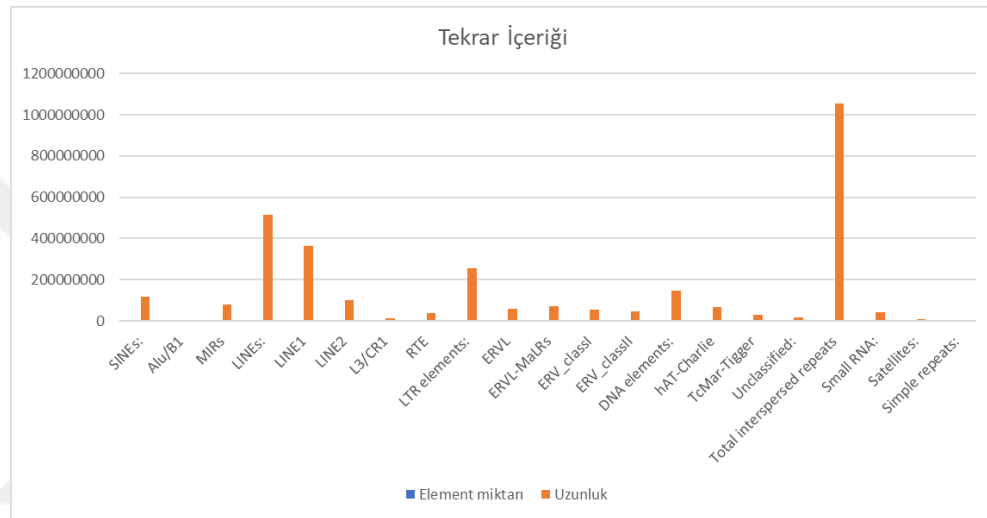
Şekil 26. Ankara keçisi genomunun tekrar içeriği grafiği

Tekrar bölgesi analizi ile yaban keçisi genomunun tekrar içeriği %42,13 olarak belirlenmiştir (Tablo 22). GC içeriği %42,13 olan 2.703.453.311 bç uzunluğundaki genomda en çok %38,93 oranla interspersed tekrarlar tespit edilmiştir (Şekil 27).

Tablo 22. Yaban keçisi genomundaki tekrar bölgelerinin anotasyonu

Toplam uzunluk: 2.703.453.311 bç			
GC içeriği: %42,13			
Maskelenen baz: 1.103.810.083 bç (%40,83)			
	Element miktarı	Uzunluk	Yüzde (%)
SINE'ler:	1.236.840	118.441.302 bç	%4,38
Alu/B1	31.082	1.608.034 bç	%0,06
MIRs	615.369	79.140.822 bç	%2,93
LINE'ler:	1.442.961	514.874.730 bç	%19,05
LINE1	851.006	363.289.290 bç	%13,44
LINE2	410.103	101.584.857 bç	%3,76
L3/CR1	63.105	12.705.594 bç	%0,47
RTE	115.675	36.857.568 bç	%1,36
LTR elementler:	1.990.517	256.481.504 bç	%9,49
ERV1	346.609	56.790.452 bç	%2,10
ERV1-MaLRs	360.234	72.143.571 bç	%2,67
ERV_classI	412.447	53.563.200 bç	%1,98
ERV_classII	640.564	47.063.442 bç	%1,74
DNA elementler:	1.364.530	147.686.851 bç	%5,46
hAT-Charlie	624.424	68.059.037 bç	%2,52
TcMar-Tigger	197.155	28.565.267 bç	%1,06
Sınıflandırılmamış:	254.899	15.037.166 bç	%0,56

Toplam dağılmış tekrarlar		1.052.521.553 bç	%38,93
Small RNA:	831.619	40.970.499 bç	%1,52
Satellitler:	172.970	9.923.937 bç	%0,37
Basit tekrarlar:	0	0 bç	%0,00
Düşük kompleksler:	0	0 bç	%0,00



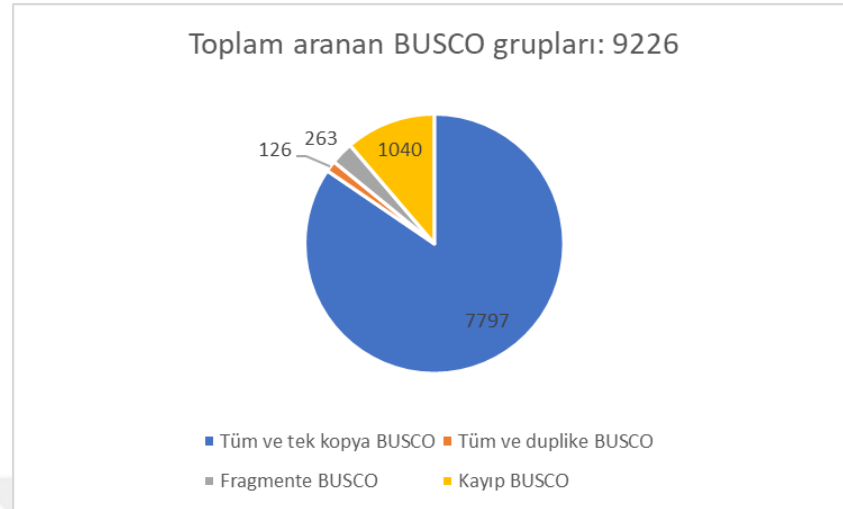
Şekil 27. Yaban keçisi genomunun tekrar içeriği grafiği

3.8.3. Gen tahmini

Ankara keçisinde homoloji temelli yaklaşımla yapılan gen tahmininde toplam 21.645 gen tahmin edilmiştir. Tahmin edilen genlerin mammalia_odb10 veri tabanındaki BUSCO analizi %85,9 bütünlük (Tek kopya: %84,5, Duplike: %1,4, Fragmente: %2,9, Kayıp: %11,2) ile sonuçlanmıştır (Tablo 23, Şekil 28). Ayrıca genomda 6 adet rRNA ve 182 adet tRNA da tahmin edilmiştir.

Tablo 23. Ankara keçisi gen tahmininde BUSCO değerleri

Veri seti: mamalia_odb10	
Toplam bütünlük	%85,9
Tüm BUSCO	7.923
Tüm ve tek kopya BUSCO	7.797
Tüm ve duplike BUSCO	126
Fragmente BUSCO	263
Kayıp BUSCO	1.040
Toplam aranan BUSCO grupları	9.226

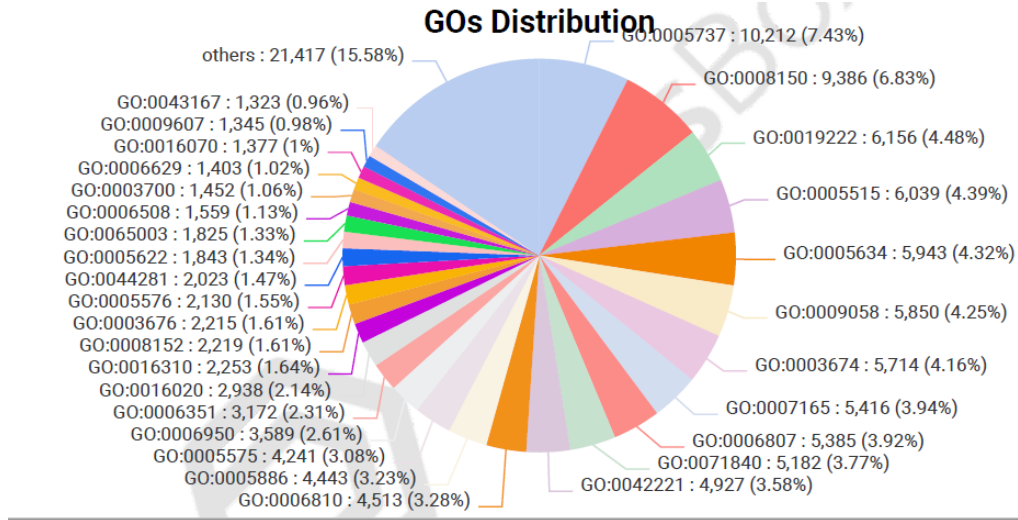


Şekil 28. Ankara keçisi genom anotasyonunun BUSCO grafiksel gösterimi

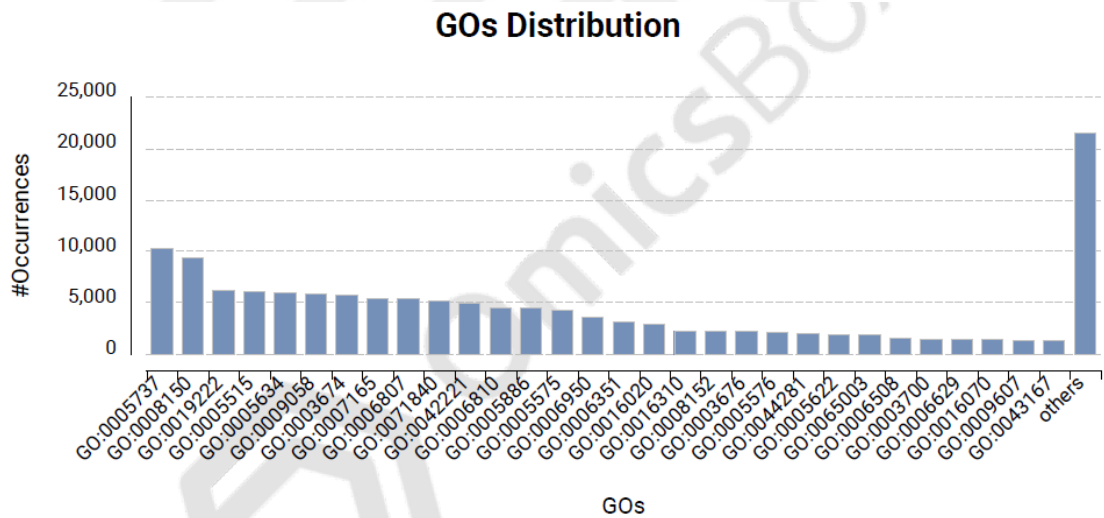
Ankara keçisinde tahmin edilen protein kodlayan genlerin fonksiyonel analizi sonucunda tahmin edilen 21.645 genden 16.117'si (%74,46) ortalama uzunluğu 495 ile fonksiyonel olarak annotate edilmiştir (Tablo 24 ve Şekil 29,30). Toplam GO anotasyonlarının sayısı 220.482 olup sekans başına düşen GO sayısı 13,68'dir. Fonksiyonel analizin 4.009'u (%18,76) bilgi depolama ve işleme kategorilerinde gerçekleşirken 7.838'i (%36,67) hücresel süreçler ve sinyalleşmede, 2.524'ü (11,81) metabolizma kategorilerinde ve 5.864'ü (%27,44) zayıf karakterize edilmiş olarak sonuçlanmıştır.

Tablo 24. Ankara keçisi genomunun fonksiyonel anotasyon verileri

Genel Dağılım
Toplam sekans miktarı: 21.645
Ortalama uzunluk: 495,0
Annote edilmiş GO sekanslarının sayısı: 16.117 / %74,46
GO anotasyonlarının sayısı: 220.482
Sekans başına ortalama GO'lar: 13,68
COG Kategori dağılımları
Bilgi Depolama ve İşleme:
Transkripsiyon (K): 1.982
Translasyon, ribozomal yapı ve biyogenez(J): 873
RNA işleme ve modifikasyon (A): 599
Replikasyon, rekombinasyon ve tamir (L): 311
Kromatin yapı ve dinamikleri (B): 244
Toplam: 4.009 / %18,76
HücreSEL Süreçler ve Sinyalleşme:
Sinyal iletim mekanizmaları (T): 4.797
Post-translasyon modifikasyon, protein dönüşümü, şaperonlar (O): 1.254
Hücre içi trafik, salgı ve veziküler taşıma (U): 617
Hücre iskeleti (Z): 467
Hücre dışı yapılar (W): 242
Hücre döngüsü kontrolü, hücre bölünmesi, kromozom bölünmesi (D): 188
Savunma mekanizması (V): 176
Hücre duvarı/zar/zarf biyogenezi (M): 66
Hücre hareketliliği (N): 22
Nükleer yapı (Y): 9
Toplam: 7.838 / %36,67
Metabolizma:
Enerji üretimi ve dönüşümü (C): 453
Lipid taşıması ve metabolizması (I): 423
İnorganik iyon taşıması ve metabolizması (P): 409
Karbonhidrat taşıması ve metabolizması (G): 392
Amino asit taşıması ve metabolizması (E): 357
Sekonder metabolitlerin biyosentezi, taşıması ve katabolizması (Q): 241
Nükleotid taşıması ve metabolizması (F): 160
Koenzim taşıması ve metabolizması (H): 89
Toplam: 2.524 / %11,81
Zayıf Karakterize Edilmiş:
Fonksiyonu bilinmeyen (S): 5.864
Yalnızca genel fonksiyon tahmini (R): 0
Total: 5.864 / %27,44



Şekil 29. Ankara keçi genomunun fonksiyonel anotasyon sonuçlarının GO dağılımları

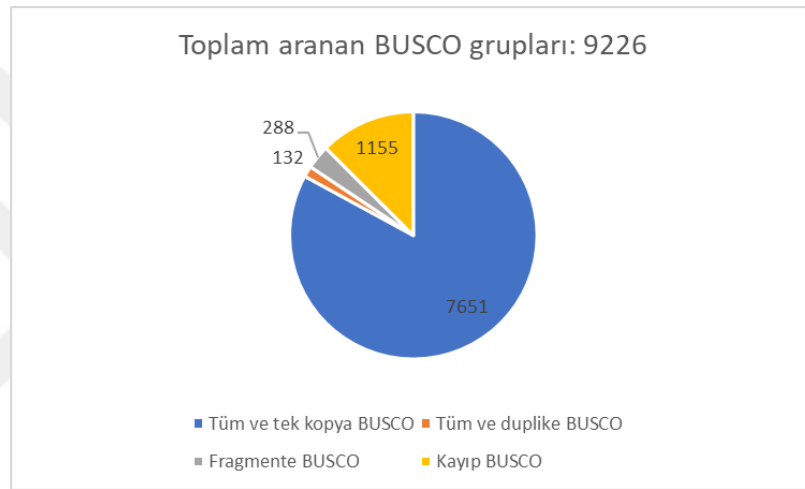


Şekil 30. Ankara keçi genomunun fonksiyonel anotasyon sonuçlarının GO dağılımları

Yaban keçisinde homoloji temelli yaklaşımla yapılan gen tahmininde toplam 21.409 gen tahmin edilmiştir. Tahmin edilen genlerin mammalia_odb10 veri tabanındaki BUSCO analizi %84,3 bütünlük (Tek kopya: %82,9, Duplike: %1,4, Fragmente: %3,1, Kayıp: %12,6) ile sonuçlanmıştır (Tablo 25, Şekil 31). Ayrıca genomda 11 adet rRNA ve 186 adet tRNA da tahmin edilmiştir.

Tablo 25. Yaban keçisi gen tahmininde BUSCO verileri

Veri seti: mamalia_odb10	
Toplam bütünlük	%84,3
Tüm BUSCO	7.783
Tüm ve tek kopya BUSCO	7.651
Tüm ve duplike BUSCO	132
Fragmente BUSCO	288
Kayıp BUSCO	1.155
Toplam aranan BUSCO grupları	9.226



Şekil 31. Yaban keçisi genom anotasyonunun BUSCO grafiksel gösterimi

Yaban keçisinde tahmin edilen protein kodlayan genlerin fonksiyonel analizi sonucunda tahmin edilen 21.409 genden 15.962'si (%74,56) ortalama uzunluğu 491 ile fonksiyonel olarak annotate edilmiştir. Toplam GO anotasyonlarının sayısı 218.033 olup sekans başına düşen GO sayısı 13,66'dır. Fonksiyonel analizin 3.959'u (%18,73) bilgi depolama ve işleme kategorilerinde gerçekleşirken 7.745'i (%36,63) hücresel süreçler ve sinyalleşmede, 2.493'ü (11,79) metabolizma kategorilerinde ve 5.812'si (%27,49) zayıf karakterize edilmiş olarak sonuçlanmıştır (Tablo 26 ve Şekil 32,33).

Tablo 26. Yaban keçisi genomunun fonksiyonel anotasyon verileri

Genel Dağılım

Toplam sekans miktarı: 21.409

Ortalama uzunluk: 491,0

Annot edilmiş GO sekanslarının sayısı: 15.962 / %74,56

GO anotasyonlarının sayısı: 218.033

Sekans başına ortalama GO'lar: 13,66

COG Kategori dağılımları

Bilgi Depolama ve İşleme:

Transkripsiyon (K): 1.956

Translasyon, ribozomal yapı ve biyogenez(J): 871

RNA işleme ve modifikasyon (A): 600

Replikasyon, rekombinasyon ve tamir (L): 294

Kromatin yapı ve dinamikleri (B): 238

Toplam: **3.959 / %18,73**

Hücrel Süreçler ve Sinyalleşme:

Sinyal iletim mekanizmaları (T): 4.728

Post-translasyon modifikasyon, protein dönüşümü, şaperonlar (O): 1.235

Hücre içi trafik, salgı ve veziküler taşıma (U): 615

Hücre iskeleti (Z): 465

Hücre dışı yapılar (W): 244

Hücre döngüsü kontrolü, hücre bölünmesi, kromozom bölünmesi (D): 185

Savunma mekanizması (V): 175

Hücre duvarı/zar/zarf biyogenezi (M): 65

Hücre hareketliliği (N): 23

Nükleer yapı (Y): 10

Total: **7.745 / %36,63**

Metabolizma:

Enerji üretimi ve dönüşümü (C): 447

Lipid taşıması ve metabolizması (I): 420

İnorganik iyon taşıması ve metabolizması (P): 409

Karbonhidrat taşıması ve metabolizması (G): 390

Amino asit taşıması ve metabolizması (E): 346

Sekonder metabolitlerin biyosentezi, taşıması ve katabolizması (Q): 235

Nükleotid taşıması ve metabolizması (F): 156

Koenzim taşıması ve metabolizması (H): 90

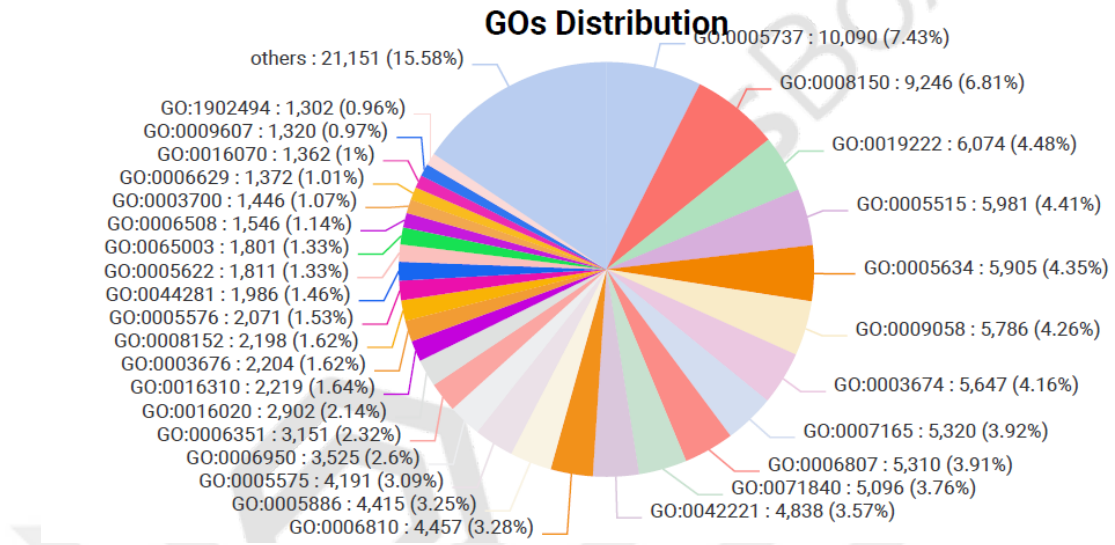
Toplam: **2.493 / %11,79**

Zayıf Karakterize Edilmiş:

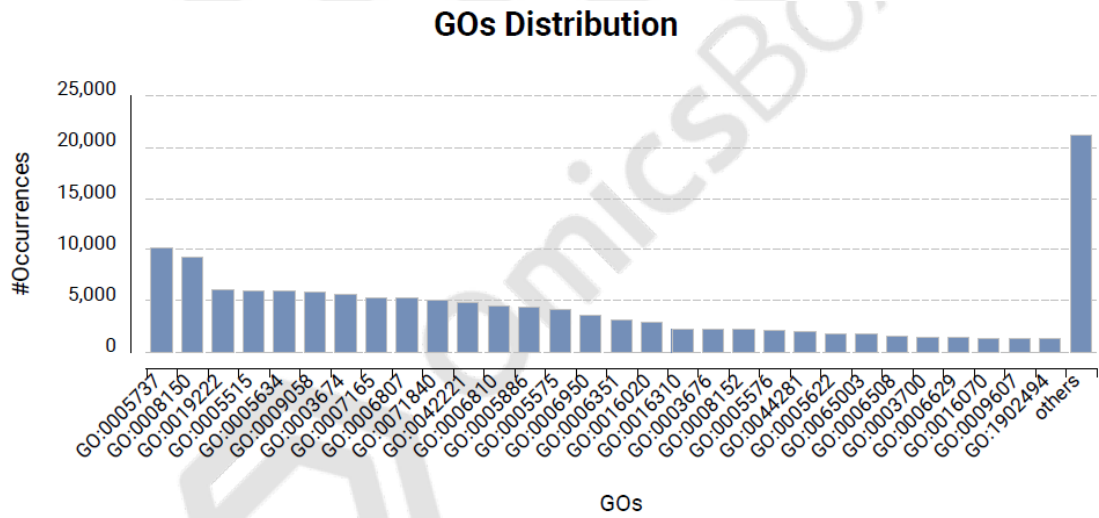
Fonksiyonu bilinmeyen (S): 5.812

Yalnızca genel fonksiyon tahmini (R): 0

Toplam: **5.812 / %27,49**



Şekil 32. Yaban keçisi genomunun fonksiyonel anotasyon sonuçlarının GO dağılımları



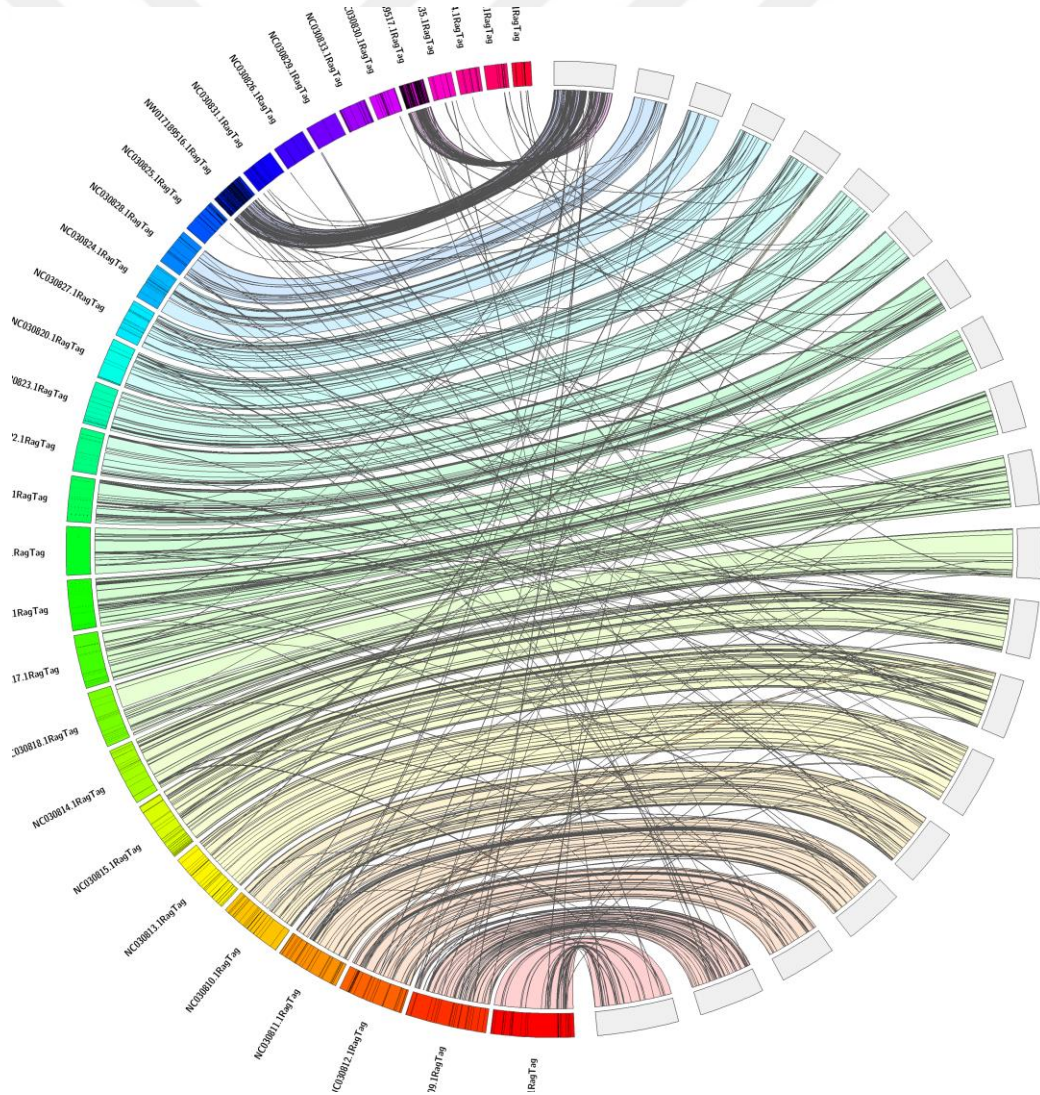
Şekil 33. Yaban keçisi genomunun fonksiyonel anotasyon sonuçlarının GO dağılımları

3.9. Ankara keçisi ve yaban keçisi genomları arasındaki sinteni

Ankara keçisi ve yaban keçisi genomlarındaki kromozomal seviyedeki 31 skafoldun karşılaştırılması ile genomlar arasında yüksek oranda korunmuş sintenik bölgeler olduğu bulunmuştur (Şekil 34). Her bir kromozomal skafold yüksek oranda diğer kromozomal skafold ile eşleşmiştir. Bunun yanı sıra farklı kromozomlar arasında küçük sintenik bloklar da gözlenmiştir (Şekil 34 ve Şekil 35).

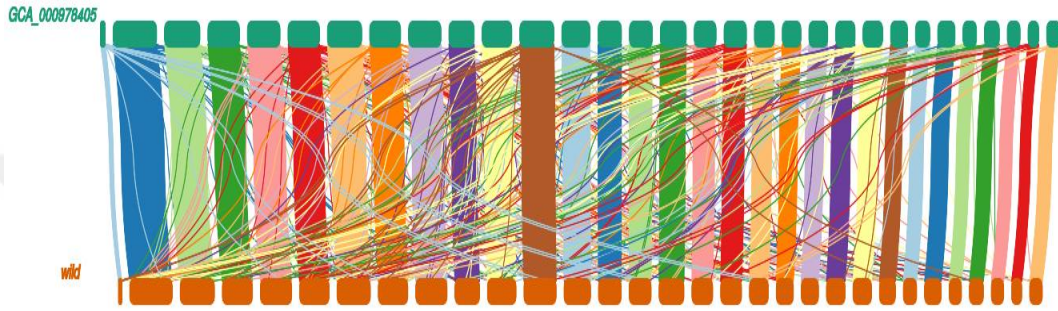


Şekil 34. Ankara keçisi ve yaban keçisi en uzun 31 skafoldu arasındaki sinteni

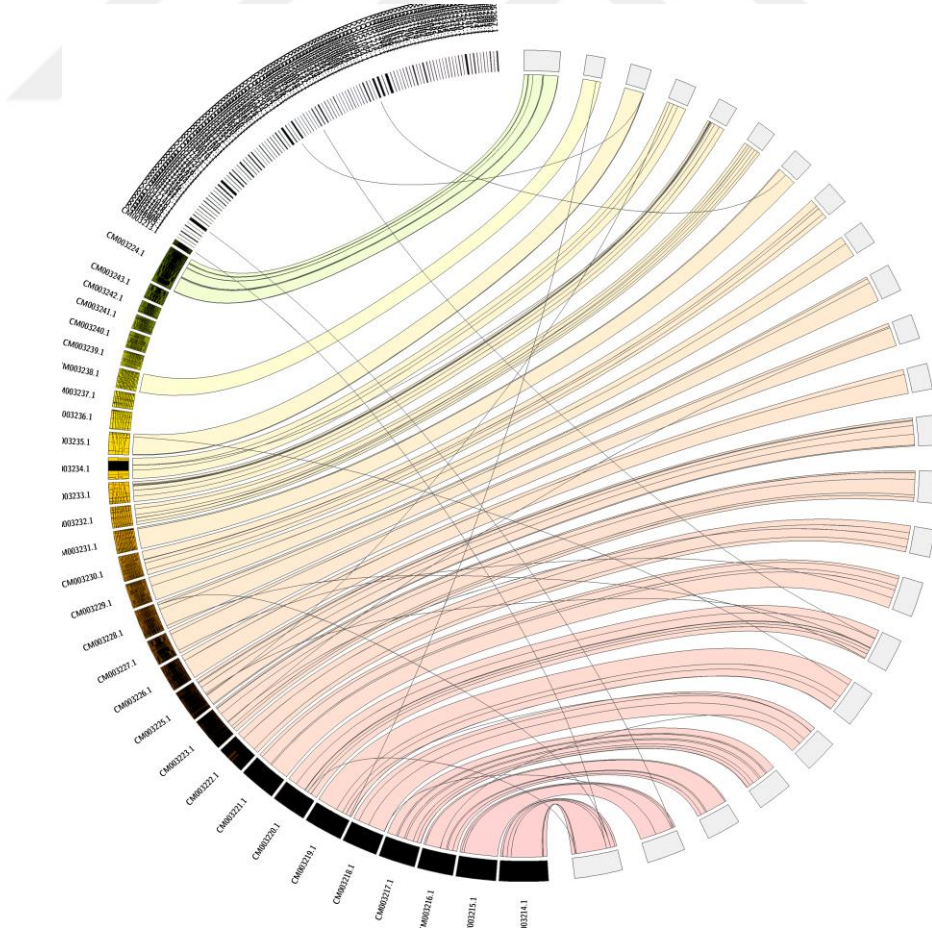


Şekil 35. Ankara keçisi ile yaban keçisi en uzun 31 skafoldları arasındaki sintenin dairesel gösterimi

Tez çalışması kapsamında elde edilen yaban keçisi kromozomal 31 skafoldu ile GenBank'dan indirilen yaban keçisi referans genomu arasında yapılan karşılaştırmada ise elde edilen kromozomal skafoldların referans yaban keçisi kromozomları ile eşleştiği gözlenmiştir (Şekil 36). Çalışmada elde edilen yaban keçisinin tüm skafoldları ile referans yaban keçisi genomu arasındaki sentenik ilişkinin dairesel görseli Şekil 37'de verilmiştir.

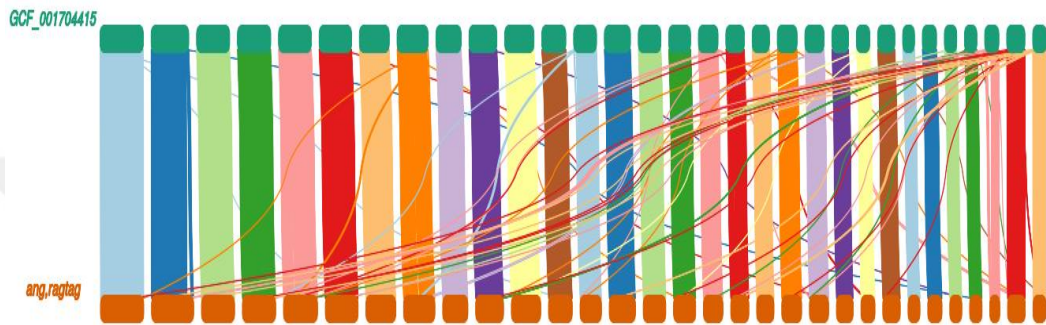


Şekil 36. Yaban keçisi ve referans yaban keçisinin en uzun 31 skafoldu arasındaki sentenik

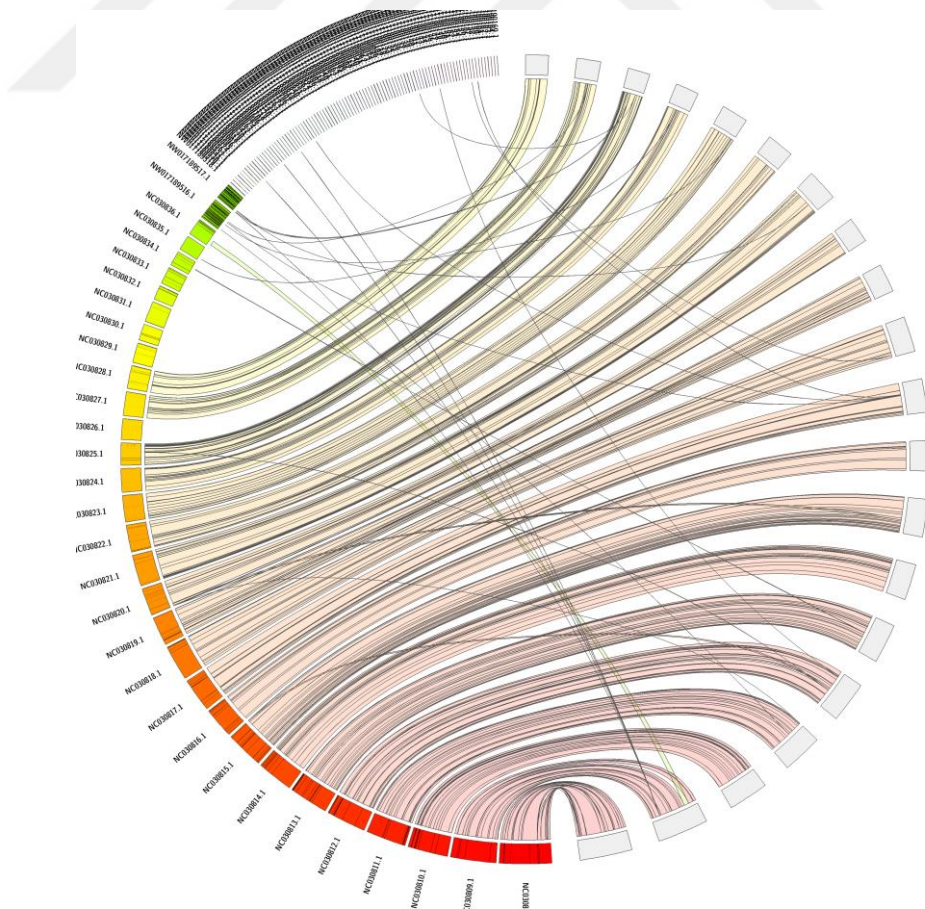


Şekil 37. Yaban keçisi ile referans yaban keçisi genomu arasındaki tüm skafoldların sentenisi

Tez çalışması kapsamında elde edilen Ankara keçisi kromozomal 31 skafoldu ile GenBank'dan indirilen evcil keçi referans genomu arasında yapılan karşılaştırmada ise elde edilen kromozomal skafoldların referans evcil keçi kromozomları ile çok yüksek bir oranda eşleştiği gözlenmiştir (Şekil 38). Çalışmada elde edilen Ankara keçisinin tüm skafoldları ile referans evcil keçi genomu arasındaki sentenik ilişkinin dairesel görseli Şekil 39'de verilmiştir.



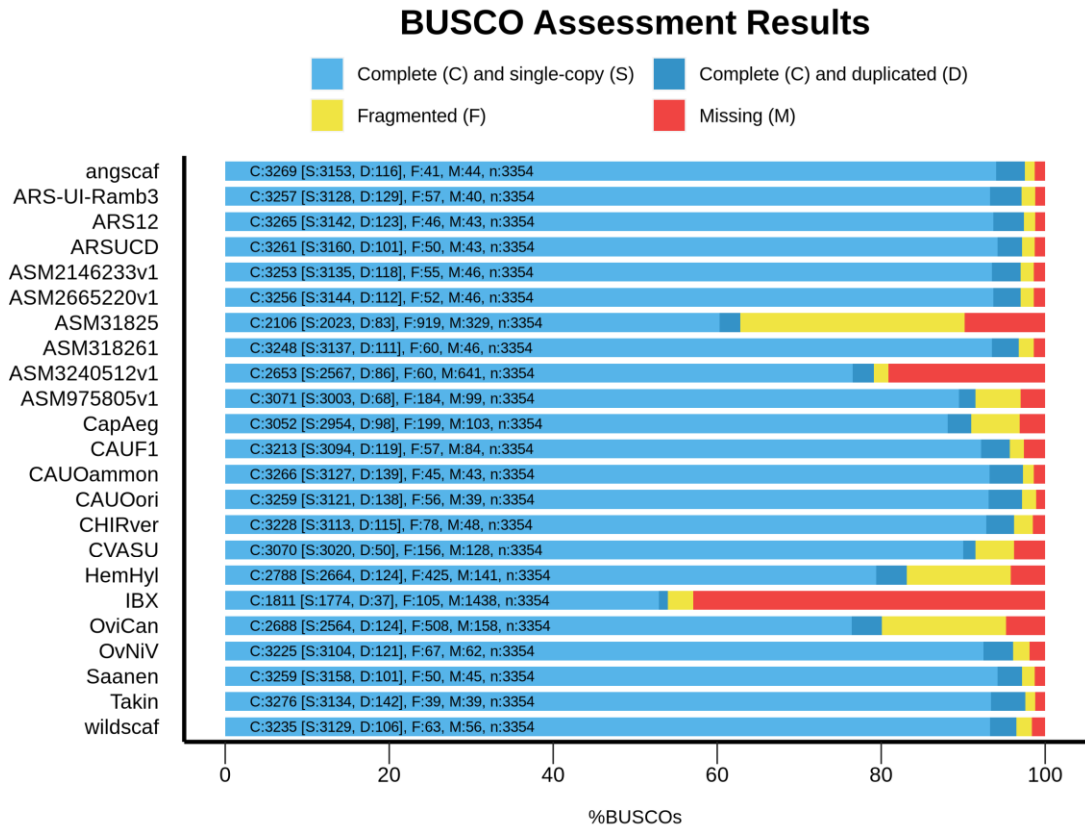
Şekil 38. Ankara keçisi ve referans evcil keçinin en uzun 31 skafoldu arasındaki sinteni



Şekil 39. Ankara keçisinin tüm skafoldları ile referans evcil keçi genomu arasındaki sinteni

3.10. Tüm genomlar kullanılarak gerçekleştirilen filogenetik analizler

Tez çalışması kapsamında elde edilen Ankara keçisi ve yaban keçisi genomları ile GenBank'dan indirilen *Caprinae* alt familyasındaki türlerin filogenetik ağacının oluşturulmasında ilk olarak tüm genomlar BUSCO ile vertebrata_odb10 (n=3.354) veri tabanına karşı değerlendirilmiş ve tek kopya genler çıkarılmıştır. Türlerin BUSCO skorları Şekil 40'ta verilmiştir. Toplamda 23 genom ile oluşturulan veri seti 346.009 uzunlukta amino asit dizisinden oluşmuştur.



Şekil 40. Filogenetik analizlerde kullanılan tüm genomların BUSCO değerlendirmeleri

PhyML kullanılarak oluşturulan 1.000 bootstrap değeri ile ML ağacında bu çalışmada elde edilen Ankara keçisi (angscf) ile referans evcil keçi genomu olan San Clemente keçisi (ARS12) birlikte kümelenmiştir. Bu çalışmada elde edilen bir diğer genom olan yaban keçisi (wildscf) ise referans yaban keçisi (CapAeg) ve Yunnan black keçisi (CHIRver) ile sister takson olan Black Bengal keçisine (CVASU) kardeş takson olarak pozisyon almıştır (Şekil 41).



Şekil 41. Tüm genomlar kullanılarak oluşturulan ML ağacı

3.11. Mitogenom birleştirme sonuçları

3.11.1. Mitogenom organizasyonu

Ankara keçisi komple mtDNA birleştirmesi sonucunda toplam uzunluk 16.644 bp'dir. Yaban keçisinin komple mtDNA uzunluğu ise 16.622 bp'dir. Elde edilen mitogenomlarda 13 adet protein kodlayan gen (PCG), 2 adet rRNA, 22 adet tRNA, 1 adet replikasyon orijini ve 1 adet kodlama yapmayan D-loop bölgesi belirlenmiştir.

Ankara keçisi mitogenomunun nükleotid kompozisyonu, adenin (A) için %33,52, timin (T) için %27,30, sitozin (C) için %26, guanin (G) için %13,10 hesaplanmıştır (Tablo 27).

Tablo 27. Ankara keçisi mitogenomunun baz kompozisyonu

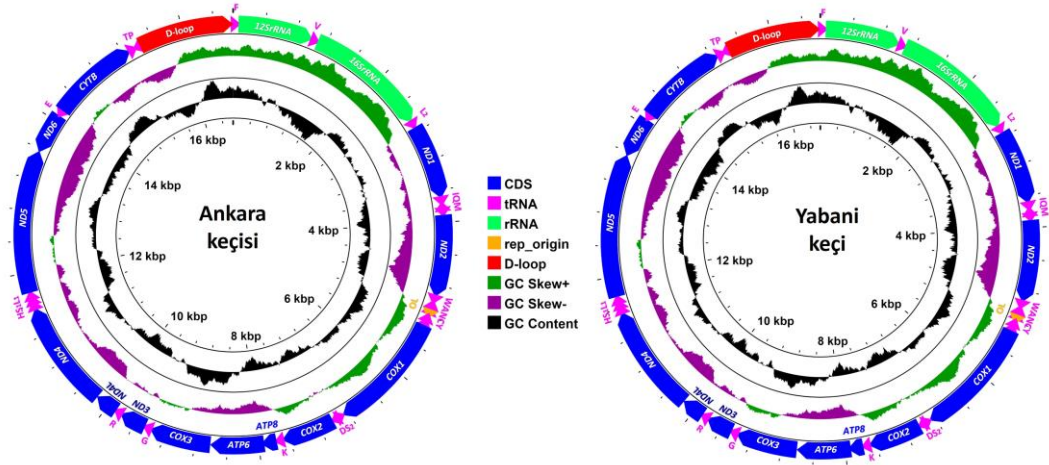
<i>Capra hircus</i> (Ankara keçisi)	Uzunluk	%	%A	%T	%G	%C	%A+T	%G+C
Mitogenom	16.644	100	33,50	27,30	13,10	26,00	60,80	39,20
PCG'ler	11.401	68,49	32,70	27,80	11,70	27,80	60,50	39,50
rRNA'lar	2.531	15,20	37,20	23,50	17,50	21,80	60,70	39,30
tRNA'lar	1.512	9,08	35,40	29,20	15,10	20,40	64,60	35,40
D-loop	1.213	7,28	31,00	29,10	14,30	25,60	60,10	39,90

Yaban keçisi mitogenomunun nükleotid kompozisyonu, adenin (A) için %33,50, timin (T) için %27,20, sitozin (C) için %26,20, guanin (G) için %13,10 hesaplanmıştır (Tablo 28).

Tablo 28. Yaban keçisi mitogenomunun baz kompozisyonu

<i>Capra aegagrus</i> (Yaban keçisi)	Uzunluk	%	%A	%T	%G	%C	%A+T	%G+C
Mitogenom	16.622	100%	33,50	27,20	13,10	26,20	60,70	39,30
PCG'ler	11.401	68,58	32,60	27,90	11,70	27,70	60,50	39,50
rRNA'lar	2.528	15,20	37,20	23,30	17,50	22,00	60,50	39,50
tRNA'lar	1.512	9,09	35,30	28,70	15,10	20,80	64,00	36,00
D-loop	1.194	7,18	31,00	28,70	13,90	26,40	59,70	40,30

Elde edilen Ankara keçisi mitogenomu ~3.890 kapsama; yaban keçisi mitogenomu ise ~5.440 X kapsama sahiptir. Mitogenomlardaki bölgelerden ND6 ve 8 tRNA geni (tRNA^{Gln}, tRNA^{Ala}, tRNA^{Asn}, tRNA^{Cys}, tRNA^{Tyr}, tRNA^{Ser}(UCN), tRNA^{Glu} ve tRNA^{Pro}) hafif zincir (L) üzerinde kodlanırken, ağır zincir (H) üzerinde 12 protein kodlayan gen (PCGs), replikasyon orijini (oL) 14 tRNA ve 2 rRNA genleri kodlanmaktadır (Şekil 42, Tablo 29 ve 30).



Şekil 42. Ankara keçisi ve yaban keçisi mitogenomlarının dairesel formu

Ankara keçisi mitogenomunda toplam 73 bç uzunluğunda (1 ile 40 bç aralığında değişen) 8 adet çakışan bölge ve toplam 29 bç uzunluğunda (1 ile 7 bç aralığında değişen) 13 adet intergenik aralık mevcuttur (Tablo 29). Yaban keçisi mitogenomunda ise Ankara keçisi ile benzer şekilde toplam 73 bç uzunluğunda (1 ile 40 bç aralığında değişen) 8 adet çakışan bölge ve toplam 29 bç uzunluğunda (1 ile 7 bç aralığında değişen) 13 adet intergenik aralık mevcuttur (Tablo 30).

Tablo 29. Ankara keçisi mitogenomun karakterizasyon tablosu

Gen/Bölge	Başlama	Bitiş	Uzunluk	Yön	Anti-kodon	Başlama-Bitiş kodonu	İntergenik bölge
tRNA-Phe	1	68	68	+	GAA		0
12S rRNA	69	1.024	956	+			0
tRNA-Val	1.025	1.091	67	+	TAC		0
16S rRNA	1.092	2.666	1.575	+			0
tRNA-Leu	2.667	2.741	75	+	TAA		2
ND1 CDS	2.744	3.699	956	+		ATG/TA-	0
tRNA-Ile	3.700	3.768	69	+	GAT		-3
tRNA-Gln	3.766	3.837	72	-	TTG		2
tRNA-Met	3.840	3.908	69	+	CAT		0
ND2 CDS	3.909	4.95	1.042	+		ATA/T--	0
tRNA-Trp	4.951	5.017	67	+	TCA		1
tRNA-Ala	5.019	5.087	69	-	TGC		1
tRNA-Asn	5.089	5.161	73	-	GTT		2
rep_orj	5.164	5.194	31	+			-1
tRNA-Cys	5.194	5.261	68	-	GCA		0
tRNA-Tyr	5.262	5.329	68	-	GTA		1
COX1 CDS	5.331	6.875	1.545	+		ATG/TAA	-3
tRNA-Ser	6.873	6.941	69	-	TGA		7
tRNA-Asp	6.949	7.016	68	+	GTC		1
COX2 CDS	7.018	7.701	684	+		ATG/TAA	3
tRNA-Lys	7.705	7.771	67	+	TTT		1
ATP8 CDS	7.773	7.97	198	+		ATG/TAA	-40
ATP6 CDS	7.931	8.611	681	+		ATG/TAA	-1
COX3 CDS	8.611	9.394	784	+		ATG/T--	0
tRNA-Gly	9.395	9.463	69	+	TCC		0
ND3 CDS	9.464	9.81	347	+		ATA/TA-	0
tRNA-Arg	9.811	9.879	69	+	TCG		0
ND4L CDS	9.88	10.176	297	+		ATG/TAA	-7
ND4 CDS	10.17	11.547	1.378	+		ATG/T--	0
tRNA-His	11.548	11.617	70	+	GTG		0
tRNA-Ser	11.618	11.677	60	+	GCT		1
tRNA-Leu	11.679	11.748	70	+	TAG		0
ND5 CDS	11.749	13.569	1.821	+		ATA/TAA	-17
ND6 CDS	13.553	14.08	528	-		ATG/TAA	0
tRNA-Glu	14.081	14.149	69	-	TTC		4
CYTB CDS	14.154	15.293	1.14	+		ATG/AGA	3
tRNA-Thr	15.297	15.366	70	+	TGT		-1
tRNA-Pro	15.366	15.431	66	-	TGG		0
D-loop	15.432	16.644	1.213	+			0

Tablo 30. Yaban keçisi mitogenomun karakterizasyon tablosu

Gen/Bölge	Başlama	Bitiş	Uzunluk	Yön	Anti-kodon	Başlama-Bitiş kodonu	İntergenik bölge
tRNA-Phe	1	68	68	+	GAA		0
s-rRNA	69	1.024	956	+			0
tRNA-Val	1.025	1.091	67	+	TAC		0
l-rRNA	1.092	2.663	1.572	+			0
tRNA-Leu	2.664	2.738	75	+	TAA		2
ND1 CDS	2.741	3.696	956	+		ATG/TA-	0
tRNA-Ile	3.697	3.765	69	+	GAT		-3
tRNA-Gln	3.763	3.834	72	-	TTG		2
tRNA-Met	3.837	3.905	69	+	CAT		0
ND2 CDS	3.906	4.947	1.042	+		ATA/T--	0
tRNA-Trp	4.948	5.014	67	+	TCA		1
tRNA-Ala	5.016	5.084	69	-	TGC		1
tRNA-Asn	5.086	5.158	73	-	GTT		2
rep_orj	5.161	5.191	31	+			-1
tRNA-Cys	5.191	5.258	68	-	GCA		0
tRNA-Tyr	5.259	5.326	68	-	GTA		1
COX1 CDS	5.328	6.872	1.545	+		ATG/TAA	-3
tRNA-Ser	6.87	6.938	69	-	TGA		7
tRNA-Asp	6.946	7.013	68	+	GTC		1
COX2 CDS	7.015	7.698	684	+		ATG/TAG	3
tRNA-Lys	7.702	7.768	67	+	TTT		1
ATP8 CDS	7.77	7.967	198	+		ATG/TAG	-40
ATP6 CDS	7.928	8.608	681	+		ATG/TAA	-1
COX3 CDS	8.608	9.391	784	+		ATG/T--	0
tRNA-Gly	9.392	9.46	69	+	TCC		0
ND3 CDS	9.461	9.807	347	+		ATA/TA-	0
tRNA-Arg	9.808	9.876	69	+	TCG		0
ND4L CDS	9.877	10.173	297	+		ATG/TAA	-7
ND4 CDS	10.167	11.544	1.378	+		ATG/T--	0
tRNA-His	11.545	11.614	70	+	GTG		0
tRNA-Ser	11.615	11.674	60	+	GCT		1
tRNA-Leu	11.676	11.745	70	+	TAG		0
ND5 CDS	11.746	13.566	1.821	+		ATA/TAA	-17
ND6 CDS	13.55	14.077	528	-		ATG/TAA	0
tRNA-Glu	14.078	14.146	69	-	TTC		4
CYTB CDS	14.151	15.29	1.14	+		ATG/AGA	3
tRNA-Thr	15.294	15.363	70	+	TGT		-1
tRNA-Pro	15.363	15.428	66	-	TGG		0
D-loop	15.429	16.622	1.194	+			0

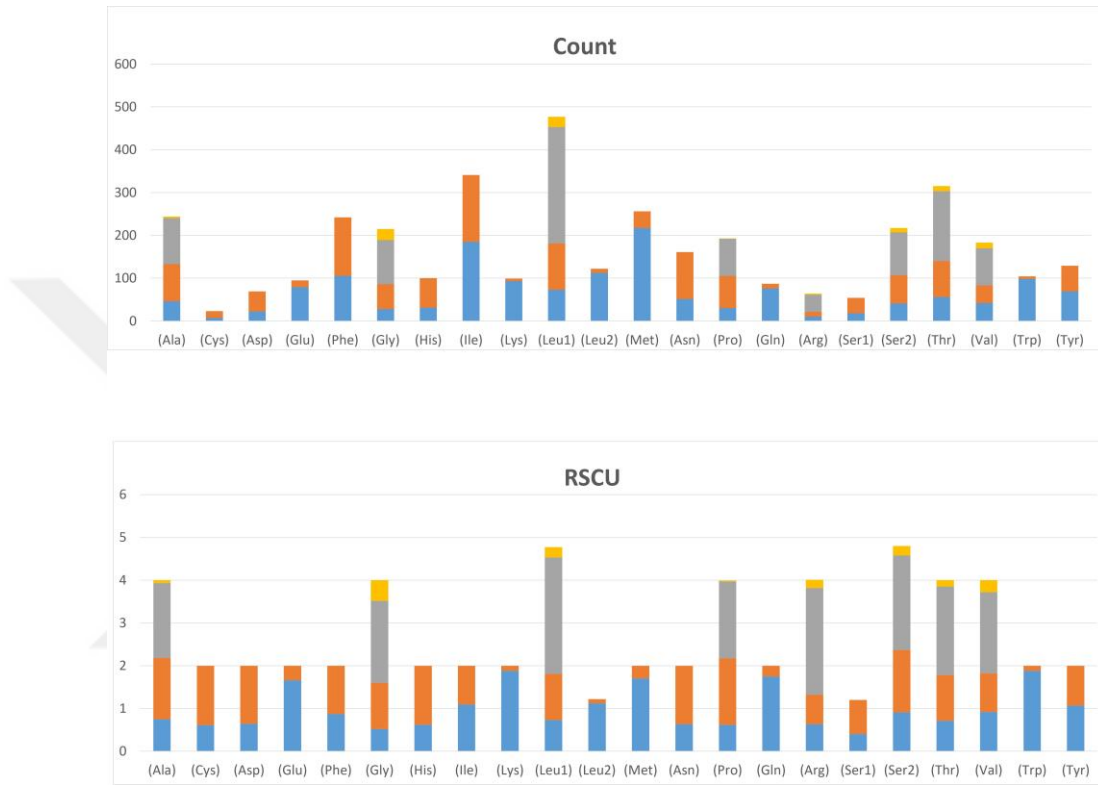
Ankara keçisi mitogenomundaki protein kodlayan genler, tüm mitogenomun %68,49'unu oluşturmakta olup, toplam 11.401 bç uzunluğundadır. PCG'lerin eksik sonlanma kodonları (5 adet) hariç toplam 3.798 amino asit kodlamaktadır. Protein kodlayan genlerden 10 tanesi standart ATG kodonu ile başlarken, ND2 ve ND3 ve ND5 genleri ATA kodonu ile başlamaktadır. Bununla birlikte 7 protein kodlayan gen TAA kodonu ve CYTB geni AGA kodonu ile sonlanırken, ND1 ve ND3 geni TA- eksik sonlanma kodonu ile ND2, ND4 ve COX3 genleri T-- eksik sonlanma kodonu ile bitmektedir.

Yaban keçisi mitogenomundaki protein kodlayan genler ise, tüm mitogenomun %68,58'ini oluşturmakta olup, toplam 11.401 bç uzunluğundadır. PCG'lerin eksik sonlanma kodonları (5 adet) hariç toplam 3.798 amino asit kodlamaktadır. Protein kodlayan genlerden 10 tanesi standart ATG kodonu ile başlarken, ND2 ve ND3 ve ND5 genleri ATA kodonu ile başlamaktadır. Bununla birlikte 5 protein kodlayan gen TAA kodonu, COX2 ile ATP8 geni TAG ve CYTB geni AGA kodonu ile sonlanırken, ND1 ve ND3 geni TA- eksik sonlanma kodonu ile ND2, ND4 ve COX3 genleri T-- eksik sonlanma kodonu ile bitmektedir.



Şekil 43. Ankara keçisi RSCU-Kodon kullanımı grafiği

Mitogenomların göreceli sinonim kodon kullanımı (Relative Synonymous Codon Usage, RSCU) ve kodon kullanım miktarları Şekil 42 ve 43’de verilmiştir. Buna göre her iki keçi için de üç amino asit, lösin, izolösin ve treonin en yüksek frekansa sahipken, aspartik asit, arginin ve sistein en düşük frekansa sahiptir.

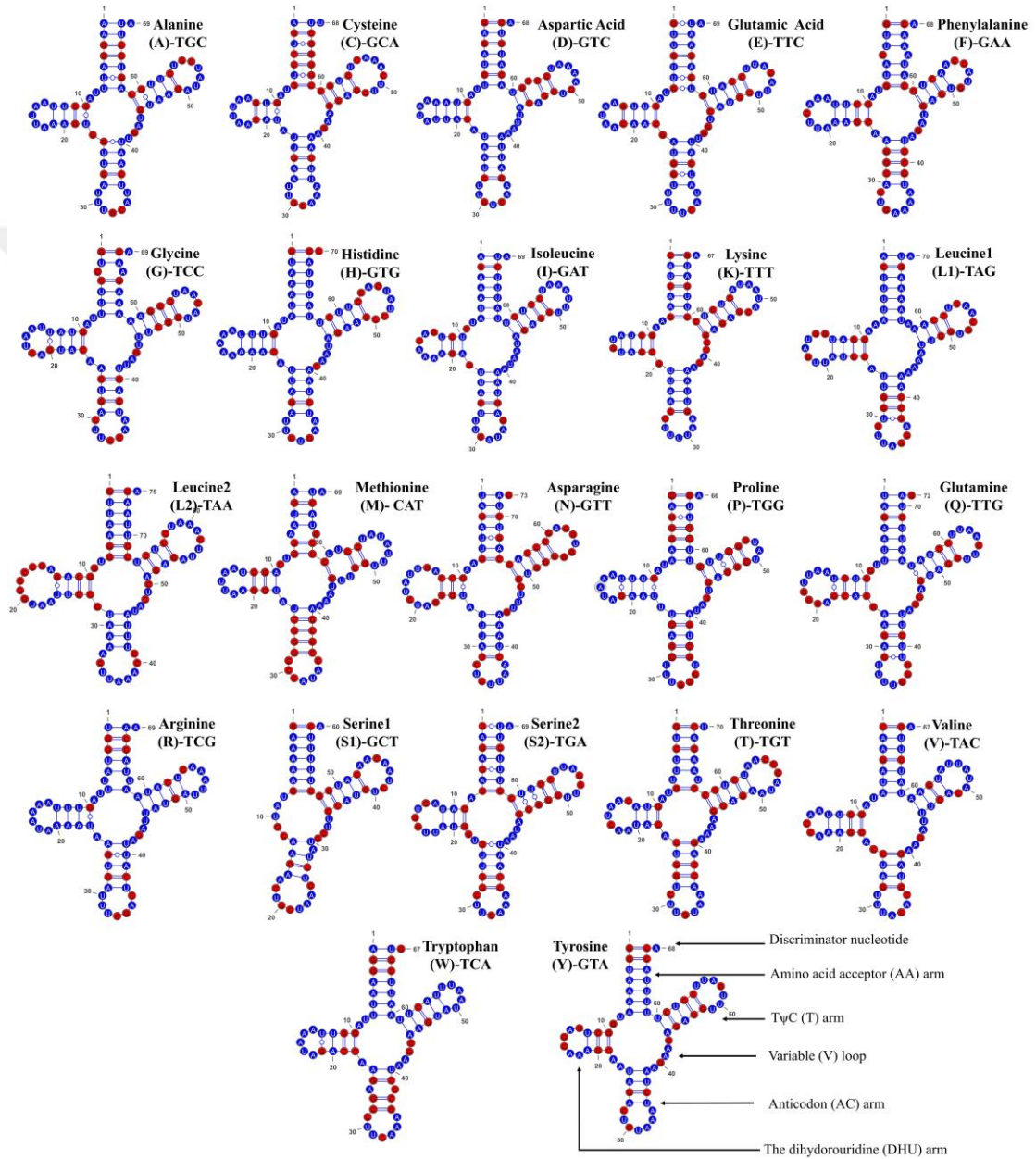


Şekil 44. Yaban keçisi RSCU-Kodon kullanımı grafiği

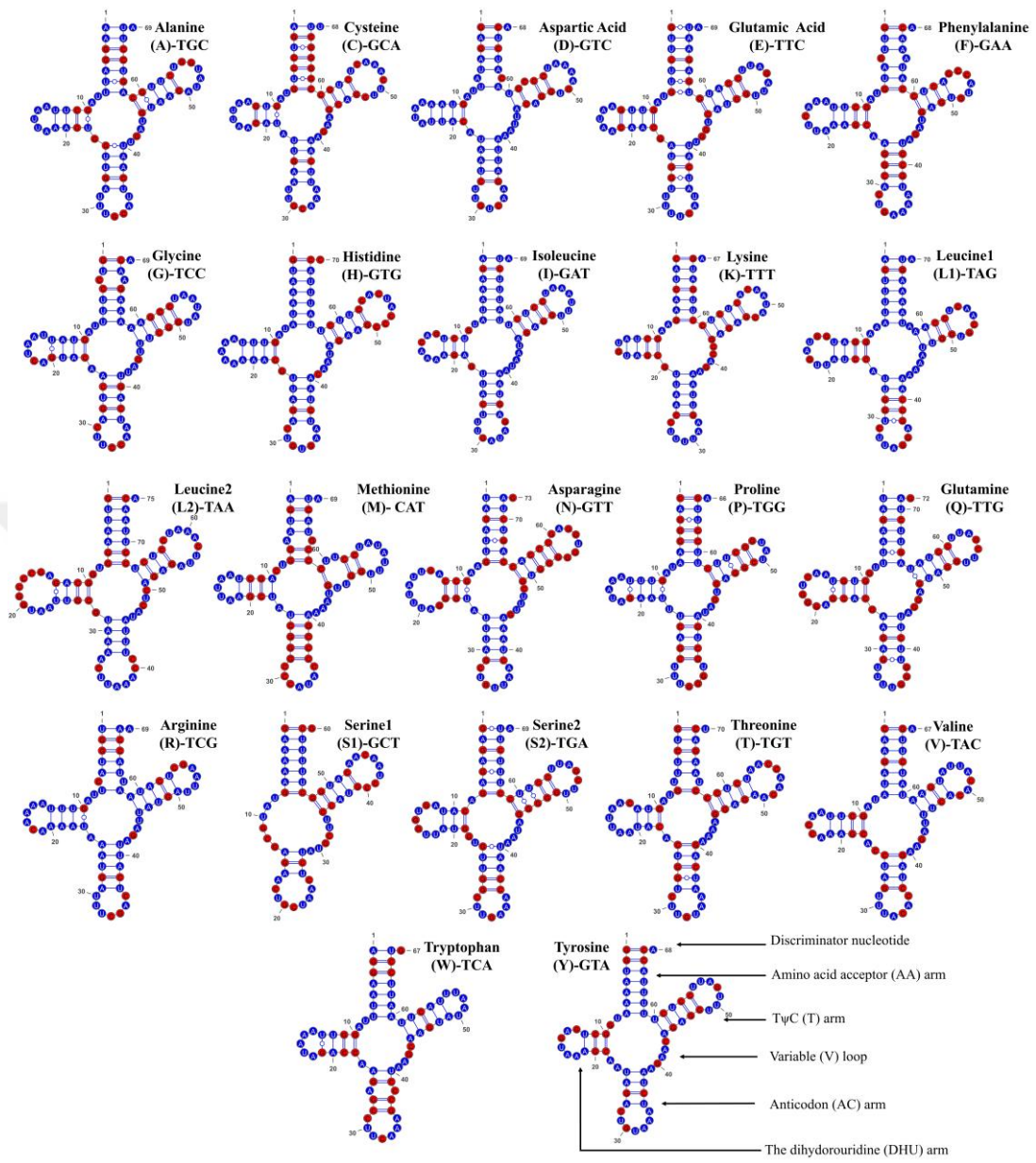
İki ribozomal RNA geni, Ankara keçisinde 956 bç (s-rRNA) ve 1.575 bç (l-rRNA) uzunlukta iken yaban keçisinde 956 bç ve 1.572 bç uzunluğundadır. İki rRNA geni birbirinden tRNA^{Val} geni ile ayrılmakta olup, tRNA^{Phe} ve tRNA^{Leu}(UUR) arasında yer almaktadır. Replikasyon orjini (OL), her iki keçide de 31 bç uzunluğunda ve WANCY tRNA bölgesinde yer almaktadır. Kodlama yapmayan kontrol bölgesi (D-loop) ise, tRNA^{Phe} ve tRNA^{Pro} tRNA genleri arasında yer almakta ve Ankara keçisi için 1.213 bç, yaban keçisi için 1.194 bç uzunluğundadır.

3.11.2. Mitogenom tRNA yapıları

Taşıyıcı RNA'lar (tRNA), Ankara keçisi ve yaban keçisi mitogenomlarında 60-75 bp uzunluğundadır. Dihidrouridin (DHU) kolu bulunmayan tRNA^{Ser1} hariç, diğer tRNA'ların ikincil yapıları tipik yonca yaprağı yapısı şeklindedir (Şekil 45 ve 46).



Şekil 45. Ankara keçisi mitogenomundaki tRNA'ların tahmini sekonder yapıları



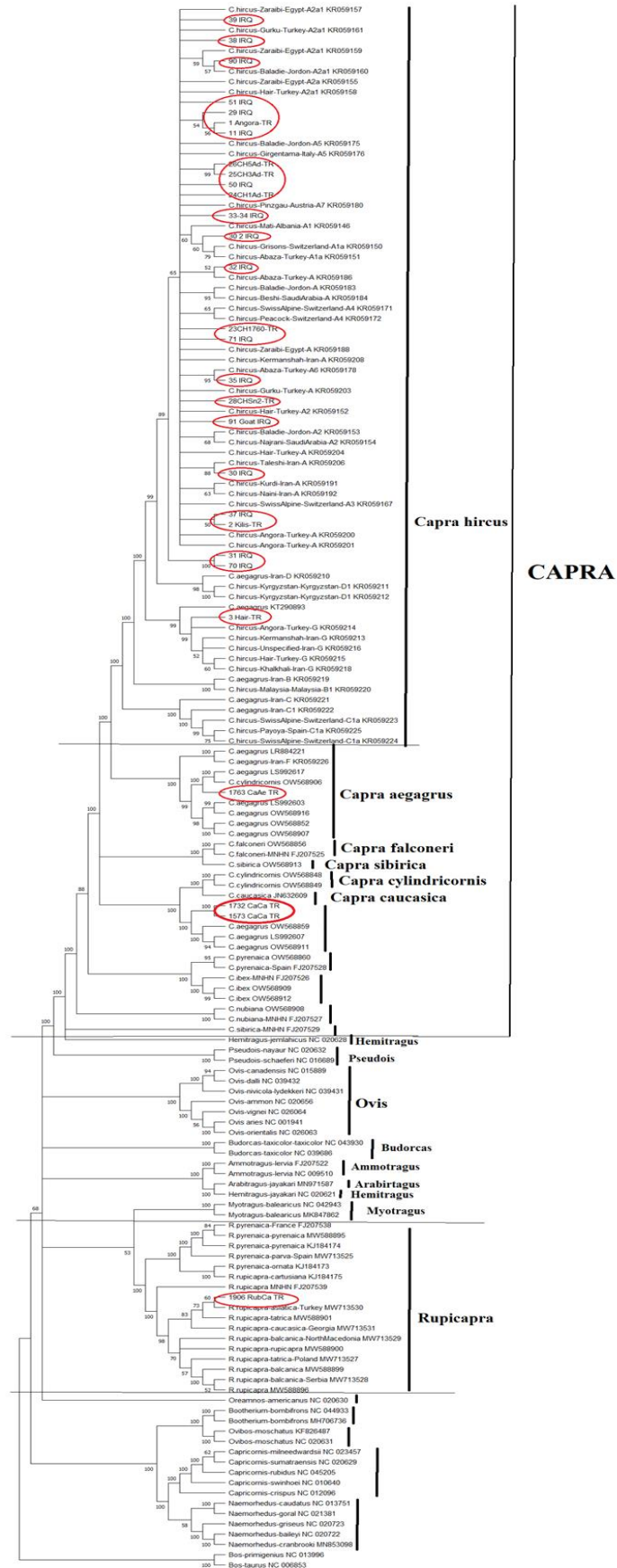
Şekil 46. Yaban keçisi mitogenomundaki tRNA'ların tahmini sekonder yapıları

3.11.3.Mitogenomlar temelli filogenetik analizler

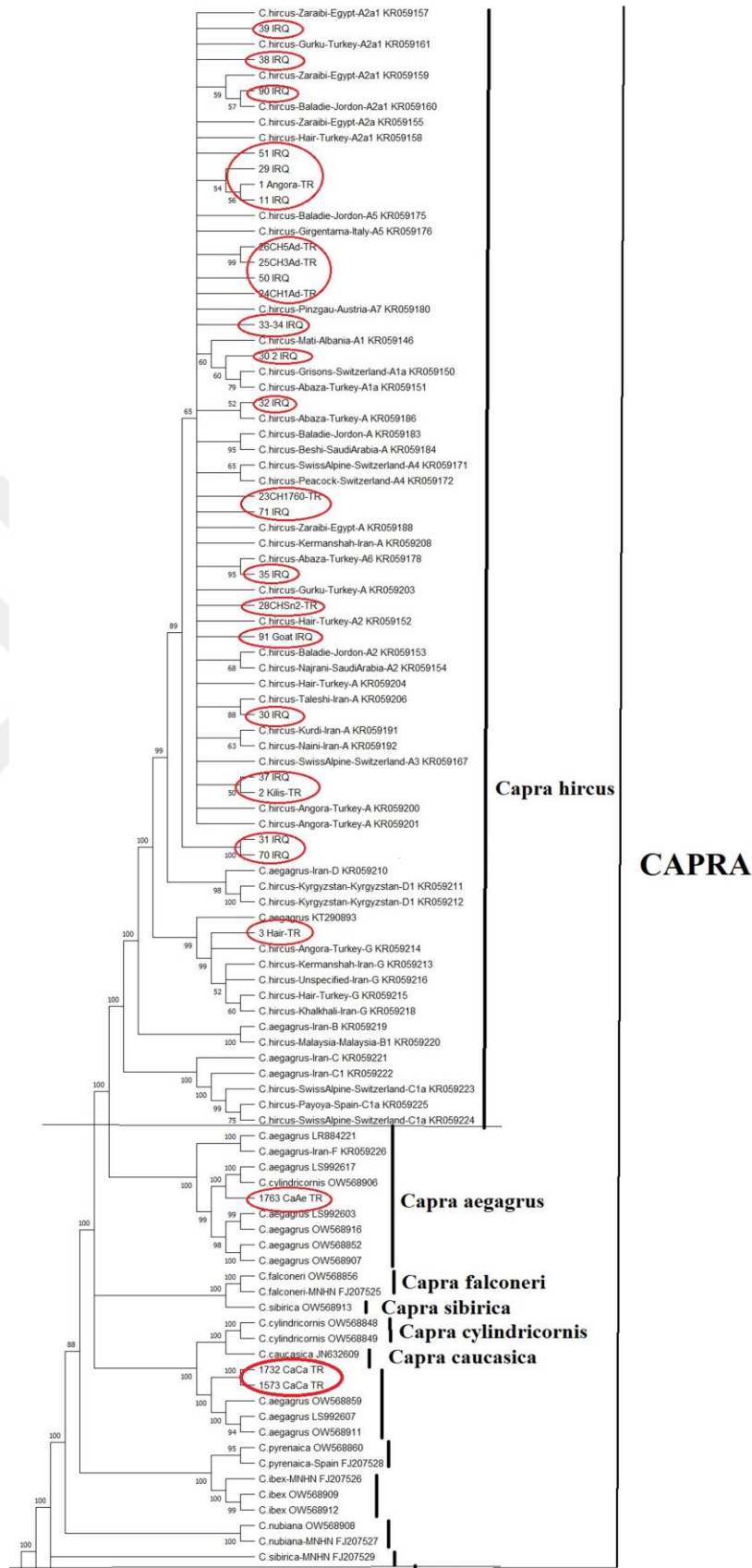
Caprinae alt familyasındaki türler (toplam 150 örnek ve 2 dış grup) kullanılarak kurulan filogenetik ağaç Şekil 47 ve 48’de gösterilmiştir. Filogenetik ağaçta tür bazlı ayrılmalar gerçekleşmiş olup, İran’dan bildirilen iki yaban keçisi aksesyonunun domestik keçilerle birlikte gruplandığı görülmüştür.

Bu çalışmada elde edilen Ankara keçisi ve diğer evcil keçi örnekleri, diğer evcil keçilerle birlikte gruplanmıştır. Ankara keçisi örneği filogenetik ağaçta Irak’tan bildirilen örnekler ile birlikte yer alırken Kıl keçisi örneği, GenBank’dan indirilen Ankara keçisi, Kıl keçisi ve İran’dan bildirilen keçiler ile; Kilis keçisi ise Irak’tan bildirilen keçi ile farklı dallarda grup oluşturmuştur (Şekil 47 ve 48).

Bu tez çalışmasında sunulan yaban keçisi örnekleri ise GenBank’daki diğer yaban keçileri ile birlikte kümelenmiş olup, filogenetik ağaçta üç farklı grubun olduğunu ortaya koymuştur (Şekil 45 ve 46).



Şekil 47. ML yöntemi kullanılarak oluşturulan filogenetik ağaç



Şekil 48. ML Filogenetik ağacın büyütülmüş hali

3.12. Ham okuma verilerinin NCBI-SRA veri tabanına yüklenmesi

Çalışma kapsamında elde edilen ham okumalar fastq dosya formatında elde edilerek diğer araştırmacıların kullanımına sunulmak üzere NCBI-SRA veri tabanına (<https://www.ncbi.nlm.nih.gov/sra>) SRR24084427 (BioProject No: PRJNA952661) erişim numarası ile kaydedilerek erişime açılmıştır.



4. BÖLÜM

TARTIŞMA-SONUÇ VE ÖNERİLER

4.1. Tartışma

Genomu ilk dizilenen model organizmalar Meyve sineği (*Drosophila melanogaster*) (Adams vd., 2000), nematod (*C. elegans*) (*C. elegans* Sequencing Consortium, 1998), *Arabidopsis* (The Arabidopsis Genome Initiative, 2000) ve insan (Venter vd., 2001) olup, son 20 yılda birçok canlının genomu dizilenmiştir. Keçiler ülkemiz için önemli ekonomik türlerdendir. Bu tez çalışması kapsamında ülkemizin önemli milli değerlerinden olan Ankara keçisi (*Capra hircus*) ve yaban hayatı bakımından oldukça önemli olan Anadolu yaban keçisi (*Capra aegagrus*) tüm genomları dizilenmiş, anotasyonları gerçekleştirilmiş, sentenik ve filogenetik ilişkileri araştırılmıştır. Çalışma kapsamında elde edilen veriler ile literatür bilgileri aşağıdaki başlıklar altında tartışılmıştır.

4.1.1. Genom birleştirme

Dong vd. (2013), dişi bir Yunnan siyah keçisi genomunu yaklaşık 2,66 Gb olarak yayınlamışlar ve bu çalışmada kısa okuma dizileme ve optik haritalama yöntemlerini kullanmışlardır. Illumina ile 191,5 Gb yüksek kaliteli okuma (tahmini genom boyutunun 65,6X kapsamında) veri elde edilmiş, genom birleştirmede SOAPdenovo version 1.03 yazılımı (Li vd., 2010) kullanılmış olup skafold N50 uzunluğu 3,06 Mb, super-skafold N50 uzunluğu 16,3 Mb olarak bildirilmiştir (Dong vd., 2013). Siyah Bengal keçisi genomu, Illumina HiSeq 2500 platformunda 150 bp uzunlukta okumalarla dizilenmiştir. Üretilen 40 Gb veri ile (14X kapsam) *de novo* birleştirme ABySS v.2.1.5 assembler algoritması (Simpson vd. 2009) kullanılarak yapılmış, genom 3,04 Gb büyüklüğünde birleştirilmiş ve BUSCO analizi birleştirmenin %82,5 bütünlükte olduğunu ortaya koymuştur (Siddiki vd., 2019). Bickhart vd. (2017), kontig formasyonu için uzun okuma,

konsensus doğrulama için kısa okuma, optik ve kromatin etkileşim haritalama ile skafoldlama ile San Clemente keçi ırkı için *de novo* genom birleştirme gerçekleştirmişlerdir. Bu yöntemle sadece 649 boşluk içeren son derece yüksek kromozom uzunluğunda skafoldlar elde etmişler ve genomu skafold N50 değeri 87 Mb ile 2,92 Gb olarak elde etmişlerdir (Bickhart vd., 2017). Dong vd. (2015) *Capra aegagrus* referans genomunu oluşturmak için Hiseq 2000 platformu ile 250 bp – 5 kb insört uzunluklarda çift uçlu okumalar ve 10 kb ile 20 kb'lik eş-çift okuma kütüphaneleri (yaklaşık 108,5X çift uçlu ve 27,7 eş-çift okuma kapsamı, toplamda 381,50 Gb veri) üretmişlerdir. Genom birleştirme SOAPdenovo (Li vd., 2010) yazılımı ile yapılmıştır. Elde edilen genom olan CapAeg_1.0'ın kontig N50 boyutunu 18.965 bp ve skafold N50 boyutunu 2.057.686 bp olarak bildirmişlerdir (Dong vd., 2015).

Boz ayı (*Ursus arctos*), genomunu elde etmek için yapılan çalışmada Hifiasm algoritması (Cheng vd., 2021), ve yak algoritması (<https://github.com/lh3/yak/releases/tag/v0.1>) kullanılarak elde ettikleri Pacific Biosciences (PacBio) Hi-Fi (32X kapsam) ile veri tabanlarındaki mevcut Hi-C verilerini birlikte analiz ederek erkek Kuzey Amerika boz ayısının kromozom düzeyinde birleştirmesini bildirmişlerdir. Final genom boyutu 2,47 Gb olan birleştirmenin skafold N50 değeri 70,08 Mb ve kontig N50 değeri 43,94 Mb'dır. BUSCO analizi, "Mammalia" veri tabanındaki tek kopya ortologların %94,5'inin genomda bulunduğunu ortaya çıkarmıştır (Armstrong vd., 2022). Edwards vd. (2021), iki Basenji köpeğinin *de novo* genom birleştirmesini gerçekleştirmiş, ilk olarak dişi Basenji köpeğinin genomunu 84,5 Gb Oxford Nanopore PromethION sistemi ile yaklaşık 2,41 Gb genom boyutu için 35X kapsamda birleştirmişlerdir. Ek olarak yaklaşık 47,7X kapsamda 115,1 Gb BGISEQ kısa okuma hata düzeltmede kullanılmıştır. Hi-C verileriyle kontigler daha uzun skafoldlara birleştirilmiştir. Erkek olan Basenji köpeğinin genomu ise PacBio dizileme ile yaklaşık 45X kapsamda dizilenmiştir.

Gonçalves vd. (2021), bir erkek karaçalı antilobun tüm genom dizisinin ilk *de novo* birleştirmesini yapmışlardır. 10X Genomics Chromium yöntemi kullanarak $1,2 \times 10^9$ çift uçlu okuma Supernova v.2.1.1 (Weisenfeld vd., 2017) yazılımıyla N50 değeri 8,42Mb olan ve %91,2'lik BUSCO tamlığına sahip 16.880 skafolddan oluşan 2,56 Gb'lik bir taslak genom (tahmini genom boyutu 3,01Gb) elde etmişlerdir. Karimi vd. (2022), PacBio (21X), Hi-C (57X) ve kısa okunmalar (BGI, 38X) kullanarak Amerikan vizonu için yüksek kaliteli bir kromozom düzeyinde genom birleştirmesi yapmışlardır. Bu

birleştirmenin ilk taslağında 2.884.047 PacBio uzun okuması kullanmışlardır. Bu genom birleştirme çalışmasında WTDBG2, Flye, Hifiasm, Hicanu ve IPA birleştirme yazılımları denenmiş ve en iyi birleştirmenin Hifiasm ile elde edildiğini raporlamışlardır. Daha sonra BGI kısa okumaları ile Polca algoritmasını kullanılarak polishing yapmışlar ve Hi-C verilerinin ilk taslağa entegrasyonu ile 183 skafold içeren 220 Mb'lık N50 değeri ile bir birleştirme elde etmişlerdir. Amerikan vizonunun bu boşluksuz genom birleştirmesi (ASM_NN_V1), tüm genomun yaklaşık %98,6'sının 15 kromozom tarafından kapsandığı 2,68 Gb uzunluğa sahiptir.

Blanchard vd. (2023) koala popülasyonlarının karşılaştırmalı genomiklerinde ve retroviral entegrasyon çalışmalarında kullanılmak üzere Güney Avustralya koalas olan "Wilpena"nın genomunu elde etmişlerdir. 58 GB ONT verisi (ortalama okuma uzunluğu 24 kb ve ortalama Q skoru 13,7 olan 2.572.260 okuma) ve 1,289 milyon (2x150bp >Q30) Illumina okuması ile Flye algoritması kullanılarak birleştirilmiştir. Birleştirilen genom 3.233.824.327 bp olup, Medaka algoritması kullanılarak ilk polishing ve Illumina verileri kullanılarak Polca algoritması ile son polishing yapılmıştır. Sonuç olarak 48.800.306 N50 değerine sahip, 1.265 adet kontigden oluşan toplam genom boyutu 3.234.982.288 bp'lık bir uzunluk ile yüksek kaliteli final genom birleştirmesi elde etmişlerdir. Palmada-Flores vd. (2022) yaptıkları çalışmada *L. catta* (mLemCat1) için Omurgalı Genomları Projesi standartlarına uygun yeni bir yüksek kaliteli referans genom birleştirmesi oluşturmuşlardır. Tahmin edilen genom boyutu 3,21 Gb olarak oluşturulan mLemCat1 genomunun birleştirilmesi için 86,43X kapsama sahip 10X bağlantılı okuma verisi, 66,68X kapsama sahip Arima verisi, 154,57X kapsama sahip Bionano verisi ve 62,88X kapsama sahip PacBio verisi üretilmiştir. 90,982 Mb skafold N50 değeri ve 10,570 Mb'lık kontig N50 değeri ile birleştirmenin bitişikliği ve bütünlüğü son derece yüksektir (Palmada-Flores vd., 2022).

Li vd. (2022) Illumina dizileme, PacBio dizileme platformları ve kromozom yapı yakalama tekniğini kullanarak *Tamias sibiricus*'un kromozom ölçeğinde genomu birleştirmiştir. Bu çalışmada 132,39 Gb Illumina, 111,63 Gb PacBio, 217,38 Gb Hi-C, 3,89 Gb RNA-seq verisi kullanmışlardır. PacBio okumalarının birleştirilmesinde Canu v1.8 algoritması ve Illumina okumalarının polish edilmesinde ise Pilon v1.23 algoritmasını kullanmışlardır. 3D-DNA ve Juicebox v1.11.08 algoritmaları ile Hi-C okumaları kullanılarak yapılan atamalarda 172,61 Mb skafold N50 uzunluğu ile 2,64 Gb

boyutunda genom elde edilmiştir. Bu çalışmada toplam 2,59 Gb genom verisi, %98,03'e varan bir oranla 19 kromozoma (28,70 ila 222,90 Mb aralığında) atanmış ve oryante edilmiştir. Jebb vd. (2020), altı yaras türünün (*Rhinolophus ferrumequinum*, *Rousettus aegyptiacus*, *Phyllostomus discolor*, *Myotis myotis*, *Pipistrellus kuhlii* ve *Molossus molossus*) ilk referans kalitesinde genomlarını oluşturmak için uzun okuma verileri (100X PacBio okuması) ve “state of-the-art scaffolding” protokolleri (Teeling vd., 2018) için Illumina 10x Genomiks ve Hi-C konfirmasyon yakalama yöntemlerini kullanmışlardır. Çalışma sonucunda altı yaras türünün tamamı için yüksek bitişikliğe sahip birleştirme sonuçları elde etmişlerdir. Bu kapsamda kromozom seviyesinde skafoldlarda her bir birleştirme için %96-99 oranında kromozom bütünlüğü (N50 değeri 92–171,1 Mb) elde etmişlerdir.

Genom birleştirme aşaması, genom çalışmalarının en temel aşamasıdır. Yüksek oranda bitişikliğe, bütünlüğe ve doğruluğa sahip genomlar, daha sonra yapılacak olan genom dizileme çalışmalarında referans olarak kullanılmaları bakımından oldukça önemlidir. Bu tez çalışması kapsamında elde edilen Ankara keçisi genomu, yaklaşık 300 Gb Illumina çift uçlu kısa okuması ve yaklaşık 32 Gb Oxford Nanopore uzun okuması ile gerek keçiler gerekse de diğer evcil ve yabani memeli hayvan genom çalışmalarına uyumlu olarak birleştirilmiştir. Yapılan birleştirme sonucunda elde edilen Ankara keçisi genomunun uzunluğu yaklaşık 2,6 Gb olarak bulunmuş olup, literatürde rapor edilen diğer keçilerle (Yunnan siyah keçisi genomu 2,66 Gb (Dong vd. 2013), Siyah Bengal keçisi genomu 3,04 Gb (Siddiki vd., 2019), San Clemente keçisi 2,92 Gb (Bickhart vd., 2017)) uyumludur.

Çalışma kapsamında yapılan birleştirme toplam 1.738 kontigden oluşmaktadır. N50 değeri 10,94 Mb ve en uzun fragman 57 Mb olup N50 değeri, birleştirilen keçi ve memeli (Yunnan siyah keçisi genomu skafold N50 değeri 3,06 Mb, super-skafold N50 değeri 16,3 Mb (Dong vd., 2013), yaban keçisi genomu kontig N50 boyutu 18.965 bç ve skafold N50 boyutu 2.057.686 bç (Dong vd., 2015), Basenji genomu skafold N50 değeri 8,42Mb (Edwards vd., 2021)) genomları ile benzerdir. Yine BUSCO ile yapılan genomun bütünlüğünün değerlendirilmesinde %92,5'lik bir bütünlük görülmüştür; bu da birleştirilen memeli genomlarının bütünlük değerleri (Siyah Bengal keçisi genomu için BUSCO bütünlük değeri %82,5 (Siddiki vd., 2019), karaçalı antilobu genomu için BUSCO bütünlük değeri %91,2 (Gonçalves vd., 2021), boz ayı genomu BUSCO değeri

%94,5 (Armstrong vd., 2022)) ile benzerdir. Ankara keçisinin skafoldlaması sonucu 2.604.753.723 bç uzunluğunda 710 skafold elde edilmiştir. Skafoldlanan genomun N50 değeri ~95 Mb olup BUSCO ile değerlendirilmiş ve %95,7 bütünlük değeri göstermiştir. 710 skafolddan kromozomal seviyede olan 31 skafold BUSCO ile değerlendirilmiş ve genom bütünlüğü %94,9 olarak elde edilmiştir. Ankara keçisi genomunun N50 değeri ve BUSCO sonuçları, diğer keçi (Yunnan siyah keçisi (Dong vd. 2013), Siyah Bengal keçisi (Siddiki vd., 2019), San Clemente keçisi (Bickhart vd., 2017)) ve memeli hayvanlar (Luchuan domuzu (Yang vd., 2019a), Karaçalı antilobu (Gonçalves vd., 2021), beyaz kuyruklu geyik (London vd., 2022), koala (Blanchard vd., 2022)) ile karşılaştırıldığında oldukça yüksektir.

Tez çalışmasındaki bir diğer keçi olan yaban keçisi ise yaklaşık 210 GB Illumina çift uçlu kısa okuması kullanılarak *de novo* yöntemle birleştirilmiştir. 181,575 kontigden oluşan 2,7 Gb boyutunda bir birleştirmede uzun kontig 332.737 bç ve N50 değeri 36.402 bç olarak yüksek kontig sayısı ve nispeten düşük N50 değeri ile elde edilmiştir. Yaban keçisinin kontig seviyedeki BUSCO değerlendirmesi ise %67,5 bütünlük değeri ile sonuçlanmıştır. Yaban keçisi genomunun skafoldlanması ile toplam uzunluğu 2.634.549.116 bç olan 114.564 skafoldun (N50 değeri ~90 Mb) BUSCO değerlendirmesi %94,7 olarak elde edilmiştir. Ek olarak kromozomal seviyede olan 31 skafoldun BUSCO bütünlüğü %90,7 olarak bulunmuştur. Bu kapsamda elde edilen yaban keçisi genom birleştirmesi diğer keçi (Yunnan siyah keçisi (Dong vd. 2013), Siyah Bengal keçisi (Siddiki vd., 2019), San Clemente keçisi (Bickhart vd., 2017)) ve memeli hayvanlar ((Luchuan domuzu (Yang vd., 2019a), Karaçalı antilobu (Gonçalves vd., 2021), beyaz kuyruklu geyik (London vd., 2022), koala (Blanchard vd., 2022)) ile tutarlıdır.

4.1.2. Genom anotasyonu

Yunnan siyah keçisinin genomunu dizilemek için yapılan bir çalışmada Dong vd. (2013), elde ettikleri genom birleştirmede 10 dokudan izole edilen RNA verileriyle birlikte 22.175 protein kodlayan gen annotate etmişlerdir (Dong vd., 2013). *Bubalus depressicornis* genomunun elde edilmesi için yapılan bir çalışmada Porrelli vd. (2022), yaptıkları genom birleştirmenin GC içeriğini %41,74 olarak rapor etmişler ve genomun %42,12'sinin tekrar bölgesi olduğunu belirtmişlerdir (Porrelli vd., 2022). Low vd. (2019) manda (*Bubalus bubalis*) genomunu dizilemişler ve anotasyonda 50 farklı dokudan yaklaşık 15

milyar RNA-seq okuma verisi kullanarak 20.801 protein kodlayan gen, 8.443 kodlanmayan gen ve 4.465 pseudogen belirlemişlerdir (Low vd., 2019). Gonalves vd. (2021), Karaalı antilobu (*Hippotragus equinus*) genomunun dizilemek amacıyla yaptıkları bir alıřmada genomun 1,1 Gb (%42,2) tekrarlayan diziler ierdiğini ve *de novo* anotasyon ile Augustus v.3.3.2 algoritması (Stanke vd., 2008) kullanılarak 20.518 protein kodlayan gen tanımlanmışlardır (Gonalves vd., 2021).

Boz ayı (*Ursus arctos*) genomunun elde edilmesi iin yapılan bir alıřmada tekrarlayan elementlerin genomun %44,48'ini oluřturduėu ve toplam 20.480 protein kodlayan gen ierdiğini belirlenmiştir (Armstrong vd., 2022). Kuzey Amerika'ya özgü etil bir memeli olan Amerikan vizonu (*Neogale vison*) üzerine yapılan bir alıřmada NCBI Ökaryotik Genom Anotasyon hattı kullanılarak toplamda 25.377 gen tahmin edilmiştir (Karimi vd., 2022). Orta Afrika kökenli eski bir köpek ırkı olan ve Afrika Kongo'daki kabilelerle birlikte yařayıp avlanan Basenji köpeklerinde yapılan bir genom alıřmasında Edwards vd. (2021), iki Basenji köpeğinin genomunda homoloji temelli yaklařımla GeMoMa v1.6.2beta algoritmasını kullanarak 27.129 ve 27.783 protein kodlayan gen tahmin etmişlerdir (Edwards vd., 2021).

Amerika'da yaygın olarak yayılıř gösteren beyaz kuyruklu geyik (*Odocoileus virginianus*) genomunun dizilenmesi iin yapılan bir alıřmada BRAKER algoritması kullanılarak 20.651 gen tahmin edilmiştir (London vd., 2022). Jebb vd. (2020), altı yarasa türünün genomlarını oldukça yüksek kalitede dizilemiş ve genomları birleřtirerek elde ettikleri genomlarda *de novo*, homoloji ve transkriptomik verilerle birlikte 19.122 ila 21.303 protein kodlayan geni annotate ettiklerini bildirmişlerdir (Jebb vd., 2020). Keseli memeli türlerinden olan Koala genomu üzerine yapılan bir alıřmada EggNOG mapper algoritması kullanılarak fonksiyonel anotasyon iřlemiyle transkripsiyon desteėine sahip toplam 27.669 gen tanımlanmıştır (Blanchard vd., 2022).

Li vd. (2022), Sibiry sincabı (*Tamias sibiricus*) genomunu elde etmek iin yaptıkları bir alıřmada 25.311 protein kodlayan tahmin etmişler ve bu genlerin %94,73'ünün fonksiyonel anotasyonunu gerekleřtirmişlerdir. Ayrıca alıřmada genomun %38,87'sini oluřturan ve toplam boyutu 1,03 Gb olan tekrar dizileri de tespit edilmiştir (Li vd., 2022). Sarı gerdanlı sansar (*Martes flavigula*) genomu üzerine yapılan bir alıřmada *ab initio* tahmin, homoloji temelli tahmin ve transkriptom-temelli yöntemler kullanılarak 20.464

protein kodlayan tahmin edilmiş olup bunlardan 20.322'sinin (%99,31) fonksiyonel anotasyonu başarıyla gerçekleştirilmiştir (Mei vd., 2023). Andreotti vd. (2022) yaptıkları bir çalışmada *Phodopus roborovskii* genomunu dizileyerek elde ettikleri genomda RNA-seq ve GeMoMa algoritması kullanılarak 21.836 transkript tespit etmişler ve 17.844 gen lokusunu annotate etmişlerdir (Andreotti vd., 2022).

Genom birleştirme aşamasından sonra genom anotasyonu altında yapılan tekrar elementlerinin belirlenmesi ve bunların maskelenmesi işlemi genlerin belirlenmesinde doğru bir şekilde gerçekleştirilebilmesi için elzemdir. Organizmadaki genlerin belirlenmesi ve bunların fonksiyonel olarak tanımlanması, genom çalışmalarının bir diğer önemli basamağıdır. Bu tez çalışması kapsamında yapılan Ankara keçisi genom birleştirmesinde GC içeriği %41,98 olup, tekrar içeriği %41,66 olarak belirlenmiştir. Belirlenen tekrar içeriğinin toplam uzunluğu 1.089.319.018 bp'dir. Tekrar içeriği maskelenen Ankara keçisi genomunda homoloji temelli yöntem kullanılarak 21.645 gen tahmin edilmiştir. Tahmin edilen genlerin %74,46'lık kısmını oluşturan 16.117 genin fonksiyonel anotasyonu gerçekleştirilmiştir. Tez çalışmasına dahil edilen yaban keçisinde ise yapılan genom birleştirmesinde GC içeriği %42,13 olup, tekrar içeriği %40,83 olarak belirlenmiştir. Belirlenen tekrar içeriğinin toplam uzunluğu 1.103.810.083 bp'dir. Tekrar içeriği maskelenen yaban keçisi genomunda homoloji temelli yöntem kullanılarak 21.409 gen tahmin edilmiştir. Tahmin edilen genlerin %74,56'lık kısmını oluşturan 15.962 genin fonksiyonel anotasyonu gerçekleştirilmiştir. Elde edilen Ankara keçisi ve yaban keçisi genom anotasyonları yukarıda rapor edilen diğer memeli türler ile uyumludur.

4.1.3. Mitogenom organizasyonu

Tez çalışması kapsamında elde edilen mitogenomların uzunlukları 16.622 ila 16.644 bp arasında olup daha önceden bildirilen memeli mitogenomları ile (Hiendleder vd., 1998; Colli vd., 2015; Ran vd., 2016; İbiş, 2019) benzerdir. Elde edilen mitogenomlarda tipik memeli mitogenomları gibi A+T içeriği, G+C içeriğinden daha yüksektir. Ayrıca bu çalışmada elde edilen mitogenomlardaki çakışan bölgeler ve intergenik tekrarlar da diğer memeliler ile (İbiş, 2019; Pacha vd., 2021; Skorupski, 2022) benzerdir.

Mitogenomlardaki eksik sonlanma kodonları daha önceden bildirilmiş olan çalışmalar ile benzer olup eksik sonlanma kodonlarının da işlevsel olduğu bildirilmiştir (Boore, 1999;

Ojala vd., 1981). Çalışmada elde edilen keçi mitogenomlarındaki tRNA'lar ise tRNA-Ser1 hariç tipik yonca yaprağı yapısındadır ve diğer memeli mtDNA tRNA'ları ile (İbiş, 2019; Arbizu vd., 2022; Skorupski, 2022) uyumludur.

4.1.4. Varyant

Fernandes Júnior vd. (2020), 52 Nellore (*Bos primigenius indicus*) boğasında yaptıkları bir çalışmada örnek başına 18,50X kapsamında genom dizileme gerçekleştirmişler ve sığır referans genomunu (ARS_UCD1.2) ve Ensembl VEP yazılımını kullanarak toplam 35.753.707 SNP ve 4.492.636 indel tespit ederek fonksiyonel anotasyonlarını gerçekleştirmişlerdir. Gonçalves vd. (2021), Karaçalı antilobu (*Hippotragus equinus*) genomu dizilemişler ve elde ettikleri genoma ek olarak yaptıkları çalışmada beş yabani bireyin 9,8X kapsamda yeniden dizilenmesiyle birlikte filtrelenmiş halde $3,4 \times 10^6$ biallelik SNV kümesi rapor etmişlerdir (Gonçalves vd., 2021). Yerli ırklarımızdan biri olan Norduz keçisinde yapılan bir çalışmada birey başına 33,1X kapsamla keçinin genomunda 9.757.980 SNP, 1.536.715 indel ve 290 CNV tespit edilmiştir (Arslan, 2023).

Bu tez çalışmasında Ankara keçisi kısa okumalarının referans evcil keçi genomuna haritalanması ile toplamda 11.168.617 SNP tespit edilmiş ve bunlardan 10.775.084'ünün anotasyonu yapılmıştır. Bunlardan 231,463'ü indel, 431,767'ü substitution, 522,054'ü insersiyon, 543,088'i delesyon ve 9,046,712'si SNV olarak tespit edilmiştir. Yaban keçisi Illumina kısa okumalarının referans yaban keçisi genomuna haritalanması sonucu toplamda 10.935.456 varyant tespit edilmiştir. Bu varyantlardan 9.256.756'sı SNP, 575.054'ü insersiyon ve 559.390'ı delesyon olarak annotate edilmiştir.

4.1.5. Sinteni

Bu tez çalışması kapsamında Ankara keçisi ile yaban keçisi genomları arasında yüksek sintenik ilişki bulunmuştur. Benzer şekilde Dong vd. (2015) tarafından yaban keçisi genomu üzerine yapılan bir çalışmada da benzer şekilde evcil keçi ile yaban keçisi arasında yüksek oranda sinteni rapor edilmiştir (Dong vd., 2015).

4.1.6. Filogeni

Tüm genomlar ile oluşturulan filogenetik ağaç türler bazında ayrılmış olup, bu tez çalışmasında elde edilen Ankara keçisi örneği ile NCBI GenBank referans keçi genomu

olan San Clemente keçisi birlikte gruplanmıştır. Çalışmaya dahil edilen yaban keçisi ise NCBI GenBank referans yaban keçisi örneğine yakın yer almış fakat GenBank'daki daha önceden bildirilen diğer evcil keçilerle birlikte kümelenmiş olup tam bir ayırım gerçekleşmemiştir.

Mitogenomlar kullanılarak elde edilen filogenetik ağaçta bu çalışma kapsamında elde edilen mitogenomlardan olan Ankara keçisi, daha önceden Irak bölgesinden elde edilen evcil keçilerle birlikte kümelenmiş ve A haplogruplarına yakın olarak pozisyon almıştır. Diğer bir keçi olan Kıl keçisi ise G haplogrubunda yer alırken Kilis keçisi ise A haplogrubunda yer almıştır. Oluşturulan filogenetik ağaç, Colli vd. (2015) tarafından yapılan çalışmadaki filogenetik ağaç ve haplogruplar bakımından uyumludur. Çalışmaya dahil olan yaban keçileri ise NCBI GenBank'dan indirilen diğer yaban keçileri ile birlikte kümelenmiş olup, yaban keçileri arasında farklı üç grubun olduğunu göstermiştir. 1763 numaralı örnek diğer *Capra aegagrus*'lar ile ağacın farklı bir pozisyonunda kümelenirken 1732 ve 1573 numaralı örnekler birlikte başka bir *C. aegagrus* grubuyla birlikte kümelenmiştir. Yaban keçileri arasındaki bu durum Colli vd. (2015) tarafından yapılan çalışmadaki farklı haplogruplara (B, C, D ve F haplogrupları) benzer olup, tez çalışmasındaki sonuçlar ile uyumludur.

4.2. Sonuç ve Öneriler

Bu tez çalışması kapsamında Ankara keçisi ve yaban keçisi tüm genomlarının dizilenmesi, birleştirmesi, anotasyonu, sentenik ilişkileri, filogenetik ve varyant analizleri yürütülmüştür. Ülkemizde yetiştiriciliği yapılan ve milli değerlerimizden olan Ankara keçisi ile yaban hayatı ve çeşitlilik bakımından oldukça önemli olan Anadolu yaban keçisinin yüksek oranda bitişik genomları; moleküler ıslah ve koruma biyolojisine ek olarak evcilleştirme, yaban hayatı çeşitliliği, karşılaştırmalı genomik gibi ileriki çalışmalar için önemli bir kaynak oluşturmaktadır.

Ankara keçisi ve yaban keçisi genomlarına ek olarak ülkemizde yetiştiriciliği yapılan diğer yerli keçi ırklarında yapılacak olan tüm genom çalışmaları, elde edilecek genomların karşılaştırılmasıyla mevcut genetik çeşitliliğin belirlenmesi bakımından önemli olacaktır.

KAYNAKÇA

1. Adams, M. D., 2000. The genome sequence of *Drosophila melanogaster*. **Science**, **287** (5461): 2185–2195.
2. Akçapınar, H. 1994. “Keçi Yetiştiriciliği”, A.Ü. Veteriner Fakültesi Zootečni Anabilim Dalı, Ders Notu.
3. Alonge, M., Lebeigle, L., Kirsche, M., Jenike, K., Ou, S., Aganezov, S., Wang, X., Lippman, Z. B., Schatz, M. C., Soyk, S., 2022. Automated assembly scaffolding using RagTag elevates a new tomato system for high-throughput genome editing. **Genome biology**, **23** (1): 1-19.
4. Amarasinghe, S. L., Su, S., Dong, X., Zappia, L., Ritchie, M. E., Gouil, Q., 2020. Opportunities and challenges in long-read sequencing data analysis. **Genome Biology**, **21** (1): 1-16.
5. Andreotti, S., Altmüller, J., Quedenau, C., Borodina, T., Nouailles, G., Alves, L. C., Landthaler, M., Bieniara, M., Trimpert, J., Wyler, E., 2022. *De novo*-whole genome assembly of the Roborovski dwarf hamster (*Phodopus roborovskii*) genome: an animal model for severe/critical COVID-19. **Genome Biology and Evolution**, **14** (7): evac100.
6. Arbizu, C. I., Ferro-Mauricio, R. D., Chávez-Galarza, J., Vásquez, H. V., Maicelo, J. L., Poemape, C., Gonzales, J., Quilcate, C., Corredor, F.-A., 2022. The complete mitochondrial genome of a neglected breed, the Peruvian creole cattle (*Bos taurus*), and its phylogenetic analysis. **Data**, **7** (6), 76.
7. Armstrong, E. E., Perry, B. W., Huang, Y., Garimella, K. V., Jansen, H. T., Robbins, C. T., Tucker, N. R., Kelley, J. L., 2022. A beary good genome: Haplotype-resolved, chromosome-level assembly of the brown bear (*Ursus arctos*). **Genome Biology and Evolution**, **14** (9): evac125.
8. Arslan, M., 2023. Whole-genome sequencing and genomic analysis of Norduz goat (*Capra hircus*). **Mammalian Genome**, 1-12.
9. Bachtrog, D., Mank, J. E., Peichel, C. L., Kirkpatrick, M., Otto, S. P., Ashman, T.-L., Hahn, M. W., Kitano, J., Mayrose, I., Ming, R., Perrin, N., Ross, L., Valenzuela, N., Vamosi, J. C., The Tree of Sex Consortium, 2014. Sex

determination: why so many ways of doing it? **PLoS biology**, **12** (7): e1001899.

10. Bankevich, A., Nurk, S., Antipov, D., Gurevich, A. A., Dvorkin, M., Kulikov, A. S., Lesin, V. M., Nikolenko, S. I., Pham, S., Prjibelski, A. D., Pyshkin, A. V., Sirotkin, A. V., Vyahhi, N., Tesler, G., Alekseyev, M. A., Pevzner, P. A., 2012. Spades: a new genome assembly algorithm and its applications to single-cell sequencing. **Journal of computational biology**, **19** (5): 455-477.
11. Bao, E., Jiang, T., Girke, T., 2014. AlignGraph: algorithm for secondary *de novo* genome assembly guided by closely related references. **Bioinformatics**, **30** (12): i319-i328.
12. Barra, V., Fachinetti, D., 2018. The dark side of centromeres: types, causes and consequences of structural abnormalities implicating centromeric DNA. **Nature Communications**, **9** (1): 4340.
13. Batu, S., 1951. Türkiye Keçi Irkları ve Keçi Yetiştirme Bilgisi. Ankara Üniversitesi Veteriner Fakültesi yayınları: 4, Ankara.
14. Bayat, A., Gamaarachchi, H., Deshpande, N. P., Wilkins, M. R., Parameswaran, S., 2020. Methods for *de-novo* genome assembly. **Preprints**, 2020060324.
15. Beckmann, J. S., Estivill, X., Antonarakis, S. E., 2007. Copy number variants and genetic traits: closer to the resolution of phenotypic to genotypic variability. **Nature Reviews Genetics**, **8** (8): 639-646.
16. Bickhart, D. M., Rosen, B. D., Koren, S., Sayre, B. L., Hastie, A., Chan, S., Lee, J. M., Lam, E. T., Liachko, I., Sullivan, S., Burton, J. N., Huson, H. J., Nystrom, J. W., Kelley, C. M., Hutchison, J. L., Zhou, Y., Sun, J., Crisà, A., Ponce, A., Schwartz, J. S., 2017. Single-molecule sequencing and chromatin conformation capture enable *de novo* reference assembly of the domestic goat genome. **Nature Genetics**, **49**: 643–650.
17. Bilgin, R., 2011. Back to the suture: the distribution of intraspecific genetic diversity in and around Anatolia. **International Journal of Molecular Sciences**, **12** (6): 4080-4103.
18. Black, A. N., Bondo, K. J., Mularo, A., Hernandez, A., Yu, Y., Stein, C. M., Gregory, A., Fricke, K. A., Prendergast, J., Sullins, D., Haukos, D., Whitson,

- M., Grisham, B., Lowe, Z., DeWoody, J. A., 2023. A highly contiguous and annotated genome assembly of the lesser Prairie-Chicken (*Tympanuchus pallidicinctus*). **Genome biology and evolution**, **15** (4): evad043.
19. Blanchard, A. M., Emes, R. D., Greenwood, A. D., Holmes, N., Loose, M. W., McEwen, G. K., Meers, J., Speight, N., Tarlinton, R. E., 2023. Genome reference assembly for bottlenecked Southern Australian koalas. **Genome Biology and Evolution**, **15** (1): evac176.
 20. Blaxter, M., Archibald, J. M., Childers, A. K., Coddington, J. A., Crandall, K. A., Di Palma, F., Durbin, R., Edwards, S. V., Graves, J. A. M., Hackett, K. J., Hall, N., Jarvis, E. D., Johnson, R. N., Karlsson, E. K., Kress, W. J., Kuraku, S., Lawniczak, M. K. N., Lindblad-Toh, K., Lopez, J. V., Moran, N. A., 2022. Why sequence all eukaryotes? **Proceedings of the National Academy of Sciences**, **119** (4): e2115636118.
 21. Bongaerts, P., Cooke, I. R., Ying, H., Wels, D., den Haan, S., Hernandez-Agreda, A., Brunner, C. A., Dove, S., Englebert, N., Eyal, G., Forêt, S., Grinblat, M., Hay, K. B., Harii, S., Hayward, D. C., Lin, Y., Mihaljević, M., Moya, A., Muir, P., Sinniger, F., 2021. Morphological stasis masks ecologically divergent coral species on tropical reefs. **Current Biology**, **31** (11): 2286-2298.
 22. Boore, J.L., 1999. Animal mitochondrial genomes. **Nucleic Acids Research**, **27** (8): 1767-1780.
 23. Botigué, L. R., Song, S., Scheu, A., Gopalan, S., Pendleton, A. L., Oetjens, M., Taravella, A. M., Seregély, T., Zeeb-Lanz, A., Arbogast, R.-M., Bobo, D., Daly, K., Unterländer, M., Burger, J., Kidd, J. M., Veeramah, K. R., 2017. Ancient European dog genomes reveal continuity since the Early Neolithic. **Nature communications**, **8** (1): 16082.
 24. Botstein, D., Cherry, J. M., Ashburner, M., Ball, C. A., Blake, J. A., Butler, H., Davis, A. P., Dolinski, K., Dwight, S. S., Eppig, J. T., Harris, M. A., Hill, D. P., Issel-Tarver, L., Kasarskis, A., Lewis, S., Matese, J. C., Richardson, J. E., Ringwald, M., Rubin, G. M., Sherlock, G., 2000. Gene Ontology: tool for the unification of biology. **Nature genetics** **25** (1): 25-29.

25. Bourque, G., Burns, K. H., Gehring, M., Gorbunova, V., Seluanov, A., Hammell, M., Imbeault, M., Izsvák, Z., Levin, H. L., Macfarlan, T. S., Mager, D. L., Feschotte, C., 2018. Ten things you should know about transposable elements. **Genome biology**, **19**: 1-12.
26. Bradnam, K. R., Fass, J. N., Alexandrov, A., Baranay, P., Bechner, M., Birol, I., Boisvert, S., Chapman, J. A., Chapuis, G., Chikhi, R., Chitsaz, H., Chou, W.-C., Corbeil, J., Del Fabbro, C., Docking, T. R., Durbin, R., Earl, D., Emrich, S., Fedotov, P., Fonseca, N. A., 2013. Assemblathon 2: evaluating *de novo* methods of genome assembly in three vertebrate species. **Gigascience**, **2**, 1–31.
27. Brent, M. R., 2005. Genome annotation past, present, and future: How to define an ORF at each locus. **Genome Reseource**, **15**: 1777–1786
28. Cantalapiedra, C. P., Hernández-Plaza, A., Letunic, I., Bork, P., Huerta-Cepas, J., 2021. eggNOG-mapper v2: functional annotation, orthology assignments, and domain prediction at the metagenomic scale. **Molecular biology and evolution**, **38** (12): 5825-5829.
29. Cao, L., Zhao, X., Chen, Y., Sun, C., 2021. Chromosome-scale genome assembly of the high royal jelly-producing honeybees. **Scientific Data**, **8** (1): 302.
30. Castresana, J., 2000. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. **Molecular biology and evolution**, **17** (4): 540-552.
31. Challis, R., Richards, E., Rajan, J., Cochrane, G., Blaxter, M., 2020. BlobToolKit – Interactive Quality Assessment of Genome Assemblies, **G3 Genes|Genomes|Genetics**, **10** (4): 1361–1374.
32. Chan, P.P., Lin, B., Lowe, T. M., 2021. tRNAscan-SE 2.0: improved detection and functional classification of transfer RNA genes. **Nucleic Acids Research**, **49** (16): 9077-9096.
33. Chen, S., Zhou, Y., Chen, Y., Gu, J., 2018. fastp: an ultra-fast all-in-one FASTQ preprocessor. **Bioinformatics**, **34** (17): i884-i890.

34. Cheng, H., Concepcion, G. T., Feng, X., Zhang, H., Li, H., 2021. Haplotype-resolved *de novo* assembly using phased assembly graphs with hifiasm. **Nature methods**, **18** (2): 170-175.
35. Cieslak, M., Reissmann, M., Hofreiter, M., Ludwig, A., 2011. Colours of domestication. **Biological Reviews**, **86** (4): 885–899.
36. Colli, L., Lancioni, H., Cardinali, I., Olivieri, A., Capodiferro, M. R., Pellecchia, M., Rzepus, M., Zamani, W., Naderi, S., Gandini, F., S. Vahidi, S. M. F., Agha, S., Randi, E., Battaglia, V., Sardina, M. T., Portolano, B., Rezaei, H. R., Lymberakis, Boyer, F., Coissac, E., 2015. Whole mitochondrial genomes unveil the impact of domestication on goat matrilineal variability. **BMC genomics**, **16**: 1-12.
37. Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., Handsaker, R. E., Lunter, G., Marth, G. T., Sherry, S. T., McVean, G., Durbin, R., 1000 Genomes Project Analysis Group, 2011. The variant call format and VCFtools. **Bioinformatics**, **27** (15): 2156-2158.
38. Danecek, P., Bonfield, J. K., Liddle, J., Marshall, J., Ohan, V., Pollard, M. O., Whitwham, A., Keane, T., McCarthy, S. A., Davies, R. M., Li, H., 2021. Twelve years of SAMtools and BCFtools. **Gigascience**, **10** (2): giab008.
39. Darty, K., Denise, A., Ponty, Y., 2009. VARNA: Interactive drawing and editing of the RNA secondary structure. **Bioinformatics**, **25** (15): 1974.
40. Darwin Tree of Life Project Consortium, 2022. Sequence locally, think globally: The Darwin tree of life project. **Proceedings of the National Academy of Sciences**, **119** (4): e2115642118.
41. Davenport, K. M., Bickhart, D. M., Worley, K. C., Murali, S. C., Mazdak Salavati, Clark, E. L., Cockett, N. E., Heaton, M. P., Timothy, Murdoch, B. M., Rosen, B. D., 2022. An improved ovine reference genome assembly to facilitate in-depth functional annotation of the sheep genome. **Gigascience**, **11**: giab096.
42. De Coster, W., D’hert, S., Schultz, D. T., Cruts, M., Van Broeckhoven, C., 2018. NanoPack: visualizing and processing long-read sequencing data. **Bioinformatics**, **34** (15): 2666-2669.

43. de Koning, A. J., Gu, W., Castoe, T. A., Batzer, M. A., Pollock, D. D., 2011. Repetitive elements may comprise over two-thirds of the human genome. **PLoS genetics**, **7** (12): e1002384.
44. De Mitcheson, Y. S., Liu, M., 2008. Functional hermaphroditism in teleosts. **Fish and Fisheries**, **9** (1): 1-43.
45. Del Angel, V. D., Hjerde, E., Sterck, L., Capella-Gutierrez, S., Notredame, C., Vinnere Pettersson, O., Amselem, J., Bouri, L., Bocs, S., Klopp, C., Gibrat, J.-F., Vlasova, A., Leskosek, B. L., Soler, L., Binzer-Panchal, M., Lantz, H. 2018. Ten steps to get started in Genome Assembly and Annotation. **F1000Research**, **7**: 148.
46. Del Hoyo, J., Elliott, A., Sargatal, J., 1992. Handbook Of The Birds Of The World (Vol. 1, No. 8). Lynx Ed.
47. Ding, R., Yang, M., Wang, X., Quan, J., Zhuang, Z., Zhou, S., Li, S., Xu, Z., Zheng, E., Cai, G., Liu, D., Huang, W., Yang, J., Wu, Z., 2018. Genetic architecture of feeding behavior and feed efficiency in a Duroc pig population. **Frontiers in genetics**, **9**: 220.
48. Donath, A., Jühling, F., Al-Arab, M., Bernhart, S. H., Reinhardt, F., Stadler, P. F., Middendorf, M., Bernt, M., 2019. Improved annotation of protein-coding genes boundaries in metazoan mitochondrial genomes. **Nucleic acids research**, **47** (20): 10543-10552.
49. Dong, Y., Xie, M., Jiang, Y., Xiao, N., Du, X., Zhang, W., Tosser-Klopp, G., Wang, J., Yang, S., Liang, J., Chen, W., Chen, J., Zeng, P., Hou, Y., Bian, C., Pan, S., Li, Y., Liu, X., Wang, W., Servin, B., 2013. Sequencing and automated whole-genome optical mapping of the genome of a domestic goat (*Capra hircus*). **Nature biotechnology**, **31** (2): 135-141.
50. Dong, Y., Zhang, X., Xie, M., Arefnezhad, B., Wang, Z., Wang, W., Feng, S., Huang, G., Guan, R., Shen, W., Bunch, R., McCulloch, R., Li, Q., Li, B., Zhang, G., Xu, X., Kijas, J. W., Salekdeh, G. H., Wang, W., Jiang, Y., 2015. Reference genome of wild goat (*Capra aegagrus*) and sequencing of goat breeds provide insight into genic basis of goat domestication. **BMC genomics**, **16** (1): 1-11.

51. Du, H., Diao, C., Zhao, P., Zhou, L., Liu, J. F., 2021. Integrated hybrid *de novo* assembly technologies to obtain high-quality pig genome using short and long reads. **Briefings in Bioinformatics**, **22** (5): bbaa399.
52. Ebert, K. M., Selgrath, J. P., DiTullio, P., Denman, J., Smith, T. E., Memon, M. A., Schindler, J. E., Monastersky, G. M., Vitale, J. A., Gordon, K., 1991. Transgenic production of a variant of human tissue-type plasminogen activator in goat milk: generation of transgenic goats and analysis of expression. **Biotechnology**, **9**: 835–838.
53. Eda, M., Itahashi, Y., Kikuchi, H., Sun, G., Hsu, K. H., Gakuhari, T., Yoneda, M., Jiang, L., Yang, G., Nakamura, S., 2022. Multiple lines of evidence of early goose domestication in a 7,000-y-old rice cultivation village in the lower Yangtze River, China. **Proceedings of the National Academy of Sciences of the United States of America**, **119** (12): e2117064119.
54. Edwards, R. J., Field, M. A., Ferguson, J. M., Dudchenko, O., Keilwagen, J., Rosen, B. D., Johnson, G. S., Rice, E. S., Hillier, L. D., Hammond, J. M., Towarnicki, S. G., Omer, A., Khan, R., Skvortsova, K., Bogdanovic, O., Zammit, R. A., Aiden, E. L., Warren, W. C., Ballard, J. W. O., 2021. Chromosome-length genome assembly and structural variations of the primal Basenji dog (*Canis lupus familiaris*) genome. **BMC genomics**, **22** (1): 1-19.
55. Ejigu, G.F., Jung, J., 2020. Review on the computational genome annotation of sequences obtained by Next-generation sequencing. **Biology (Basel)**, **9** (9): 295.
56. Ekblom, R., Wolf, J. B. W., 2014. A field guide to whole-genome sequencing, assembly and annotation. **Evolutionary Applications**, **7** (9): 1026–1042.
57. Erol, H., 2015. Erkek ve Kastre Ankara Keçisi Oğlaklarında Farklı Kesim Ağırlıklarında Besi Performansı, Kesim, Karkas ve Bazı Et Kalite Özellikleri. Ankara Üniversitesi Sağlık Bilimleri Enstitüsü, Doktora Tezi, Ankara 135s.
58. Ewing, B., Green, P., 1998. Base-calling of automated sequencer traces using phred. II. Error probabilities. **Genome Research**, **8** (3): 186-194.
59. Fabregat, A., Jupe, S., Matthews, L., Sidiropoulos, K., Gillespie, M., Garapati, P., Haw, R., Jassal, B., Korninger, F., May, B., Milacic, M., Roca, C. D., Rothfels,

- K., Sevilla, C., Shamovsky, V., Shorser, S., Varusai, T., Viteri, G., Weiser, J., Wu, G., 2018. The reactome pathway knowledgebase. **Nucleic acids research**, **46** (D1): D649-D655.
60. Fang, S. M., Hu, B. L., Zhou, Q. Z., Yu, Q. Y., Zhang, Z., 2015. Comparative analysis of the silk gland transcriptomes between the domestic and wild silkworms. **BMC genomics**, **16** (1): 1-12.
 61. Fernandes Junior, G. A., de Oliveira, H. N., Carvalheiro, R., Cardoso, D. F., Fonseca, L. F. S., Ventura, R. V., de Albuquerque, L. G., 2020. Whole-genome sequencing provides new insights into genetic mechanisms of tropical adaptation in Nellore (*Bos primigenius indicus*). **Scientific reports**, **10** (1): 9412.
 62. Fisher, R., O'Leary, R. A., Low-Choy, S., Mengersen, K., Knowlton, N., Brainard, R. E., Caley, M. J., 2015. Species richness on coral reefs and the pursuit of convergent global estimates. **Current Biology**, **25**: 500–505.
 63. Fredman, D., Siegfried, M., Yuan, Y. P., Bork, P., Lehv  slaiho, H., Brookes, A. J., 2002. HGVbase: a human sequence variation database emphasizing data quality and a broad spectrum of data sources. **Nucleic Acids Research**, **30** (1): 387-391.
 64. Fu, W., Wang, R., Yu, J., Hu, D., Cai, Y., Shao, J., Jiang, Y., 2021. GGVD: a goat genome variation database for tracking the dynamic evolutionary process of selective signatures and ancient introgressions. **Journal of Genetics and Genomics**, **48** (3): 248-256.
 65. Garrison E, Marth G., 2012. Haplotype-based variant detection from short-read sequencing. **arXiv preprint**, arXiv:1207.3907 [q-bio.GN]
 66. Geneva, A. J., Park, S., Bock, D. G., Pietro, Fatih Sarigol, Tollis, M., Donihue, C. M., R. Graham Reynolds, Feiner, N., Rasys, A. M., Lauderdale, J. D., Minchey, S. G., Alcala, A. J., Infante, C. R., Kolbe, J. J., Schluter, D., Menke, D. B., Losos, J. B., 2022. Chromosome-scale genome assembly of the brown anole (*Anolis sagrei*), an emerging model species. **Communications Biology**, **5** (1): 1126.

67. Gonçalves, M., Siegismund, H. R., Vuuren, B. J., Koepfli, K. P., Ferrand, N., Godinho, R., 2021. *De novo* whole-genome assembly and resequencing resources for the roan (*Hippotragus equinus*), an iconic African antelope. **G3 Genes|Genomes|Genetics**, **11** (2): jkab002.
68. Gordon, D. L., Huddleston, J., Chaisson, M., Hill, C., Kronenberg, Z. N., Munson, K. M., Malig, M., Raja, A., Fiddes, I. T., Hillier, L. W., Dunn, C. J., Baker, C., Armstrong, J., Diekhans, M., Paten, B., Shendure, J., Wilson, R. K., Haussler, D., Chin, C.-S., Eichler, E. E., 2016. Long-read sequence assembly of the gorilla genome. **Science**, **352** (6281): aae0344.
69. Grant, J. R., Enns, E., Marinier, E., Mandal, A., Herman, E. K., Chen, C.-Y., Graham, M., Domselaar, G. V., Stothard, P., 2023. Proksee: in-depth characterization and visualization of bacterial genomes. **Nucleic Acids Research**, **51** (W1): W484–W492.
70. Grassman Jr, L. I., Tewes, M. E., Silvy, N. J., 2005. Ranging, habitat use and activity patterns of binturong *Arctictis binturong* and yellow-throated marten *Martes flavigula* in north-central Thailand. **Wildlife biology**, **11** (1): 49-57.
71. Groenen, M. A. M., Archibald, A. L., Uenishi, H., Tuggle, C. K., Takeuchi, Y., Rothschild, M. F., Rogel-Gaillard, C., Park, C., Milan, D., Megens, H.-J., Li, S., Larkin, D. M., Kim, H., Frantz, L. A. F., Caccamo, M., Ahn, H., Aken, B. L., Anselmo, A., Anthon, C., Auvil, L., 2012. Analyses of pig genomes provide insight into porcine demography and evolution. **Nature**, **491** (7424): 393-398.
72. Guindon, S., Dufayard, J.F., Lefort, V., Anisimova, M., Hordijk, W., Gascuel, O., 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. **Systematic Biology**, **59** (3), 307-21.
73. Gurevich, A., Saveliev, V., Vyahhi, N., Tesler, G., 2013. Quast: quality assessment tool for genome assemblies. **Bioinformatics**, **29** (8): 1072-1075.
74. Gündoğdu, E., 2006. Isparta Yöresinde Yaban Keçisi *Capra aegagrus* erksleben 1777'nin Popülasyon Ekolojisi. Süleyman Demirel Üniversitesi Fen Bilimleri Enstitüsü, Doktora Tezi, Isparta.

75. Haghshenas, E., Hach, F., Sahinalp, S. C., Chauve, C., 2016. Colormap: Correcting long reads by mapping short reads. **Bioinformatics**, **32** (17): i545–i551.
76. Han, L., Zhu, S., Ning, C., Cai, D., Wang, K., Chen, Q., Hu, S., Yang, J., Shao, J., Zhu, H., Zhou, H., 2014. Ancient DNA provides new insight into the maternal lineages and domestication of Chinese donkeys. **BMC Evolutionary Biology**, **14** (1): 1-10.
77. Hastings, J., Owen, G., Dekker, A., Ennis, M., Kale, N., Muthukrishnan, V., Turner, S., Swainston, N., Mendes, P., Steinbeck, C., 2016. ChEBI in 2016: Improved services and an expanding collection of metabolites. **Nucleic acids research**, **44** (D1): D1214-D1219.
78. He, K., Zhao, L., Yuan, Z., Canario, A., Liu, Q., Chen, S., Guo, J., Luo, W., Yan, H., Zhang, D., Li, L., Yang, S., 2022. Chromosome-level genome assembly of largemouth bass (*Micropterus salmoides*) using PacBio and Hi-C technologies. **Scientific Data**, **9** (1): 482.
79. He, W., Yang, J., Jing, Y., Xu, L., Yu, K., Fang, X., 2023. NGenomeSyn: an easy-to-use and flexible tool for publication-ready visualization of syntenic relationships across multiple genomes. **Bioinformatics**, **39** (3): btad121.
80. Henrichsen, C. N., Chaignat, E., Reymond, A., 2009. Copy number variants, diseases and gene expression. **Human molecular genetics**, **18** (R1): R1-R8.
81. Hiendleder, S., Lewalski, H., Wassmuth, R., Janke, A., 1998. The complete mitochondrial DNA sequence of the domestic sheep (*Ovis aries*) and comparison with the other major ovine haplotype. **Journal of Molecular Evolution**, **47**: 441-448.
82. Ho Ko, J., Lee, C.-S., Kim, K. H., Pang, M.-G., Koo, J. S., Fang, N., Koo, D.-B., Oh, K. B., Youn, W.-S., Zheng, G. D., Park, J. S., Kim, S. J., Han, Y.-M., Choi, I. Y., Lim, J., Shin, S. T., Jin, S. W., Lee, K.-K., Yoo, O. J., 2000. Production of biologically active human granulocyte colony stimulating factor in the milk of transgenic goat. **Transgenic research**, **9**: 215-222.
83. Hu, X., Jiang, Z., Ming, Y., Jian, J., Jiang, S., Zhang, D., Zhang, J., Zheng, S., Fang, X., Yang, Y., Zheng, R., 2022. A chromosomal level genome sequence for

- Quasipaa spinosa* (Dicroglossidae) reveals chromosomal evolution and population diversity. **Molecular Ecology Resources**, **22** (4): 1545-1558.
84. Islam, M. S., Hibino, M., Tanaka, M., 2007. Distribution and diet of the roughskin sculpin, *Trachidermus fasciatus*, larvae and juveniles in the Chikugo River estuary, Ariake Bay, Japan. **Ichthyological Research**, **54**: 160-167.
 85. İbiş, O., 2019. Türkiye Anadolu Eşeği (*Equus asinus*)'nin Mitogenom Karakterizasyonu ve Filogenetik İlişkileri. **Türkiye Tarımsal Araştırmalar Dergisi**, **6** (3): 257-267.
 86. Jain, M., Olsen, H. E., Paten, B., Akeson, M., 2016. The oxford nanopore minion: delivery of nanopore sequencing to the genomics community. **Genome biology**, **17** (1): 239.
 87. Jarne, P., Auld, J. R., 2006. Animals mix it up too: the distribution of self-fertilization among hermaphroditic animals. **Evolution**, **60** (9): 1816-1824.
 88. Jebb, D., Huang, Z., Pippel, M., Hughes, G. M., Lavrichenko, K., Devanna, P., Winkler, S., Jermiin, L. S., Skirmuntt, E. C., Katzourakis, A., Burkitt-Gray, L., Ray, D. A., Sullivan, K. A. M., Roscito, J. G., Kirilenko, B. M., Dávalos, L. M., Corthals, A. P., Power, M. L., Jones, G., Ransome, R. D., 2020. Six reference-quality genomes reveal evolution of bat adaptations. **Nature**, **583** (7817): 578-584.
 89. Jin, J. J., Yu, W. B., Yang, J. B., Song, Y., DePamphilis, C. W., Yi, T. S., Li, D. Z., 2020. GetOrganelle: a fast and versatile toolkit for accurate de novo assembly of organelle genomes. **Genome biology**, **21**: 1-31.
 90. Kajitani, R., Toshimoto, K., Noguchi, H., Toyoda, A., Ogura, Y., Okuno, M., Yabana, M., Harada, M., Nagayasu, E., Maruyama, H., Kohara, Y., Fujiyama, A., Hayashi, T., Itoh, T., 2014. Efficient *de novo* assembly of highly heterozygous genomes from whole-genome shotgun short reads. **Genome research**, **24** (8): 1384-1395.
 91. Kalyanamoorthy, S., Minh, B.Q., Wong, T. K., Von Haeseler, A., Jermiin, L. S., 2017. ModelFinder: fast model selection for accurate phylogenetic estimates. **Nature methods**, **14** (6), 587-589.

92. Kanehisa, M., Goto, S., 2000. KEGG: kyoto encyclopedia of genes and genomes. **Nucleic acids research**, **28** (1): 27-30.
93. Karataş A., Filiz H., Erciyas-Yavuz K., Özeren, S. C., Tok, C. V., 2021. The vertebrate biodiversity of Turkey. In biodiversity, conservation and sustainability in Asia: Volume 1: Prospects and Challenges in West Asia and Caucasus (pp. 175-274). Cham: Springer International Publishing
94. Karimi, K., Do, D. N., Wang, J., Easley, J., Borzouie, S., Sargolzaei, M., Plastow, G., Wang, Z., Miar, Y., 2022. A chromosome-level genome assembly reveals genomic characteristics of the American mink (*Neogale vison*). **Communications Biology**, **5** (1): 1381.
95. Katoh, K., Standley, D. M., 2016. A simple method to control over-alignment in the MAFFT multiple sequence alignment program. **Bioinformatics**, **32** (13): 1933-1942.
96. Keilwagen, J., Hartung, F., Grau, J. 2019. GeMoMa: Homology-based gene prediction utilizing intron position conservation and RNA-seq data. **Methods in Molecular Biology**, **1962**: 161-177.
97. Keilwagen, J., Hartung, F., Paulini, M., Twardziok, S. O. Grau, J., 2018. Combining RNA-seq data and homology-based gene prediction for plants, animals and fungi. **BMC Bioinformatics**. **19**: 189.
98. Kingan, S. B., Heaton, H., Cudini, J., Lambert, C. C., Baybayan, P., Galvin, B. D., Durbin, R., Korlach, J., Lawniczak, M. K. N., 2019. A High-Quality *De novo* Genome Assembly from a Single Mosquito Using PacBio Sequencing. **Genes (Basel)**, **10** (1): 62.
99. Kingsford, C., Schatz, M. C., Pop, M., 2010. Assembly complexity of prokaryotic genomes using short reads. **BMC Bioinformatics**, **11**: 1-11.
100. Klees, S., Heinrich, F., Schmitt, A. O., Gültas, M., 2021. AgReg-SNPdb: A database of regulatory SNPs for agricultural animal species. **Biology**, **10** (8): 790.
101. Kolmogorov, M., Yuan, J., Lin, Y., Pevzner, P. A., 2019. Assembly of long, error-prone reads using repeat graphs. **Nature Biotechnology**, **37** (5): 540–546.

102. Koren, S., Walenz, B. P., Berlin, K., Miller, J. R., Bergman, N. H., Phillippy, A. M., 2017. Canu: scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. **Genome research**, **27** (5): 722-736.
103. Kryštufek, B., Vohralik, V., 2001. Mammals of Turkey and Cyprus, Introduction, Checklist, Insectivora. University of Primorska, Koper, Slovenia, 138 pp.
104. Kryštufek, B., Vohralik, V., 2009. Mammals of Turkey and Cyprus, Introduction, Checklist, Rodentia II. University of Primorska, Koper, Slovenia, 374 pp.
105. Lappalainen, I., Lopez, J., Skipper, L., Hefferon, T., Spalding, J. D., Garner, J., Chen, C., Maguire, M., Corbett, M., Zhou, G., Paschall, J., Ananiev, V., Flicek, P., Church, D. M., 2012. DbVar and DGVA: public archives for genomic structural variation. **Nucleic acids research**, **41**(D1): D936-D941.
106. Larson, G., Dobney, K., Albarella, U., Fang, M., Matisoo-Smith, E., Robins, J., Lowden, S., Finlayson, H., Brand, T., Willerslev, E., Rowley-Conwy, P., Andersson, L., Cooper, A., 2005. Worldwide phylogeography of wild boar reveals multiple centers of pig domestication. **Science**, **307** (5715): 1618-1621.
107. Li, C., Zuo, W., Tong, X., Han, M., Gao, R., Hu, H., Lu, K., Luan, Y., Zhang, B., Liu, Y., Dai, F., 2021. Whole-genome resequencing reveals loci under selection during silkworm improvement. **Journal of Animal Breeding and Genetics**, **138** (3): 278-290.
108. Li, D., Liu, C. M., Luo, R., Sadakane, K., Lam, T. W., 2015. MEGAHIT: An ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. **Bioinformatics**, **31** (10): 1674-1676.
109. Li, F., Zhao, X., Li, M., He, K., Huang, C., Zhou, Y., Li, Z., Walters, J. R., 2019. Insect genomes: Progress and challenges. **Insect Molecular Biology**, **28**: 739–758.
110. Li, H., 2013. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. **arXiv preprint**, arXiv:1303.3997.
111. Li, H., 2018. Minimap2: pairwise alignment for nucleotide sequences. **Bioinformatics**, **34** (18): 3094-3100.
112. Li, L. L., Ma, S. K., Peng, W., Fang, Y. G., Duo, H. R., Fu, H. Y., Jia, G. X., 2021. Genetic diversity and population structure of Tibetan sheep breeds determined

- by whole genome resequencing. **Tropical Animal Health and Production**, **53** (1): 174.
113. Li, R., Zhang, M., Cha, M., Xiang, J., Yi, X., 2022. Chromosome-level genome assembly of the Siberian chipmunk (*Tamias sibiricus*). **Scientific Data**, **9** (1): 783.
 114. Li, R., Zhang, W., Lu, J., Zhang, Z., Mu, C., Song, W., Migaud, H., Bekaert, M., 2020. The whole-genome sequencing and hybrid assembly of *Mytilus coruscus*. **Frontiers in Genetics**, **11**: 440.
 115. Li, R., Zhu, H., Ruan, J., Qian, W., Fang, X., Shi, Z., Li, Y., Li, S., Shan, G., Kristiansen, K., Li, S., Yang, H., Wang, J., Wang, J., 2010. *De novo* assembly of human genomes with massively parallel short read sequencing. **Genome Research**, **20**: 265–272.
 116. Li, W., Du, J., Yang, L., Liang, Q., Yang, M., Zhou, X., Du, W., 2023. Chromosome-level genome assembly and population genomics of Mongolian racerunner (*Eremias argus*) provide insights into high-altitude adaptation in lizards. **BMC Biology**, **21**(1): 40.
 117. Li, Y., Gao, G., Lin, Y., Hu, S., Luo, Y., Wang, G., Jin, L., Wang, Q., Wang, J., Tang, Q., Li, M., 2020. Pacific Biosciences assembly with Hi-C mapping generates an improved, chromosome-level goose genome. **Gigascience**, **9** (10): giaa114.
 118. London, E. W., Roca, A. L., Novakofski, J. E., Mateus-Pinilla, N. E., 2022. A *de novo* chromosome-level genome assembly of the white-tailed deer, *Odocoileus virginianus*. **Journal of Heredity**, **113** (4): 479-489.
 119. Losos, J. B., 2011. Lizards in an evolutionary tree: ecology and adaptive radiation of anoles (Vol. 10). Univ of California Press.
 120. Low, W. Y., Tearle, R., Bickhart, D. M., Rosen, B. D., Kingan, S. B., Swale, T., Thibaud-Nissen, F., Murphy, T. D., Young, R., Lefevre, L., Hume, D. A., Collins, A., Ajmone-Marsan, P., Smith, T. P. L., Williams, J. L., 2019. Chromosome-level assembly of the water buffalo genome surpasses human and goat genomes in sequence contiguity. **Nature communications**, **10** (1): 260.

121. Lu, H., Giordano, F., Ning, Z., 2016. Oxford Nanopore MinION sequencing and genome assembly. **Genomics, Proteomics & Bioinformatics**, **14**: 265–279.
122. Lu, S., Wang, J., Chitsaz, F., Derbyshire, M. K., Geer, R. C., Gonzales, N. R., Gwadz, M., Hurwitz, D. I., Marchler, G. H., Song, J. S., Thanki, N., Yamashita, R. A., Yang, M., Zhang, D., Zheng, C., Lanczycki, C. J., Marchler-Bauer, A., 2020. CDD/SPARCLE: The conserved domain database in 2020. **Nucleic Acids research**, **48**: D265–D268.
123. Luikart, G., Gielly, L., Excoffier, L., Vigne, J. D., Bouvet, J., Taberlet, P., 2001. Multiple maternal origins and weak phylogeographic structure in domestic goats. **Proceedings of the National Academy of Sciences**, **98** (10): 5927–5932.
124. Ma, L., Li, Z., Cai, Y., Xu, H., Yang, R., Lan, X., 2018. Genetic variants in fat- and short-tailed sheep from high-throughput RNA-sequencing data. **Animal Genetics**, **49** (5): 483–487.
125. MacDonald, J. R., Ziman, R., Yuen, R. K., Feuk, L., Scherer, S. W., 2014. The Database of Genomic Variants: a curated collection of structural variation in the human genome. **Nucleic acids research**, **42** (D1): D986–D992.
126. MacHugh, D. E., Bradley, D. G., 2001. Livestock genetic origins: goats buck the trend. **Proceedings of the National Academy of Sciences**, **98** (10): 5382–5384.
127. Machugh, D. E., Bradley, D. G., 2001. Livestock genetic origins: Goats buck the trend. **Proceedings of the National Academy of Sciences**, **98**: 5382–5384.
128. MacHugh, D. E., Shriver, M. D., Loftus, R. T., Cunningham, P., Bradley, D. G., 1997. Microsatellite DNA variation and the evolution, domestication and phylogeography of taurine and zebu cattle (*Bos taurus* and *Bos indicus*). **Genetics**, **146** (3): 1071–1086.
129. Manni, M., Berkeley, M. R., Seppey, M., Simão, F. A., Zdobnov, E. M., 2021. BUSCO update: novel and streamlined workflows along with broader and deeper phylogenetic coverage for scoring of eukaryotic, prokaryotic, and viral genomes. **Molecular biology and evolution**, **38** (10): 4647–4654.

130. Mathé, C., Sagot, M. F., Schiex, T., Rouzé, P., 2002. Current methods of gene prediction, their strengths and weaknesses. **Nucleic acids research**, **30** (19): 4103-4117.
131. Matyot, P., 2004. The establishment of the crested tree lizard, *Calotes versicolor* (Daudin, 1802) (Squamata: Agamidae) in Seychelles. **Phelsuma**, **12**: 35–47.
132. McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M., DePristo, M. A., 2010. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. **Genome research**, **20** (9): 1297-1303.
133. McLaren, W., Gil, L., Hunt, S. E., Riat, H. S., Ritchie, G. R. S., Thormann, A., Flicek, P., Cunningham, F., 2016. The ensembl variant effect predictor. **Genome biology**, **17** (1): 1-14.
134. Mei, X., Liu, G., Yan, J., Zhao, C., Wang, X., Zhou, S., Wei, Q., Zhao, S., Liu, Z., Sha, W., Zhang, H., 2023. A chromosome-level genome assembly of the yellow-throated marten (*Martes flavigula*). **Scientific Data**, **10** (1): 216.
135. Metzker, M. L., 2010. Sequencing technologies the next generation. **Nature reviews genetics**, **11**: 31–46.
136. Meurens, F., Summerfield, A., Nauwynck, H., Saif, L., Gerdts, V., 2012. The pig: a model for human infectious diseases. **Trends Microbiology**, **20**: 50–57.
137. Meyer, R. S., Purugganan, M. D., 2013. Evolution of crop species: genetics of domestication and diversification. **Nature Reviews Genetics**, **14** (12): 840–852.
138. Miao, X., Yu, Y., Zhao, Z., Wang, Y., Qian, X., Wang, Y., Li, S., Wang, C., 2022. Chromosome-level haplotype assembly for *Equus asinu*. **Frontiers in genetics**, **13**: 738105.
139. Mikheenko, A., Prjibelski, A., Saveliev, V., Antipov, D., Gurevich, A., 2018. Versatile genome assembly evaluation with QUAST-LG. **Bioinformatics**, **34** (13): i142-i150.
140. Mills, R. E., Walter, K., Stewart, C., Handsaker, R. E., Chen, K., Alkan, C., Abyzov, A., Yoon, S. C., Ye, K., Cheetham, R. K., Chinwalla, A., Conrad, D. F., Fu, Y.,

- Grubert, F., Hajirasouliha, I., Hormozdiari, F., Iakoucheva, L. M., Iqbal, Z., Kang, S., Kidd, J. M., 1000 Genomes Project, 2011. Mapping copy number variation by population-scale genome sequencing. **Nature**, **470** (7332): 59-65.
141. Mishra, S., Rastogi, Y. P., Jabin, S., Kaur, P., Amir, M., Khatoon, S. A., 2020. bacterial phyla dataset for protein function prediction. **Data Brief**, **28**: 105002.
142. Mohr, D. W., Gaughran, S. J., Paschall, J., Naguib, A., Pang, A. W. C., Dudchenko, O., Aiden, E. L., Church, D. M., Scott, A. F., 2022. A chromosome-length assembly of the Hawaiian monk seal (*Neomonachus schauinslandi*): A history of "Genetic purging" and genomic stability. **Genes (Basel)**, **13** (7): 1270.
143. Morgat, A., Lombardot, T., Axelsen, K. B., Aimo, L., Niknejad, A., Hyka-Nouspikel, N., Coudert, E., Pozzato, M., Pagni, M., Moretti, S., Rosanoff, S., Onwubiko, J., Bougueleret, L., Xenarios, I., Redaschi, N., Bridge, A., 2016. Updates in Rhea—an expert curated resource of biochemical reactions. **Nucleic acids research**, **45** (D1): D415–D418.
144. Mostovoy, Y., Levy-Sakin, M., Lam, J., Lam, E. T., Hastie, A. R., Marks, P., Lee, J., Chu, C., Lin, C., Džakula, Ž., Cao, H., Schlebusch, S. A., Giorda, K., Schnall-Levin, M., Wall, J. D., Kwok, P. Y., 2016. A hybrid approach for *de novo* human genome sequence assembly and phasing. **Nature Methods**, **13** (7): 587-90.
145. Moyle, R. G., Filardi, C. E., Smith, C. E., Diamond, J., 2009. Explosive Pleistocene diversification and hemispheric expansion of a ‘great speciator’. **Proceedings of the National Academy of Sciences**, **106** (6): 1863–1868.
146. Murray, C., Huerta-Sanchez, E., Casey, F., Bradley, D. G., 2010. Cattle demographic history modelled from autosomal sequence variation. **Philosophical Transactions of the Royal Society B: Biological Sciences**, **365** (1552): 2531-2539.
147. Nagahama, Y., Chakraborty, T., Paul-Prasanth, B., Ohta, K., Nakamura, M., 2021. Sex determination, gonadal sex differentiation, and plasticity in vertebrate species. **Physiological reviews**, **101** (3): 1237–1308.
148. Ni, H., Zhang, Y., Yang, Y., Yin, Y., Xie, H., Zheng, J., Dong, L., Diao, J., Wei, M., Zhichao Lv, Yan, S., Li, Y., Sun, H., Sun, X., 2023. Whole genome

- sequencing of a wild swan goose population. **Frontiers in Genetics**, **14**: 1038606.
149. Nowak, R. M., Jastrzębski, J. P., Kuśmirek, W., Sałamatin, R., Rydzanicz, M., Sobczyk-Kopciół, A., Sulima-Celińska, A., Paukszto, Ł., Makowczenko, K. G., Płoski, R., Tkach, V. V., Basałaj, K., Młocicki, D., 2019. Hybrid *de novo* whole-genome assembly and annotation of the model tapeworm *Hymenolepis diminuta*. **Scientific data**, **6** (1): 302.
 150. Ojala, D., Montoya, J., Attardi, G., 1981. tRNA punctuation model of RNA processing in human mitochondria. **Nature**, **290**: 470-474.
 151. Ouyang, J., Zheng, S., Huang, M., Tang, H., Qiu, X., Chen, S., Wang, Z., Zhou, Z., Gao, Y., Xiong, Y., Zeng, G., Huang, J., He, J., Ren, J., Chen, H., Yan, X., 2022. Chromosome-level genome and population genomics reveal evolutionary characteristics and conservation status of Chinese indigenous geese. **Communications Biology**, **5** (1): 1191.
 152. Pacha, A.S., Nigam, P., Pandav, B., Mondol, S., 2021. Sequencing and annotation of the endangered wild buffalo (*Bubalus arnee*) mitogenome for taxonomic assessment. **Molecular Biology Reports**, **48** (2), 1995-2003.
 153. Palmada-Flores, M., Orkin, J. D., Haase, B., Mountcastle, J., Bertelsen, M. F., Fedrigo, O., Kuderna, L. F. K., Jarvis, E. D., Marques-Bonet, T., 2022. A high-quality, long-read genome assembly of the endangered ring-tailed lemur (*Lemur catta*). **GigaScience**, **11**: giac026.
 154. Park, S. D., Magee, D. A., McGettigan, P. A., Teasdale, M. D., Edwards, C. J., Lohan, A. J., Murphy, A., Braud, M., Donoghue, M. T., Liu, Y., Chamberlain, A. T., Rue-Albrecht, K., Schroeder, S., Spillane, C., Tai, S., Bradley, D. G., Sonstegard, T. S., Loftus, B. J., MacHugh, D. E., 2015. Genome sequencing of the extinct Eurasian wild aurochs, *Bos primigenius*, illuminates the phylogeography and evolution of cattle. **Genome biology**, **16**: 1-15.
 155. Paşalı, H., 2014. Türkiye’de Yaban Keçisi *Capra aegagrus aegagrus*. **Animal Health Production and Hygiene**, **3** (1): 245-247.
 156. Payne, A., Holmes, N., Rakyan, V., Loose, M., 2019. BulkVis: a graphical viewer for Oxford nanopore bulk FAST5 files. **Bioinformatics**, **35**: 2193–2198.

157. Pitt, D., Sevane, N., Nicolazzi, E. L., MacHugh, D. E., Park, S. D. E., Colli, L., Martinez, R., Bruford, M. W., Orozco-terWengel, P., 2019. Domestication of cattle: Two or three events? **Evolutionary applications**, **12** (1): 123-136.
158. Pop M, Salzberg S. L., 2008. Bioinformatics challenges of new sequencing technology. **Trends in Genetics**, **24** (3): 142-149.
159. Porrelli, S., Gerbault-Seureau, M., Rozzi, R., Chikhi, R., Curaudeau, M., Ropiquet, A., Hassanin, A., 2022. Draft genome of the lowland anoa (*Bubalus depressicornis*) and comparison with buffalo genome assemblies (Bovidae, Bubalina). **G3**, **12** (11): jkac234.
160. Porto-Neto, L. R., Sonstegard, T. S., Liu, G. E., Bickhart, D. M., V.B, M., Machado, M. A., Yuri Tani Utsunomiya, José Fernando Garcia, Gondro, C., Van Tassell, C. P., 2013. Genomic divergence of zebu and taurine cattle identified through high-density SNP genotyping. **BMC genomics**, **14**: 1-12.
161. Posada, D., 2008. jModelTest: phylogenetic model averaging. **Molecular biology and evolution**, **25** (7): 1253-1256.
162. Ran, M. L., Chen, B., He, C. Q., Li, Z., Dong, L. H., 2016. The complete mitochondrial genome of Jintang black goat (*Capra hircus*). **Mitochondrial DNA Part A**, **27** (6), 4151-4152.
163. Rhie, A., McCarthy, S. A., Fedrigo, O., Damas, J., Formenti, G., Koren, S., Uliano-Silva, M., Chow, W., Fungtammasan, A., Kim, J., Lee, C., Ko, B. J., Chaisson, M., Gedman, G. L., Cantin, L. J., Thibaud-Nissen, F., Haggerty, L., Bista, I., Smith, M., Haase, B., 2021. Towards complete and error-free genome assemblies of all vertebrate species. **Nature**, **592** (7856): 737-746.
164. Salzberg, S. L., Phillippy, A. M., Zimin, A., Puiu, D., Magoc, T., Koren, S., Treangen, T. J., Schatz, M. C., Delcher, A. L., Roberts, M., Marçais, G., Pop, M., Yorke, J. A., 2012. GAGE: a critical evaluation of genome assemblies and assembly algorithms. **Genome Research**, **22** (3): 557–567.
165. Schneeberger, K., Ossowski, S., Ott, F., Klein, J. D., Wang, X., Lanz, C., Smith, L. M., Cao, J., Fitz, J., Warthmann, N., Henz, S. R., Huson, D. H., Weigel, D., 2011. Reference-guided assembly of four diverse *Arabidopsis thaliana*

- genomes. **Proceedings of the National Academy of Sciences USA**, **108** (25): 10249–10254.
166. Seyfi, E., Bulut, Ş., Karataş, A., 2021. Türkiye'nin Tehlike Altındaki Memeli Türleri. **Doğanın Sesi**, (7): 54-72.
 167. Siddiki, A. Z., Baten, A., Billah, M., Alam, M. A. U., Shawrob, K. S. M., Saha, S., Chowdhury, M., Rahman, A. H., Stear, M., Miah, G., Kumkum, M., Islam, M. S., Hossain, M. A., Mollah, A. K. M. M., Khan, Md. K. I., 2019. The genome of the Black Bengal goat (*Capra hircus*). **BMC research notes**, **12** (1): 1-3.
 168. Silanikove, N., 2000. The physiological basis of adaptation in goats to harsh environments. **Small Ruminant Research**, **35** (3): 181–193.
 169. Simpson, J. T., Wong, K., Jackman, S. D., Schein, J. E., Jones, S. J., Birol, I., 2009. ABySS: a parallel assembler for short read sequence data. **Genome research**, **19** (6): 1117–1123.
 170. Skorupski, J., 2022. Characterisation of the complete mitochondrial genome of critically endangered *Mustela lutreola* (Carnivora: Mustelidae) and its phylogenetic and conservation implications. **Genes**, **13** (1): 125.
 171. Smith, D. G., Pearson, R. A., 2005. A review of the factors affecting the survival of donkeys in semi-arid regions of sub-Saharan Africa. **Tropical Animal Health and Production**, **37**: 1-19.
 172. Smith, S. R., Normandeau, E., Djambazian, H., Nawarathna, P. M., Berube, P., Muir, A. M., Ragoussis, J., Penney, C. M., Scribner, K. T., Luikart, G., Wilson, C. C., Bernatchez, L., 2022. A chromosome-anchored genome assembly for Lake Trout (*Salvelinus namaycush*). **Molecular Ecology Resources**, **22** (2): 679-694.
 173. Stanke, M., Diekhans, M., Baertsch, R., Haussler, D., 2008. Using native and syntenically mapped cDNA alignments to improve *de novo* gene finding. **Bioinformatics**, **24** (5): 637–644.
 174. Stanke, M., Schöffmann, O., Morgenstern, B., Waack, S., 2006. Gene prediction in eukaryotes with a generalized hidden Markov model that uses hints from external sources. **BMC Bioinformatics** **7**: 62.

175. Stenson, P. D., Ball, E. V., Mort, M., Phillips, A. D., Shaw, K., Cooper, D. N., 2012. The Human Gene Mutation Database (HGMD) and its exploitation in the fields of personalized genomics and molecular evolution. **Current protocols in bioinformatics**, **39** (1): 1-13.
176. Sudasinghe, H., Somaweera, R., 2015. *Calotes versicolor* (oriental garden lizard). **Herpetological Review**, **46**: 625–629.
177. Sullivan, T. P., Eaglstein, W. H., Davis, S. C., Mertz, P., 2001. The pig as a model for human wound healing. **Wound repair and regeneration**, **9** (2): 66-76.
178. Sun, T., Pei, S., Liu, Y., Hanif, Q., Xu, H., Chen, N., Lei, C., Yue, X., 2023. Whole genome sequencing of simmental cattle for SNP and CNV discovery. **BMC genomics**, **24** (1): 179.
179. TAGEM, Türkiye Evcil Hayvan Genetik Kaynakları, **TAGEM Yayını**, Ankara, 2009.
180. Tamura, K., Stecher, G., Kumar, S., 2021. MEGA11: molecular evolutionary genetics analysis version 11. **Molecular biology and evolution**, **38** (7): 3022-3027.
181. Tan, M. H., Austin, C. M., Hammer, M. P., Lee, Y. P., Croft, L. J., Gan, H. M., 2018. Finding Nemo: hybrid assembly with Oxford Nanopore and Illumina reads greatly improves the clownfish (*Amphiprion ocellaris*) genome assembly. **Gigascience**, **7** (3): gix137.
182. Tang, B., Wang, Z., Liu, Q., Wang, Z., Ren, Y., Guo, H., Qi, T., Li, Y., Zhang, H., Jiang, S., Ge, B., Xuan, F., Sun, Y., She, S., Tin Yam Chan, Sha, Z., Jiang, H., Li, H., Jiang, W., Qin, Y., 2021. Chromosome-level genome assembly of *Paralithodes platypus* provides insights into evolution and adaptation of king crabs. **Molecular ecology resources**, **21** (2): 511-525.
183. Teeling, E. C., Vernes, S. C., Dávalos, L. M., Ray, D. A., Gilbert, M. T. P., Myers, E., 2018. Bat biology, genomes, and the Bat1K project: to generate chromosome-level genomes for all living bat species. **Annual Review of Animal Biosciences**, **6**: 23–46.
184. The Arabidopsis Genome Initiative, 2000. Analysis of the genome sequence of the flowering plant *arabidopsis thaliana*. **Nature**, **408** (6814): 796–815.

185. The *C. elegans* Sequencing Consortium, 1998. Genome Sequence of the Nematode *C. elegans*: A Platform for Investigating Biology. **Science**, **282** (5396): 2012–2018.
186. Tosser-Klopp, G., Bardou, P., Bouchez, O., Cabau, C., Crooijmans, R., Dong, Y., Donnadieu-Tonon, C., Eggen, A., Heuven, H. C. M., Jamli, S., Jiken, A. J., Klopp, C., Lawley, C. T., McEwan, J., Martin, P., Moreno, C. R., Mulsant, P., Nabihoudine, I., Pailhoux, E., Palhière, I., International Goat Genome Consortium, 2014. Design and characterization of a 52K SNP chip for goats. **PloS one**, **9** (1): e86227.
187. Treangen, T. J., Abraham, A. L., Touchon, M., Rocha, E. P., 2009. Genesis, effects and fates of repeats in prokaryotic genomes. **FEMS microbiology reviews**, **33** (3): 539-571.
188. Turan N, 1987. Antalya-Termesos Yaban Keçisi (*Capra aegagrus aegagrus*) Populasyonunun Gelişimi, Bugünkü Durumu ve Sorunları, s. 83-105. *Türkiye ve Balkan Ülkelerinde Yaban Hayatı*, 16-20 Eylül, İstanbul, Türkiye.
189. Venkatraman, M., Fleischer, R. C., Tsuchiya, M. T. N., 2021. Comparative analysis of annotation pipelines using the first Japanese White-Eye (*Zosterops japonicus*) genome. **Genome Biology and Evolution**, **13** (5): evab063.
190. Venter, J. C., Adams, M. D., Myers, E. W., Li, P. W., Mural, R. J., Sutton, G. G., Smith, H. O., Yandell, M., Evans, C. A., Holt, R. A., Gocayne, J. D., Amanatides, P., Ballew, R. M., Huson, D. H., Wortman, J. R., Zhang, Q., Kodira, C. D., Zheng, X. H., Chen, L., Skupski, M., 2001. The Sequence of the Human Genome. **Science**, **291** (5507): 1304–1351.
191. Vezzi, F., Cattonaro, F., Policriti, A., 2011. E-RGA: enhanced reference guided assembly of complex genomes. **EMBnet Journal**, **17** (1): 46–54.
192. Walker, B. J., Abeel, T., Shea, T., Priest, M., Abouelliel, A., Sakthikumar, S., Cuomo, C. A., Zeng, Q., Wortman, J., Young, S. K., Earl, A. M., 2014. Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. **PloS one**, **9** (11): e112963.

193. Wang, B. A., Ekblom, R., Bunikis, I., Siitari, H., Hoglund, J., 2014. Whole genome sequencing of the black grouse (*Tetrao Tetrix*): reference guided assembly suggests faster-Z and MHC evolution. **BMC Genomics**, **15**: 1-13.
194. Wang, C., Li, H., Guo, Y., Huang, J., Sun, Y., Min, J., Wang, J., Fang, X., Zhao, Z., Wang, S., Zhang, Y., Liu, Q., Jiang, Q., Wang, X., Guo, Y., Yang, C., Wang, Y., Tian, F., Zhuang, G., Fan, Y., 2020. Donkey genomes provide new insights into domestication and selection for coat color. **Nature communications**, **11** (1): 6014.
195. Wang, H., Wan, H. T., Wu, B., Jian, J., Ng, A. H. M., Chung, C. Y.-L., Chow, E. Y.-C., Zhang, J., Wong, A. O. L., Lai, K. P., Chan, T. F., Zhang, E. L., Wong, C. K. C., 2022. A chromosome-level assembly of the Japanese eel genome, insights into gene duplication and chromosomal reorganization. **Gigascience**, **11**: giac120.
196. Wang, J., Chen, G., 2010. The historical variance and causes of geographical distribution of a roughskin sculpin (*Trachidermus fasciatus Heckel*) in Chinese territory. **Acta Ecologica Sinica**, **30**: 6845–6853.
197. Wang, K., Li, M., Hakonarson, H., 2010. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. **Nucleic acids research**, **38** (16): e164-e164.
198. Wang, Q., He, F., Huang, R.Y., Yang, X., Yang, D., Ngatia, J. N., Gong, Y., Xu, Y., Huang, S., Liu, H., 2023. Draft genome of the oriental garden lizard (*Calotes versicolor*). **Frontiers in Genetics**, **14**: 1091544.
199. Weisenfeld, N. I., Kumar, V., Shah, P., Church, D. M., Jaffe, D. B., 2017. Direct determination of diploid genome sequences. **Genome Research**, **27** (5): 757–767.
200. Wiener, P., Wilkinson, S., 2011. Deciphering the genetic basis of animal domestication. **Proceedings Biological Science/Royal Society B**, **278** (1722): 3161–3170.
201. Williams, E. E., 1969. The ecology of colonization as seen in the zoogeography of anoline lizards on small islands. **The Quarterly Review of Biology**, **44** (4): 345-389.

202. Xie, G., Zhang, X., Lv, F., Sang, M., Hu, H., Wang, J., Liu, D., 2021. Nanopore sequencing and Hi-C based *de novo* assembly of *Trachidermus fasciatus* genome. **Genes (Basel)**, **12** (5): 692.
203. Xie, Q., Zhan, W., Shi, J., Liu, F., Niu, B., He, X., Liu, M., Wang, J., Liang, Q., Xie, Y., Xu, P., Wang, X., Lou, B., 2023. Whole-genome assembly and annotation for the little yellow croaker (*Larimichthys polyactis*) provide insights into the evolution of hermaphroditism and gonochorism. **Molecular Ecology Resources**, **23** (3): 632-658.
204. Xiong, J., Bao, J., Hu, W., Shang, M., Zhang, L., 2022. Whole-genome resequencing reveals genetic diversity and selection characteristics of dairy goat. **Frontiers in Genetics**, **13**: 1044017.
205. Xu, T., Li, Y., Chu, Q., Zheng, W., 2021. A chromosome-level genome assembly of the red drum, *Sciaenops ocellatus*. **Aquaculture and Fisheries**, **6** (2): 178-185.
206. Yandell, M., Ence, D., 2012. A beginner's guide to eukaryotic genome annotation. **Nature Reviews Genetics**, **13** (5): 329-342.
207. Yang, L., Wang, Y., Wang, T., Duan, S., Dong, Y., Zhang, Y., He, S. A., 2019b. Chromosome-scale reference assembly of a Tibetan Loach, *Triplophysa siluroides*. **Frontiers in Genetics**, **10**: 991.
208. Yang, Y., Lian, J., Xie, B., Chen, M., Niu, Y., Li, Q., Liu, Y., Yi, G., Fan, X., Tang, Y., Li, J., Liachko, I., Sullivan, S. T., Nelson, B., Zuo, E., Tang, Z., 2019a. Chromosome-scale *de novo* assembly and phasing of a Chinese indigenous pig genome. **BioRxiv**, 770958.
209. Ye, C., Hill, C. M., Wu, S., Ruan, J., Ma, Z. S., 2016. DBG2OLC: efficient assembly of large genomes using long erroneous reads of the third generation sequencing Technologies. **Scientific Reports**, **6**: 31900.
210. Yu, Y., Santat, L. A., Choi, S., 2006. Bioinformatics packages for sequence analysis. In Applied mycology and biotechnology (Vol. 6, pp. 143-160). Elsevier.

211. Zeder, M. A., 2008. Domestication and early agriculture in the Mediterranean Basin: Origins, diffusion, and impact. **Proceedings of the national Academy of Sciences**, **105** (33): 11597-11604.
212. Zeder, M. A., Bradley, D. G., Smith, B. D., Emshwiller, E. (Eds.). 2006. Documenting domestication: new genetic and archaeological paradigms, Univ of California Press.
213. Zeder, M. A., Hesse, B., 2000. The initial domestication of goats (*Capra hircus*) in the Zagros mountains 10,000 years ago. **Science**, **287** (5461): 2254–2257.
214. Zhang, D., Zhang, X., Li, F., Liu, T., Hu, Z., Gao, N., Yuan, L., Li, X., Zhao, Y., Zhao, L., Zhang, Y., Xu, D., Song, Q., Cheng, J., Wang, W., 2021: Whole-genome resequencing identified candidate genes associated with the number of ribs in Hu sheep. **Genomics**, **113** (4): 2077-2084.
215. Zhao, Q., Lin, Z., Chen, J., Xie, Z., Wang, J., Feng, K., Lin, W., Li, H., Hu, Z., Chen, W., Chen, F., Junaid, M., Zhang, H., Xie, Q., Zhang, X., 2023. Chromosome-level genome assembly of goose provides insight into the adaptation and growth of local goose breeds. **GigaScience**, **12**: giad003.
216. Zimin, A. V., Marçais, G., Puiu, D., Roberts, M., Salzberg, S. L., Yorke, J. A., 2013. The MaSuRCA genome assembler. **Bioinformatics**, **29**: 2669–2677.
217. Zimin, A. V., Puiu, D., Luo, M.-C., Zhu, T., Koren, S., Marçais, G., Yorke, J. A., Dvořák, J., Salzberg, S. L., 2017. Hybrid assembly of the large and highly repetitive genome of *Aegilops tauschii*, a progenitor of bread wheat, with the MaSuRCA mega-reads algorithm. **Genome research**, **27** (5): 787-792.
218. Zimin, A. V., Salzberg, S. L., 2020. The genome polishing tool POLCA makes fast and accurate corrections in genome assemblies. **PLoS computational biology**, **16** (6): e1007981.

ÖZGEÇMİŞ

KİŞİSEL BİLGİLER

Adı Soyadı: Saffet Teber

e-mail:

EĞİTİM

Derece	Kurum	Mezuniyet Tarihi
Doktora	Erciyes Üniversitesi, Tarımsal Biyoteknoloji Bölümü	2024
Yüksek Lisans	Erciyes Üniversitesi, Tarımsal Biyoteknoloji Bölümü	2019
Lisans	Erciyes Üniversitesi, Tarımsal Biyoteknoloji Bölümü	2016
Lise	75. Yıl Cumhuriyet Anadolu Lisesi, Kayseri	2011

YABANCI DİL

İngilizce

YAYINLAR

SCI:

1. Karademir, G.K., Teber, S., Caner Kulig, C., Toroslu, A.M., Ibis, O., Yildirim, A. (2023). Complete mitochondrial genome analyses of *Forcipomyia pulchrithorax* (Diptera:Ceratopogonidae): genome orientation and phylogenetic implications. *Kafkas Univ Vet Fak Derg*, 30(2). (Article in Press). DOI: 10.9775/kvfd.2023.30825.
2. Şeker, P.S., Selçuk, A.Y., Selvi, E., Baran, M., Teber, S., et al. (2022). Complete mitochondrial genomes of *Chionomys roberti* and *Chionomys nivalis* (Mammalia: Rodentia) from Turkey: Insight into their phylogenetic position within Arvicolinae. *Organisms Diversity & Evolution*, 22, 821–841.

3. İbiş, O., Selçuk, A. Y., Teber, S., Baran, M., Kaya, A., Özcan, S., ... & Tez, C. (2022). Complete mitogenomes of Turkish tree squirrels, *Sciurus anomalus* and *S. vulgaris*, (Sciuridae: Rodentia: Mammalia) and their phylogenetic status within the tribe Sciurini. *Gene*, 841, 146773.
4. Teber, S., Gürcan, K., Brevet, M., Palmisano, F., Candresse, T., Dallot, S. (2022). High genetic diversity suggests a long evolutionary history of plum pox virus-M in Turkey. *Plant Pathology*, 00, 1– 12. <https://doi.org/10.1111/ppa.13663>.
5. Gürcan, K., Turan, S., Teber, S., Kılınçer, İ., Uz, İ., Tamisier, L., Massart, S., Çağlayan, K. (2021). Molecular and biological characterization of a new mulberry Idaeovirus. *Virus Research*, 298, 198411.
6. Gürcan, K., Teber, S., Akbulut, M. et al. (2021). Genetic diversity and a long evolutionary history of plum pox virus strain rec in Turkey. *European Journal of Plant Pathology* 161, 453–461.
7. Gürcan, K., Teber, S. and Candresse, T. (2020), Genetic analysis suggests a long and largely isolated evolutionary history of plum pox virus strain D in Turkey. *Plant Pathology*, 69, 370-378.
8. Gürcan, K., Teber, S., Çağlayan, K. (2019). Further investigation of a genetically divergent group of plum pox virus-M strain in Turkey. *Journal of Plant Pathology*, 101, 385–391.
9. Teber, S., Ceylan, A., Gürcan, K., Candresse, T., Ulubaş Serçe, Ç., Akbulut, M., Kaymak, S., Akbaş, B. (2019). Genetic diversity and molecular epidemiology of the T strain of Plum pox virus. *Plant Pathology*, 68, 755–763.
10. Gürcan, K., Teber, S., Ercisli, S., Yilmaz, K.U. (2016). Genotyping by Sequencing (GBS) in Apricots and Genetic Diversity Assessment with GBS-Derived Single-Nucleotide Polymorphisms (SNPs). *Biochemical Genetics*, 54, 854-885.

Tam metin bildirir:

1. Teber, S., Gürcan, K., Akbulut, M., Abbasov, M. and Ercişli, S. (2020). High-throughput whole genome sequencing of apricot (*Prunus armeniaca*) cultivar 'Hacıhaliloğlu'. *Acta horticulture*, 1290, 53-58.
2. Gürcan, K. and Teber, S. (2020). Molecular epidemiology of Plum pox virus in Turkey. *Acta horticulture*, 1290, 83-88.
3. Canbulat, O., Gürcan, K., Teber, S. (2020). Analysis of RNAseq reads from orchard trees and seed-borne seedlings for virus detection in apricot (*Prunus armeniaca*). *Acta horticulture*, 1290, 123-128.
4. Turan, S., Teber, S., Canbulat, O., Gürcan, K. (2020). The complete chloroplast genome of *Prunus armeniaca*. *Acta horticulture*, 1290, 41-46.
5. Gürcan, K., Teber, S., Yılmaz, K.U. (2019). Susceptibility of SEO/'Harlayne' type resistant apricot cultivars to PPV-T. *Acta horticulture*, 1242, 521-526.
6. Teber, S., Gürcan, K. (2016). Recombination analysis of 51 PPV isolates including 10 genomes of PPV M İstanbul. *Acta horticulture*, 1163, 85-92.

Sempozyum ve Kongreler:

1. Gürcan, K., Ceylan, A., Teber, S. (2015). Plum Pox Virus with a High Diversity are Prevalent in Turkey. International Conference on Virus and Other Graft Transmissible Disease of Fruit Crops. Japonya, 8-12 Haziran 2015, Abstract book ss 49.
2. Teber, S., Gürcan, K., Koç, A., Bostancı, D., Akçadağ, M., (2015). Natural Stone Fruit Hosts of Plum Pox Virus strain Turkey. International Conference on Virus and Other Graft Transmissible Disease of Fruit Crops. Japonya, 8-12 Haziran 2015, Abstract book ss 48.
3. Gürcan, K., Teber, S. (2016). Türkiye Plum pox virus PPV Haritası. Uluslararası Katılımlı Türkiye VI. Bitki Koruma Kongresi. Konya, Özet Kitaplığı ss 765.

4. Gürcan K., Teber S., Canbulat O., Çağlayan K. (2017). Molecular characterization and genetic diversity of Plum pox virus strains in Thrace region of Turkey. 24th International Conference on Virus and Other Graft Transmissible Diseases of Fruit Crops. June 2017, Thessaloniki (Greece), Abstract book ss 97.
5. Gürcan, K., Teber, S., Ceylan, A., Candresse, T., Akbulut, M., Ulubaş Serçe, Ç., Kaymak, S., Akbaş, B., Akbaş, L. (2018). PPV-T izolatlarının tüm genom analizi; Türkiye'de PPV-T izolatlarının kökeni ve yayılma rotası. Türkiye VII. Bitki Koruma Kongresi (Uluslararası Katılımlı), 14-17 Kasım 2018, Muğla, Türkiye, Özet Kitapcığı ss 103.
6. Teber, S., Uzun, A., Gürcan, K. (2018). Kutdiken limon (Citrus lemon) çeşidinde virom analizi. Türkiye VII. Bitki Koruma Kongresi (Uluslararası Katılımlı), 14-17 Kasım 2018, Muğla, Türkiye, Özet Kitapcığı ss 91.
7. Canbulat, O., Gürcan, K., Teber, S., Olmos, A., Ruiz García, A.B., Çağlayan, K. (2018). Kayısı (Prunus armeniaca) transkriptom veri analizi ile hastalık etmenlerinin belirlenmesi. Türkiye VII. Bitki Koruma Kongresi (Uluslararası Katılımlı), 14-17 Kasım 2018, Muğla, Türkiye, Özet Kitapcığı ss 88.
8. Çetinsay, N., Canbulat, O., Köse, M.A., Teber, S., Caymaz, F., Gürcan, K. (2018). Yerel Kayısı (Prunus armeniaca) Popülasyonlarında Plum Pox Virüsü Dayanıklılığı ve Kendine Uyumazlığın Moleküler Markörler ile Belirlenmesi. Niğde Ömer Halisdemir Üniversitesi 1. Bitki Islahı ve Genetiği Öğrenci Kongresi, Niğde (Tukey), Özet Kitapcığı ss 20.
9. Teber, S., Gürcan, K. (2018) Hacıhaliloğlu Kayısı (Prunus armeniaca) Çeşidinin Tüm Kloroplast Genom Analizi. Niğde Ömer Halisdemir Üniversitesi 1. Bitki Islahı ve Genetiği Öğrenci Kongresi Niğde (Tukey), Özet Kitapcığı ss 18.
10. Canbulat, O., Çetinsay, N., Köse, M.A., Teber, S., Caymaz, F., Gürcan, K. (2018). Goji berry 'de in vitro Bitki Rejenerasyon Protokolünün Oluşturulması. Niğde Ömer Halisdemir Üniversitesi 1. Bitki Islahı ve Genetiği Öğrenci Kongresi Niğde (Tukey), Özet Kitapcığı ss 25.

11. Teber, S. and Gürcan, K. (2021). Plum Pox Virüs Marcus Istanbul (PPV-MIs) İzolatlarının Tüm Genom Analizi, 8. Uluslararası Katılımlı Bitki Koruma Kongresi, 24-28 Ağustos 2021, Online, Özet kitapçığı ss 263.

12. Şenol, Ö., İbiş, O., Tez, C., Görür, G., Teber, S. (2021). Molecular Investigation of Lachnus spp. (Hemiptera:Aphidoidae) Distributed in Turkey. International Congress on Biological and Health Sciences, 26-28 February 2021, Online, Abstract book ss 105.

13. Baran, M., Karataş, M. H., Teber, S., Selçuk, A. Y., İbiş, O., Tez, C. (2023). Complete Mitochondrial Genomes of Otis tarda (Otidiformes: Otididae) From Türkiye, With The Taxonomic implications. VII. International Eurasian Ornithology Congress, 18-21 October 2023, İzmir, Türkiye, Özet Kitapçığı ss 96.

14. Toroslu, A. M., Teber, S., Yıldırım, A., Pekmezci, G. Z., Uslu, S., Kızgın, A., Arslanhan, B., İnci, A., Alto, B. (2023). Phylogenetic Characterization of Aedes albopictus generations in Turkey's

Eastern Black Sea Region Based on Complete Mitochondrial Genome Analyses. 23. Parazitoloji Kongresi, 30 Ekim – 3 Kasım 2023, Antalya, Türkiye, Özet Kitapçığı ss 150-151.

15. Baran, M., İbiş, O., Selçuk, A. Y., Tez, C., Teber, S., Kaya, A., Ulutürk, S. (2023). Whole mitochondrial genome of striped hyena, Hyaena hyaena (Hyaenidae, Carnivora), from Türkiye; insights for phylogenetic relationships. 12th International Molecular Biology and Biotechnology Congress, 13-15 December 2023, Kayseri, Türkiye, Özet Kitapçığı ss XX

Yürüttüğü Projeler:

1. Üniversite Öğrencileri Araştırma Projeleri Destekleme Programı, Kayseri İli Plum Pox Virüs İzolatlarında P3-6k1 Bölgesi Genetik Analizi, 2209-A 2014.

2. 122O844, Ankara Keçisi (Capra hircus) Genomunun Uzun Okuma Platformu İle Dizilenmesi ve Birleştirilmesi, 1002B, 15.10.2022 – 15.04.2023

Yer Aldığı Projeler:

Devam eden projeler:

1. 10916, Türkiye *Microtus* (Arvicolinae: Rodentia: Mammalia) Türlerinin Komple Mitogenom Karakterizasyonu ve Moleküler Filogenisi, BAP-Doktora Sonrası Araştırmacı Projesi, Araştırmacı.
2. 11535, Türkiye Keçilerinde (*Capra hircus*) Tüm Genomun Dizilenmesi ve Biyoinformatik Analizleri, BAP-Doktora Tez Projesi, Araştırmacı.
3. 12262, Türkiye Memelilerinin Mitokondriyal Genom Karakterizasyonu, Filogenisi ve Mitogenom Veri Tabanının Oluşturulması, BAP-ADEP Projesi, Araştırmacı.
4. Enhancing arbovirus surveillance and risk management in the public health systems of Georgia, Turkey, and Ukraine Uluslararası Araştırma Projesi (USA-DTRA), Bursiyer.

Sonuçlanan projeler:

1. 214O400, Shotgun DNA Dizileme ile Kayısı Tüm Genom Fiziksel Harita (Genom) Analizi ve RNAseq Transkriptom Profilleme Yoluyla Kayısı Tüm Gen Setinin Anotasyonu, Uluslararası, Burslu, Sonuçlandı, ARDEB, TOVAG - Tarım, Ormancılık Ve Veterinerlik Araştırma Destek Grubu, Projeye Katılma/Ayrılma Tarihleri: 02.03.2017 - 01.09.2018, Proje Başlangıç/Bitis Tarihleri: 01.04.2016 - 01.10.2018.
2. 118O548, Endemik Plum Pox Virüsü (PPV) Dideron, Marcus, ve Rec İzolatlarında Ful Genom, Rekombinasyon, Filogenetik Analizler ve Yeni Irkların Önerilmesi İle PPV-T Irkında Konukçu ve Patajonite Analizleri, 1001 - Araştırma, Burslu, Sonuçlandı, ARDEB, TOVAG - Tarım, Ormancılık Ve Veterinerlik Araştırma Destek Grubu, Projeye Katılma/Ayrılma Tarihleri: 15.08.2019 - 01.11.2021: 07.11.2018 - 23.07.2019, Proje Başlangıç/Bitis Tarihleri: 01.11.2018 - 01.11.2021.

Sertifikalar:

1. Assembling the zebrafish genome with long-read sequencing, Online, 27 Nisan 2022
2. Workshop on Genomics 2023, Cesky Krumlov/Çekya, 14-27 Mayıs 2023 (Tam burslu)
3. TÜBİTAK 2237 Analitik Doğa – Kümeleme ve Ordinasyon Teknikleri Antalya/Türkiye, 6-12 Aralık 2023