

**T.C.
AKDENİZ ÜNİVERSİTESİ**



**PAMUK (*GOSSYPIUM* L.) GENOMLARINDA MİKROSATELLİT DNA
TİPLERİ VE DAĞILIMLARI**

Damla AKMAN

FEN BİLİMLERİ ENSTİTÜSÜ

TARLA BİTKİLERİ

ANABİLİM DALI

YÜKSEK LİSANS TEZİ

ŞUBAT 2025

ANTALYA

T.C.
AKDENİZ ÜNİVERSİTESİ



**PAMUK (*GOSSYPIUM* L.) GENOMLARINDA MİKROSATELLİT DNA
TİPLERİ VE DAĞILIMLARI**

Damla AKMAN
FEN BİLİMLERİ ENSTİTÜSÜ
TARLA BİTKİLERİ
ANABİLİM DALI
YÜKSEK LİSANS TEZİ

ŞUBAT 2025

ANTALYA

T.C.
AKDENİZ ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

PAMUK (*GOSSYPIMUM L.*) GENOMLARINDA MİKROSATELLİT DNA
TIPLERİ VE DAĞILIMLARI

Damla AKMAN

TARLA BİTKİLERİ

ANABİLİM DALI

YÜKSEK LİSANS TEZİ

Bu tez 22/01/2025 tarihinde jüri tarafından Oybirliği / Oyçokluğu ile kabul edilmiştir.

Prof. Dr. Mehmet KARACA (Danışman)

Prof. Dr. Bayram ÇEVİK

Prof. Dr. Ayşe Gül İNCE

ÖZET

Pamuk (*Gossypium* L.) Genomlarında Mikrosatellit DNA Tipleri ve Dağılımları

Damla AKMAN

Yüksek Lisans Tezi, Tarla Bitkileri Anabilim Dalı

Danışman: Prof. Dr. Mehmet KARACA

Şubat 2025; 149 sayfa

Hızla gelişen tekstil sektörünün en önemli doğal hammaddesi olan pamuk (*Gossypium* L.) üretiminin dünya nüfusunun artmasıyla birlikte artması ve insanların çeşitli ihtiyaçlarını karşılaması ve diğerleriyle birlikte istihdam olanağı yaratması gerekmektedir. Pamuk dünyada Hindistan, Çin, ABD'nin de ilk sıralarda yer aldığı yaklaşık 80 ülkede yetiştiriciliği yapılmaktadır. Pamuk lifleri tekstil için ana kaynak olarak kullanılırken, tohumu yağı için kullanılmakta ve küspesi de proteince zengin hayvan yemi olarak kullanıldığından ülkemiz ve dünyada önemli bir endüstri bitkisidir. Organizmalar genomları çerçevesinde yaşamlarını idame ettiklerinden dolayı genomun yapısı ve işlevi insanoğlunun dikkatini çekme nedeni genomların yapısı ile fonksiyonlarının ilişkili olmalarından kaynaklanmaktadır. Günümüzde özellikle son 10-15 yıllık dönemde Yeni Nesil Dizileme (NGS) teknolojileri yardımıyla çok sayıda genom dizilenmekte ("sequencing") veya tekrar dizileme ("resequencing") yapılabilmektedir. Pamuk genomu da bu dizilemelerden faydalanmıştır.

Bu çalışmanın ana amacı kültürü yapılan tetraploit Upland pamuğun (*Gossypium hirsutum* L.) ve Upland pamuğun oluşumunda ata olan iki diploit pamuğun temsilcileri olan (*G. arboreum* L. ve *G. raimondii* L.) Oryantal (Asya) ve Peru pamuklarının bütün genom dizileri kullanılarak mikrosatellit (SSR) DNA'nın tipleri ve dağılımının türlerin genomları, kromozomları, 5' ve 3' UTR bölgeleri, upstream, downstream, ekson, intron ve intergen dizilerinde ortaya konulması, mikrosatellitlerin fonksiyonu konularında daha fazla bilgi edinmek için selüloz ve selüloz sentaz benzeri genlerinin mikrosatellit analizlerinin yapılması, filogenetik çalışmalar ve nihayetinde de mikrosatellit DNA'nın protein kodlama potansiyelleri üzerine bilgi edinmek amacıyla bu çalışma yürütülmüştür.

Çalışmada allotetraploit bir tür olan ve yaygın olarak Upland pamuk olarak adlandırılan *G. hirsutum*, diploit türler olan ve Oryantal (Asya/Ağaç) pamuk olarak adlandırılan *G. arboreum*, Peru pamuğu olarak bilinen *G. raimondii*, ve bu türlerin bütün genom dizilerine ait referans dizileri kullanılmıştır. Bütün genom dizileri ve genom bilgilerini içeren dosyalar GenBank veri tabanından FASTA, GFF ve GTF formatında indirilmiş olup mikrosatellit içerikleri, genomal dağılımları ve tipleri PERL tabanlı GMATA ve farklı linux tabanlı yazılım programları yardımları ile ortaya konulmuştur. GFF/GTF verileri kullanılarak selüloz sentaz genlerinin konumsal bilgileri AWK, SED ve GREP skriptleri yardımıyla belirlenmiştir.

Analizlerde 3 pamuk türünün genomlarının, alt genomlarının, kromozomlarının mikrosatellit içerikleri ve tekrar sayılarının da türlerde farklı olduğu bulunmuştur. Tri-, tetra-, penta- ve hekza-nükleotitlerin *G. hirsutum* da daha yüksek iken; di-nükleotit tekrarların ise *G. arboreum* ve *G. raimondii* de daha yüksek olduğu tespit edilmiştir. Di-nükleotitler arasında *G. arboreum*'un mikrosatellitleri *G. raimondii*'den daha fazla olduğu sonucuna varılmıştır. Üç türün her kromozomunda mikrosatellit sayıları değişiklik göstermekte olup en yüksek değerler Oryantal pamukta bulunmuştur.

Üç pamuk türünün gen (genik) ve genler arası (intergenik) SSR sıklığı karşılaştırıldığında genik-SSR yönünden en fazla sıklığa sahip olan Peru pamuğu bulunmuş ve bunu Upland pamuğu ve Oryantal pamuğu takip etmiştir. Genler arası SSR yönünden ise en yüksek sıklığa sahip olan Oryantal pamuğu olup bunu Peru pamuğu ve Upland pamuğu takip etmiştir. Her bir pamuk kromozomunda farklı SSR tiplerinin sıklığı ve dağılımı incelendiğinde SSR lokuslarının sıklığı kromozom boyutuyla ilişkili olmadığı belirlenmiştir. Üç pamuk türünün SSR yoğunluklarının farklı olduğu ortaya konmuştur. *G. hirsutum*'da genik-SSR ve intergenik SSR daha fazla iken; *G. raimondii*'de genik ve intergenik SSR sayısının daha az olduğu sonucuna varılmıştır.

Bu çalışma mikrosatellitlerin gen elementlerinin önemli bir bileşeni olduğunu doğrulamıştır. Oryantal pamuk (A genom) ve Upland pamukta (AD genom) tri-nükleotit tekrarları kodlama bölgelerinde daha yaygın olurken; Peru pamukta di-nükleotit tekrarlarının daha fazla olduğu sonucuna ulaşılmıştır. Mikrosatellit dizilerin üç farklı okuma çerçevesinde kodladıkları amino asitler belirlendi ve SSR dizilerinden tri-nükleotit ve penta-nükleotitlerin mono-peptit veya di-peptit tekrarlı amino asitleri kodladığı, di-nükleotit tekrarların ise di-peptit amino asitleri kodladığı belirlenirken, tetra- ve penta-nükleotit tekrarların tetra- ve penta-peptit tekrarlı amino asitleri sentezledikleri belirlenmiştir.

Canlılar aleminin en önemli polisakkariti olan selülozu sentezleyen selüloz ve selüloz sentaz benzeri genler mikrosatellit içerenler ve içermeyenler olarak gruplandırıldı ve mikrosatellit içeren selüloz ve selüloz sentaz benzeri genleri içerdikleri mikrosatellitlerin onların fonksiyonları ve filogenetik pozisyonları yönünden etkili oldukları sonucuna varılmıştır. Mikrosatellitlerin verici diploit genomlardaki sayıları ile tetraploit genomdaki sayıları karşılaştırıldığında evrim sürecinde bir kısım mikrosatellitlerin kayıp olduğu ancak bazı yeni mikrosatellitlerin ise *de novo* oluştuğu tespit edilmiştir.

Çalışmanın diğer bir sonucu olarak NGS tabanlı referans genom bilgileri kullanılarak herhangi bir organizmanın tüm mikrosatellit türlerini tanımlamak ve genik ile intergenik mikrosatellitleri tür içerisinde veya türler arası karşılaştırmaların mümkün olduğu görülmüştür. Bu sayede gen, gen bölgelerine özgü fonksiyonel mikrosatellit (SSR) markırlarının elde edilebileceği bu tez çalışmasıyla teyit edilebilmiştir.

ANAHTAR KELİMELELER: *Gossypium hirsutum*, *Gossypium arboreum*, *Gossypium raimondii*, *In silico*, Mikrosatellit, Pamuk, Referans genom, Selüloz sentaz

JÜRİ: Prof. Dr. Mehmet KARACA

Prof. Dr. Bayram ÇEVİK

Prof. Dr. Ayşe Gül İNCE



ABSTRACT

Types and Distributions of Microsatellite DNA in the Cotton (*Gossypium* L.) Genome

Damla AKMAN

M.Sc. Thesis in the Department of Field Crops

Supervisor: Prof. Dr. Mehmet KARACA

February 2025; 149 pages

Production of Cotton (*Gossypium* L.), which is the most important natural raw material of the rapidly developing textile industry, needs to increase with the increase in the world population and should meet the various needs of people and create employment opportunities with its other by-products. Cotton is grown in approximately 80 countries in the world, including India, China and the USA. While cotton fibers are used as the main source for textiles, its seeds are used for its oil and its cake is used as protein-rich animal feed, so it is an important industrial plant in our country and in the world. Since organisms survive within the framework of their genomes, the structure and function of the genome attract human attention because the structure and functions of genomes are related. Nowadays, especially in the last 10-15 years, many genomes have been sequenced or resequenced with the help of Next Generation Sequencing (NGS) technologies. The cotton genome also benefited from these sequencings.

The main objective of this study was to use the whole genome sequences of cultured tetraploid Upland cotton (*Gossypium hirsutum* L.) and Oriental (Asian) and Peruvian cottons, representatives of the two diploid cottons (*G. arboreum* L. and *G. herbaceum* L.) that are the ancestors of Upland cotton. This study was carried out to obtain information regarding the types and distribution of microsatellite (SSR) DNA in species genomes, chromosomes, 5' UTR, 3' UTR, Gene, Upstream, Downstream, Exon, Intron and Intergene sequences. Furthermore, microsatellite analysis of cellulose and cellulose synthase-like genes to obtain more information on the function of microsatellites, phylogenetic studies and protein coding of microsatellite DNA were studied.

In the study, reference sequences constructed using whole genome sequences of *G. hirsutum*, which is an allotetraploid species and commonly called Upland cotton, *G. arboreum*, which is a diploid species and called Oriental (Asian/Tree) cotton, and *G. raimondii*, known as Peruvian cotton, were used. Whole genome sequences and files containing genome information were downloaded from the GenBank database in FASTA, GFF and GTF formats, and their microsatellite contents, genomic distributions and types of microsatellites were revealed using PERL-based GMATA and Linux-based software programs. Using GFF/GTF data, positional information of cellulose synthase genes was determined with the help of AWK, SED and GREP scripts.

In this study, it was found that the microsatellite contents and repeat types of the genomes, sub-genomes and chromosomes of 3 cotton species were different among the species. It was determined that tri-, tetra-, penta- and hexa-nucleotides were higher in *G. hirsutum*, while di-nucleotide repeats were higher in *G. arboreum* and *G. raimondii*. It was concluded that *G. arboreum* had more microsatellites than *G. raimondii* among di-nucleotides. Microsatellite numbers vary in each chromosome of the three species. Among species, the highest number of microsatellite values were found in Oriental cotton.

When the genic and intergenic SSR frequencies of three cotton species were compared, Peruvian cotton was found to have the highest frequency in terms of genic-SSR (SSRs within 5' UTR, 3' UTR, gene, upstream, downstream and exon), followed by Upland cotton and Oriental cotton. In terms of intergenic SSR, Oriental cotton had the highest frequency, followed by Peruvian cotton and Upland cotton. When the frequency and distribution of different SSR types in each cotton chromosome were examined, it was determined that the frequency of SSR loci was not related to chromosome size. It has been revealed that the SSR densities of the three cotton types are different. It was concluded that while genic-SSR and intergenic SSR were more in *G. hirsutum*, genic and intergenic SSR were less in *G. raimondii*.

This study confirmed that microsatellites are important components of gene elements. In Oriental cotton (A genome) and Upland cotton (AD genome), tri-nucleotide repeats were more common in coding regions; It was concluded that di-nucleotide repeats were higher in Peruvian cotton. The amino acids encoded by microsatellite sequences in three different reading frames were determined and the amino acid sequences they encoded were identified. It was determined that tri-nucleotide and penta-nucleotides code for mono-peptide or di-peptide repeated amino acids, dinucleotide repeats code for di-peptide amino acids, while tetra- and penta-nucleotide repeats code for tetra-peptide amino acids.

Cellulose and cellulose synthesis-like genes that synthesize cellulose, the most important polysaccharide of the living kingdom, were grouped as those containing microsatellites and those not containing microsatellites, and it was concluded that microsatellites with cellulose and cellulose synthesis-like genes existed, and they were effective in terms of the functions and phylogenetic positions of the microsatellites they contain. When the numbers of microsatellites in the donor diploid genomes were compared with the numbers in the tetraploid genome, it was determined that some microsatellites were probably lost during the evolution process, but some new microsatellites were formed. In this study, it was also found that by using NGS reference genome information, it is possible to identify all microsatellite types of any organism and compare genic and intergenic microsatellites within and between genomes. In this way, it has been confirmed in this thesis that functional microsatellite (SSR) markers specific to genes and gene regions can be obtained.

KEY WORDS: Cellulose synthases, Cotton, *Gossypium hirsutum*, *Gossypium arboreum*, *Gossypium raimondii*, *In silico*, Microsatellite, Reference genome

COMMITTEE: Prof. Dr. Mehmet KARACA

Prof. Dr. Bayram EVİK

Prof. Dr. Ayşe Gül İNCE



ÖNSÖZ

Dünyada ve ülkemizde önemli bir endüstri bitkisi olan pamuk birinci derecede lif ve ikinci derecede önemli bir yağ bitkisi olup ülkemiz ekonomisinde de stratejik bir öneme sahiptir. Pamuk işlenmesi açısından çırçır sanayisinin, lifi ile tekstil sanayisinin, tohumu ile yağ ve yem sanayisinin, linteri ile de kâğıt sanayisinin hammaddesi durumundadır. Pamuk ayrıca plastik, cila, barut, kozmetik, boya, diş macunu, araç döşeme, mobilya ve para kâğıdı gibi değişik alanlara hammadde sunan bir bitkidir.

Günümüzde mikrosatellitler (SSR) yaygın olarak genom karakterizasyonu ve gen haritalama çalışmalarında moleküler markır olarak kullanılmaktadır. Pamuk bitkisinde bütün genom dizileri kullanılarak türler arası ve türler içi karşılaştırmalı olarak mikrosatellitlerin dağılımı ile ilgili çalışma sayısı sınırlıdır. Bu nedenle elde edilecek her türlü bilgi pamuk ıslahı ve üretimine katkı sağlama potansiyeli bulunmaktadır. Pamuk bitkisi önemli bir yağ bitkisi olmasının yanında birinci derecede lif bitkisidir. Bu bağlamda lif gelişimi ile ilgili mikrosatellitlerin tespitinde kullanılacak yöntemler lif kalite ve miktarının artırılmasında önemlilik göstermektedir.

Çalışmamın her aşamasında beni destekleyen, bilgi, beceri ve deneyimlerini hiçbir zaman benden esirgemeyen, desteğini her zaman hissetmekten onur duyduğum değerli danışman hocam Prof. Dr. Mehmet KARACA'ya tez çalışmama olan katkılarından dolayı teşekkürlerimi ve saygılarımı sunarım. Yüksek lisans tez konumun belirlenmesinden bitiş sürecine kadar manevi desteğini esirgemeyen Sayın Prof. Dr. Ayşe Gül İNCE hocama en içten teşekkürlerimi sunarım. Ayrıca sayın Prof. Dr. Bayram ÇEVİK hocama katkılarından dolayı teşekkür ederim.

Eğitim-öğretimim süresince, her daim yanımda olan maddi-manevi desteklerini esirgemeyen sevgili babam Ali AKMAN ve sevgili annem Hamiyet AKMAN'a sonsuz teşekkürlerimi sunarım.

İÇİNDEKİLER

ÖZET.....	i
ABSTRACT.....	iv
ÖNSÖZ	vii
AKADEMİK BEYAN	ix
SİMGELER VE KISALTMALAR.....	x
ŞEKİLLER DİZİNİ.....	xiii
ÇİZELGELER DİZİNİ	xviii
1. GİRİŞ	1
2. KAYNAK TARAMASI	6
2.1. Pamuk Bitkisinin Genomu.....	6
2.2. Mikrosatellitler (SSR) Tanımı, Oluşumu ve Kullanım Alanları.....	7
2.3. Mikrosatellitlerin Belirlenmesi	8
2.4. Genom Dizilemeleri.....	9
2.5. Selüloz Sentaz Gen Ailesi.....	10
2.6. Pamuk Bitkisinde Mikrosatellit Çalışmaları.....	11
3. MATERYAL VE METOT	13
3.1. Genom DNA Dizileri.....	13
3.2. Metot.....	15
3.2.1. Pamuk Türlerinin Mikrosatellit Analizleri	15
3.2.2. Pamuk Türlerine ait Selüloz Sentaz Gen Ailesinin Mikrosatellit İçerikleri ve Filogenetik Analizleri.....	17
3.2.3. Pamuk Türlerine ait Selüloz Sentaz Mikrosatellitlerinin Kodladığı Aminoasitlerin Belirlenmesi	22
4. BULGULAR VE TARTIŞMA	23
4.1. Bütün Genom Düzeyinde Mikrosatellitler (SSR).....	23
4.1.1. Genom ve Kromozomlarda SSR Dağılımları	23
4.1.2. SSR Motif Uzunlukları ve Motif İçerikleri.....	28
4.1.3. Genomlarda SSR Motif Sıklıkları.....	41
4.1.4. Gen Elementleri Dizilerinde SSR Dağılımları.....	43
4.2. Mikrosatellit İçeren Selüloz Sentaz ve Selüloz Sentaz Benzeri Genler ve Filogenetik İlişkiler.....	83
4.3. Selüloz Sentaz ve Selüloz Sentaz Benzeri Genlerdeki Mikrosatellit Dizilerin Kodladığı Aminoasitler.....	101
5. SONUÇLAR	135
6. KAYNAKLAR	137
ÖZGEÇMİŞ	

AKADEMİK BEYAN

Yüksek Lisans Tezi olarak sunduğum “Pamuk (*Gossypium* L.) Genomlarında Mikrosatellit DNA Tipleri ve Dağılımları” başlıklı bu çalışmanın akademik kurallar ve etik değerlere uygun olduğunu belirtir, bu tez çalışmasında bana ait olmayan tüm bilgilerin kaynağını gösterdiğimi beyan ederim.

22/01/2025

Damla AKMAN

SİMGELER VE KISALTMALAR

Simgeler

% : yüzde

° : derece

°C : santigrat derece

bp : “base pair”, baz çifti

Gb : gigabaz

kb : kilobaz

Mb : megabaz

β : beta

Kısaltmalar

ABD : Amerika Birleşik Devletleri

BED : “Browser Extensible Data”

cDNA : “Complementary DNA”

cesA : “Cellulose Synthase A”

cpSSR : kloroplast SSR

Csl : “Cellulose Synthase-like”

da : dekar

DNA : Deoksiribonükleik Asit

EST : “Expressed Sequence Tag”

EST-SSR : “Expressed Sequence Tag-Simple Sequence Repeat”

ExPASy : “Expert Protein Analysis System”

gDNA : genomik DNA

GFF : “General Feature Format”

GMATA : “Genome-wide Microsatellite Analyzing Toward Application”

gSSR : genomik SSR

GTF : “General Transfer Format”

GWAS : “Genome-Wide Association Study”

ha : hektar

ICAC : “International Cotton Advisory Committee”

ISSR : “Inter-Simple Sequence Repeat”

MAS : “Marker Assisted Selection”

miRNA : mikro-RNA

mtSSR : mitokondriyal SSR

nuSSR	: nükleer SSR
PIC	: “Polymorphic Information Content”
RAPD	: “Random Amplified Polymorphic DNA”
snRNA	: “Small nuclear RNA”
SSCP	: “Single-Strand Conformation Polymorphism”
SSRLP	: “Simple Sequence Repeat Length Polymorphism”
SSRs	: “Simple Sequence Repeats”
TM-1	: “Texas-Marker-1”
tRNA	: transfer RNA
TÜİK	: Türkiye İstatistik Kurumu

ŞEKİLLER DİZİNİ

Şekil 1.1. Pamuğun (<i>Gossypium</i> L.) bazı kullanım alanlarının şematik gösterimi.....	1
Şekil 2.1. Pamuğun evcilleştiği bölgeler.....	6
Şekil 2.2. Tekrarlı motif dizileri.	7
Şekil 2.3. Kısa okumalardan elde edilen dizilerin montajı	10
Şekil 3.1. Çalışmada kullanılan gen bölgeleri	16
Şekil 4.1. Üç pamuk türünün SSR sayısı ve genom büyüklükleri yönünden karşılaştırılması	27
Şekil 4.2. Oryantal pamuğa ait en sık görülen ilk üç motif içerikleri.....	29
Şekil 4.3. Oryantal pamuğa ait kromozom düzeyinde en sık görülen SSR motif içerikleri	31
Şekil 4.4. Peru pamuğuna ait en sık görülen ilk üç SSR motif içerikleri	33
Şekil 4.5. Peru pamuğuna ait kromozom düzeyinde en sık görülen SSR motif içerikleri	34
Şekil 4.6. Upland pamuğuna ait en sık görülen ilk üç motif içerikleri	37
Şekil 4.7. Upland pamuğa ait kromozom düzeyinde en sık görülen SSR motif içerikleri	38
Şekil 4.8. Üç pamuk türünün SSR dağılımları.....	39
Şekil 4.9. Üç pamuk türünün SSR motif içerikleri	40
Şekil 4.10. Oryantal pamuk A genomunda motif sıklıkları	41
Şekil 4.11. Peru pamuk D genomunda SSR motif sıklıkları	42
Şekil 4.12. Upland pamuk A alt genomunda SSR motif sıklıkları	42
Şekil 4.13. Upland pamuk D alt genomunda SSR motif sıklıkları	43
Şekil 4.14. Oryantal pamuk, Peru pamuk ve Upland pamukta gen element dizilerinin SSR yönünden karşılaştırılması.....	66
Şekil 4.15. Oryantal pamukta (A genom) artı sarmal için gen elementlerinin karşılaştırılması	68
Şekil 4.16. Oryantal pamukta (A genom) eksi sarmal için gen elementlerinin karşılaştırılması	69
Şekil 4.17. Peru pamukta (D genom) artı sarmal için gen elementlerinin karşılaştırılması	70
Şekil 4.18. Peru pamukta (D genom) eksi sarmal için gen elementlerinin karşılaştırılması	71
Şekil 4.19. Upland pamukta (A alt genom) artı sarmal için gen elementlerinin karşılaştırılması	72
Şekil 4.20. Upland pamukta (A alt genom) eksi sarmal için gen elementlerinin karşılaştırılması	73

Şekil 4.21. Upland pamukta (D alt genom) artı sarmal için gen elementlerinin karşılaştırılması	74
Şekil 4.22. Upland pamukta (D alt genom) eksi sarmal için gen elementlerinin karşılaştırılması	75
Şekil 4.23. Oryantal pamukta mikrosatellit içeren selüloz sentaz ve selüloz sentaz benzeri genlerin filogenetik analizi.....	85
Şekil 4.24. Peru pamukta için mikrosatellit içeren selüloz sentaz ve selüloz sentaz benzeri genlerin filogenetik analizi.....	88
Şekil 4.25. Upland pamukta mikrosatellit içeren selüloz sentaz ve selüloz sentaz benzeri genlerin filogenetik analizi.....	92
Şekil 4.26. Maksimum olasılık yöntemi ve Tamura-Nei modelini kullanılarak tahmin edilen evrimsel analiz.....	93
Şekil 4.27. Maksimum olasılık yöntemi ve Tamura-Nei modeli kullanılarak SSR içeren ve SSR içermeyen <i>CesA</i> ve <i>Csl</i> genlerinin filogenetik analizi.....	94
Şekil 4.28. Oryantal pamukta mikrosatellit içeren selüloz sentaz ve selüloz sentaz benzeri genlerin protein için filogenetik analizi.....	96
Şekil 4.29. Peru pamukta mikrosatellit içeren selüloz sentaz ve selüloz sentaz benzeri genlerin protein için filogenetik analizi.....	98
Şekil 4.30. Upland pamukta mikrosatellit içeren selüloz sentaz ve selüloz sentaz benzeri genlerin protein için filogenetik analizi.....	100
Şekil 4.31. Oryantal pamuğa ait ID numarası 108469566 olan <i>CesA3-like</i> genine ait [GA] ₂₂ motifinin kodladığı aminoasitler.....	101
Şekil 4.32. Oryantal pamuğa ait ID numarası 108457518 olan <i>CslE1</i> genine ait [AT] ₁₀ motifinin kodladığı aminoasitler	102
Şekil 4.33. Oryantal pamuğa ait ID numarası 108464531 olan <i>CslE1</i> genine ait [AG] ₁₂ motifinin kodladığı aminoasitler.....	102
Şekil 4.34. Oryantal pamuğa ait ID numarası 108464531 olan <i>CesA2</i> genine ait [TA] ₉ motifinin kodladığı aminoasitler	103
Şekil 4.35. Oryantal pamuğa ait ID numarası 108476974 olan <i>CslG3</i> genine ait [TA] ₁₁ motifinin kodladığı aminoasitler.....	103
Şekil 4.36. Oryantal pamuğa ait ID numarası 108475592 olan <i>CesA7-like</i> genine ait [AT] ₈ motifinin kodladığı aminoasitler.....	104
Şekil 4.37. Oryantal pamuğa ait ID numarası 108475592 olan <i>CesA7-like</i> genine ait [AG] ₁₉ motifinin kodladığı aminoasitler.....	104
Şekil 4.38. Oryantal pamuğa ait ID numarası 108475592 olan <i>CesA7-like</i> genine ait [AG] ₁₉ motifinin kodladığı aminoasitler.....	105
Şekil 4.39. Oryantal pamuğa ait ID numarası 108469578 olan <i>CesA3</i> genine ait [AT] ₂₈ motifinin kodladığı aminoasitler	105
Şekil 4.40. Oryantal pamuğa ait ID numarası 108469868 olan <i>CesA4-like</i> genine ait [TTC] ₈ motifinin kodladığı aminoasitler.....	106

Şekil 4.41. Oryantal pamuğa ait ID numarası 108454833 olan <i>CsIG3</i> genine ait [AAT] ₅ motifinin kodladığı aminoasitler	107
Şekil 4.42. Oryantal pamuğa ait ID numarası 108450932 olan <i>CsIG3</i> genine ait [TACA] ₁₈ motifinin kodladığı aminoasitler	107
Şekil 4.43. Oryantal pamuğa ait ID numarası 108479552 olan <i>CsID5</i> genine ait [TAT] ₅ motifinin kodladığı aminoasitler	108
Şekil 4.44. Oryantal pamuğa ait ID numarası 108479552 olan <i>CsID5</i> genine ait [ATTT] ₅ motifinin kodladığı aminoasitler	108
Şekil 4.45. Oryantal pamuğa ait ID numarası 108479552 olan <i>CsID5</i> genine ait [TTTTA] ₄ motifinin kodladığı aminoasitler	109
Şekil 4.46. Oryantal pamuğa ait ID numarası 108469859 olan <i>CsID3</i> genine ait [AGAAAG] ₅ motifinin kodladığı aminoasitler.....	109
Şekil 4.47. Peru pamuğuna ait ID numarası 105800592 olan <i>CsIG2</i> genine ait [AT] ₁₄ motifinin kodladığı aminoasitler	110
Şekil 4.48. Peru pamuğuna ait ID numarası 105804459 olan <i>CsIE1</i> genine ait [TA] ₉ motifinin kodladığı aminoasitler	111
Şekil 4.49. Peru pamuğuna ait ID numarası 105765323 olan <i>CsIG3</i> genine ait [TA] ₉ motifinin kodladığı aminoasitler	111
Şekil 4.50. Peru pamuğuna ait ID numarası 105781668 olan <i>CsIE6</i> genine ait [TA] ₁₅ motifinin kodladığı aminoasitler	112
Şekil 4.51. Peru pamuğuna ait ID numarası 128039937 olan <i>CesA8-like</i> genine ait [AT] ₁₈ motifinin kodladığı aminoasitler	112
Şekil 4.52. Peru pamuğuna ait ID numarası 128039937 olan <i>CesA8-like</i> genine ait [TA] ₁₁ motifinin kodladığı aminoasitler	113
Şekil 4.53. Peru pamuğuna ait ID numarası 128039937 olan <i>CesA8-like</i> genine ait [TGG] ₇ motifinin kodladığı aminoasitler.....	113
Şekil 4.54. Peru pamuğuna ait ID numarası 105767346 olan <i>CesA1</i> genine ait [AT] ₁₃ motifinin kodladığı aminoasitler	114
Şekil 4.55. Peru pamuğuna ait ID numarası 105767346 olan <i>CesA1</i> genine ait [TTA] ₅ motifinin kodladığı aminoasitler	114
Şekil 4.56. Peru pamuğuna ait ID numarası 105767346 olan <i>CesA1</i> genine ait [TTTA] ₄ motifinin kodladığı aminoasitler.....	115
Şekil 4.57. Peru pamuğuna ait ID numarası 105767346 olan <i>CesA1</i> genine ait [CGACC] ₅ motifinin kodladığı aminoasitler.....	115
Şekil 4.58. Peru pamuğuna ait ID numarası 105769839 olan <i>CesA1</i> genine ait [TGC] ₅ motifinin kodladığı aminoasitler	116
Şekil 4.59. Peru pamuğuna ait ID numarası 105800594 olan <i>CsIE1</i> genine ait [ATT] ₅ motifinin kodladığı aminoasitler.....	116
Şekil 4.60. Peru pamuğuna ait ID numarası 105785384 olan <i>CesA2</i> genine ait [TTA] ₇ motifinin kodladığı aminoasitler.....	117

Şekil 4.61. Peru pamuğuna ait ID numarası 105800206 olan <i>CsIG3</i> genine ait [TAA] ₁₇ motifinin kodladığı aminoasitler	117
Şekil 4.62. Peru pamuğuna ait ID numarası 105764976 olan <i>CsID3</i> genine ait [AAT] ₅ motifinin kodladığı aminoasitler	118
Şekil 4.63. Peru pamuğuna ait ID numarası 105781684 olan <i>Prob-CesA5</i> genine ait [AAAT] ₄ motifinin kodladığı aminoasitler	118
Şekil 4.64. Peru pamuğuna ait ID numarası 105800591 olan <i>CsIG3</i> genine ait [GGTGC] ₄ motifinin kodladığı aminoasitler.....	119
Şekil 4.65. Peru pamuğuna ait ID numarası 105800591 olan <i>CsIG3</i> genine ait [TGCAA] ₄ motifinin kodladığı aminoasitler	119
Şekil 4.66. Upland pamuğuna ait ID numarası 107928590 olan <i>CsIG2</i> genine ait [AT] ₈ motifinin kodladığı aminoasitler.....	120
Şekil 4.67. Upland pamuğuna ait ID numarası 107914099 olan <i>CesA5</i> genine ait [TC] ₈ motifinin kodladığı aminoasitler	121
Şekil 4.68. Upland pamuğuna ait ID numarası 107914099 olan <i>CesA5</i> genine ait [TGC] ₅ motifinin kodladığı aminoasitler	121
Şekil 4.69. Upland pamuğuna ait ID numarası 107900907 olan <i>CsII1</i> genine ait [AG] ₁₁ motifinin kodladığı aminoasitler	122
Şekil 4.70. Upland pamuğuna ait ID numarası 107896452 olan <i>CsIG3</i> genine ait [TA] ₉ motifinin kodladığı aminoasitler	122
Şekil 4.71. Upland pamuğuna ait ID numarası 107907451 olan <i>CesA7</i> genine ait [TA] ₁₂ motifinin kodladığı aminoasitler	123
Şekil 4.72. Upland pamuğuna ait ID numarası 107907451 olan <i>CesA7</i> genine ait [TAA] ₆ motifinin kodladığı aminoasitler	123
Şekil 4.73. Upland pamuğuna ait ID numarası 107907451 olan <i>CesA7</i> genine ait [TTTA] ₄ motifinin kodladığı aminoasitler	124
Şekil 4.74. Upland pamuğuna ait ID numarası 107953681 olan <i>CesA7-like</i> genine ait [AC] ₈ motifinin kodladığı aminoasitler	124
Şekil 4.75. Upland pamuğuna ait ID numarası 107935226 olan <i>CsIE1</i> genine ait [TA] ₈ motifinin kodladığı aminoasitler.....	125
Şekil 4.76. Upland pamuğuna ait ID numarası 121229536 olan <i>CesA5-like</i> genine ait [TGC] ₅ motifinin kodladığı aminoasitler	125
Şekil 4.77. Upland pamuğuna ait ID numarası 107955410 olan <i>CesA4</i> genine ait [AAG] ₅ motifinin kodladığı aminoasitler	126
Şekil 4.78. Upland pamuğuna ait ID numarası 107898377 olan <i>CsII1</i> genine ait [ATT] ₅ motifinin kodladığı aminoasitler	126
Şekil 4.79. Upland pamuğuna ait ID numarası 107937655 olan <i>CesA4-like</i> genine ait [AAG] ₅ motifinin kodladığı aminoasitler	127
Şekil 4.80. Upland pamuğuna ait ID numarası 107911480 olan <i>CesA2</i> genine ait [TTA] ₇ motifinin kodladığı aminoasitler.....	127

Şekil 4.81. Upland pamuğuna ait ID numarası 107949226 olan <i>CsID4</i> genine ait [ATT] ₅ motifinin kodladığı aminoasitler	128
Şekil 4.82. Upland pamuğuna ait ID numarası 107958687 olan <i>CesA1</i> genine ait [TAT] ₅ motifinin kodladığı aminoasitler	128
Şekil 4.83. Upland pamuğuna ait ID numarası 107907453 olan <i>CesA1</i> genine ait [TAA] ₉ motifinin kodladığı aminoasitler	129
Şekil 4.84. Upland pamuğuna ait ID numarası 107901762 olan <i>CesA5</i> genine ait [TAA] ₈ motifinin kodladığı aminoasitler	129
Şekil 4.85. Upland pamuğuna ait ID numarası 107935218 olan <i>CsIE1</i> genine ait [AAT] ₆ motifinin kodladığı aminoasitler	130
Şekil 4.86. Upland pamuğuna ait ID numarası 107935218 olan <i>CsIE1</i> genine ait [AAAT] ₄ motifinin kodladığı aminoasitler	130
Şekil 4.87. Upland pamuğuna ait ID numarası 107896451 olan <i>CsIG2</i> genine ait [TTTC] ₄ motifinin kodladığı aminoasitler	131
Şekil 4.88. Upland pamuğuna ait ID numarası 107935224 olan <i>CsIE1</i> genine ait [ACAT] ₄ motifinin kodladığı aminoasitler	131
Şekil 4.89. Upland pamuğuna ait ID numarası 107945371 olan <i>Prob-CesA8</i> genine ait [ATTTT] ₄ motifinin kodladığı aminoasitler	132
Şekil 4.90. Upland pamuğuna ait ID numarası 107950592 olan <i>CesA2</i> genine ait [AGAAA] ₄ motifinin kodladığı aminoasitler	132
Şekil 4.91. Upland pamuğuna ait ID numarası 107950592 olan <i>CesA2</i> genine ait [AAAAG] ₅ motifinin kodladığı aminoasitler	133
Şekil 4.92. Upland pamuğuna ait ID numarası 107912646 olan <i>CsIH1</i> genine ait [CCATGC] ₄ motifinin kodladığı aminoasitler	134

ÇİZELGELER DİZİNİ

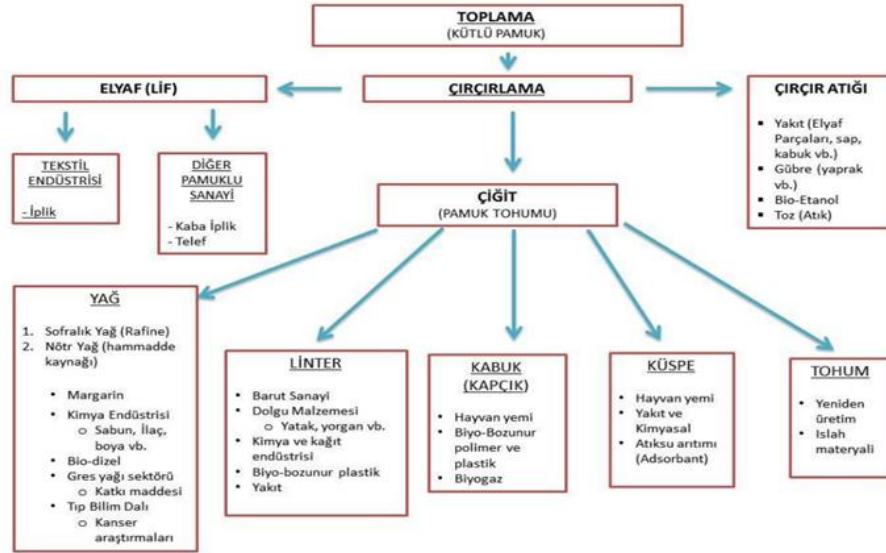
Çizelge 1.1. TÜİK 2019-2023 illere göre pamuk ekim alanı (bin da).....	2
Çizelge 1.2. TÜİK 2019-2023 illere göre kütlü pamuk üretim miktarı (bin ton)	3
Çizelge 1.3. TÜİK 2019-2023 illere göre lif pamuk üretim miktarı (bin ton).....	3
Çizelge 1.4. TÜİK 2019-2023 illere göre pamuk çiğdi üretim miktarı (bin ton)	4
Çizelge 3.1. Upland pamuk referans genomuna ait bilgiler.....	13
Çizelge 3.2. Oryantal pamuk referans genomuna ait bilgiler.....	14
Çizelge 3.3. Peru pamuk referans genomuna ait bilgiler.....	15
Çizelge 3.4. Oryantal pamukta kullanılan selüloz sentaz A ve selüloz sentaz benzeri genler.....	17
Çizelge 3.5. Peru pamukta kullanılan selüloz sentaz A ve selüloz sentaz benzeri genler.....	18
Çizelge 3.6. Upland pamuğa ait selüloz sentaz A ve selüloz sentaz benzeri genler.....	20
Çizelge 4.1. Oryantal pamuğa ait kromozom ve genlerin SSR dağılımları.....	23
Çizelge 4.2. Peru pamuğa ait kromozom ve genlerin SSR dağılımları.....	25
Çizelge 4.3. Upland pamuğa ait kromozom ve genlerin SSR dağılımları.....	26
Çizelge 4.4. Oryantal pamuğa ait kromozom büyüklükleri ve SSR dağılımları.....	28
Çizelge 4.5. Peru pamuğuna ait kromozom büyüklükleri ve SSR dağılımları.....	32
Çizelge 4.6. Upland pamuğa ait A alt genomu kromozom büyüklükleri ve SSR dağılımları.....	35
Çizelge 4.7. Upland pamuğa ait D alt genomu kromozom büyüklükleri ve SSR dağılımları.....	36
Çizelge 4.8. Oryantal pamuk +/- sarmal için upstream dizilerinde SSR dağılımları.....	44
Çizelge 4.9. Oryantal pamuk +/- sarmal için downstream dizilerinde SSR dağılımları.....	45
Çizelge 4.10. Oryantal pamuk +/- sarmal için ekson dizilerinde SSR dağılımları.....	46
Çizelge 4.11. Oryantal pamuk +/- sarmal için intron dizilerinde SSR dağılımları.....	46
Çizelge 4.12. Oryantal pamuk +/- sarmal için 5'-UTR dizilerinde SSR dağılımları.....	47
Çizelge 4.13. Oryantal pamuk +/- sarmal için 3'-UTR dizilerinde SSR dağılımları.....	48
Çizelge 4.14. Oryantal pamuk +/- sarmal için intergen dizilerinde SSR dağılımları.....	49
Çizelge 4.15. Oryantal pamuk +/- sarmal için tüm gen dizilerinde SSR dağılımları.....	49

Çizelge 4.16. Peru pamukta +/- sarmal için upstream dizilerinde SSR dağılımları.....	50
Çizelge 4.17. Peru pamukta +/- sarmal için downstream dizilerinde SSR dağılımları.....	51
Çizelge 4.18. Peru pamukta +/- sarmal için ekson dizilerinde SSR dağılımları.....	52
Çizelge 4.19. Peru pamukta +/- sarmal için intron dizilerinde SSR dağılımları.....	53
Çizelge 4.20. Peru pamukta +/- sarmal için 5'-UTR dizilerinde SSR dağılımları.....	53
Çizelge 4.21. Peru pamukta +/- sarmal için 3'-UTR dizilerinde SSR dağılımları.....	54
Çizelge 4.22. Peru pamukta +/- sarmal için intergen dizilerinde SSR dağılımları.....	55
Çizelge 4.23. Peru pamukta +/- sarmal için tüm gen dizilerinde SSR dağılımları.....	56
Çizelge 4.24. Upland pamukta +/- sarmal için upstream dizilerinde SSR dağılımları.....	57
Çizelge 4.25. Upland pamukta +/- sarmal için downstream dizilerinde SSR dağılımları.....	58
Çizelge 4.26. Upland pamukta +/- sarmal için ekson dizilerinde SSR dağılımları.....	59
Çizelge 4.27. Upland pamukta +/- sarmal için intron dizilerinde SSR dağılımları.....	60
Çizelge 4.28. Upland pamukta +/- sarmal için 5'-UTR dizilerinde SSR dağılımları.....	61
Çizelge 4.29. Upland pamukta +/- sarmal için 3'-UTR dizilerinde SSR dağılımları.....	62
Çizelge 4.30. Upland pamukta +/- sarmal için intergen dizilerinde SSR dağılımları.....	63
Çizelge 4.31. Upland pamukta +/- sarmal için tüm gen dizilerinde SSR dağılımları.....	64
Çizelge 4.32. Oryantal pamukta + ve- sarmal için kromozomlara göre 2000up ve 200down gen elementlerinin SSR ve sıklık dağılımları.....	76
Çizelge 4.33. Oryantal pamukta + ve - sarmal için kromozomlara göre ekson ve intron gen elementlerinin SSR ve sıklık dağılımları.....	76
Çizelge 4.34. Oryantal pamukta + ve - sarmal için kromozomlara göre 5'-UTR ve 3'-UTR gen elementlerinin SSR ve sıklık dağılımları.....	77
Çizelge 4.35. Peru pamukta artı ve eksi sarmal için kromozomlara göre 2000up ve 200down gen elementlerinin SSR ve sıklık dağılımları.....	78
Çizelge 4.36. Peru pamukta + ve - sarmal için kromozomlara göre ekson ve intron gen elementlerinin SSR ve sıklık dağılımları.....	78

Çizelge 4.37. Peru pamukta artı ve eksi sarmal için kromozomlara göre 5'-UTR ve 3'-UTR gen elementlerinin SSR ve sıklık dağılımları.....	79
Çizelge 4.38. Upland pamukta artı ve eksi sarmal için kromozomlara göre 2000up ve 200down gen elementlerinin SSR ve sıklık dağılımları.....	80
Çizelge 4.39. Upland pamukta artı ve eksi sarmal için kromozomlara göre ekson ve intron gen elementlerinin SSR ve sıklık dağılımları.....	81
Çizelge 4.40. Upland pamukta +/- sarmal için kromozomlara göre 5'-UTR ve 3'-UTR gen elementlerinin SSR ve sıklık dağılımları.....	82
Çizelge 4.41. <i>Gossypium arboreum</i> 'a ait <i>CesA</i> ve <i>Csl</i> genlerini, motif içeriklerini ve sarmal bilgilerini içeren mikrosatellitlerin listesi.....	83
Çizelge 4.42. <i>Gossypium raimondii</i> 'ye ait <i>CesA</i> ve <i>Csl</i> genlerini, motif içeriklerini ve sarmal bilgilerini içeren mikrosatellitlerin listesi.....	86
Çizelge 4.43. <i>Gossypium hirsutum</i> 'a ait <i>CesA</i> ve <i>Csl</i> genlerini, motif içeriklerini ve sarmal bilgilerini içeren mikrosatellitlerin listesi.....	89

1. GİRİŞ

Malvaceae familyasının *Gossypium* cinsine dahil edilen pamuk (*Gossypium* L.) tropik ve subtropik bölgelerde yetiştirilen bir kültür bitkisidir. Pamuk işlenmesi açısından çırçır sanayisinin, lifi ile tekstil sanayisinin, tohumu ile yağ sanayisinin, küspesi ile yem sanayisinin, linteri ile de kâğıt sanayisinin hammaddesini oluşturmaktadır (Şekil 1.1). Birçok sanayinin hammaddesini oluşturduğundan üretici ülkelere istihdam olanakları sağlaması ile büyük öneme sahiptir.



Şekil 1.1. Pamuğun (*Gossypium* L.) bazı kullanım alanlarının şematik gösterimi (Keskinlik 2017)

Pamuk üretimi dünyada 45° Kuzey Yarım Küreden, 32° Güney Küre enlemlerine uzanmaktadır. Dünyada pamuk tarımının %90'ı içinde Türkiye'nin de bulunduğu az sayıda ülke tarafından gerçekleştirilmektedir. Pamuk tarımının en az yapıldığı ülke toplam 13 ha ile Gine iken; en fazla yapıldığı ülke Hindistan'dır. Türkiye coğrafi konum olarak 36°-42° Kuzey enlemleri arasında yer aldığı için pamuk tarımı yapan ülkeler arasında önemli bir yere sahiptir. Türkiye'de pamuk üretimi Ege Bölgesi, Güneydoğu Anadolu Bölgesi, Antalya ve Çukurova yörelerinde yapılmaktadır. Ayrıca Iğdır, Elazığ ve Malatya illerinde de pamuk üretimi yapılmaktadır.

Uluslararası Pamuk Danışma Komitesi ("International Cotton Advisory Committee") (ICAC)'nin verilerine göre; son 5 yıllık dönemin verileri incelendiğinde, dünya pamuk ekimi ortalaması yaklaşık 33 milyon hektar, üretim ortalaması ise 25 milyon tondur. Uzun yıllar boyunca dünya pamuk ekim alanlarında Çin başı çekmekteyken son yıllarda Hindistan ve ABD, Çin'i geride bırakmıştır. Pamuk üretiminde ise sıralama Çin, Hindistan ve ABD şeklinde olmuştur. Ancak bugün ABD, dünyadaki pamuk ihracatının neredeyse üçte birini yapmaktadır ve bu alanda lider konumdadır (Uğurlu 2020; USDA 2020). Ülkemiz ise yıllara göre değişmekle beraber 450 bin hektar ekim alanı ile 14'üncü sıradadır.

2023/2024 sezonuna göre dünyada en çok pamuk ekim alanına sahip ilk 5 ülke sırasıyla 12469000 ha ile Hindistan, 2872000 ha ile Çin, 2606000 ha ile ABD, 2370000 ha ile Pakistan, 1650000 ha ile Brezilya olmuştur. Bu dönemde dünyada en çok pamuk üreten ilk 5 ülke sırasıyla 5600000 ton ile Çin, 5493000 ton ile Hindistan, 3200000 ton ile Brezilya, 2627000 ton ile ABD ve 1420000 ton ile Pakistan takip etmiştir. Tüketimde ise ilk 5 sırayı; 8300000 ton ile Çin, 5389000 ton ile Hindistan, 2300000 ton ile Pakistan, 1600000 ton ile Bangladeş, 1500000 ton ile Türkiye yer almaktadır.

2023/2024 sezonunda birim alandan elde edilen verimde ilk 5 ülke; Avustralya, Çin, Brezilya, İsrail ve Meksika olmuştur. Bu dönemde en çok pamuk ithalatı yapan ilk 5 ülkenin; Çin, Bangladeş, Vietnam, Pakistan ve Türkiye olduğu, en çok ihracat yapan ilk 5 ülkenin ise sırasıyla; Brezilya, ABD, Avustralya, Hindistan ve Türkiye olduğu rapor edilmektedir (ICAC 2023).

Uluslararası Pamuk Danışma Komitesi'nin verilerine göre; içinde bulunduğumuz sezonda Türkiye'nin pamuk ekim alanı yönünden 450000 ha ile dünyada on dördüncü, birim alandan elde edilen lif pamuk verimi yönünden 1556 kg/ha ile altıncı, lif pamuk üretim miktarı yönünden 700000 ton ile yedinci, lif pamuk tüketimi yönünden 1500 ton ile beşinci, pamuk ithalatı yönünden 640000 ton ile beşinci, pamuk ihracatı yönünden ise 400000 ton ile beşinci ülke olduğu rapor edilmektedir.

Çizelge 1.1. verileri dikkate alındığında ülkemizde 454000 hektar alanda pamuk ekimi yapılmıştır. Ekim alanlarının genişliği bakımından ise ilk sırayı Güneydoğu Anadolu Bölgesi alırken; 2. sırada Ege Bölgesi yer almaktadır. Türkiye'de iller düzeyinde pamuk ekim alanı incelendiğinde ilk 5 sırayı; 2034000 da ile Şanlıurfa ili birinci, Diyarbakır ili 705000 da ile ikinci, Aydın 565000 da ile üçüncü, Hatay 389000 da ile dördüncü, İzmir 249000 da ile beşinci sıradadır. Ülkemiz pamuk ekim alanlarının %87'si ilk beş ilimizde bulunmaktadır (Çizelge 1.1).

Çizelge 1.1. TÜİK 2019-2023 illere göre pamuk ekim alanı (bin da)

İller	2019	2020	2021	2022	2023
Şanlıurfa	2088	1287	1835	2425	2034
Diyarbakır	477	403	552	829	705
Aydın	465	545	493	574	565
Hatay	456	332	390	460	389
İzmir	253	277	262	333	249
Adana	375	230	219	303	183
Manisa	90	92	124	174	145
Denizli	72	95	97	124	123
Adıyaman	72	55	61	82	65
Mardin	130	74	60	102	82
Toplam	4478	3390	4093	5406	4540

Çizelge 1.2. verileri incelendiğinde 2023 yılında 2005000 ton kütlü pamuk üretilmiştir. Türkiye’de kütlü pamuk üretiminin büyük çoğunluğu Güneydoğu Anadolu Bölgesi’nde gerçekleştirilmektedir. Bu üretimde ilk beş sırayı; 881000 ton ile Şanlıurfa, 303000 ton ile Diyarbakır, 243000 ton ile Aydın, 185000 ton ile Hatay, 115000 ton ile İzmir yer almaktadır. Ülkemiz kütlü pamuk üretim miktarının %86’sı ilk beş ilimizde bulunmaktadır (Çizelge 1.2).

Çizelge 1.2. TÜİK 2019-2023 illere göre kütlü pamuk üretim miktarı (bin ton)

İller	2019	2020	2021	2022	2023
Şanlıurfa	813	567	893	1102	881
Diyarbakır	234	218	309	408	303
Aydın	246	272	265	265	243
Hatay	220	175	210	239	185
İzmir	142	149	151	173	115
Adana	206	123	114	144	85
Manisa	55	56	73	92	72
Denizli	38	50	53	62	58
Adıyaman	35	27	31	31	27
Mardin	65	40	33	54	36
Toplam	2054	1677	2132	2570	2005

Türkiye’de 2023 yılında 742000 ton lif pamuk üretilmiştir. Bu üretimde 326000 ton ile Şanlıurfa ili birinci sırada yer alırken; 112000 ton ile Diyarbakır ikinci, 90000 ton ile Aydın üçüncü, 68000 ton ile Hatay dördüncü, 43000 ton ile İzmir beşinci, 32000 ton ile Adana altıncı, 27000 ton ile Manisa yedinci, 21000 ton ile Denizli sekizinci, 10000 ton ile Adıyaman dokuzuncu, 13000 ton ile Mardin onuncu sırada yer almaktadır (Çizelge 1.3).

Çizelge 1.3. TÜİK 2019-2023 illere göre lif pamuk üretim miktarı (bin ton)

İller	2019	2020	2021	2022	2023
Şanlıurfa	301	210	331	408	326
Diyarbakır	86	81	114	151	112
Aydın	91	101	98	105	90
Hatay	81	65	78	88	68
İzmir	52	55	56	64	43
Adana	76	46	42	53	32
Manisa	20	21	27	34	27
Denizli	14	18	19	23	21
Adıyaman	13	10	11	14	10
Mardin	24	15	12	20	13
Toplam	758	622	788	960	742

Türkiye’de 2023 yılında 1204000 ton pamuk çiğdi üretilmiştir. Bu üretimde 529000 ton ile Şanlıurfa ili birinci sırada yer alırken; 182000 ton ile Diyarbakır ikinci, 146000 ton ile Aydın üçüncü, 111000 ton ile Hatay dördüncü, 69000 ton ile İzmir beşinci, 51000 ton ile Adana altıncı, 43000 ton ile Manisa yedinci, 35000 ton ile Denizli sekizinci, 16000 ton ile Adıyaman dokuzuncu, 22000 ton ile Mardin onuncu sırada yer almaktadır (Çizelge 1.4).

Çizelge 1.4. TÜİK 2019-2023 illere göre pamuk çiğdi üretim miktarı (bin ton)

İller	2019	2020	2021	2022	2023
Şanlıurfa	488	340	536	662	529
Diyarbakır	140	131	186	245	182
Aydın	148	163	159	170	146
Hatay	132	105	126	143	111
İzmir	85	89	90	104	69
Adana	123	74	68	87	51
Manisa	33	33	44	55	43
Denizli	23	30	32	37	35
Adıyaman	21	16	19	23	16
Mardin	39	24	20	33	22
Toplam	1232	1005	1280	1559	1204

Pamuk (*Gossypium* L.) yaklaşık 52 türe sahiptir ve bu türlerden dördü (iki diploit ve iki allotetraploit) 80’den fazla ülkede üretimi yapılmaktadır (Karaca vd. 2019). Bunlardan ikisi Afrika-Asya türleri olan; diploit genoma sahip *Gossypium arboreum* (Ağaç pamuk / Oryantal pamuk, Hindistan) ve *Gossypium herbaceum* (Levant pamuğu, Afrika) iken; diğer ikisi allotetraploit genoma sahip yeni dünya pamukları *Gossypium hirsutum* (Upland pamuğu / Yayla pamuğu, Güney Amerika) ve *Gossypium barbadense* (Pima pamuğu / Mısır pamuğu, Sea Island pamuğu, Güney Amerika) pamuklarından oluşmaktadır (Karaca 2001; Mert 2007).

Organizmalar genomlarında şifrelenmiş olan bilgiler çerçevesinde yaşamlarını sürdürmektedirler. Özellikle son 10-15 yıllık dönemde Yeni Nesil Dizileme yardımıyla çok sayıda genom dizilemesi tamamlanmıştır (Smith 2022). Günümüzde mikrosatellitler (SSR) yaygın olarak genom karakterizasyonu ve gen haritalama çalışmalarında moleküler markır olarak kullanılmaktadır. Ancak pamuk bitkisinde bütün genomda karşılaştırmalı olarak mikrosatellitlerin dağılımı ile ilgili çalışma sayısı sınırlıdır. Bu nedenle elde edilecek her türlü bilgi pamuk ıslahı ve üretimine katkı sağlama potansiyeli bulunmaktadır. Tarımsal araştırmalara önemli katkılar sağlayan moleküler markırlar, başta genetik çeşitlilik olmak üzere genetik haritalama gibi konularda katkı sağlamıştır. Mikrosatellit veya SSR (Simple Sequence Repeats) bütün bir genom boyunca dağılmış genellikle 1 ile 6 nükleotit uzunluğundaki bazların ardışık olarak tekrarlandığı bir markır sistemidir (Tautz ve Renz 1984; Karaca 2001).

Bu tez çalışmasında bütün genom dizilerine sahip ve aynı zamanda referans genom belirlenmiş olan pamuk türlerine ait genomik bilgiler kullanılmıştır. Bu genom dizileri tetraploit pamuğun alt genomu olan A ve D genomunu taşıyan iki farklı türün

genom bilgileri ile AD genomunu bir arada bulunduran tetraploit genom bilgileridir. A genomu Oryantal pamuk olarak bilinen *Gossypium arboreum*, D genomu Peru pamuğu olarak bilinen *Gossypium raimondii* ve AD genomu ise Upland pamuğu olarak bilinen *Gossypium hirsutum* tarafından temsil edilmiştir. Böylelikle projenin ana amacı olan bütün genom dizilerinde mikrosatellit DNA'nın tipleri ve dağılımının ortaya konulması ve aynı çalışma verileri kullanılarak pamuk lifinin ana kaynağı ve doğada da en fazla bulunan selülozun kaynağı olan selüloz sentaz genlerinin mikrosatellit (SSR) içerikleri belirlenmesi hedeflenmiştir.

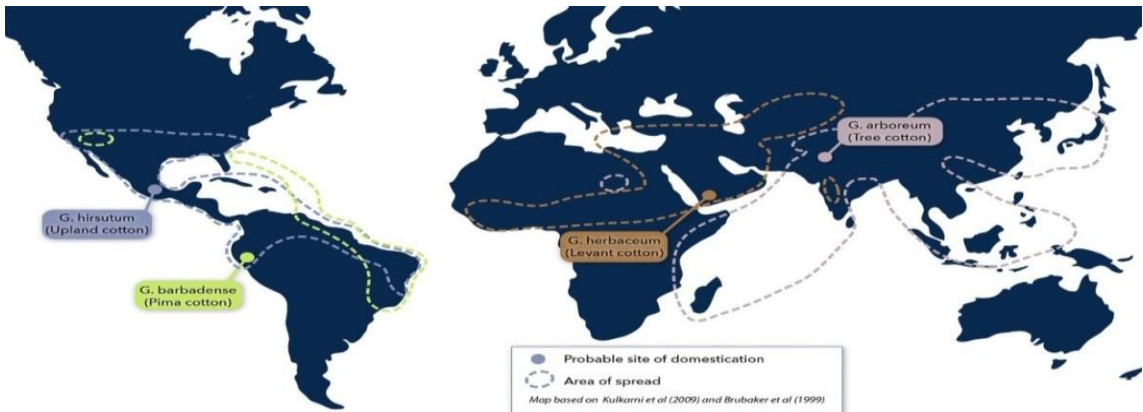
Projenin bağlı olduğu ana hipotez genom büyüklükleri ile mikrosatellit içerikleri, tipleri ve motifleri arasında bir ilişkinin varlığı, mikrosatellitlerin varlıklarının diploit ve tetraploit türlerde, türlerin kromozomlarında ve gen elementlerinde (promotör, 5'-UTR, 3'-UTR, ekson, intron ve intergenik diziler) dağılımları arasında farklılık durumlarının oluşabildiğidir. Bütün bu soruların yanıtları ise birbirleri ile ilişkili türlerin bütün genomların aynı yöntemlerle çalışılmalarıyla ortaya konulabileceği hipotezi de eklenmiştir. Ayrıca selüloz sentaz geni özelinde evrim sürecinde mikrosatellitlerin değişimi ve sonraki nesillerdeki durumlarının değişebileceği de başka bir hipotez olarak düşünülmüştür.

2. KAYNAK TARAMASI

2.1. Pamuk Bitkisinin Genomu

Dünyadaki doğal tekstil liflerinin birincil ana kaynağı olan pamuk (*Gossypium* L.), aynı zamanda insanlar için önemli bir yağ, hayvancılık için ise protein kaynağıdır (Ince ve Karaca 2023). Yaklaşık 12,5 milyon yıl önce *Gossypium*'un Miyosen döneminde (günümüzden 5.3-23.03 milyon yıl öncesi dönem) *Kokia* ve *Gossypoides*'ten ayrıldığı tahmin edilmektedir (Cronn vd. 2002). Malvaceae familyası *Gossypium*, *Houzingenia*, *Karpas* ve *Sturtia* olmak üzere dört alt cins içermektedir. Pamuk *Gossypium* cinsine ait olup bu cinsten yaklaşık 45'i diploit ($2n = 2x = 26$) ve yedisi tetraploit ($2n = 4x = 52$) olmak üzere 52 tür bulunmaktadır (Karaca 2001; Wendel ve Cronn 2003; Wendel ve Grover 2015). Diploit türler A, B, C, D, E, F, G ve K olmak üzere sekiz monofiletik genom grubundan oluşmaktadır (Wendel ve Cronn 2003; Grover vd. 2007; Wendel vd. 2010). Amerika kıtalarında 13 D genom türü, Asya ve Afrika kıtalarında 15 tür (A-, B-, E- ve F- genomları) ve Avustralyada 18 tür (C-, G- ve K- genomları) bulunmaktadır (Wendel ve Cronn 2003). Dolayısıyla pamuk türleri, farklı kıtalar arasındaki evrimi, kültüre alınma, genetik introgresyonu ve gen akışını incelemek için büyüleyici bir model sistemi sağlar (Fryxell 1969, 1978; Wendel 1989; Wendel ve Grover 2015; Chen vd. 2016, 2017a, 2017b). Tetraploid pamuklar tropikal ve subtropikal Amerika kıtalarında bulunmaktadır. Bunlar yedi türe ayrılmış olup *G. hirsutum* L. (AD₁), *G. barbadense* L. (AD₂), *G. tomentosum* Nuttalex Seemann (AD₃), *G. mustelinum* Miersex Watt (AD₄), *G. darwinii* Watt (AD₅), *G. ekmanianum* Wittmack, M.C.L. (AD₆) ve *G. stephensii* (AD₇) olarak adlandırılmıştır (Wendel ve Cronn 2003; Wendel ve Grover 2015).

Gossypium cinsinin eğrilebilir lif (elyaf) üreten dört türünün kültürü yapılmaktadır. Şekil 2.1'de pamuğun kültürü yapılmakta olan 4 pamuk türüne ait orijinleri ve kültüre alındığı kıtalar gösterilmiştir. Bunlar iki allotetraploit ($2n = 4x = 52$) tür olan *G. hirsutum* L. ve *G. barbadense* L. ile iki diploit ($2n = 2x = 26$) tür olan *Gossypium herbaceum* L. ve *Gossypium arboreum* L.'dir (Wendel ve Cronn 2003; Wendel ve Grover 2015).



Şekil 2.1. Pamuğun evcilleştiği bölgeler [<https://learn.genetics.utah.edu/cotton> Son erişim tarihi: 16.10.2023]

Günümüze değin yapılmış çalışmaların kültürü yapılan iki tetraploit pamuğun *G. hirsutum* (AD₁) ve *G. barbadense* (AD₂)'nin yaklaşık 1,5 milyon yıl önce ortaya çıkmış olduğu düşünülmüş olup bunların en az 4000 ile 5000 yıl önce insanlar tarafından kültüre alındığı kabul görmektedir (Wendel 1989; Fang vd. 2017; Wang vd. 2017). Bu türlerin genomlarda A ve D alt genomların varlığı ve D genom öncüsü türün Amerika kıtasında A genom öncüsü türün ise Asya kıtasında bulunması nedeniyle allotetraploit pamukların kıtaların oluşumundan önceki bir dönemde oluşmuş olduğunu göstermektedir (Karaca 2001). Diğer taraftan *G. hirsutum* (AD₁) ve *G. barbadense* (AD₂) türlerinin ise farklı zamanlarda oluştuğu düşünülmektedir. Allotetraploit pamuğu iki bileşenli genomunun birincisi Meksika'da (*G. hirsutum*) ikincisinin ise Peru'da (*G. barbadense*) ortaya çıkmış olup evrimleşmenin gerçekleştiğini göstermektedir (Wendel ve Grover 2015; Fang vd. 2017; Du vd. 2018; Huang vd. 2020).

Günümüzde iki tetraploit pamuk türünden başka iki diploit türünde tarımı yapılmaktadır. Bu diploit türlerin (*G. arboreum* ve *G. herbaceum*) birkaç bin yıldır tarımı yapılmaktadır (Simon vd. 2016). Bu iki türün başlangıçta Madagaskar'da veya İndus Vadisi'nde (Mohenjo Daro) kültüre alındığı ve daha sonra Afrika'ya ve Asya'nın diğer bölgelerine dağılmış olduğu tahmin edilmektedir (Wendel ve Grover 2015; Du vd. 2018; Huang vd. 2020). Günümüzde yetiştirilen pamuk türleri arasında *G. hirsutum* en yaygın olarak kültüre alınan pamuk türü olup Upland pamuk olarak bilinmekte ve dünya çapındaki üretimin %90'ından fazlasını karşılamaktadır (Cronn vd. 2003; Karaca vd. 2020).

2.2. Mikrosatellitler (SSR) Tanımı, Oluşumu ve Kullanım Alanları

SSR'lar diğer bir isimle mikrosatellitler genomlarda yaygın olarak bulunan 1 ile 6 nükleotit bazlı kısa ardışık (tandem) tekrar birimleri olarak tanımlanmaktadır (Tautz 1989; Jurka ve Pethiyagoda 1995; Chen vd. 2009; Zhang vd. 2020). İlk olarak Litt ve Luty (1989) tarafından adlandırılan bu basit tekrarlı diziler motif uzunlukları mono-, di-, tri-, tetra-, penta- ve hekza-nükleotit dizilerinden oluşmaktadır (Karaca 2001; Karaca vd. 2005; Ince vd. 2010; Ashley 2010) (Şekil 2.2).



(a)

```

---GGTACCAAAAA(A)nCGTCCA---
---TGCGATCGAGAGA(GA)nATTCGT---
---TTAGCTAGACGAC(GAC)nCCAGTGA---
---AAGTTATGGCC(GGCC)nTGACAGT---
---GCTACTAGGGTTC(GGTTC)nTTAGCAT---
---GCATTCCGGTAAC(GGTAAC)nACTGTCA--

```

(b)

Şekil 2.2. Tekrarlı motif dizileri. **(a)** genomda mikrosatellitlerin dağılımı, **(b)** mono-, di-, tri-, tetra-, penta- ve hekza-nükleotit motifler

Mikrosatellit (SSR) dizilerin uzunluğu genellikle 200 bp'nin altında olup tekrar sayılarına 10-100 olup, bu diziler genomun farklı yerlerinde rastgele dağıldıkları yönünde bilgiler yaygındır (Tautz 1989; Akkaya vd. 1992; Bagshaw 2017). Bu diziler moleküler markırlar olarak, bitki türlerinin genetik çeşitliliğini incelemek için kullanılan en etkili genetik araçlardan biridir (Joshi vd. 1999; Govindaraj vd. 2015; Kundu ve Ghosh 2015; Nadeem vd. 2018; Bhandari vd. 2021).

SSR markırları, sırasıyla genomik DNA (gDNA) ve cDNA dizileri veya eksprese (ifade) edilmiş dizi etiketleri (EST'ler) ile geliştirilen genomik (gSSR) ve genik (EST-SSR) SSR'lar olarak sınıflandırılabilir (Vasek vd. 2020). SSR'lar hem genik hem de genik olmayan bölgelerde bulunur. Eksprese edilmiş dizi etiketlerinde (EST'ler) ve daha sıklıkla genik olmayan DNA dizilerinde tespit edilebilirler (Mason 2015). Bu diziler (mikrosatellitler) kodominant özellikte olması, genomda bol miktarda bulunması, oldukça polimorfik, multi-allelilik ve multi-lokus özellikleri yanında, farklı araştırmacılar veya denemelerde aynı sonuçları vermesi bakımından tekrarlanabilir olmaları onları güvenilir bir moleküler markır tekniği olarak öne çıkarmaktadır (Ince vd. 2017). Bu markırlar başta genetik çeşitlilik, genetik haritalama, hibrit tanımlama, popülasyon genetiği, markır destekli seleksiyon da dahil olmak üzere birçok uygulamada yaygın olarak kullanılmaktadır (Ince vd. 2008; Azpilicueta vd. 2013; Dufresnes vd. 2014; Karaca vd. 2015).

Mikrosatellitler içerdikleri tekrar dizi türüne göre sınıflandırılabilir (Wang vd. 2009). Ayrıca SSR'ları basit mükemmel, basit kusurlu, bileşik mükemmel veya bileşik kusurlu olarak sınıflandırmıştır. SSR'lar, genomdaki konumlarına göre nükleer (nuSSR), mitokondriyal (mtSSR) (Soranzo vd. 1999) veya kloroplast SSR'lar (cpSSR) (Weising ve Gardner 1999) olarak sınıflandırılabilir. Genomik SSR'ların çoğu nükleer SSR'lardır. Mikrosatellitlerin oluşumuna ve polimorfizmine iki mekanizmanın sebep olduğu bilinmektedir. DNA polimerazın replikasyon sırasında dizi üzerinde kayması ve eşit olmayan rekombinasyonlar sonucu polimorfizm oluşmaktadır (Wang vd. 2017; Zhang vd. 2020; Shrestha vd. 2023).

2.3. Mikrosatellitlerin Belirlenmesi

Mikrosatellitler genetik bir markır olarak son çeyrek yüzyıldır DNA tabanlı genetik analizlerde yaygın olarak kullanılmaktadır. Yeni mikrosatellitler geleneksel Sanger dizileme veya yüksek verimli yeniden dizileme verileri kullanılarak *in vitro* veya *in silico* analizleriyle tespit edilebilmektedir. *In silico* yöntemler olarak çok sayıda algoritma ve yazılım geliştirilmiştir. Bunlara örnek olarak "Tandem Repeats Finder" [TRF] (Benson 1999), "REPuter" (Kurtz vd. 2001), "Simple Sequence Repeat Identification Tool" [SSRIT] (Kantety vd. 2002), "Simple Sequence Repeat Finder" [SSRF] (Sreenu vd. 2003), "Search for TRs IN Genomes" [STRING] (Parisi vd. 2003), "Microsatellite Search" [MISA] (Thiel vd. 2003), "Tandem Repeats Analyzer 1.5" [TRA] (Bilgen vd. 2004), "Search for Tandem Approximate Repeats" [STAR] (Delgrange ve Rivals 2004), "Exact Tandem Repeats Analyzer 1.0" [E-TRA] (Karaca vd. 2005), "SSR Classification and Investigation by Robert Kofler" [SciRoKo] (Kofler vd. 2007), "Imperfect Microsatellite Extractor" [IMEx] (Mudunuri ve Nagarajaram 2007), "Genome Microsatellite Analyzing Tool Files" [GMATo] (Wang vd. 2013), "Genome-wide Microsatellite Analyzing Tool Package" [GMATA] (Wang ve Wang 2016) bulunmaktadır. Yeni nesil dizileme yöntemlerinin yaygınlaşması ile

mikrosatellitler bütün bir genomda tespit edilebilmekte ve kromozomsal lokasyonları ile ilişkili genler ve gen elementleri arasındaki fiziksel ve lokasyon etki ilişkileri ve dağılımları ortaya konulabilmektedir.

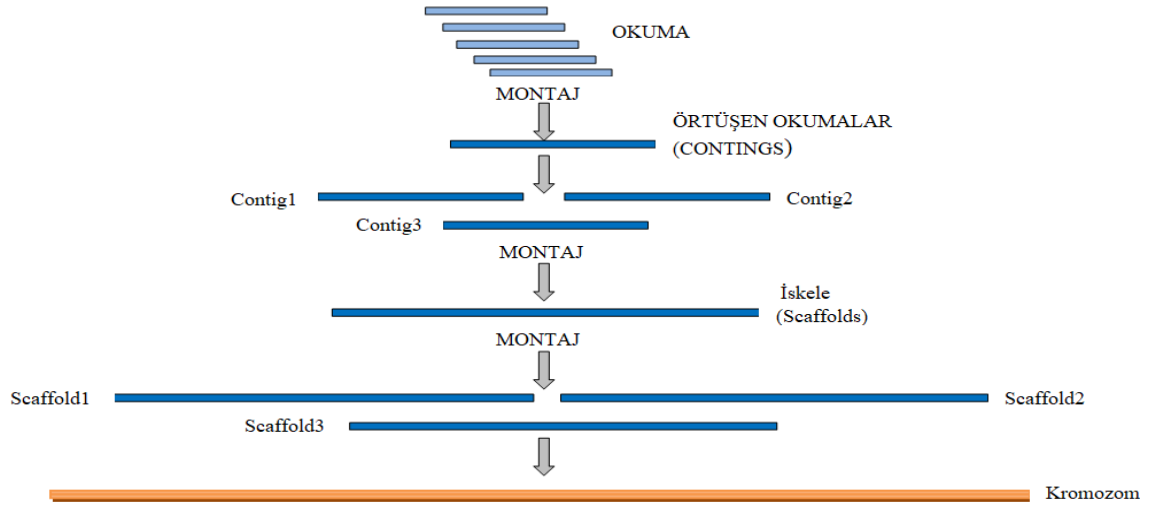
2.4. Genom Dizilemeleri

Genom dizileme teknolojisinin gelişimi ile aralarında *Gossypium*'unda bulunduğu çok sayıda bitkide bütün genomun dizi bilgilerinin belirlenebilmesine imkân sunmuştur. Peru pamuğu (*G. raimondii*) yeni nesil dizileme yöntemleriyle dizilenen ilk pamuk genomlarından biri olmuştur (Paterson vd. 2012; Wang vd. 2012). Bunu D genomunun bazı türleri için genom dizilemeleri ve genom betimlemeleri (Udall vd. 2019) ve 13 D-genom türünün tümü için yeniden dizileme çalışmaları izlemiştir (Grover vd. 2019). Benzer şekilde, A-genom diploitleri *G. arboreum* (A₂) (Li vd. 2014; Du vd. 2018; Huang vd. 2020) ve *G. herbaceum* (A₁) için genom derlemeleri ve yeniden dizileme veri setleri yayınlanmıştır (Huang vd. 2020). *G. hirsutum* (AD₁) genomunun dokuz genomu (Li vd. 2015; Zhang vd. 2015; Hu vd. 2019; Yang vd. 2019; Wang vd. 2019; Chen vd. 2020; Huang vd. 2020) ve *G. barbadense* (AD₂) türü içinde dördü (Hu vd. 2019; Wang vd. 2019; Chen vd. 2020) de dahil olmak üzere allopoliploidler için genomik kaynaklar da mevcuttur (Chen vd. 2020).

Yeni nesil dizileme teknolojilerinin kullanım alanlarında yeniden dizileme (“resequencing”), dizileme ile genotipleme (“genotyping by sequencing, GBS”) kullanılarak çok yönlü analizler ortaya konulabilmektedir. Örneğin kültür bitkilerinde genom çapında ilişkilendirme çalışmaları (GWAS) ile tarımsal özelliklere ait haritalama çalışmaları GBS desteği ile ortaya konulabilmektedir. Metilom olarak adlandırılan bütün genomda metilasyon farklılığının ortaya konulması veya bütün bir hücrede ifade edilen genlerin ortaya konulması RNA-Seq veya protein kodlayan genlerin belirlenmesi transkriptomik bu teknolojilerin gelişimiyle daha düşük maliyet ve sürede tamamlanabilmekte böylelikle önemli tarımsal özelliklerle ilişkili genler dahil olmak üzere önemli genomik varyasyonların tanımlanmasını hızlandırmıştır (Huang vd. 2010, 2012; Jia vd. 2013; Li vd. 2013; Mace vd. 2013; Huang ve Han 2014; Yano vd. 2016). Yeni nesil dizileme teknolojilerinin pamukta bilimsel çalışmalarda kullanımı da giderek yaygınlaşmaktadır (Tyagi vd. 2014; Zhao vd. 2014; Goodwin vd. 2016; İslam vd. 2016; Su vd. 2016; Abdullaev vd. 2017; Cai vd. 2017; Fang vd. 2017a, 2021; Huang vd. 2017; Sun vd. 2017; Wang vd. 2017a; Li vd. 2018; Liu vd. 2018; Ma vd. 2018a, 2018b; Dong vd. 2019; Yuan vd. 2021; Karaca ve Ince 2024).

Dizilemenin azalan maliyetine rağmen, okumaları yüksek kaliteli genomlara yeniden birleştirmek hala zor ve zaman alıcıdır (Salzberg vd. 2012; Bradnam vd. 2013). Tüm genomların mevcudiyeti, birçok biyolojik uygulamada sonraki analizlerde (downstream dizi analizi) ve yorumlanması için aslında önemlidir. Bu yöntemlerde biyoinformatik kapasiteli bilgisayarlar, küçük DNA dizi parçalarını daha büyük DNA parçalarına birleştirerek dönüştürmek ve bunları daha sonra daha büyük bitişik parçalar halinde birleştirerek dönüştürmek için kullanılır. Daha sonra “contig”ler (örtüşen okumalar), iskeleler (“scaffold”) halinde ve son olarak da kromozomlar halinde birleştirilir. Dolayısıyla bir contig, sürekli bir nükleotit dizisidir; bir iskele ise contiglerden oluşan bir genomun parçasıdır. Hem contig hem de iskele, yeniden yapılandırılmış genomik dizilerdir (Şekil 2.3).

Diploit veya poliploit organizmalar genellikle yüksek derecede heterozigotluk içerir (You vd. 2013; Paape vd. 2016) ve bu durum, heterozigot bölgelerin sıklıkla birden fazla bitişik parçaya bölündüğü montaj sürecinde sorunlara neden olmaktadır (Pryszcz ve Gabaldon 2016). Bu nedenle; genom düzenekleri, yanlış birleştirilmiş bölgeler ve hatalar içeren eksik ve parçalanmış bitişik yapılara / iskelelere neden olabilir (Ekblom ve Wolf 2014). Son çalışmalar, tekrarlayan bölgeleri çözmek ve daha uzun yapı iskeleleri oluşturmak için daha uzun okumalar (örneğin, Pacific BioSciences tarafından tek molekülli gerçek zamanlı sıralama ve Bionano tarafından tek molekülli optik haritalama kullanılarak) kullanmaya başlamıştır (Bradnam vd. 2013; Koren ve Phillippy 2015; Gordon vd. 2016; Yang vd. 2016; Hatakeyama vd. 2018). Bununla birlikte, kompleks DNA kısımlarının ekstraksiyonunun daha zor olması, yüksek miktarda hata ve daha yüksek maliyetler ek sorunları beraberinde getirmekte ve kullanımlarını hâlâ sınırlamaktadır (Bradnam vd. 2013; Chakraborty vd. 2016; Goodwin vd. 2016). Giderek daha fazla tür dizilindikçe, okumaların önemli bir kısmının haritalanabileceği, farklı ancak ilişkili bir türün genomunun zaten mevcut olması ihtimali ortaya çıkıyor. Yakın akraba türler olarak adlandırdığımız böyle bir türün genomu, daha sonra hedef türün bir araya gelmesine yardımcı olmak için kullanılabilir. Referans kılavuzlu yaklaşımlar olarak adlandırılan bu yaklaşımlar, ek bilgi elde etmek için hedef ve referans türler arasındaki benzerlikten yararlanır ve bu da genellikle daha eksiksiz ve gelişmiş bir genoma yol açar (Pop vd. 2004; Schneeberger vd. 2011; Bao vd. 2014). Düşük kapsamda dizilenen genomlar bile, eğer bir referans genom tarafından yönlendirilirse yararlı genomik kaynaklar sağlayabilirler (Vezzi vd. 2011; Card vd. 2014).



Şekil 2.3. Kısa okumalardan elde edilen dizilerin montajı

2.5. Selüloz Sentaz Gen Ailesi

Selüloz, β -1,4 bağlı glukanlardan oluşan ve bitkiler, bakteriler, funguslar ve birtakım protistalarda oldukça bol miktarda bulunan doğal bir polisakkarit polimerdir (Matthysse ve Thomas 1995; Grimson vd. 1996; Saxena ve Brown 2005; Keegstra

2010). Selüloz, bitkilere mekanik destek sağlayan ve onları biyotik ve abiyotik streslerden koruyan bitki hücre duvarlarının önemli bir bileşenidir (Cantu vd. 2008; Wang vd. 2016). Selüloz ve hemiselüloz, selüloz sentezinde yer alan selüloz sentaz (*CesA*) ve selüloz sentaz benzeri (*Csl*) selüloz sentaz gen süper ailesi tarafından sentezlenir. Önemli bir doğal tekstil lifi ürünü olan pamuk liflerinin %96'lık kısmı selülozdur ve bu da onları diğer lif bitkilerinden farklı kılar (Lee vd. 2007; Ning vd. 2014; Shao vd. 2014; Delhom vd. 2022).

Ham pamuk lifi β -1,4-glikozidik bağlarla bağlanmış glikoz polimeri olan selüloz, hemiselüloz (üronik asit, ksiloz, arabinoz, mannoz, glikoz, galaktoz vb), pektin ve az miktarda da ligninden oluşmaktadır. Primer ve sekonder hücre duvarının ana bileşeni selüloz olup lifin %96 kadarını oluşturabilmektedir. Pamuk lifinin primer ve sekonder hücre duvarı plazma membranında rozet benzeri yapılar oluşturan multimerik protein kompleksleri olan selüloz sentaz A (*CesA*) ve selüloz sentaz benzeri (*Csl*) proteinler tarafından sentezlenir (Zou vd. 2018; Zhang vd. 2021; Wen vd. 2022; Pedersen vd. 2023). Pamuk türlerindeki *CesA1*, *CesA3* ve *CesA6* birincil hücre duvarındaki selüloz senteziyle; *CesA4*, *CesA7* ve *CesA8* genleri ikincil hücre duvarındaki selüloz senteziyle ilişkilidir (Liu vd. 2015; Cao vd. 2019). *CesA* genleri, pamuk lifi gelişimi dışında bitki hücre büyümesi ve gelişimi sırasında selüloz biyosentezinde önemli roller oynar (Lin vd. 2012). Selüloz sentaz (*CesA*) geni ilk olarak pamuk (*G. hirsutum*) liflerinden bakteriyel bir *CesA* geniyle dizi benzerliğine göre tanımlanmıştır (Pear vd. 1996). Daha sonra *Arabidopsis* (Arioli vd. 1998), mısır (Appenzeller 2004), arpa (Burton vd. 2004) ve çeltikte (Yin vd. 2009) selüloz sentaz genleri tanımlanmıştır (Gu vd. 2010; McFarlane vd. 2014).

2.6. Pamuk Bitkisinde Mikrosatellit Çalışmaları

Erkılınç ve Karaca (2005) yaptıkları çalışmada 36 pamuk çeşidinin genetik karakterizasyonu için 25 primer çifti ile SSRLP (mikrosatellite, SSR) DNA parmak izi tekniği kullanmıştır. Kullanılan 25 SSRLP primer çiftinden sadece 4 primer çiftinin polimorfik bantlarla sonuçlandığını rapor etmişlerdir. Çalışmada kullanılan çeşitlerde alel sayısı 2 ile 4 arasında değiştiği görülerek Türkiye'de tarımı yapılan 36 çeşidin birbirlerine benzer özelliklere sahip olduğu sonucuna ulaşılmıştır. Ülkemizde yetiştirilen pamuklarda genetik tabanın son derece dar olduğu ve germplazm havuzlarındaki genetik çeşitliliği arttırmak için aralarında türler arası melezlerinde bulunduğu pamuk hatlarının kullanılması gerektiği sonucuna varılmıştır.

Bertini vd. (2006) yaptıkları çalışma ile *Gossypium hirsutum* türüne ait 53 pamuk genotipinin arasındaki genetik mesafeyi tahmin etmek ve ayrıca pamuk türleri arasında ayırım yapabilmek için 31 SSR primeri kullanmışlar ve 66 adet polimorfik bant elde etmişlerdir. Benzerlik katsayısını 0,00 ile 0,41 olarak bulduklarını belirtmişlerdir.

Guang ve Xiong-Ming (2006) yaptıkları çalışma ile Çin'deki farklı ebeveyn kökenleri, üreme dönemleri ve ekolojik yetiştirme alanlarına sahip 43 Upland pamuğun genetik kaynağının genetik çeşitliliği, SSR markırları temelinde incelenmiştir. Toplam 130 gen aleli 36 SSR primerinden %80 polimorfizm tespit etmişlerdir. Primer başına alel sayısı iki ile sekiz arasında değiştiğini, ortalama 3,6 polimorfizm bilgi içeriği (PIC) aralığı 0,278-0,865 olup ortalama 0,62 olduğunu belirlemişlerdir.

Bardak ve Bölek (2012) yaptıkları çalışmada dünyanın farklı yerlerinde yetiştirilen diploit ve tetraploit pamuk genotiplerinin genetik çeşitliliğini belirlemişlerdir. *Gossypium* türleri ve genotipleri arasında genetik ilişkileri belirlemek 39 SSR ve 5 ISSR markırları kullanarak 25 pamuk genotipinden 173 alel ürettiklerini belirtmişlerdir. İncelenen tüm genotipler arasında genetik farklılığın sadece 0,04 ile 0,58 arasında değiştiğini ve germplazm havuzlarındaki genetik çeşitliliği arttırmak için yabani tip pamukların kullanılması gerektiğini rapor etmişlerdir.

Rehman ve Iqbal (2022) yaptıkları çalışmada farklı pamuk genotiplerinin genetik çeşitliliğinin ve popülasyon yapısının DNA düzeyinde değerlendirilmesi amacıyla toplam 105 pamuk genotipi hem A hem de D genomlarına dağılmış 208 polimorfik SSR markırı kullanarak analiz etmişlerdir. Çalışma sonucunda pamuk germplazmasını dört alt popülasyona ayırıp, en yüksek moleküler varyasyonun alt popülasyonlar arasında ortaya çıktığını rapor etmişlerdir.

Shi vd. (2023) yaptıkları çalışmada Upland pamuğu germplazmalarının genetik çeşitliliğini ve popülasyon yapısını ortaya çıkarmak için 26 SSR markırı kullanarak 191 Upland pamuğu germplazm kaynağında moleküler markır araştırması olarak yürütmüşlerdir. 65 SSR markırından ortalama 3,54 olmak üzere toplam 92 tanımlanabilir bant çoğaltmışlardır. Gözlenen alellerin ortalama sayısı (N_a) 1,984; etkili alel sayısı (N_e) 1,534; gen çeşitlilik indeksi (h) 0,312; Shannon-Weaver indeksi (I) 0,470 ve primerlerin polimorfizm bilgi içeriği (PIC) 0,530 olduğunu bulmuşlardır. Çalışma sonucunda pamuk bölgelerindeki germplazması kaynaklarının farklı olduğunu kaynak kullanım kapsamının nispeten dar olduğunu ve çeşitlerin genetik temelinin uzun süre ıslah birimlerinde nispeten düşük olduğunu rapor etmişlerdir.

3. MATERYAL VE METOT

3.1. Genom DNA Dizileri

Çalışmada kullanılan materyal GenBank (www.ncbi.nlm.nih.gov) veri tabanlarından indirilen 3 farklı pamuk türüne ait genom dizileri FASTA, GFF ve GTF formatlarında indirilerek kullanılmıştır. Referans genomlarına ait ilgili bilgiler Çizelge 3.1, 3.2 ve 3.3'te sunulmuştur. Bu tez çalışmasında 3 türün seçilme nedeni Upland pamuğunun Oryantal ve Peru pamuğunun ataları arasındaki doğal melezleme ve genom katlaması sonucunda oluşmuş olması nedeniyle bu türler kullanılmıştır. “Upland” veya “Yayla pamuğu” olarak bilinen *Gossypium hirsutum*'un referans genomu v2.1 GCF_007990345.1, “Oryantal” veya “Ağaç pamuk” olarak bilinen *Gossypium arboreum*'un referans genomu GCF_025698485.1_ASM2569848v2/ ve “Peru pamuğu” olarak bilinen *Gossypium raimondii*'nin referans genomu GCF_025698545.1_ASM2569854v1 kullanılmıştır. Çizelge 3.1'de PacBio RSII dizileme teknolojisi ile elde edilen ve Mecat v. 1.0 montajlama yöntemiyle elde edilen *Gossypium hirsutum* Texas Marker-1 (TM-1)'e ait referans genom bilgileri sunulmaktadır. İlgili çizelgede görüleceği üzere Upland pamuğun referans genomu 2.3 Gb büyüklüğünde olup 1031 adet iskelelenin montajlanması ile 26 adet haploit kromozom dizisi elde edilmiştir.

Çizelge 3.1. Upland pamuk referans genomuna ait bilgiler

Özellik	Referans Genom Dizisi	GenBank
Numara	GCF_007990345.1	GCA_007990345.1
Genom boyutu	2.3 Gb	2.3 Gb
Gen sayısı	101.016	75.376
Yayın Tarihi	25 Mart 2021	25 Şubat 2021
Grup	NCBI	HudsonAlpha GSC
Toplam boşluksuz uzunluk	2.3 Gb	2.3 Gb
İskeleler arası boşluklar	6	6
Kromozom sayıları	26	26
Organel sayısı	2	0
İskele (Scaffolds) sayısı	1031	1031
İskele N50	108.1 Mb	108.1 Mb
İskele L50	10	10
Örtüşen okumalar (Contig) sayısı	6733	6733
Örtüşen okumalar N50	783.9 kb	783.9 kb
Örtüşen okumalar L50	789	789
%GC	34.5	34.5
Dizileme derinliği	94.1x	94.1x
Montajlama Düzeyi	Kromozom	Kromozom
Açıklama: Gb: Gigabaz, Mb: Megabaz. L50 veya N50 anlamı toplam uzunluğunun veya sayısının yarısını kapsayan en küçük veya kısa contig veya iskele sayısını ifade eder. Genom kapsam ve dizileme derinliği bir genom veya bir genom bölgesine hizalanan okumalarının sayısıdır. Örneğin bir bölge için bir baz okuma sayısı 30 kez ise bu 30x dizileme derinliğidir. Genom kapsamı veya dizileme kapsamı ise bir genomun bütün kısmının ne kadarının dizilendiğini gösterir.		

Upland referans genomu $2n = 4x = 52$ kromozomlu tetraploit pamuk türünü temsil etmektedir. Sekansların genom dizileme derinliği 94.1x olup bu değer ile kapsamlı bir dizilemeye sahip olduğu görülmektedir. Genom kapsam kalite analizi yapıldığında okumaların %8'nin tek bir okumadan, %91.6'lık kısmının birden fazla okumadan oluştuğu, %0.3'lük bölümün parçalı kısımlardan oluştuğu ve genom düzeyinde %0.2'lik bir bölüme ait dizileme bilgisi bulunmamaktadır. Upland referans genomuna ait mitokondri ve kloroplast genom dizileri de Upland pamuk için belirlenmiştir. 1 Ağustos, 2024 tarihi itibarıyla Upland referans genomunda 99453 gen ve bu genlerden 67584 adedinin ise protein kodlayan genler olduğu ortaya konmuştur (https://www.ncbi.nlm.nih.gov/datasets/genome/GCF_007990345.1).

Çizelge 3.2'de Oryantal (Ağaç) pamuğa ait referans genom bilgileri sunulmuştur. Bu referans genom Oxford Nanopore, Illumina dizileme tekniği kullanılarak *Gossypium arboreum* Shixiya-1 hattından elde edilmiştir. Montajlama yöntemi olarak Canu v. 1.3 versiyonu kullanılmıştır. İlgili çizelgede görüleceği gibi Oryantal (Ağaç) pamuğun referans genomu 1.6 Gb büyüklüğünde olup 946 adet iskelenin montajlanması ile 13 adet kromozom dizisi elde edilmiştir. Oryantal (Ağaç) pamuk referans genomu $2n = 2x = 26$ kromozomlu diploit pamuk türünü temsil etmektedir. Sekansların genom kapsamı 131x dizilemeye sahip olduğu anlaşılmaktadır.

Çizelge 3.2. Oryantal pamuk referans genomuna ait bilgiler

Özellik	Referans Genom Dizisi	GenBank
Numara	GCF_025698485.1	GCA_025698485.2
Genom boyutu	1.6 Gb	1.6 Gb
Gen sayısı	54.481	53.474
Yayın Tarihi	16 Aralık 2022	17 Ekim 2022
Grup	RS_2022_12	Huazhong AU
Toplam boşluksuz uzunluk	1.6 Gb	1.6 Gb
Kromozom sayıları	13	13
Organel sayısı	2	0
İskele (Scaffolds) sayısı	946	946
İskele N50	129.6 Mb	129.6 Mb
İskele L50	6	6
Örtüşen okumalar (Contig) sayısı	1.171	1.171
Örtüşen okumalar N50	11.7 Mb	11.7 Mb
Örtüşen okumalar L50	40	40
%GC	35	35
Dizileme derinliği	131.0x	131.0x
Montaj seviyesi	Kromozom	Kromozom
Açıklama: Gb: Gigabaz, Mb: Megabaz. L50 veya N50 anlamı toplam uzunluğunun veya sayısının yarısını kapsayan en küçük veya kısa contig veya iskele sayısını ifade eder. Genom kapsam ve dizileme derinliği bir genom veya bir genom bölgesine hizalanan okumalarının sayısıdır. Örneğin bir bölge için bir baz okuma sayısı 30 kez ise bu 30x dizileme derinliğidir. Genom kapsamı veya dizileme kapsamı ise bir genomun bütün kısmının ne kadarının dizilendiğini gösterir.		

Çizelge 3.3’de Oxford Nanopore, Illumina dizileme teknolojisi ile ve Canu v. 1.3 montajlama yöntemi ile elde edilen *Gossypium raimondii* GPD5lz hattına ait referans genom bilgileri sunulmaktadır. İlgili çizelgede Peru pamuğu referans genomu 750.2 Mb büyüklüğünde ve 287 adet iskelenin montajlanması ile 13 adet kromozom dizisi ve ilgili diğer bazı bilgiler bulunmaktadır. Peru pamuğu referans genomu $2n = 2x = 26$ kromozomlu diploit pamuk türüdür. Sekansların genom kapsamı ise 167.0x değerine sahip olduğu bu haliyle hem Upland hem de Oryantal pamuğa kıyasla daha kapsamlı bir dizileme görülmektedir.

Çizelge 3.3. Peru pamuk referans genomuna ait bilgiler

Özellik	Referans Genom Dizisi	GenBank
Numara	GCF_025698545.1	GCA_025698545.1
Genom boyutu	750.2 Mb	750.3 Mb
Gen sayısı	47.736	46.830
Yayın Tarihi	7 Aralık 2022	17 Ekim 2022
Grup	NCBI	Huazhong AU
Toplam boşluksuz uzunluk	750.2 Mb	750.2 Mb
Kromozom sayıları	13	13
Organel sayısı	2	0
İskele (Scaffolds) sayısı	287	287
İskele N50	57.7 Mb	57.7 Mb
İskele L50	7	7
Örtüşen okumalar (Contig) sayısı	366	366
Örtüşen okumalar N50	17 Mb	17 Mb
Örtüşen okumalar L50	16	16
%GC	33	33
Dizileme derinliği	167.0x	167.0x
Montaj seviyesi	Kromozom	Kromozom
Açıklama: Gb: Gigabaz, Mb: Megabaz. L50 veya N50 anlamı toplam uzunluğunun veya sayının yarısını kapsayan en küçük veya kısa contig veya iskele sayısını ifade eder. Genom kapsam ve dizileme derinliği bir genom veya bir genom bölgesine hizalanan okumalarının sayısıdır. Örneğin bir bölge için bir baz okuma sayısı 30 kez ise bu 30x dizileme derinliğidir. Genom kapsamı veya dizileme kapsamı ise bir genomun bütün kısmının ne kadarının dizilendiğini gösterir.		

3.2. Metot

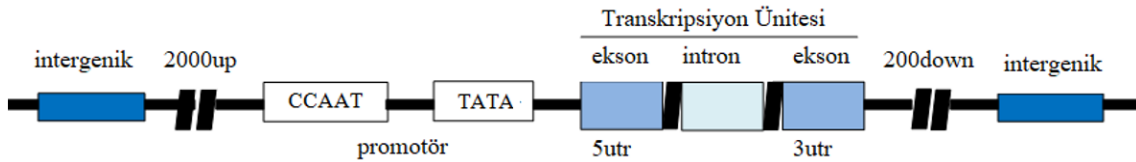
3.2.1. Pamuk Türlerinin Mikrosatellit Analizleri

Bu tez çalışmasında verilerin elde edilmesinde, ön çalışmalarında ve diğer gerekli çalışmalar Linux programları kullanılarak gerçekleştirilmiştir. Veriler “wget” veya “curl” komutları/skriptleri kullanılarak hazırlanmıştır. Türleri ait referans genom dizileri ileri (+) ve geri (-) sarmalları SeqTK yazılım programı (Li 2015) kullanılarak valide edilmiş ve fasta varsayılan formatına dönüştürülmüş ve ayrı dosyalara kaydedilmiştir. Her bir dosyadaki (+) ve (-) sarmallara ait DNA dizileri Perl tabanlı GMATA yazılım programı (Wang ve Wang 2016) kullanılarak aşağıda belirtilen kriterler doğrultusunda mikrosatellit içerikleri belirlenmiştir. Mikrosatellit tespit

kriterleri ise şu şekilde uygulanmıştır; di-nükleotit (2'li) mikrosatellitler en az 12 tekrarlı motif içerenler, tri-nükleotit (3'lü) mikrosatellitler en az 8 tekrarlı motif içerenler, tetra-nükleotit (4'lü) mikrosatellitler en az 6 tekrarlı motif içerenler, penta-nükleotit (5'li) mikrosatellitler en az 5 tekrarlı motif içerenler ve hekza-nükleotit (6'lı) mikrosatellitler 4 tekrarlı motif içerenler olmak üzere minimum 24 nükleotit düzeyinde belirlenmiştir (Ince ve Karaca 2023).

Çalışmada kullanılan GMATA yazılımı farklı motiflerin değişik tekrarlı teşhislerini bir arada analiz edemediği için her bir motif ayrı ayrı analiz edilmiştir (Karaca ve Ince 2023). Çalışmada mikrosatellitler için toplamda 30 farklı analiz gerçekleştirilmiştir. Analizler tamamlandıktan sonra veriler AWK yazılım programı (Aho vd. 1988) kullanılarak her bir tür için BED dosyaları elde edilmiştir. Oluşturulan BED dosyalarında kromozom adı ve alt genom bilgisi, genomda bulunan her bir mikrosatellit için 1 tabanlı başlama – bitiş koordinatları, mikrosatellitin tipi, içeriği ve tekrar sayısı bilgileri “tab” formatında matrilere dönüştürülmüştür.

BEDTools, INTERSECT ve AWK yazılım programları (Quinlan ve Hall 2010) kullanılarak BED formatındaki mikrosatellit dosyaları ile GFF formatındaki referans genom bilgileri karşılaştırılarak eşlenmiştir. Böylelikle her bir mikrosatellit; genom, alt genom, kromozom, 5'-UTR (mRNA'nın başında bulunan kodlamayan bölge), 3'-UTR (kodlama bölgesinden sonra gelen ve genellikle poli-A kuyruğunu içeren kodlamayan bölge), ekson (translasyona giren diziler), intron (transkripsiyona giren ancak translasyona alınmayan, iki ekson arasındaki DNA dizisi), gen bitiminden 200 bp sonra (“200 bp downstream”), gen başlamasından 2000 bp önce (“2000 bp upstream”) ve intergenik bölgeleri olmak üzere yerleştirilebilmiştir. Şekil 3.1’de çalışmada hedeflenen gen elementleri / gen bölgeleri gösterilmiştir.



Şekil 3.1. Çalışmada kullanılan gen bölgeleri

3.2.2. Pamuk Türlerine ait Selüloz Sentaz Gen Ailesinin Mikrosatellit İçerikleri ve Filogenetik Analizleri

Tez çalışmasında her üç pamuk türüne ait selüloz sentaz genleri referans genomlarda GREP programı (Thompson 1974) kullanılarak genom koordinasyonları BED dosyası ve dizileri FASTA formatında kaydedilmiştir. Selüloz sentaz ve selüloz sentaz benzeri genlerin SSR ve motif içerikleri Tandem Repeat Analyzer 1.5 (TRA 1.5) (Bilgen vd. 2004) yazılım programı kullanılarak belirlenmiştir. SSR içeren selüloz sentaz ve selüloz sentaz benzeri genler motif ve tiplerine göre listelenmiştir. Çalışılan genler Çizelgeler 3.4, 3.5 ve 3.6’da listelenmiştir. Selüloz sentaz genlerinin filogenetik analizleri ise Mega11 (Tamura vd. 2021) programı kullanılarak analiz edilmiştir. SSR içeren genlerin türler arasında ve türler içerisindeki dağılımları diğer genler ile olan yakınlık ilişkileri moleküler filogeni yöntemiyle ortaya analiz edilmiştir. Ayrıca grafiklerin görselleştirilmesinde SRPlot yazılım programı kullanılmıştır (Tang vd. 2023). Çizelge 3.4’te, Oryantal pamuğa ait selüloz sentaz A ve selüloz sentaz benzeri genler (38 lokus) gösterilmektedir.

Çizelge 3.4. Oryantal pamukta kullanılan selüloz sentaz A ve selüloz sentaz benzeri genler

UID No	Gen	Enzim/Protein	Sarmal
108482778	<i>CsIE6</i>	Cellulose synthase-like protein E6	(+)
108483025	<i>CsIE1</i>	Cellulose synthase-like protein E1	(+)
108475068	<i>CsIG2</i>	Cellulose synthase-like protein G2	(+)
108489335	<i>CesA7-like</i>	Cellulose synthase A catalytic subunit 7-like	(+)
108469566	<i>CesA3-like</i>	Cellulose synthase A catalytic subunit 3-like	(+)
108468415	<i>CesA3-like</i>	Cellulose synthase A catalytic subunit 3-like	(+)
108469868	<i>CesA4-like</i>	Cellulose synthase A catalytic subunit 4-like	(+)
108456087	<i>CsID1</i>	Cellulose synthase-like protein D1	(+)
108454873	<i>CsIE1</i>	Cellulose synthase-like protein E1	(+)
108455269	<i>CsIE1</i>	Cellulose synthase-like protein E1	(+)
108456116	<i>CsIG2</i>	Cellulose synthase-like protein G2	(+)
108457518	<i>CsIE1</i>	Cellulose synthase-like protein E1	(+)
108489366	<i>CesA8-like</i>	Cellulose synthase A catalytic subunit 8-like	(+)
128282068	<i>CsIG2</i>	Cellulose synthase-like protein G2	(+)
108450932	<i>CsIG3</i>	Cellulose synthase-like protein G3	(+)
108471157	<i>CsIE1</i>	Cellulose synthase-like protein E1	(+)
108464531	<i>CesA2</i>	Cellulose synthase A catalytic subunit 2	(+)
108476974	<i>CsIG3</i>	Cellulose synthase-like protein G3	(+)
108462694	<i>Prob-CesA5</i>	Probable cellulose synthase A catalytic subunit 5	(+)
108466975	<i>CesA2-like</i>	Cellulose synthase A catalytic subunit 2-like	(-)
108486136	<i>CsID3</i>	Cellulose synthase-like protein D3	(-)
108467006	<i>Prob-CesA8</i>	Probable cellulose synthase A catalytic subunit 8	(-)
108475592	<i>CesA7-like</i>	Cellulose synthase A catalytic subunit 7-like	(-)
108471752	<i>CsID4</i>	Cellulose synthase-like protein D4	(-)

Çizelge 3.4'ün devamı

UID No	Gen	Enzim/Protein	Sarmal
108489015	<i>CesA4-like</i>	Cellulose synthase A catalytic subunit 4-like	(-)
108480723	<i>CesA9-like</i>	Cellulose synthase A catalytic subunit 9-like	(-)
108469859	<i>CsID3</i>	Cellulose synthase-like protein D3	(-)
108469578	<i>CesA3</i>	Cellulose synthase A catalytic subunit 3	(-)
108454833	<i>CsIG3</i>	Cellulose synthase-like protein G3	(-)
108456145	<i>CsIG2</i>	Cellulose synthase-like protein G2	(-)
108454864	<i>CsIG3</i>	Cellulose synthase-like protein G3	(-)
108473927	<i>CsIB4</i>	Cellulose synthase-like protein B4	(-)
108486168	<i>CesA2</i>	Cellulose synthase A catalytic subunit 2	(-)
108472682	<i>CsIE1</i>	Cellulose synthase-like protein E1	(-)
108474219	<i>CsIE1</i>	Cellulose synthase-like protein E1	(-)
108479552	<i>CsID5</i>	Cellulose synthase-like protein D5	(-)
108479185	<i>CsID2</i>	Cellulose synthase-like protein D2	(-)
108466975	<i>CesA2-like</i>	Cellulose synthase A catalytic subunit 2-like	(-)

Çizelge 3.5'te, Peru pamuğuna ait selüloz sentaz A ve selüloz sentaz benzeri genlerin (46 lokus) UID numaraları ve sarmaldaki konumları gösterilmektedir.

Çizelge 3.5. Peru pamukta kullanılan selüloz sentaz A ve selüloz sentaz benzeri genler

UID No	Gen	Enzim/Protein	Sarmal
108482778	<i>CsIE6</i>	Cellulose synthase-like protein E6	(+)
105796705	<i>CesA7</i>	Cellulose synthase A catalytic subunit 7	(+)
105780982	<i>CsID4</i>	Cellulose synthase-like protein D4	(+)
105769076	<i>CesA8</i>	Cellulose synthase A catalytic subunit 8	(+)
105769839	<i>CesA5</i>	Cellulose synthase A catalytic subunit 5	(+)
105780427	<i>CesA7</i>	Cellulose synthase A catalytic subunit 7	(+)
105791773	<i>CesA3</i>	Cellulose synthase A catalytic subunit 3	(+)
105792680	<i>CsID3</i>	Cellulose synthase-like protein D3	(+)
105800355	<i>CsID1</i>	Cellulose synthase-like protein D1	(+)
105800594	<i>CsIE1</i>	Cellulose synthase-like protein E1	(+)
105800591	<i>CsIG3</i>	Cellulose synthase-like protein G3	(+)
105800592	<i>CsIG2</i>	Cellulose synthase-like protein G2	(+)
105776704	<i>CesA8</i>	Cellulose synthase A catalytic subunit 8	(+)
105775938	<i>CsIG2</i>	Cellulose synthase-like protein G2	(+)
105804467	<i>CsIE1</i>	Cellulose synthase-like protein E1	(+)
105804459	<i>CsIE1</i>	Cellulose synthase-like protein E1	(+)
105804460	<i>CsIE1</i>	Cellulose synthase-like protein E1	(+)

Çizelge 3.5'in devamı

UID No	Gen	Enzim/Protein	Sarmal
105785384	<i>CesA2</i>	Cellulose synthase A catalytic subunit 2	(+)
105803468	<i>CslH1</i>	Cellulose synthase-like protein H1	(+)
105762106	<i>CesA2-like</i>	Cellulose synthase A catalytic subunit 2-like	(+)
105765323	<i>CslG3</i>	Cellulose synthase-like protein G3	(+)
105781668	<i>CslE6</i>	Cellulose synthase-like protein E6	(+)
105781669	<i>CslE1</i>	Cellulose synthase-like protein E1	(+)
105783616	<i>CslE6</i>	Cellulose synthase-like protein E6	(+)
105782024	<i>CslE1</i>	Cellulose synthase-like protein E1	(-)
105782063	<i>CslE6</i>	Cellulose synthase-like protein E6	(-)
105787891	<i>CesA2</i>	Cellulose synthase A catalytic subunit 2	(-)
105787927	<i>CslD3</i>	Cellulose synthase-like protein D3	(-)
105796703	<i>CslG2</i>	Cellulose synthase-like protein G2	(-)
128039937	<i>CesA8-like</i>	Cellulose synthase A catalytic subunit 8-like	(-)
105767345	<i>CesA1</i>	Cellulose synthase A catalytic subunit 1	(-)
105767347	<i>CesA7</i>	Cellulose synthase A catalytic subunit 7	(-)
105767346	<i>CesA1</i>	Cellulose synthase A catalytic subunit 1	(-)
105773046	<i>CesA5</i>	Cellulose synthase A catalytic subunit 5	(-)
105766109	<i>CesA4</i>	Cellulose synthase A catalytic subunit 4	(-)
105790486	<i>CesA4</i>	Cellulose synthase A catalytic subunit 4	(-)
105790569	<i>CesA3</i>	Cellulose synthase A catalytic subunit 3	(-)
105800206	<i>CslG3</i>	Cellulose synthase-like protein G3	(-)
105800205	<i>CslG2</i>	Cellulose synthase-like protein G2	(-)
105800204	<i>CslG3</i>	Cellulose synthase-like protein G3	(-)
105804464	<i>CslE1</i>	Cellulose synthase-like protein E1	(-)
105800885	<i>CslG3</i>	Cellulose synthase-like protein G3	(-)
105762105	<i>CesA6</i>	Cellulose synthase A catalytic subunit 6	(-)
105764094	<i>CslD5</i>	Cellulose synthase-like protein D5	(-)
105764976	<i>CslD3</i>	Cellulose synthase-like protein D3	(-)
105781684	<i>Prob-CesA5</i>	Probable cellulose synthase A catalytic subunit 5	(-)

Çizelge 3.6’da Upland pamuğa ait selüloz sentaz A ve selüloz sentaz benzeri genlerinin (82 lokus) UID numaraları ve sarmaldaki konumları gösterilmektedir.

Çizelge 3.6. Upland pamuğa ait selüloz sentaz A ve selüloz sentaz benzeri genler

UID No	Gen	Enzim/Protein	Sarmal
121203141	<i>CesA7-like</i>	Cellulose synthase A catalytic subunit 7-like	(+)
107951817	<i>Prob-CesA8</i>	Probable cellulose synthase A catalytic subunit 8	(+)
121229536	<i>CesA5-like</i>	Cellulose synthase A catalytic subunit 5-like	(+)
107955410	<i>CesA4</i>	Cellulose synthase A catalytic subunit 4	(+)
107938343	<i>CesA3</i>	Cellulose synthase A catalytic subunit 3	(+)
107940246	<i>CsID3</i>	Cellulose synthase-like protein D3	(+)
107942314	<i>CsID1</i>	Cellulose synthase-like protein D1	(+)
107928422	<i>CsIE1</i>	Cellulose synthase-like protein E1	(+)
107928458	<i>CsIG2</i>	Cellulose synthase-like protein G2	(+)
107928590	<i>CsIG2</i>	Cellulose synthase-like protein G2	(+)
107925233	<i>CesA8</i>	Cellulose synthase A catalytic subunit 8	(+)
107898377	<i>CsII1</i>	Protein cellulose synthase interactive 1	(+)
107902334	<i>CesA2-like</i>	Cellulose synthase A catalytic subunit 2-like	(+)
107951165	<i>CsIE1</i>	Cellulose synthase-like protein E1	(+)
107923673	<i>CsIE1</i>	Cellulose synthase-like protein E1	(+)
107894409	<i>CsIG2</i>	Cellulose synthase-like protein G2	(+)
107927657	<i>CsIG2</i>	Cellulose synthase-like protein G2	(+)
107945371	<i>Prob-CesA8</i>	Probable cellulose synthase A catalytic subunit 8	(+)
107949942	<i>CesA6-like</i>	Cellulose synthase A catalytic subunit 6-like	(+)
121217085	<i>CesA8-like</i>	Cellulose synthase A catalytic subunit 8-like	(+)
107902682	<i>CesA8</i>	Cellulose synthase A catalytic subunit 8	(+)
107914099	<i>CesA5</i>	Cellulose synthase A catalytic subunit 5	(+)
107937655	<i>CesA4-like</i>	Cellulose synthase A catalytic subunit 4-like	(+)
107900907	<i>CsII1</i>	Protein cellulose synthase interactive 1	(+)
107900748	<i>CesA3</i>	Cellulose synthase A catalytic subunit 3	(+)
107909056	<i>CsID3</i>	Cellulose synthase-like protein D3	(+)
121220991	<i>CsID1</i>	Cellulose synthase-like protein D1	(+)
107928982	<i>CsIE1</i>	Cellulose synthase-like protein E1	(+)
107929047	<i>CsIG3</i>	Cellulose synthase-like protein G3	(+)
107929134	<i>CsIG2</i>	Cellulose synthase-like protein G2	(+)
107909647	<i>CesA8-like</i>	Cellulose synthase A catalytic subunit 8-like	(+)
107916051	<i>CsIG3</i>	Cellulose synthase-like protein G3	(+)
107915502	<i>CsIG2</i>	Cellulose synthase-like protein G2	(+)
107911480	<i>CesA2</i>	Cellulose synthase A catalytic subunit 2	(+)
107935220	<i>CsIE1</i>	Cellulose synthase-like protein E1	(+)
107945523	<i>CsII1</i>	Protein cellulose synthase interactive 1	(+)
107947180	<i>CsIG3</i>	Cellulose synthase-like protein G3	(+)
107922547	<i>Prob-CesA5</i>	Probable cellulose synthase A catalytic subunit 5	(+)

Çizelge 3.6'nın devamı

UID No	Gen	Enzim/Protein	Sarmal
107918220	<i>CsIE1</i>	Cellulose synthase-like protein E1	(-)
107916923	<i>CsIE6</i>	Cellulose synthase-like protein E6	(-)
107925040	<i>CesA2</i>	Cellulose synthase A catalytic subunit 2	(-)
107924712	<i>CsID3</i>	Cellulose synthase-like protein D3	(-)
107949226	<i>CsID4</i>	Cellulose synthase-like protein D4	(-)
107958681	<i>CesA1-like</i>	Cellulose synthase A catalytic subunit 1-like	(-)
107958683	<i>CesA7</i>	Cellulose synthase A catalytic subunit 7	(-)
107958687	<i>CesA1</i>	Cellulose synthase A catalytic subunit 1	(-)
121229512	<i>CsII3-like</i>	Protein cellulose synthase interactive 3-like	(-)
107962315	<i>CesA5-like</i>	Cellulose synthase A catalytic subunit 5-like	(-)
107961131	<i>CesA4-like</i>	Cellulose synthase A catalytic subunit 4-like	(-)
107896451	<i>CsIG2</i>	Cellulose synthase-like protein G2	(-)
107896452	<i>CsIG3</i>	Cellulose synthase-like protein G3	(-)
107937198	<i>CsIE1</i>	Cellulose synthase-like protein E1	(-)
121212539	<i>CsIE6</i>	Cellulose synthase-like protein E6	(-)
107894422	<i>CsIE1</i>	Cellulose synthase-like protein E1	(-)
107927704	<i>CsIG2</i>	Cellulose synthase-like protein G2	(-)
107950592	<i>CesA2</i>	Cellulose synthase A catalytic subunit 2	(-)
107950561	<i>CsID3</i>	Cellulose synthase-like protein D3	(-)
107898894	<i>CsID4</i>	Cellulose synthase-like protein D4	(-)
107907453	<i>CesA1</i>	Cellulose synthase A catalytic subunit 1	(-)
107907451	<i>CesA7</i>	Cellulose synthase A catalytic subunit 7	(-)
107942877	<i>CesA1</i>	Cellulose synthase A catalytic subunit 1	(-)
107940554	<i>CsII3</i>	Protein cellulose synthase interactive 3	(-)
107901428	<i>CsII3</i>	Protein cellulose synthase interactive 3	(-)
107901762	<i>CesA5</i>	Cellulose synthase A catalytic subunit 5	(-)
107953681	<i>CesA7-like</i>	Cellulose synthase A catalytic subunit 7-like	(-)
107916504	<i>CesA4-like</i>	Cellulose synthase A catalytic subunit 4-like	(-)
107915763	<i>CesA3</i>	Cellulose synthase A catalytic subunit 3	(-)
107890709	<i>CsIG3</i>	Cellulose synthase-like protein G3	(-)
107892837	<i>CsIG3</i>	Cellulose synthase-like protein G3	(-)
107892464	<i>CsIG3</i>	Cellulose synthase-like protein G3	(-)
107912646	<i>CsIH1</i>	Cellulose synthase-like protein H1	(-)
107913679	<i>CsII1-like</i>	Protein cellulose synthase interactive 1-like	(-)
107935224	<i>CsIE1</i>	Cellulose synthase-like protein E1	(-)
107935218	<i>CsIE1</i>	Cellulose synthase-like protein E1	(-)
107935226	<i>CsIE1</i>	Cellulose synthase-like protein E1	(-)
107946051	<i>CsID5</i>	Cellulose synthase-like protein D5	(-)
107946664	<i>CsID2</i>	Cellulose synthase-like protein D2	(-)
121203053	<i>CsIE6</i>	Cellulose synthase-like protein E6	(-)

Çizelge 3.6'nın devamı

UID No	Gen	Enzim/Protein	Sarmal
107947909	<i>CslG2</i>	Cellulose synthase-like protein G2	(-)
121207998	<i>CslE1</i>	Cellulose synthase-like protein E1	(-)
107924700	<i>CslB4</i>	Cellulose synthase-like protein B4	(-)
107925659	<i>CslD5</i>	Cellulose synthase-like protein D5	(-)

3.2.3. Pamuk Türlerine ait Selüloz Sentaz Mikrosatellitlerinin Kodladığı Aminoasitlerin Belirlenmesi

Çalışmada kullanılan *G. hirsutum* (Upland pamuk), *G. raimondii* (Peru pamuk) ve *G. arboreum* (Oryantal pamuk) genomlarına ait gen dizilerinin her bir tür için ayrı ayrı ID numaralarına, selüloz sentaz genlerine ve motiflerine göre incelenmiştir. Mikrosatellit DNA dizilerinin *in silico* translasyonu (Aminoasit dizilerine çevirisi) ExPASy (<https://web.expasy.org/translate/>) programı aracılığıyla sadece 5' => 3' yönünde 3 farklı Reading Frame (okunabilir çerçeve, (RF)) veya Açık Okunabilir Çerçeve (ORF) göre yapılmıştır.

4. BULGULAR VE TARTIŞMA

4.1. Bütün Genom Düzeyinde Mikrosatellitler (SSR)

Önemli bir endüstri bitkisi olan pamuk; lifleri ile tekstil sanayisinin, lifleri elde edildikten sonra geri kalan kısmı tohum, yağ ve küspe üretiminde önemli bir bitkidir. Dünyada 52 pamuk türünden 4'ünün kültürü yapılmaktadır. Bu tez çalışmasında kültürü yapılan 4 pamuk türünden 3'ü olan *Gossypium arboreum* (Ağaç pamuk - Oryantal pamuk, Hindistan orijinli), *Gossypium raimondii* (Peru pamuğu, Güney Amerika orijinli) ve *Gossypium hirsutum*'un (Upland pamuk, Orta Amerika orijinli) referans genomları kullanılarak mikrosatellitlerin türlere, türlerin alt genomlarına, kromozomlara ve gen elementlerine (upstream, 5'-UTR, ekson, intron, 3'-UTR, intergen, bütün gen, downstream) göre dağılımları ve içerikleri çalışılmıştır. Analizler her bir tür için ayrı ayrı ve daha sonra türler karşılaştırılmalı gerçekleştirilmiştir. Literatürde mikrosatellitlerin tanımı, dağılımları ve fonksiyonları konusunda tam bir konsensus bulunmamaktadır. Bu çalışmada mikrosatellitler 2 ile 6 nükleotit uzunluklarındaki motifler olarak tanımlanmış ve her bir motifdeki tekrar sayıları di- için 12 tekrar, tri- için 8 tekrar, tetra- için 6 tekrar, penta- için 5 tekrar ve hekza-nükleotitler için 4 tekrar olacak şekilde GMATA programı kullanılarak saptanmıştır. Ayrıca bu çalışmada yerkürede en fazla üretilen polisakkarit olan selülozu sentezleyen selüloz sentaz genlerinin mikrosatellit içerikleri ve filogenik yönünden çalışmalarına da yer verilmiştir.

4.1.1. Genom ve Kromozomlarda SSR Dağılımları

Çizelge 4.1'de *Gossypium arboreum*'a ait mikrosatellit dağılımları gösterilmektedir. Çizelgede de görüleceği üzere toplam 34561 adet SSR belirlenmiştir. Veriler incelendiğinde bu diploit türün genomunun her bir kilo bazında (kb) yaklaşık 46,11 SSR belirlenmiştir (Çizelge 4.1).

Çizelge 4.1. Oryantal pamuğa ait kromozom ve genlerin SSR dağılımları

Krom	Büyüklik (Mb)	SSR sayısı	Genik SSR	Sıklık	İntergenik SSR	Sıklık
A01	123,41	2693	144	1,17	2658	21,54
A02	115,15	2128	96	0,83	2095	18,19
A03	139,24	2903	118	0,85	2867	20,60
A04	122,51	2308	106	0,87	2272	18,55
A05	93,16	2586	168	1,80	2513	26,98
A06	143,37	2862	168	1,17	2809	19,60
A07	106,03	2562	134	1,26	2525	23,81
A08	142,30	2766	150	1,05	2725	19,15
A09	89,40	2208	150	1,68	2168	24,25
A10	131,76	3066	163	1,24	3013	22,87
A11	137,85	3130	195	1,41	3057	22,18
A12	119,98	2757	155	1,29	2708	22,57
A13	129,59	2592	138	1,06	2559	19,75
Σ	1593,75	34561	1885	1,18	33969	21,31

*Sıklık: Uzunluk (Mb) / SSR sayısı

En fazla SSR sayısının A11 kromozomunda (3130) ve en az SSR sayısının ise A02 kromozomunda (2128) bulunduğu belirlenmiştir. *Gossypium arboreum*'un A01 numaralı kromozomunun en büyük, A13 kromozomunun ise en küçük kromozomlar olmasına rağmen çalışmamızda A11 kromozomunun en fazla ve A02 kromozomunun en az SSR'a sahip olması nedeniyle SSR sayısının ve kromozom büyüklüğü arasında doğrudan bir ilişki olmadığı görülmüştür. Diğer bir ifade ile mikrosatellitler (SSR) genomda dağılımları tam olarak rastlantısal olmadığı görülmüştür (Bilgen vd. 2004).

Referans genom bilgileri incelendiğinde Oryantal pamukta (*Gossypium arboreum*) tanımlanmış 41330 protein kodlayan ve 20803 protein kodlamayan gen olmak üzere toplam 62133 gen bulunmaktadır (Li vd. 2014). Analiz sonucunda Oryantal pamukta 1885 adet genin mikrosatellit içerdiği tespit edilmiştir (Çizelge 4.1). Bu veri *Gossypium arboreum*'un genlerinde yaklaşık %3'nün mikrosatellit içerdiğini göstermektedir. Protein kodlamayan genlerin uzunlukları protein kodlayan genlerden oldukça kısadır. Örneğin miRNA ve snRNA genleri 22 nükleotit, tRNA genleri 60-90 nükleotit, lncRNA genleri ise 300 nükleotitten daha kısadır (Karaca ve Ince 2024). Bu nedenle bu tür protein kodlamayan genler daha az düzeyde SSR içermeleri, büyüklükleri ile ilgili olabilir.

Çizelge 4.1'de görüleceği üzere SSR'lar intergenik bölgede genik bölgelere oranla daha yüksek oranda bulunmaktadır. Örneğin A03'te toplam SSR sayısı 2903 olup bunun 118 adeti genlerde bulunurken, 2867 adeti intergenik bölgede bulunmaktadır. SSR'ların intergenik bölgelerde bulunduğu daha önceki çalışmalarda da ortaya konmuş olup bu sonuçlar tez çalışmasıyla uyumluluk göstermektedir. *Gossypium arboreum*'da en fazla genik SSR içeren kromozom A11'de 195 iken; en az A02'de 96 adet SSR içermektedir. SSR bulundurma frekansı incelendiğinde ise en fazla A05'te her megabaz başına düşen SSR sayısı 1,80 olup en az A02'de 0,83 SSR bulunmaktadır. Veriler birlikte incelendiğinde Oryantal pamuktaki kromozomlar SSR yönünden farklılıklar göstermektedir. İlgili çizelgede de görüleceği üzere toplam SSR sayısı genik SSR sayısı ve intergenik SSR sayısının toplamalarının altında kalmaktadır. Mikrosatellitlerin intergenik bölgelerde sayısının çok yüksek olması her ne kadar ilk keşfedildiklerinde "çöp" DNA olarak nitelenmelerini anımsatsa dahi fonksiyonel olmadıkları anlamına gelmemektedir. Örneğin genik olmayan mikrosatellitlerin genom stabilitesi, nükleozom oluşumunda veya mutasyon absorpsiyonuna katkı sağlamayacağı anlamına gelmemektedir (M. Karaca, Kişisel Görüşmeler).

Çizelge 4.2'de Peru pamuğa (*Gossypium raimondii*) ait mikrosatellitlerin bütün genomda dağılımları gösterilmektedir. Çizelgede de görüleceği üzere toplam 20699 adet mikrosatellit belirlenmiştir. Veriler incelendiğinde bu diploit türün genomunun her bir kilo bazında yaklaşık 40,40 mikrosatellit belirlenmiştir. Oryantal pamukta bu değer 46,11 olduğu göz önüne alındığında Peru pamuğunun Oryantal pamuktan daha az sıklıkta mikrosatellit içerdiğini ortaya koymuştur. Peru pamuğunun kromozomları incelendiğinde en fazla mikrosatellit sayısının D11 kromozomunda 1760, en az ise D02'de 1269 olarak belirlenmiştir (Çizelge 4.2). *Gossypium raimondii*'nin D01 numaralı kromozomunun en büyük, D13 kromozomunun ise en küçük kromozomlar olmasına rağmen çalışmamızda D11 kromozomunun en fazla ve D02 kromozomunun en az mikrosatellite sahip olması nedeniyle mikrosatellit sayısının ve kromozom büyüklüğü arasında doğrudan bir ilişki olmadığı görülmüştür. Peru pamuğuna ait 40976

protein kodlayan ve 3036 protein kodlamayan gen olmak üzere toplam 44012 gen bulunmaktadır (Wang vd. 2012).

Çizelge 4.2. Peru pamuğa ait kromozom ve genlerin SSR dağılımları

Krom	Büyük­lük (Mb)	SSR sayısı	Genik SSR	Sıklık	İntergenik SSR	Sıklık
D01	52,42	1474	44	0,84	1430	27,28
D02	48,27	1269	35	0,73	1221	25,30
D03	63,02	1547	30	0,48	991	15,73
D04	50,65	1479	48	0,95	1431	28,25
D05	57,14	1729	42	0,74	1672	29,26
D06	60,40	1676	16	0,26	1278	21,16
D07	53,50	1597	45	0,84	1506	28,15
D08	62,61	1721	40	0,64	1434	22,90
D09	49,93	1471	53	1,06	1382	27,68
D10	61,94	1650	54	0,87	1580	25,51
D11	72,82	1760	56	0,77	1682	23,10
D12	63,26	1685	21	0,33	1051	16,61
D13	65,10	1641	63	0,97	1554	23,87
Σ	836,33	20699	547	0,65	18212	21,77

*Sıklık: Uzunluk (Mb) / SSR sayısı

Çalışmada 547 adet *G. raimondii* geninde mikrosatellit bulunduğu tespit edilmiştir (Çizelge 4.2). Bu veri *G. raimondii*'nin genlerinin yaklaşık %8'nin mikrosatellit içerdiğini göstermektedir. Bu rakam Oryantal pamuk ile karşılaştırıldığında yaklaşık iki kez daha fazla olduğu anlamına gelmektedir. Çizelge 4.2'de görüleceği üzere mikrosatellitler intergenik bölgede genik bölgelere oranla daha yüksek oranda bulunmaktadır. Örneğin D07'de toplam mikrosatellit sayısı 1597 olup bunun sadece 45 adeti genlerde bulunurken, 1506 adeti ise intergenik bölgede bulunmaktadır. Mikrosatellitlerin intergenik bölgelerde bulunduğu daha önceki çalışmalarda da ortaya konmuş olup bu sonuçlar tez çalışmasıyla uyumluluk göstermektedir (Bilgen vd. 2004). *G. raimondii*'de en fazla genik mikrosatellit içeren kromozom D13'te 63 iken; en az D06'da 16 adet mikrosatellit içermektedir. Mikrosatellit bulundurma frekansı (sıklığı) incelendiğinde ise en fazla D09'da her megabaz başına düşen mikrosatellit sayısı 1,06 olup, en az D06'da 0,26 mikrosatellit bulunmaktadır. Veriler birlikte incelendiğinde kromozomların mikrosatellit içerikleri yönünden karşılaştırıldığında önemli farklılıklar göstermektedir. Çizelge 4.2'de görüleceği üzere toplam mikrosatellit sayısı, genik ve intergenik mikrosatellit sayısının toplamlarından fazladır.

Bu tez çalışmasında araştırılan diğer bir tür Upland pamuğudur. Dünya ve ülkemiz genelinde yetiştirilen pamuk türleri arasında Upland pamuğu (*Gossypium hirsutum*) en yaygın olarak yetiştirilen pamuk türüdür (Cronn vd. 2003; Karaca vd. 2020). Çizelge 4.3'te *G. hirsutum*'a ait mikrosatellit dağılımları gösterilmektedir. Çizelgede de görüleceği üzere toplam 27717 adet SSR belirlenmiştir. Veriler incelendiğinde bu tetraploit türün genomunun her bir kilo bazında yaklaşık 82,35 SSR

belirlenmiştir. Her iki türün, Oryantal ve Peru, SSR sıklığı ile karşılaştırıldığında bu sıklığın her iki genomun yaklaşık iki katı olduğu, evrimsel süreçte mikrosatellit varlığında bir azalış yaşanmadığı anlaşılmıştır.

Çizelge 4.3. Upland pamuğa ait kromozom ve genlerin SSR dağılımları

Krom	Büyüklik (Mb)	SSR sayısı	Genik SSR	Sıklık	İntergenik SSR	Sıklık
A01	119,76	1053	104	0,87	1013	8,46
A02	108,14	957	86	0,80	921	8,52
A03	113,69	1111	93	0,82	1068	9,39
A04	89,18	800	43	0,48	786	8,81
A05	111,10	1453	151	1,36	1374	12,37
A06	129,20	1170	70	0,55	1138	8,88
A07	98,90	1093	100	1,01	1048	10,60
A08	127,50	1219	94	0,74	1179	9,25
A09	85,34	972	110	1,29	925	10,84
A10	118,18	1286	91	0,77	1246	10,54
A11	124,18	1350	129	1,04	1286	10,36
A12	109,47	1176	97	0,89	1134	10,36
A13	111,65	1111	90	0,81	1072	9,60
Σ	1446,29	14751	1258	0,87	14190	9,81
D01	65,21	1073	127	1,95	1020	15,64
D02	72,19	1001	110	1,52	959	13,29
D03	54,96	775	85	1,55	737	13,41
D04	58,23	835	74	1,27	807	13,86
D05	66,48	1210	175	2,63	1137	17,10
D06	66,68	978	105	1,57	930	13,95
D07	59,44	961	108	1,82	905	15,23
D08	69,43	1089	114	1,64	1044	15,04
D09	54,44	892	109	2,00	843	15,48
D10	68,09	1063	114	1,67	1018	14,95
D11	72,82	1122	153	2,10	1058	14,53
D12	63,26	1026	143	2,26	953	15,07
D13	65,10	941	108	1,66	891	13,69
Σ	836,33	12966	1525	1,82	12302	14,71
GenelΣ	2282,62	27717	2783	1,22	26492	11,61

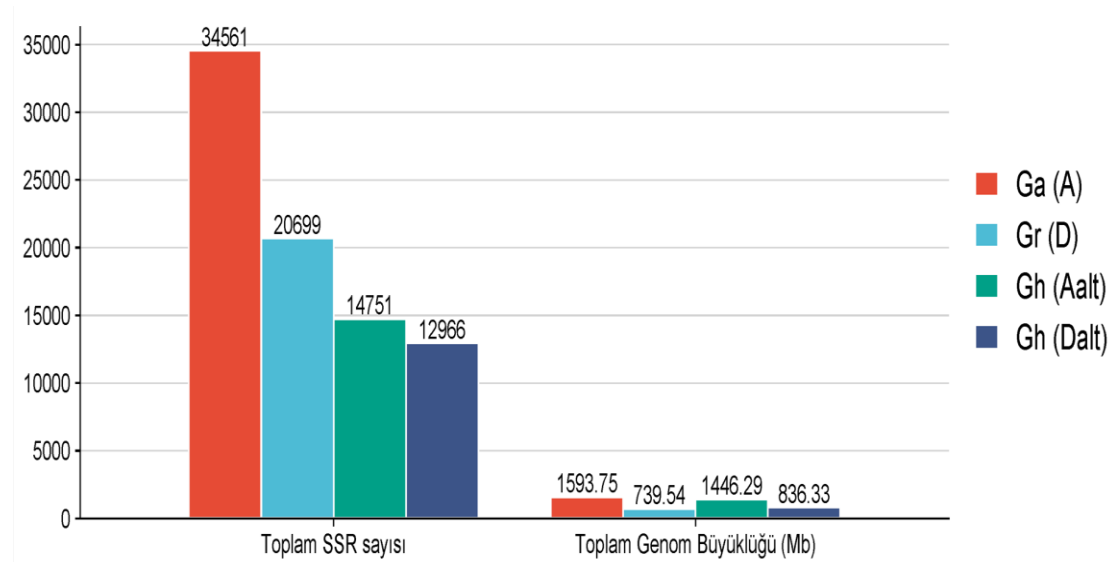
*Sıklık: Uzunluk (Mb) / SSR sayısı

Bu türün, (Upland pamuk), kromozomları incelendiğinde en fazla SSR sayısının A05 kromozomunda 1453, en az ise D03 kromozomunda 775 olarak belirlenmiştir. Referans genom bilgilerine göre *G. hirsutum*'a ait 70478 protein kodlayan ve 4473 protein kodlamayan gen olmak üzere toplam 74951 gen bulunmaktadır (Zhang vd. 2015). Çalışmada 2783 adet *G. hirsutum* geninde SSR bulunduğu tespit edilmiştir (Çizelge 4.3). Gen sayıları yönünden incelendiğinde en fazla genin Upland türünde (74951), Oryantal pamuğun bunu (62133) izlediği ve Peru pamukta ise en az gen

sayısının (44012) bulunduğu anlaşılmaktadır. Upland pamuğun genlerinin yaklaşık %2'sinin SSR içerdiğini göstermektedir. Bu yönden bakıldığında ise gen sayısı SSR oranı yönünden Upland pamuğun her iki türden daha fazla gen içermesine rağmen genlerinin daha az sayıda mikrosatellit içerdiği tespit edilmiştir.

Çizelge 4.3'te görüleceği üzere SSR'lar intergenik bölgede genik bölgelere oranla daha yüksek oranda bulunmaktadır. Örneğin A08'de toplam SSR sayısı 1219 olup bunun sadece 94 adeti genlerde bulunurken; 1179 adeti intergenik bölgede bulunmaktadır. SSR'ların intergenik bölgelerde bulunduğu daha önceki çalışmalarda da ortaya konmuş olup bu sonuçlar tez çalışmasıyla uyumluluk göstermektedir. Upland pamukta en fazla genik SSR içeren kromozom D05'te 175 iken; en az A04'te 43 adet SSR içermektedir. SSR bulundurma frekansı incelendiğinde ise en fazla D05'te her megabaz başına düşen SSR sayısı 2,63 olup en az A04'te 0,48 SSR bulunmaktadır. Veriler birlikte incelendiğinde kromozomlar SSR yönünden farklılıklar göstermektedir. Çizelge 4.3'te görüleceği üzere toplam SSR sayısı genik SSR sayısı ve intergenik SSR sayısının toplamalarının altında kalmaktadır.

Genetik, genomik ve evrimsel çalışmalar için mikrosatellitlerin (SSR'ların) genom çapında analizi ile elde edilen veriler etkin olarak kullanılabilir. Bu tez çalışmasında üç pamuk türünün genom büyüklükleri, SSR sayısı ile karşılaştırıldığında genom büyüklüğü açısından en fazla 2282,62 Mb ile tetraploit tür olan Upland pamuğa ait olmasına rağmen SSR sayısına bakıldığında ise en fazla 34561 ile Oryantal pamuğa aittir. A genomu içeren Oryantal pamuk ile A alt genomuna sahip Upland pamukları genom büyüklükleri ve D genomuna sahip Peru pamuğu ile D alt genomuna sahip Upland pamukları genom büyüklükleri açısından sayıca yakın olsa da SSR sayısı açısından farklılık göstermektedir. Bu da SSR sayısı ile toplam genom büyüklüğü açısından doğrudan bir ilişki olmadığını göstermektedir (Şekil 4.1).



Şekil 4.1. Üç pamuk türünün SSR sayısı ve genom büyüklükleri yönünden karşılaştırılması

Çalışma sonuçları değerlendirildiğinde Upland pamuktaki SSR'ların kromozomlar arasındaki dağılımları tam anlamda rastlantısal olmayıp değişkendir. SSR'ların eşit şekilde dağılmaması bir taraftan SSR markırların aslında genetik markır olarak kullanılmasında dikkatli olunması gerektiğini ortaya koyarken diğer taraftan rastlantısal olmaması ise bilinen fonksiyonlarının ötesinde farklı ödevlerinin de olabileceğini ortaya koymuştur. Detaylı incelemede mikrosatellitlerin kromozomlar arasında eşit dağılmamasının yanı sıra alt genomlar arasında da eşit dağılmadığını ortaya koymaktadır. Bu yönüyle mikrosatellitlerin kromozom veya alt genomsal içerikleri, sayı ve motifsel özellikleri onların genomsal bir özellikleri olarak ortaya konulma özellikleri olabileceği görüşünü ortaya koymaktadır. Hem kromozomda eşit dağılmaması hem de alt genomda eşit dağılmaması bu tür SSR'ların markır olarak kullanılmasının sakıncalı olduğunu ortaya koymaktadır. Diğer taraftan bu tür SSR'ların genomda farklı dağılımları bunların genom organizasyonunda, genomun gen ifadesinde etkili olduğunu göstermektedir.

4.1.2. SSR Motif Uzunlukları ve Motif İçerikleri

Çizelge 4.4'te *Gossypium arboreum*'a ait mikrosatellit sayıları ve mikrosatellitlerin kromozomdaki dağılımları gösterilmektedir. Tanımlanan SSR'ların tekrarları di-, tri-, tetra-, penta- ve hekza-nükleotitler olmak üzere sınıflandırılmıştır. Çizelgede de görüldüğü üzere en çok bulunan mikrosatellit, 19576 (SSR'ların %57'si) sayısı ile di-nükleotitlerdir. Bunu sırasıyla tri-nükleotitler 7662 (SSR'ların %22'si), hekza-nükleotitler 3623 (SSR'ların %10'u), penta-nükleotitler 2191 (SSR'ların %6'sı) ve tetra-nükleotitler 1509 ile (SSR'ların %4'ü) takip etmektedir.

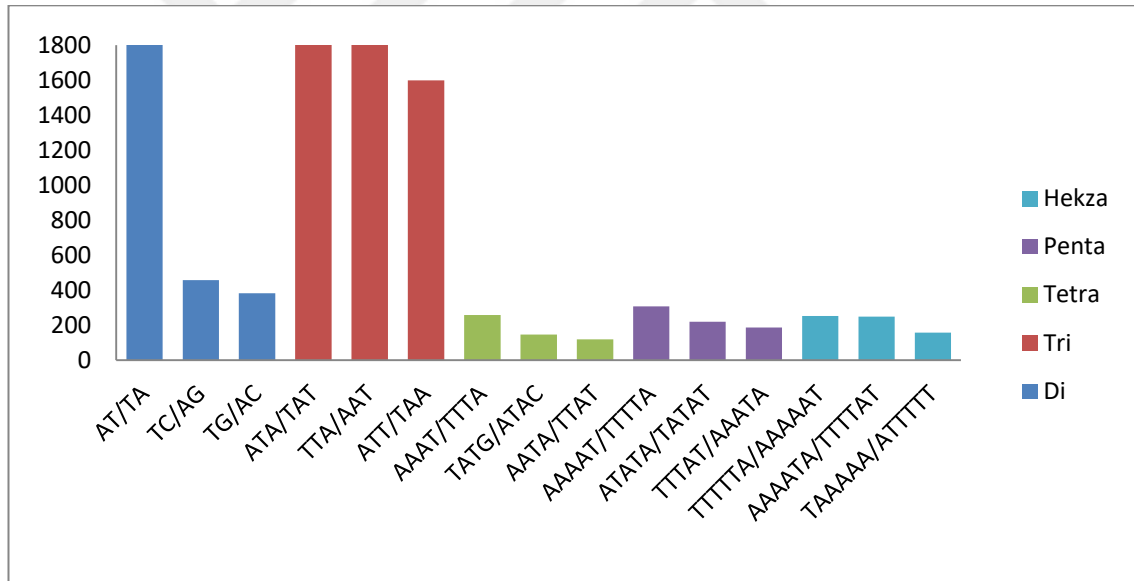
Çizelge 4.4. Oryantal pamuğa ait kromozom büyüklükleri ve SSR dağılımları

Krom	Büyüklük (Mb)	Di	Tri	Tetra	Penta	Hekza	Toplam
A01	123,41	1533	636	108	157	259	2693
A02	115,15	1269	415	111	131	202	2128
A03	139,24	1550	635	121	186	411	2903
A04	122,51	1357	476	98	146	231	2308
A05	93,16	1416	611	112	196	251	2586
A06	143,37	1705	600	121	173	263	2862
A07	106,03	1442	582	116	168	254	2562
A08	142,30	1625	577	109	180	275	2766
A09	89,40	1194	532	93	139	250	2208
A10	131,76	1677	705	142	202	340	3066
A11	137,85	1746	736	146	191	311	3130
A12	119,98	1541	597	129	180	310	2757
A13	129,59	1521	560	103	142	266	2592
Σ	1593,75	19576	7662	1509	2191	3623	34561

Oryantal pamuğun toplam genom büyüklüğü 1593,75 Mb'dir. Bu türün kromozomları incelendiğinde en fazla SSR'ın ilk sırada yer alan 1746 sayısı ile A11 kromozomunda bulunan di-nükleotitler olurken; 2. sırada yer alan tri-nükleotitlerde de

A11 kromozomunda bulunmaktadır. 3. sırada bulunan hekza-nükleotitler en fazla SSR A03 kromozomunda, 4. sırada yer alan penta-nükleotitler ise A10 kromozomunda yer almaktadır. En az bulunan tetra-nükleotitlerde ise en yoğun olarak A11 kromozomundadır. Bu sonuçlar göz önüne alındığında SSR sayısı ile kromozom büyüklüğü arasında doğrudan bir ilişki bulunmamaktadır. Toplam SSR açısından bakıldığında 3130 sayısı ile A11 kromozomunda en fazla değere sahip olmaktadır. Basit korelasyon analizine göre di-nükleotitlerde 0,80 sayısı ile kromozom büyüklükleri ve tekrarlanan motif arasında yüksek bir ilişki bulunmakta olup; tri-nükleotitlerde 0,36 değeri, tetra-nükleotitlerde 0,46 değeri ve penta-nükleotitlerde 0,30 değeri ve hekza-nükleotitlerde 0,47 ile zayıf ilişki bulunmaktadır.

Di-nükleotit tekrarları birçok türde en yaygın olanıdır; ancak kodlama bölgesinde kodlamayan bölgelere göre çok daha az sıklıkta görülür (Wang vd. 1994; Li vd. 2002). A genomuna sahip *Gossypium arboreum*'un toplam motif içeriklerine bakıldığında di-nükleotitlerden en sık görülen motif AT/TA olurken; bunu sırasıyla TC/AG ve TG/AC izlemiştir. Tri-nükleotitlerden en sık görülen motifler sırasıyla ATA/TAT, TTA/AAT, ATT/TAA motifleri olmuştur. Tetra-, penta- ve hekza-nükleotitler arasında en çok bulunan motifler ise sırasıyla AAAT/TTTA, TATG/ATAC, AATA/TTAT; AAAAT/TTTTA, ATATA/TATAT, TTTAT/AAATA; TTTTAA/AAAAAT, AAAATA/TTTTAT ve TAAAAA/ATTTTT olmuştur (Şekil 4.2).



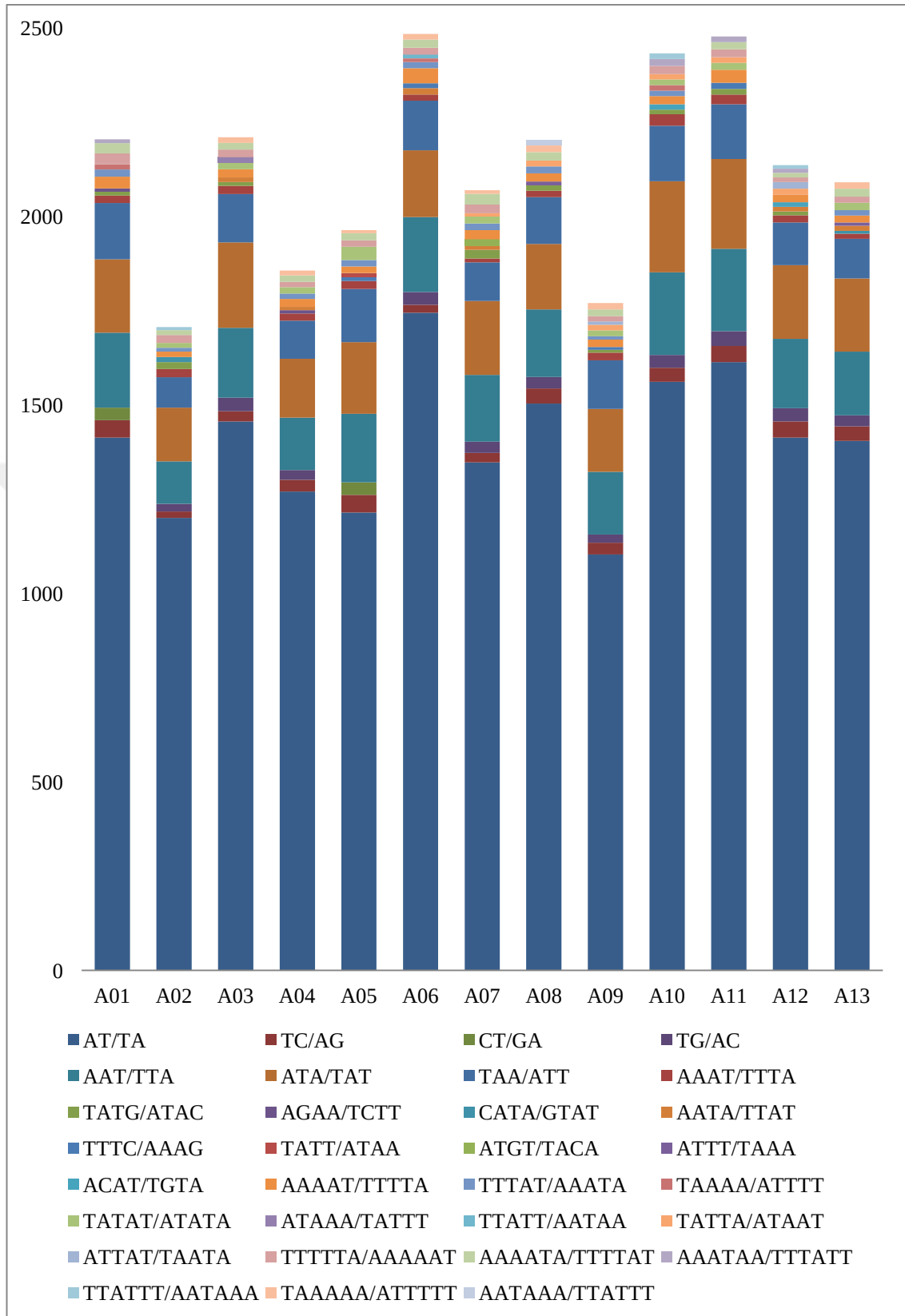
Şekil 4.2. Oryantal pamuğa ait en sık görülen ilk üç motif içerikleri

Tüm SSR verileri birlikte incelendiğinde di-nükleotit tekrarların kromozomlarda en fazla görülen motif olduğu görülmektedir. Kromozomlarda en çok görülen ilk üç di-nükleotit sırasıyla AT/TA, TC/AG ve TG/AC motifleri olmuştur. AT/TA ve TC/AG motifleri bütün kromozomlarda yer almakla birlikte TG/AC motifleri A01 ve A05 dışındaki kromozomlarda daha yoğun bulunmaktadır. Tri-nükleotitlerde en çok görülen motifler AAT/TTA, ATA/TAT ve TAA/ATT motifleri tüm kromozomlarda yer

almaktadır. Tetra-nükleotitler içinde yer alan AAAT/TTTA, TATG/ATAC ve AATA/TTAT kromozomlarda görülen en fazla motiflerdir.

AAAT/TTTA motifleri bütün kromozomlarda yer alıp TATG/ATAC motifleri A01, A02, A03, A07, A08, A09, A10, A11 ve A12 kromozomlarında yaygın iken; AATA/TTAT motifleri ise A03, A04, A06, A07, A12, A13 kromozomlarında bulunmaktadır. AAAAT/TTTTA, TTTAT/AAATA, TATAT/ATATA motifleri penta-nükleotitler içinde olmakla birlikte AAAAT/TTTTA motifi tüm kromozomlarda bulunup TTTAT/AAATA motifi A01, A02, A04, A05, A06, A07, A08, A09, A10, A13 kromozomunda; TATAT/ATATA motifi ise A02, A03, A04, A05, A07, A09, A10, A11 ve A13 kromozomunda yer almaktadır. 6'lı motif olan hekza-nükleotitler arasında en fazla bulunan ilk üç motif ise sırasıyla AAAATA/TTTTAT, TTTTAA/AAAAAT, TAAAAA/ATTTTT olmuştur. AAAATA/TTTTAT motifi A10, TTTTAA/AAAAAT A08 dışındaki bütün kromozomlarda bulunurken; TAAAAA/ATTTTT motifi ise A03, A04, A05, A06, A07, A08, A09 ve A13 kromozomlarında yer almaktadır (Şekil 4.3).





Şekil 4.3. Oryantal pamuğa ait kromozom düzeyinde en sık görülen SSR motif içerikleri

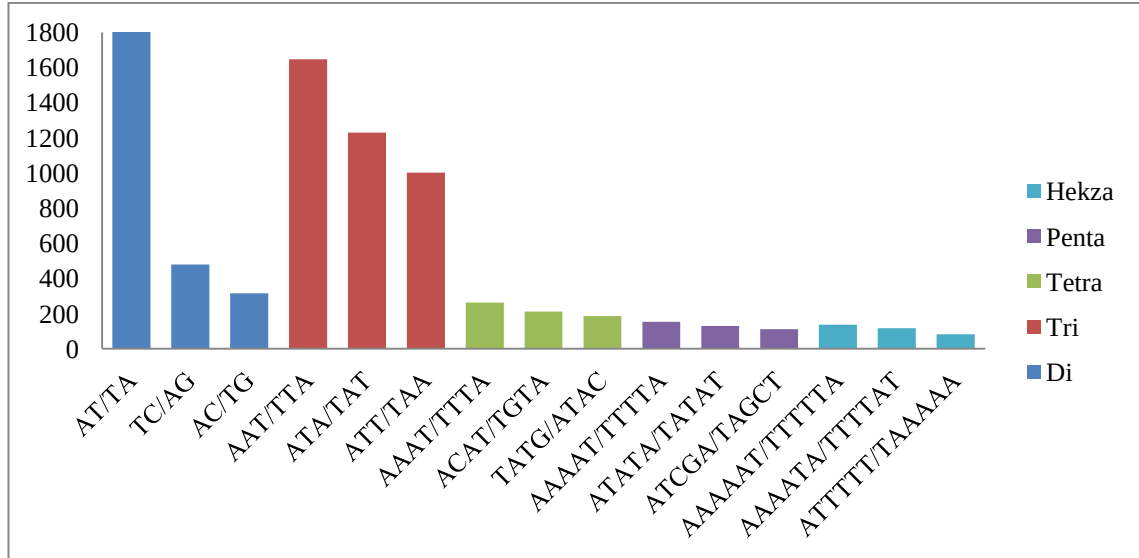
A genomuna sahip *Gossypium arboreum*'un (Ağaç pamuk) genomu (1.687 Mb), D genomuna sahip *Gossypium raimondii*'nin (Peru pamuğu, 880 Mb) iki katı büyüklüğündedir (Hendrix ve Stewart 2005). Çizelge 4.5'te görüleceği üzere Peru pamuğu için di-nükleotitler toplam 8979 (SSR'ların %43'ü) sayısı ile 1. sırada yer alırken; bunu sırasıyla tri-nükleotitler 5435 (SSR'ların %26'sı), hekza-nükleotitler 2837 (SSR'ların %14'ü), penta-nükleotitler 1755 (SSR'ların %8'i) ve son olarak tetra-nükleotitler 1693 (SSR'ların %8'i) izlemektedir.

Çizelge 4.5. Peru pamuğuna ait kromozom büyüklükleri ve SSR dağılımları

Krom	Büyüklük (Mb)	Di	Tri	Tetra	Penta	Hekza	Toplam
D01	52,42	660	395	99	127	193	1474
D02	48,27	558	291	103	118	199	1269
D03	63,02	640	410	140	150	207	1547
D04	50,65	639	341	187	122	190	1479
D05	57,14	824	440	106	145	214	1729
D06	60,40	685	431	169	123	268	1676
D07	53,50	699	425	126	138	209	1597
D08	62,61	768	453	125	128	247	1721
D09	49,93	625	425	102	111	208	1471
D10	61,94	722	421	141	133	233	1650
D11	63,19	764	482	135	159	220	1760
D12	57,72	720	437	144	146	238	1685
D13	58,75	675	484	116	155	211	1641
Σ	739,54	8979	5435	1693	1755	2837	20699

Peru pamuğunun toplam genom büyüklüğü 739,54 Mb olarak belirlenmiştir. Peru pamuğunun kromozomları incelendiğinde en fazla SSR'ın di-nükleotitlerde D05 kromozomunda, tri-nükleotitlerde D13 kromozomunda, tetra-nükleotitlerde D04 kromozomunda, penta-nükleotitlerde D11 kromozomunda ve hekza-nükleotitlerde ise D06 kromozomunda bulunmaktadır. Toplam SSR ise en fazla D11 kromozomunda yer almaktadır. Basit korelasyon analizine göre kromozom büyüklükleri ve tekrarlanan motif arasında di-nükleotitlerde 0,62 değeri, penta-nükleotitlerde 0,64 değeri ve hekza-nükleotitlerde 0,63 değeri ile orta düzeyde ilişki, tri-nükleotitlerde 0,70 değeri ile yüksek ilişki, tetra-nükleotitlerde ise 0,27 değeri ile çok zayıf ilişki olduğu anlaşılmaktadır.

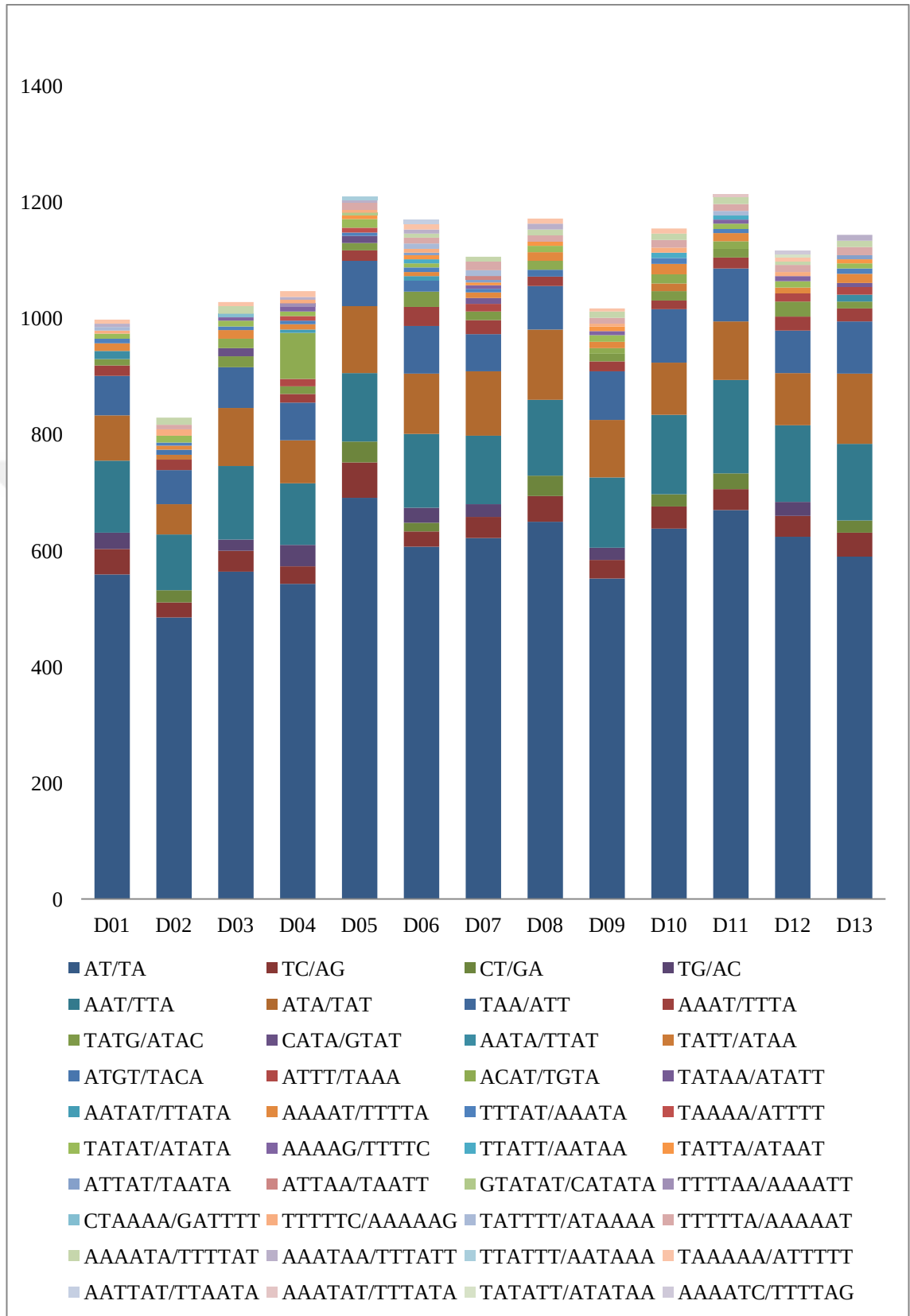
Peru pamuğuna ait toplam SSR motif içerikleri ele alındığında di-nükleotitlerden en sık görülen motif AT/TA olurken; bunu sırasıyla TC/AG ve TG/AC izlemiştir. Tri-nükleotitlerden en sık görülen motifler sırasıyla TTA/AAT, ATA/TAT ve ATT/TAA motifi olmuştur. Tetra-, penta- ve hekza-nükleotitler arasında en çok bulunan motifler ise sırasıyla AAAT/TTTA, ACAT/TGTA, TATG/ATAC; AAAAT/TTTTA, ATATA/TATAT, ATCGA/TAGCT olurken; hekza-nükleotitler içinde yer alan TTTTAA/AAAAAT, AAAATA/TTTTAT ve TAAAAA/ATTTTT olmuştur (Şekil 4.4).



Şekil 4.4. Peru pamuğuna ait en sık görülen ilk üç SSR motif içerikleri

Peru pamuğunun diploit genomunda her bir kromozomdaki SSR'lar karakterize edilmiştir. Son zamanlarda, *G. raimondii*'nin tüm genom dizisinden 136.345 mikrosatellit ve 112.177 primer çifti tasarlanmış olmakla birlikte SSR'ların madenciliği için yeterli derinliğe inilememiştir (Wang vd. 2012; Zou vd. 2012). Oysaki genom boyunca dağıtılan çok geniş bir dizi kromozoma özgü SSR'ın mevcudiyeti, pamuk araştırma ve yetiştirme konularında çalışan araştırmacılara fayda sağlamakla kalmayıp SSR fonksiyonları konularına da katkı sağlayabilecektir.

Peru pamuğu verileri incelendiğinde di-nükleotitler kromozomlarda ilk sırada yer alan motifler olduğu görülmektedir. Kromozomlarda en çok görülen ilk üç di-nükleotit sırasıyla AT/TA, TC/AG ve TG/AC motifleri olmuştur. AT/TA ve TC/AG motifleri bütün kromozomlarda yer almakta olup TG/AC motifleri D01, D03, D04, D06, D07, D09 ve D12 kromozomlarında bulunmaktadır. Tri-nükleotitlerde en çok görülen motifler AAT/TTA, ATA/TAT ve TAA/ATT motifleri tüm kromozomlarda yer almaktadır. Tetra-nükleotitler içinde yer alan AAAT/TTTA, TATG/ATAC ve ACAT/TGTA kromozomlarda görülen en fazla motiflerdir. AAAT/TTTA motifleri D03 haricindeki kromozomlarda yer alırken; TATG/ATAC motifleri D02 ve D08 dışındaki kromozomlarda olup ACAT/TGTA motifleri ise D03, D04, D08, D09, D10, D11 kromozomlarında bulunmaktadır. AAAAT/TTTTA, TTTAT/AAATA, TATAT/ATATA motifleri penta-nükleotitler içinde olmakla birlikte AAAAT/TTTTA motifi D05 kromozomunda, TTTAT/AAATA motifi D08, D09 ve D12 kromozomlarında, TATAT/ATATA motifi ise D07 ve D10 dışındaki bütün kromozomlarda yer almaktadır. 6'lı motif olan hekza-nükleotitler arasında en fazla bulunan ilk üç motif ise sırasıyla ATTTTT/TAAAAA, AAAATA/TTTTAT ve TAAAAA/ATTTTT olmuştur. ATTTTT/TAAAAA motifi D01, D03, D04 kromozomları hariç bütün kromozomlarda, AAAATA/TTTTAT motifi D01, D04, D05 dışındaki bütün kromozomlarda bulunurken; TAAAAA/ATTTTT motifi ise D01, D03, D04, D06, D08, D09, D10 ve D12 kromozomlarında yer almaktadır (Şekil 4.5).



Şekil 4.5. Peru pamuğuna ait kromozom düzeyinde en sık görülen SSR motif içerikleri

Di-nükleotit tekrarları arasında AT/TA motifi her iki genomda da çarpıcı biçimde temsil edilmiştir. Ağaç pamuktaki di-nükleotit motiflerinin %96'sını, Peru pamukta di-nükleotit motiflerinin ise %91'ini oluşturmaktadır. AT/TA motifi ayrıca A ve D genomunda en yoğun bulunan motif olmuştur. Tri-nükleotit tekrarları arasında *G. arboreum*'da en fazla bulunan ATA/TAT motifi %39'unu oluştururken, *G. raimondii*'de AAT/TTA motifi %42'sini kapsamaktadır. Tetra-nükleotit tekrarları arasında AAAT/TTTA motifi her iki genomda en bol bulunan motifler olmuştur. A genomu olan Ağaç pamukta %49'unu, D genomu olan Peru pamukta ise %40'ını temsil etmektedir. AAAAT/TTTTA, penta-nükleotitler içinde yer alan en fazla motif içeriği olup *G. arboreum*'da %43, *G. raimondii*'de %39'unu oluşturmaktadır. Hekza-nükleotitler arasında AT açısından zengin motifler her iki genomda da baskın olup AAAAAT/TTTTTA motifi Ağaç pamuğun %38'ini, Peru pamuğun ise %41'ini oluşturmaktadır.

Çizelge 4.6'da görüldüğü üzere Upland pamuğa ait A alt genomu için toplam genom büyüklüğü 1446,29 Mb olup tri-nükleotitler toplam 5009 sayısı ile (SSR'ların %34'ü) 1. sırada bulunmaktadır. Bunu sırasıyla hekza-nükleotitler 3919 (SSR'ların %27'si), penta-nükleotitler 2136 (SSR'ların %14'ü), tetra-nükleotitler 1950 (SSR'ların %13'ü) ve di-nükleotitler 1737 sayısı (SSR'ların %12'si) izlemektedir.

Çizelge 4.6. Upland pamuğa ait A alt genomu kromozom büyüklükleri ve SSR dağılımları

Krom	Büyüklük (Mb)	Di	Tri	Tetra	Penta	Hekza	Toplam
A01	119,76	113	362	137	141	300	1053
A02	108,14	100	300	136	137	284	957
A03	113,69	134	376	152	167	282	1111
A04	89,18	73	288	107	120	212	800
A05	111,10	184	480	178	236	375	1453
A06	129,20	140	379	147	196	308	1170
A07	98,90	130	380	143	151	289	1093
A08	127,50	166	394	135	174	350	1219
A09	85,34	118	321	117	139	277	972
A10	118,18	126	466	174	188	332	1286
A11	124,18	156	508	204	176	306	1350
A12	109,47	145	391	173	165	302	1176
A13	111,65	152	364	147	146	302	1111
Σ	1446,29	1737	5009	1950	2136	3919	14751

Upland pamuğa ait A alt genomunun kromozomları incelendiğinde en fazla SSR'ın di-nükleotitlerde A05 kromozomunda, tri- ve tetra-nükleotitlerde A11 kromozomunda, penta- ve hekza-nükleotitlerde A05 kromozomunda bulunmaktadır. Toplam SSR göz önünde alındığında ise en fazla A05 kromozomunda yer alırken; A03 ve A13 kromozomlarında 1111 sayısı ile eşit değerler çıkmaktadır. Basit korelasyon analizine göre kromozom büyüklükleri ve tekrarlanan motif arasında di-nükleotitlerde

0,54; tri-nükleotitlerde 0,56; tetra- ve penta-nükleotitlerde 0,54; hekza-nükleotitlerde 0,62 değeri ile orta ilişki bulunmaktadır.

Çizelge 4.7’de ise Upland pamuğa ait D alt genomu için toplam genom büyüklüğü 836,33 Mb olup tri-nükleotitler 4574 sayısı ile (SSR’ların %35’i) 1. sırada yer almaktadır. Bunu sırasıyla hekza-nükleotitler 3047 (SSR’ların %23’ü), tetra-nükleotitler 1980 (SSR’ların %15’i), penta-nükleotitler 1691 (SSR’ların %13’ü) ve di-nükleotitler 1674 sayısı (SSR’ların %13’ü) izlemektedir.

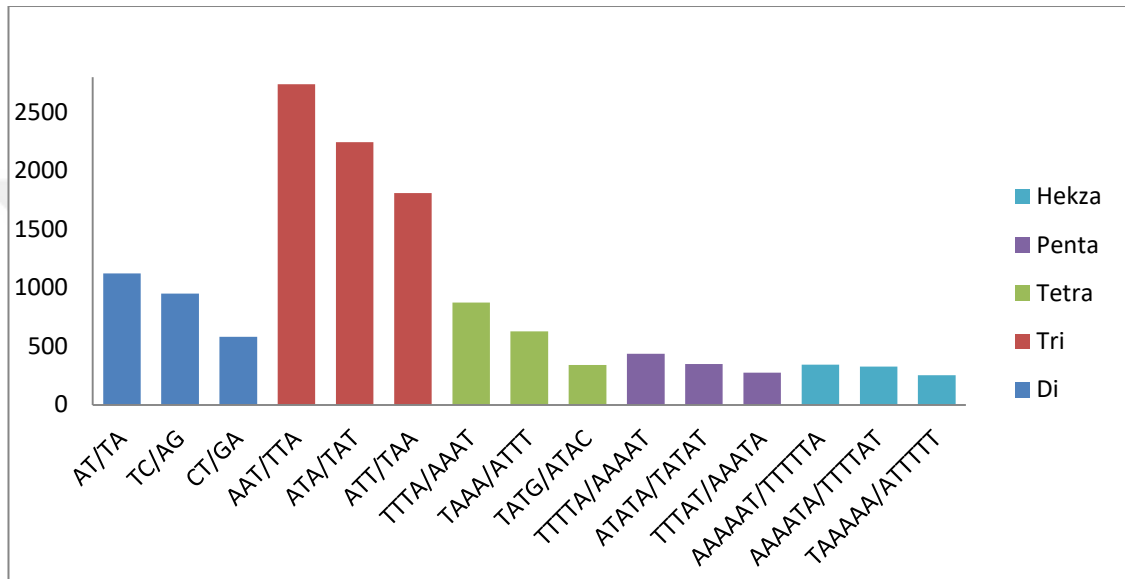
Çizelge 4.7. Upland pamuğa ait D alt genomu kromozom büyüklükleri ve SSR dağılımları

Krom	Büyüklük (Mb)	Di	Tri	Tetra	Penta	Hekza	Toplam
D01	65,21	140	390	159	134	250	1073
D02	72,19	120	314	181	142	244	1001
D03	54,96	101	256	120	102	196	775
D04	58,23	117	306	119	106	187	835
D05	66,48	197	443	168	146	256	1210
D06	66,68	110	329	132	146	261	978
D07	59,44	117	349	149	127	219	961
D08	69,43	139	373	188	144	245	1089
D09	54,44	125	313	105	110	239	892
D10	68,09	126	401	165	133	238	1063
D11	72,82	142	412	155	140	273	1122
D12	63,26	130	327	188	133	248	1026
D13	65,10	110	361	151	128	191	941
Σ	836,33	1674	4574	1980	1691	3047	12966

Upland pamuğa ait D alt genomunun kromozomları incelendiğinde ise en fazla SSR’ın di- ve tri-nükleotitlerde D05 kromozomunda, tetra-nükleotitlerde 188 sayısı ile hem D08 hem de D12 kromozomunda, penta-nükleotitlerde de 146 sayısı ile D05 ve D06 kromozomunda en fazla bulunurken; hekza-nükleotitlerde D11 kromozomunda 273 sayısı ile ilk sırada yer almaktadır. Basit korelasyon analizi incelendiğinde kromozom büyüklükleri ve tekrarlanan motif arasında di-nükleotitlerde 0,35 ile zayıf ilişki olup, tri-nükleotitlerde 0,62 değeri ile orta ilişki, tetra- ile penta-nükleotitlerde ise sırasıyla 0,73 ve 0,87 bulunarak çok yüksek düzeyde ilişki, hekza-nükleotitlerde ise 0,61 ile orta ilişkili olduğu tespit edilmiştir.

Upland pamukta A ve D alt genomları birlikte incelendiğinde toplam genom büyüklüğü 2282,62 Mb’dır. Bu türe ait en fazla SSR 9583 sayısı ile (SSR’ların %35’i) tri-nükleotitler olurken; bunu sırasıyla hekza-nükleotitler 6966 (SSR’ların %25’i), tetra-nükleotitler 3930 (SSR’ların %14’ü), penta-nükleotitler 3827 (SSR’ların %14’ü) ve di-nükleotitler 3411 (SSR’ların %12’si) izlemektedir. Upland pamuğunun kromozomları incelendiğinde ise en fazla SSR’ın di-nükleotitlerde D05 kromozomunda, tri- ve tetra-nükleotitlerde A11 kromozomunda, penta- ve hekza-nükleotitlerde ise A05 kromozomunda bulunmaktadır. Basit korelasyon analizine göre kromozom büyüklükleri

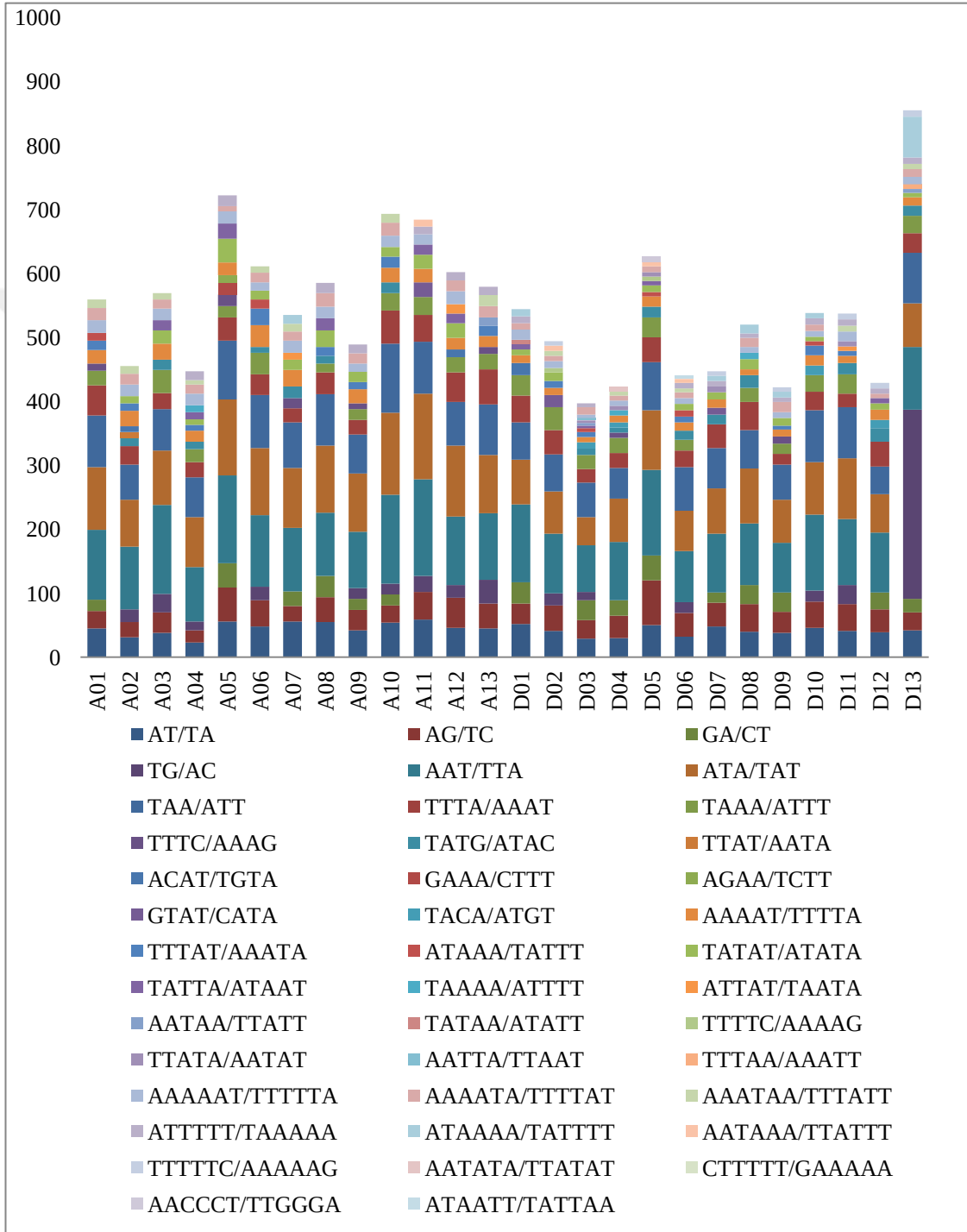
ve tekrarlanan motif arasında di-nükleotitlerde ve tri-nükleotitlerde sırasıyla 0,27 ve 0,47 ile zayıf ilişki, tetra-nükleotitlerde 0,18 ile çok zayıf ilişki, penta- ile hekza-nükleotitlerde ise sırasıyla 0,74 ve 0,82 olup yüksek ilişki bulunmaktadır. Upland pamuğuna ait toplam SSR motif içerikleri ele alındığında di-nükleotitlerden en sık görülen motif AT/TA olmakla beraber bunu sırasıyla TC/AG ve CT/GA izlemiştir. Tri-nükleotitlerden en sık görülen motifler sırasıyla AAT/TTA, ATA/TAT ve ATT/TAA motifi olmuştur. Tetra-, penta- ve hekza-nükleotitler arasında en çok bulunan motifler ise sırasıyla AAAT/TTTA, TAAA/ATTT, TATG/ATAC; AAAAT/TTTTA, ATATA/TATAT, TTTAT/AAATA olurken; hekza-nükleotitler içinde yer alan TTTTAA/AAAAAT, AAAATA/TTTTAT ve TAAAAA/ATTTTT'dir.



Şekil 4.6. Upland pamuğuna ait en sık görülen ilk üç motif içerikleri

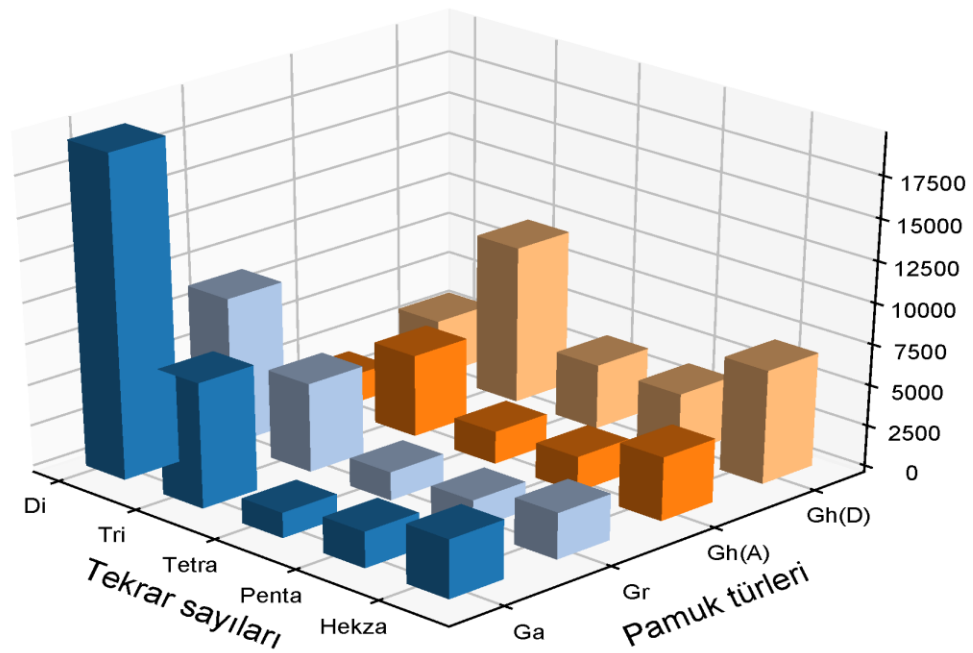
Upland pamuğuna ait veriler incelendiğinde kromozomlarda en çok görülen ilk üç di-nükleotit sırasıyla AT/TA, AG/TC ve GA/CT motifleri olmuştur. AT/TA ve AG/TC motifleri bütün kromozomlarda yer almakla birlikte GA/CT motifleri A alt genomunda A01, A05, A07, A08, A09 ve A10 kromozomları ile D alt genomunda D01, D03, D04, D05, D07, D08, D09, D12 ve D13 kromozomlarında bulunmaktadır. Tri-nükleotitlerde en çok görülen motifler AAT/TTA, ATA/TAT ve TAA/ATT motifleri tüm kromozomlarda yer almaktadır. Tetra-nükleotitler içinde yer alan TTTA/AAAT, TAAA/ATTT ve TATG/ATAC kromozomlarda görülen en fazla motiflerdir. TTTA/AAAT motifleri bütün kromozomlarda yer alırken; TAAA/ATTT motifleri A02, A07, D07 ve D12 dışındaki tüm kromozomlarda yer almakla birlikte, TATG/ATAC motifleri ise A02, A03, A04, A06, A07, A08, A10, D03, D04, D05, D06, D07, D08, D11, D12 ve D13 kromozomlarında bulunmaktadır. AAAAT/TTTTA, TATAT/ATATA ve TTTAT/AAATA motifleri penta-nükleotitler içinde olmakla birlikte AAAAT/TTTTA motifi A08 haricindeki kromozomlarda bulunup TATAT/ATATA motifi A alt genomunda A01 ve A13; D alt genomunda ise D03, D04 ve D11 dışındaki tüm kromozomlarda yer almaktadır. TTTAT/AAATA motifi ise A01, A02, A04, A06, A08, A09, A10, A13, D02, D03, D06, D09, D10 ve D11

kromozomlarında bulunmaktadır. Altılı motif olan hekza-nükleotitler arasında en fazla bulunan ilk üç motif ise sırasıyla AAAAAT/TTTTTA, AAAATA/TTTTAT, ATTTTT/TAAAAA olmuştur. AAAAAT/TTTTTA motifi A13, D05, D07, D12; AAAATA/TTTTAT A11, D07, D11 dışındaki bütün kromozomlarda bulunurken, TAAAAA/ATTTTT motifi ise A04, A05, A08, A09, A11, A12, A13, D01, D03, D06, D07, D08, D09, D10, D11, D12 ve D13 kromozomlarında bulunmaktadır (Şekil 4.7).



Şekil 4.7. Upland pamuğa ait kromozom düzeyinde en sık görülen SSR motif içerikleri

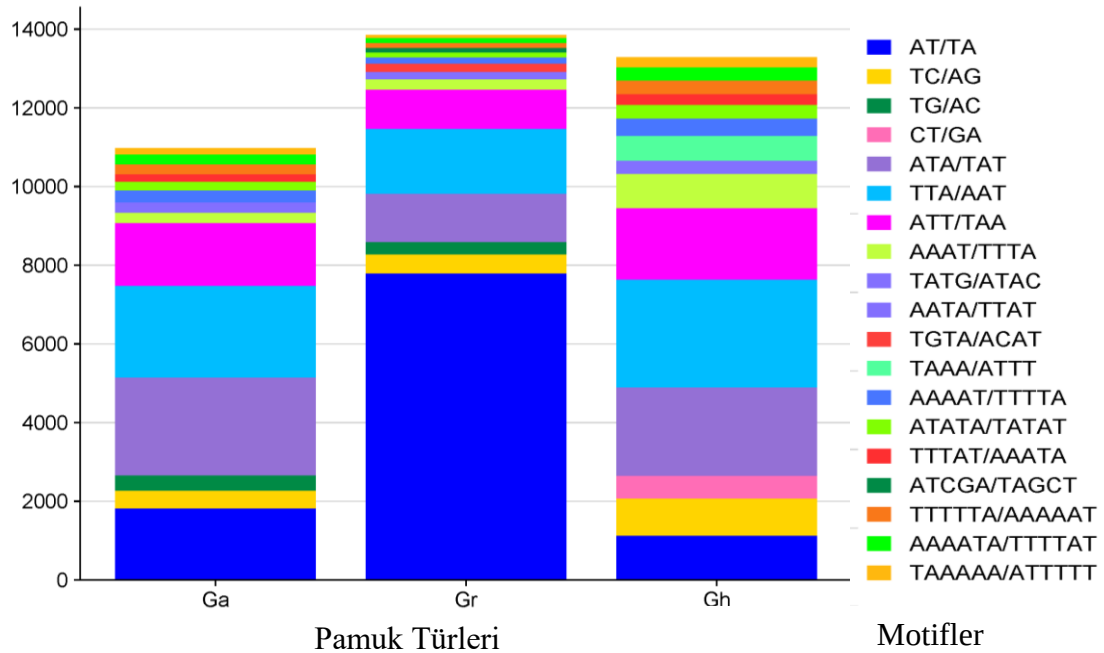
SSR'ların di-nükleotit tipleri, genom ve alt genom türleri arasında farklılık göstermektedir. Di-nükleotitlerde en fazla SSR sırasıyla Ağaç pamuk, Peru pamuğu, A alt genomuna sahip Upland pamuğu ve D alt genomuna sahip Upland pamuğu yer almaktadır. Tri-nükleotitlerde diploit türler kendi arasında karşılaştırıldığında A genomuna sahip *G. arboreum*'un ve tetraploit tür olan *G. hirsutum*'un A alt genomunda yüksek sayıda tri-nükleotit içermektedir. Türlerde daha yüksek sayıda tri-nükleotit tekrarların fazla olması tri-nükleotitlerin genom ifadesindeki fonksiyonel rollerini olabileceğini gösterebilir. *G. hirsutum*'un tetra-nükleotitleri, D alt genomunda baskınken, *G. arboreum*'un A genomunda en düşük sayıda tetra-nükleotit içermektedir. Bu gözlemlere dayanarak pamuk türlerinin genomlarının ve alt genomlarının farklı sayıda tetra-nükleotit içerdiği sonucuna varılabilir. *G. hirsutum*'un D alt genomunda daha fazla hekza-nükleotit bulunurken; D genomuna sahip *G. raimondii* en az hekza-nükleotitler içermektedir (Şekil 4.8). Mikrosatellit tekrar sayıları tetraploit pamuk türü olan *G. hirsutum* da tri-, tetra-, penta- ve hekza-nükleotitlerde daha yüksek iken, di-nükleotit tekrarları diploit türlerde daha yüksektir.



Şekil 4.8. Üç pamuk türünün SSR dağılımları

G. arboreum L., *G. raimondii* L. ve *G. hirsutum* L.'un birlikte incelendiğinde genomik dizilerindeki mikrosatellitlerin dağılımları tri-, tetra-, penta- ve hekza-nükleotit tekrarlarının çok benzer oranlara karşılık geldiğini, di-nükleotit tekrarlarının ise nispeten bunlardan farklı olduğunu göstermiştir. Evrimsel zaman çizelgeleri boyunca, genomdaki farklı SSR tekrarlarının frekanslarında önemli farklılıklara yol açabilir.

Çalışmada di-nükleotit SSR'ların bolluğunun türler arasında farklılık gösterdiği görülmüştür (Gao vd. 2009). Memelilerde ve *Drosophila*'da GT/CA en yaygın di-nükleotit iken; bitkilerde AT/TA di-nükleotitleri yaygın olanıdır (Langercrantz vd. 1993). Şekil 4.9'da görüleceği üzere AT/TA motifi en yaygın di-nükleotit olup en fazla Ağaç pamukta yer alırken; en az Upland pamukta bulunmaktadır. TC/AG motifi en fazla Upland pamukta bulunup, en az Ağaç pamukta yer almaktadır. TG/AC motifi Ağaç pamuk ve Peru pamukta; CT/GA motifi ise sadece Upland pamukta yer almaktadır. Tri-nükleotitler arasında ATA/TAT motifi Ağaç pamukta en fazla bulunup, TTA/AAT motifi Upland pamuk ve Peru pamukta en yüksektir. AAAT/TTTA ve TATG/ATAC tetra-nükleotit motifi en fazla sırasıyla *G. hirsutum*, *G. raimondii* ve *G. arboreum*'da yer alırken, AATA/TTAT motifi sadece *G. arboreum*'da, TGTA/ACAT motifi *G. raimondii*'de ve TAAA/ATTT motifi ise yalnız *G. hirsutum*'da bulunmaktadır. Penta-nükleotitler arasında yer alan AAAAT/TTTTA ve ATATA/TATAT motifi en fazla sırasıyla *G. hirsutum*, *G. arboreum* ve *G. raimondii*'de; TTTAT/AAATA motifi *G. hirsutum* ve *G. arboreum*'da; ATCGA/TAGCT motifi ise sadece *G. raimondii*'de bulunmaktadır. Hekza-nükleotitler içinde bulunan TTTTAA/AAAAAT, ATATA/TATAT, TAAAAA/ATTTTT motifleri en fazla sırasıyla *G. hirsutum*, *G. arboreum* ve *G. raimondii*'de yer almaktadır (Şekil 4.9).



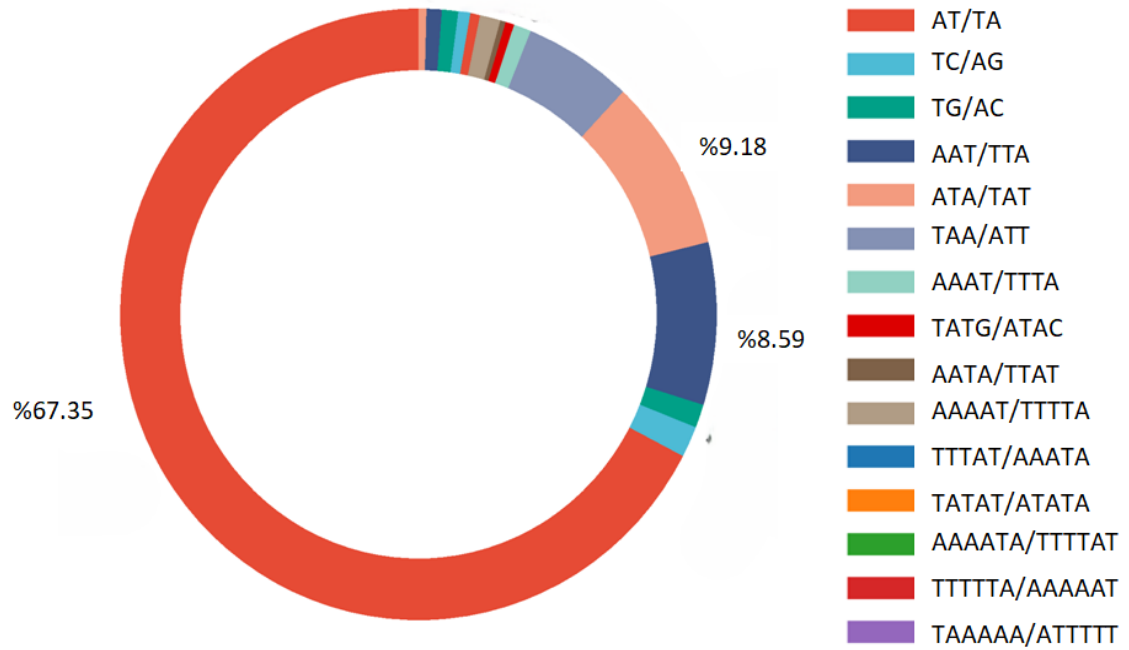
Şekil 4.9. Üç pamuk türünün SSR motif içerikleri

Lu vd. (2015) *Gossypium arboreum* ve *Gossypium raimondii* arasında yaptıkları çalışmada penta-nükleotit tekrarları en yaygın tekrar olduğunu, di-nükleotitlerde AT motifi, tri-nükleotitlerde AAT, tetra-nükleotitlerde AAAT, penta-nükleotitlerde AAAAT, hekza-nükleotitler arasında ise en yaygın AAAAAT motifini bulmuşlardır. Bu tez çalışmasında ise en sık bulunan tekrar di-nükleotitler iken; en yaygın motifler di-

nükleotitlerde AT/TA, tri-nükleotitlerde AAT/TTA, tetra-nükleotitlerde AAAT/TTTA, penta-nükleotitlerde AAAAT/TTTTA ve hekza-nükleotitlerde AAAAAT/TTTTTA bulunmuştur. Her iki çalışma arasındaki benzerliğin bulunması yapılan bu tez çalışmasının daha önceki çalışmalarla benzer sonuçlar ortaya koyduğunu göstermektedir.

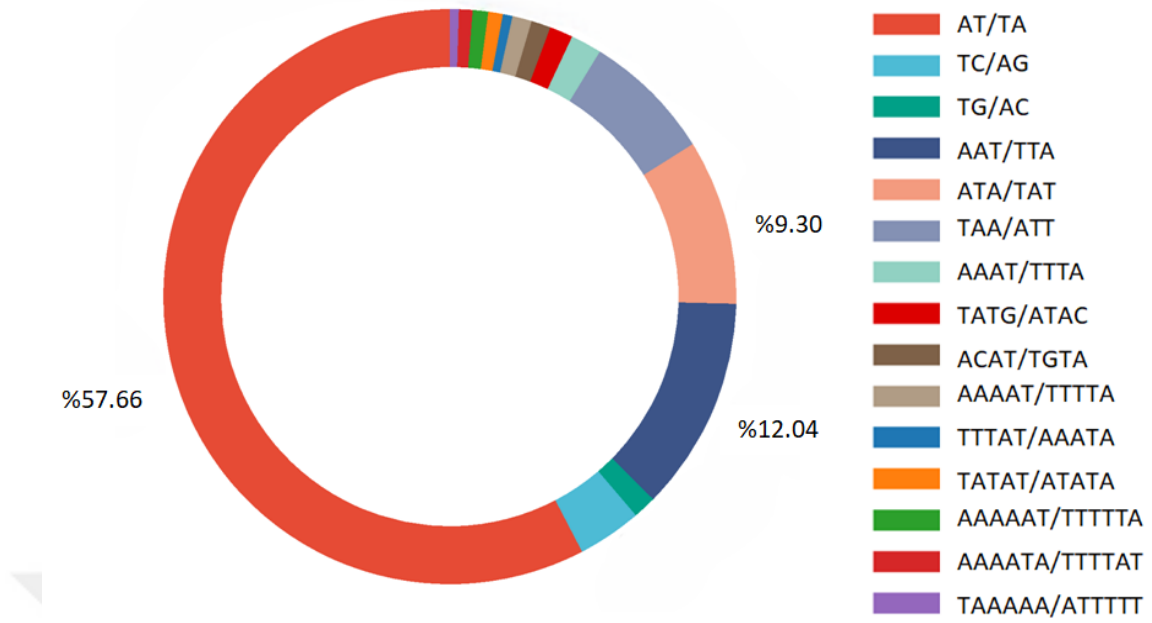
4.1.3. Genomlarda SSR Motif Sıklıkları

Şekil 4.10'da Oryantal pamuk A genomunda bulunan motiflerin sıklığı incelenmiştir. SSR sıklığı ile SSR sayıları arasında benzerlik beklenen bir durum olmakla birlikte tüm genom ve kromozom düzeyinde farklılıklar ortaya koyabilir ve böylelikle çalışılan pamuk türlerinin SSR içeriklerindeki benzerlik ve farklılıkları daha net olarak ortaya konulabilmesi hedeflenmiştir. Tüm SSR'lar arasında di-nükleotit motifi olan AT/TA motifi %67,35 sıklık ile en fazla olan motif olup bunu sırasıyla tri-nükleotit motifleri olan %9,18 ile ATA/TAT ve %8,59 AAT/TTA motifi izlemiştir. Bu türe ait SSR dizilerinin ezici çokluğunun di-nükleotit tekrarlara ait olduğu görülmektedir.



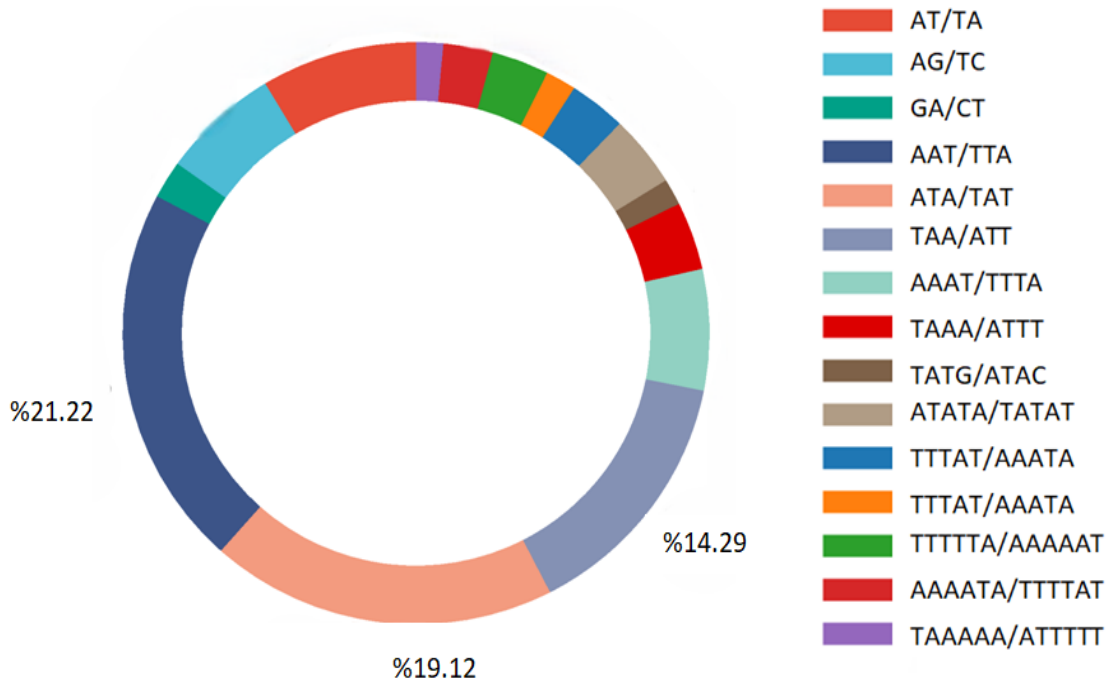
Şekil 4.10. Oryantal pamuk A genomunda motif sıklıkları

Şekil 4.11'de Peru pamuğun D genomunda bulunan motiflerin sıklığı ele alınmıştır. Tüm motifler arasında di-nükleotit motifi olan AT/TA motifi %57,66 sıklık ile en fazla olan motif olup bunu sırasıyla tri-nükleotit motifleri olan %12,04 ile AAT/TTA ve %9,30 ile ATA/TAT motifi izlemiştir.



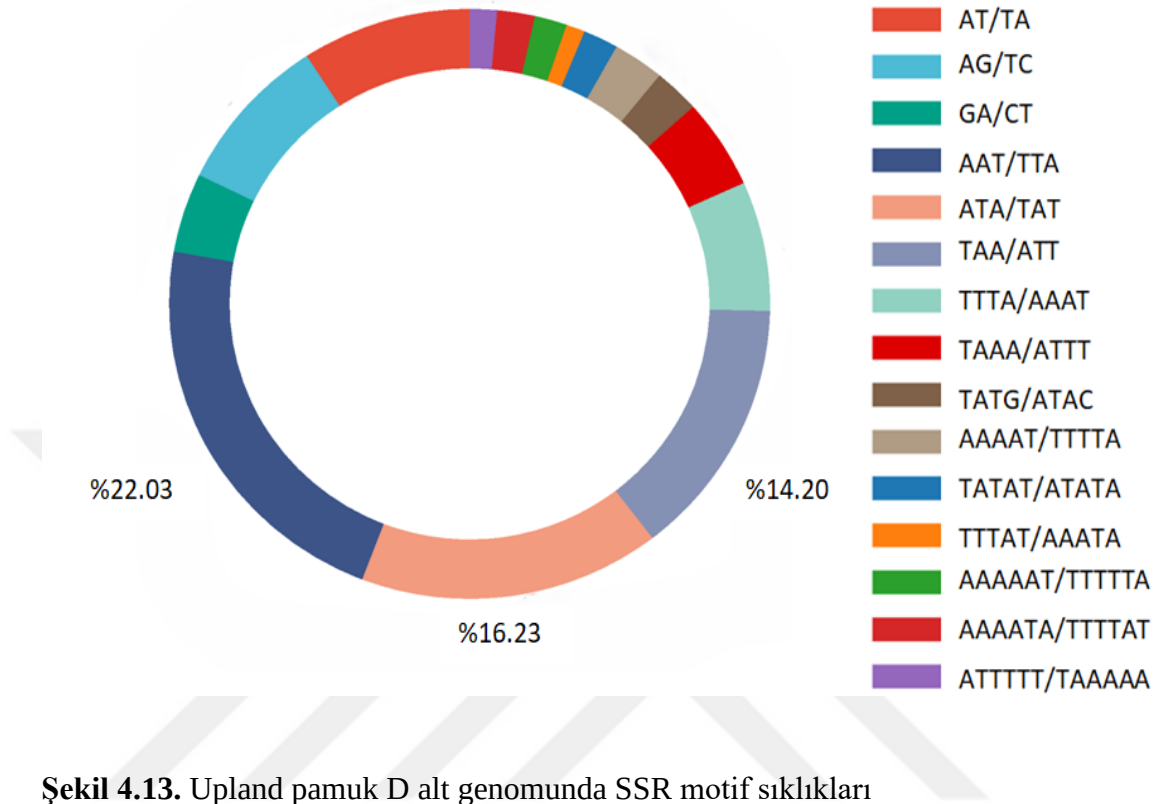
Şekil 4.11. Peru pamuk D genomunda SSR motif sıklıkları

Şekil 4.12’de Upland pamuğa ait A alt genomu için motif sıklıkları incelenmiştir. Tüm motifler arasında tri-nükleotit motifi olan AAT/TTA motifi %1.01 ile en sık görülen motiftir. Bunu sırasıyla %0,9 ile ATA/TAT ve %0,7 ile TAA/ATT motifi izlemiştir.



Şekil 4.12. Upland pamuk A alt genomunda SSR motif sıklıkları

Şekil 4.13'te Upland pamuğa ait D alt genomu için motif sıklıkları incelenmiştir. Tüm motifler arasında tri-nükleotit motifi olan AAT/TTA motifi %1.5 ile en sık görülen motiftir. Bunu sırasıyla %1.1 ile ATA/TAT ve %0.1 ile TAA/ATT motifi izlemiştir.



Şekil 4.13. Upland pamuk D alt genomunda SSR motif sıklıkları

4.1.4. Gen Elementleri Dizilerinde SSR Dağılımları

Basit dizi tekrarları (SSR'lar), ökaryotik genomlarda hem kodlayan hem de kodlamayan bölgelerde bulunan genom alanlarıdır. Her ne kadar SSR'lar SNP ile karşılaştırıldığında genellikle düşük polimorfizm ile karakterize edilse de bunların DNA-yan dizileri, ilgilenilen özellikleri kontrol eden genlerle ilişkili olmaları yanında kodominant özellikte olmaları, homozigot dominant ve homozigot resesif bireyleri ve hibritleri ayırt edebildikleri için genetik araştırmalara ve üremeye yardımcı olabilecek yararlı bir DNA markır kaynaklarıdır (Karaca vd. 2005).

Evrimsel açıdan bakıldığında SSR'lar, kodlama bölgelerinde düşük frekansta dağılmış olmalarına rağmen, bitkiler de dahil olmak üzere tüm ökaryotik genomlarda mevcut olmaları nedeniyle ilgi çekicidir. Ayrıca kodlamayan bölgelerde bulunan SSR'lar nötr iken; kodlama bölgelerinde bulunanlar nötr değildir (Powell vd. 1996; Varshney vd. 2005). Genik SSR'lar ayrıca popülasyon yapısını tahmin etmek ve ilişki analizinde yaygın olarak kullanılır; çünkü ilgilenilen belirli özellikler için fenotip değişkenliğinden sorumlu genlerin spesifik allelik varyantlarını tanımlayabilirler. Spesifik allelik varyantların tanımlanması, markır destekli seleksiyonla spesifik özelliklerin seçilmesini sağlayan yetiştirme programlarını destekleyebilen moleküler araçların geliştirilmesi için temel bir ön koşuldur.

Tüm hücrelerde bulunan bir protein sınıfı olan transkripsiyon faktörleri (TF'ler), hedef genlerin yakınındaki DNA'ya bağlanarak hedef gen transkripsiyonunu yönetir (Wray vd. 2003). Transkripsiyon faktörleri, kromatin yeniden modelleme ve proteinleri değiştirme gibi bazı transkripsiyonel düzenleyicilerle etkileşime girerek, RNA polimerazın hedef genlerin destekleyicilerine yaklaşmasını kolaylaştırır veya engeller. Dolayısıyla ifadeyi etkinleştirir veya baskılar. Bir gen içindeki mikrosatellitlerin yeri ve türü, onun işlevsel rolünü belirlemektedir. 3'-UTR'lerdeki mikrosatellitler gen susturulması ve transcription slippage (transkripsiyon kayması)'de rol oynar. 5'-UTR'lerdeki mikrosatellitler gen transkripsiyonu veya düzenlenmesi üzerinde etkilere sahiptir ve intronlardaki mikrosatellitler gen transkripsiyonu, düzenlenmesi, mesajcı RNA (mRNA) eklenmesi, gen susturulması üzerinde etkilere sahiptir. Kodlama bölgelerindeki mikrosatellitler genleri inaktif veya aktif hale getirebilir veya proteinleri kesebilir (Ince vd. 2010). Bu nedenlerden dolayı bu tez çalışmasında Oryantal pamuk, Peru pamuk ve Upland pamuktaki gen elementleri (upstream, 5'-UTR, ekson, intron, 3'UTR, gen, downstream) artı ve eksi sarmal üzerinde incelenmiştir.

Çizelge 4.8'de Oryantal pamukta artı ve eksi sarmal için upstream bölgesinde SSR dağılımları görülmektedir. Oryantal pamukta upstream dizi bölgeleri toplam artı sarmal için 435, eksi sarmal için 307 SSR içermektedir. Çizelgede de görüldüğü üzere en fazla bulunan SSR di-nükleotitler olup bunu sırasıyla tri-, hekza-, tetra- ve penta-nükleotitler izlemektedir. Bu türün kromozomları incelendiğinde ise artı sarmal üzerinde en fazla SSR di-nükleotitlerde A11 kromozomunda yer alırken; tri-nükleotitlerde A10 ve A11 kromozomu, tetra-nükleotitlerde A12 kromozomu, penta-nükleotitlerde A01, A02, A05, A09 kromozomunda, hekza-nükleotitlerde ise A04 ve A08 kromozomunda yer almaktadır. Eksi sarmal üzerinde en fazla SSR di-nükleotitlerde A05 kromozomunda, tri-nükleotitlerde A11 kromozomunda, tetra-nükleotitlerde A03, A05, A06, A09, A10 ve A12 kromozomunda, penta-nükleotitlerde A05 ve A11 kromozomunda olup, hekza-nükleotitlerde A03 kromozomunda bulunmaktadır (Çizelge 4.8).

Çizelge 4.8. Oryantal pamuk +/- sarmal için upstream dizilerinde SSR dağılımları

Krom	Di	Tri	Tetra	Penta	Hekza	Toplam
A01	18/13	6/4	1/1	3/2	1/1	29/21
A02	12/10	3/1	2/0	3/0	2/2	22/13
A03	13/8	4/1	3/2	2/1	6/33	28/45
A04	13/9	5/4	2/1	0/1	8/3	28/18
A05	19/18	8/6	3/2	3/3	5/1	38/30
A06	22/11	5/2	2/2	0/1	3/0	32/16
A07	15/14	7/1	1/0	2/0	6/2	31/17
A08	22/15	8/5	1/1	2/2	5/4	38/27
A09	18/10	10/3	0/2	3/1	5/0	36/16
A10	20/6	12/5	2/2	1/2	6/2	41/17
A11	24/15	12/14	1/1	1/3	8/5	46/38
A12	19/14	6/10	6/2	2/0	3/1	36/27
A13	18/14	5/4	2/1	0/0	5/3	30/22

Çizelgede "/" işaretinin sol tarafı "+ sarmalı" sağ tarafı ise "- sarmalı" nitelemektedir.

Çizelge 4.9’da Oryantal pamukta artı ve eksi sarmal için downstream dizilerinde SSR dağılımları görülmektedir. Oryantal pamukta downstream bölgeleri toplam artı sarmal için 5, eksi sarmal için 12 SSR içermektedir. Çizelgede de görüldüğü üzere en fazla bulunan SSR artı sarmal üzerinde bulunan di-nükleotitler olup bunu tri-nükleotitler takip etmektedir. Tetra-, penta- ve hekza-nükleotitlerde ise SSR bulunmamaktadır. Eksi sarmal üzerinde ise en fazla SSR di- ve tri-nükleotitlerde iken; sırasıyla hekza- ve tetra-nükleotitler izlemektedir. Ancak, penta-nükleotitlerde SSR bulunmamaktadır. Bu türün kromozomları incelendiğinde ise artı sarmal üzerinde en fazla SSR di-nükleotitlerde A02, A03 ve A05 kromozomunda yer alırken; tri-nükleotitlerde A08 ve A10 kromozomu, tetra-nükleotitlerde, penta-nükleotitlerde ve hekza-nükleotitlerde tüm kromozomlarda SSR bulunmamaktadır. Eksi sarmal üzerinde ise en fazla SSR di-nükleotitlerde A05, A08, A10 ve A11 kromozomunda, tri-nükleotitlerde A09 kromozomunda, tetra-nükleotitlerde A10 kromozomunda, penta-nükleotitlerde kromozomlarda SSR bulunmamakla beraber hekza-nükleotitlerde A12 kromozomunda bulunmaktadır (Çizelge 4.9).

Çizelge 4.9. Oryantal pamuk +/- sarmal için downstream dizilerinde SSR dağılımları

Krom	Di	Tri	Tetra	Penta	Hekza	Toplam
A01	0/0	0/0	0/0	0/0	0/0	0/0
A02	1/0	0/0	0/0	0/0	0/0	1/0
A03	1/0	0/0	0/0	0/0	0/0	1/0
A04	0/0	0/0	0/0	0/0	0/0	0/0
A05	1/1	0/1	0/0	0/0	0/0	1/2
A06	0/0	0/1	0/0	0/0	0/0	0/1
A07	0/0	0/0	0/0	0/0	0/0	0/0
A08	0/1	1/0	0/0	0/0	0/0	1/1
A09	0/0	0/2	0/0	0/0	0/0	0/2
A10	0/1	1/0	0/1	0/0	0/0	1/2
A11	0/1	0/0	0/0	0/0	0/0	0/1
A12	0/0	0/0	0/0	0/0	0/3	0/3
A13	0/0	0/0	0/0	0/0	0/0	0/0

Çizelgede “/” işaretinin sol tarafı “+ sarmalı” sağ tarafı ise “- sarmalı” nitelemektedir.

Çizelge 4.10’da Oryantal pamukta artı ve eksi sarmal için ekson dizilerinde SSR dağılımları görülmektedir. Oryantal pamukta ekson bölgeleri toplam artı sarmal için 725, eksi sarmal için 1162 SSR içermektedir. Çizelgede de görüldüğü üzere en fazla bulunan SSR artı sarmal üzerinde bulunan tri-nükleotitler olup bunu sırasıyla di-nükleotitler, hekza-nükleotitler ve 42 sayısı ile tetra- ve penta-nükleotitler izlemektedir. Eksi sarmal üzerinde ise en fazla bulunan SSR’lar sırasıyla di-nükleotitler, tri-nükleotitler, hekza-nükleotitler, penta-nükleotitler ve tetra-nükleotitler izlemektedir. Bu türün kromozomları incelendiğinde ise artı sarmal üzerinde en fazla SSR di-nükleotitlerde A05 kromozomunda yer alırken; tri-nükleotitlerde ve tetra-nükleotitlerde A11 kromozomu, penta-nükleotitlerde A05 ve A11 kromozomunda, hekza-nükleotitlerde A11 kromozomunda yer almaktadır. Eksi sarmal üzerinde ise en fazla SSR di-nükleotitlerde ve tri-nükleotitlerde A11 kromozomunda, tetra-nükleotitlerde

A10 kromozomunda, penta-nükleotitlerde A05 kromozomunda, hekza-nükleotitlerde A03 kromozomunda bulunmaktadır (Çizelge 4.10).

Çizelge 4.10. Oryantal pamuk +/- sarmal için ekson dizilerinde SSR dağılımları

Krom	Di	Tri	Tetra	Penta	Hekza	Toplam
A01	20/43	12/22	2/1	1/3	11/9	46/71
A02	6/23	16/15	4/3	4/4	8/6	38/51
A03	13/35	9/34	3/5	5/7	15/47	45/128
A04	11/37	9/19	2/3	3/3	13/7	38/69
A05	29/71	23/36	5/5	6/11	20/7	83/130
A06	26/47	13/15	2/1	2/3	23/8	66/74
A07	13/35	18/12	3/3	4/9	15/10	53/69
A08	15/46	15/17	2/6	1/6	19/11	52/86
A09	13/30	21/18	4/3	2/1	14/8	54/60
A10	14/38	28/22	3/10	2/6	18/14	65/90
A11	13/72	29/42	8/5	6/10	27/8	83/137
A12	19/46	24/25	1/5	3/5	13/46	60/127
A13	22/36	11/23	3/3	3/2	13/6	42/70

Çizelgede “/” işaretinin sol tarafı “+ sarmalı” sağ tarafı ise “- sarmalı” nitelemektedir.

Çizelge 4.11’de Oryantal pamukta artı (+) ve eksi (-) sarmal için intron dizilerinde SSR dağılımları görülmektedir. Oryantal pamukta intron bölgeleri toplam artı sarmal için 277, eksi sarmal için 233 SSR içermektedir. Çizelgede de görüldüğü üzere en fazla bulunan SSR +/- sarmal üzerinde bulunan di-nükleotitler olup bunu sırasıyla tri-nükleotitler, hekza-nükleotitler, tetra-nükleotitler ve penta-nükleotitler takip etmektedir. Bu türün kromozomları incelendiğinde ise artı sarmal üzerinde en fazla SSR di-nükleotitlerde A06, A10 ve A11 kromozomunda yer alırken; tri-nükleotitlerde A11 kromozomu, tetra-nükleotitlerde A09 ve A11 kromozomu, penta-nükleotitlerde A05 kromozomunda, hekza-nükleotitlerde A13 kromozomunda yer almaktadır. Eksi sarmal üzerinde ise en fazla SSR di-nükleotitlerde A01 ve A11 kromozomunda, tri-nükleotitlerde A13 kromozomunda, tetra-nükleotitlerde A10 kromozomunda, penta-nükleotitlerde A07 kromozomunda, hekza-nükleotitlerde A12 kromozomunda bulunmaktadır (Çizelge 4.11).

Çizelge 4.11. Oryantal pamuk +/- sarmal için intron dizilerinde SSR dağılımları

Krom	Di	Tri	Tetra	Penta	Hekza	Toplam
A01	9/14	1/4	0/0	0/3	1/1	11/22
A02	11/6	8/0	3/1	0/0	2/0	24/7
A03	9/4	7/5	2/1	0/0	1/0	19/10
A04	7/10	1/1	1/1	1/1	2/1	12/14
A05	6/9	4/4	1/2	4/1	2/2	17/18
A06	15/2	3/2	0/0	0/1	1/0	19/5
A07	6/9	9/5	2/1	1/4	5/2	23/21

Çizelge 4.11'in devamı

Krom	Di	Tri	Tetra	Penta	Hekza	Toplam
A08	13/10	12 /5	1/0	2/1	2/1	30/17
A09	12/5	8/0	4 /0	1/1	2/1	27/7
A10	15 /6	3/5	0/ 9	1/1	3/1	22/22
A11	15/14	8/3	4 /1	1/1	3/2	31 /21
A12	9/11	5/4	1/0	0/1	3/31	18/ 47
A13	13/5	3/12	0/2	2/0	6 /3	24/22

Çizelgede “/” işaretinin sol tarafı “+ sarmalı” sağ tarafı ise “– sarmalı” nitelemektedir.

Çizelge 4.12’de Oryantal pamukta +/- sarmal için 5'-UTR dizilerinde SSR dağılımları görülmektedir. Oryantal pamukta 5'-UTR bölgeleri toplam artı sarmal için 346, eksi sarmal için 62 SSR içermektedir. Çizelgede de görüldüğü üzere en fazla bulunan SSR artı sarmal üzerinde bulunan di-nükleotitler olup bunu sırasıyla hekza-nükleotitler, tri-nükleotitler, penta ve tetra-nükleotitler takip etmektedir. Eksi sarmal üzerinde en fazla bulunan SSR’lar ise sırasıyla di-nükleotitler, tri-nükleotitler, hekza-nükleotitler, tetra- ve penta-nükleotitler izlemektedir. Bu türün kromozomları incelendiğinde ise artı sarmal üzerinde en fazla SSR di-nükleotitlerde A05 kromozomunda yer alırken, tri-nükleotitlerde A10 kromozomu, tetra-nükleotitlerde A02, A05 ve A11 kromozomu, penta- ve hekza-nükleotitlerde A05 kromozomunda yer almaktadır. Eksi sarmal üzerinde ise en fazla SSR di-nükleotitlerde A11 ve A12 kromozomunda, tri-nükleotitlerde A11 kromozomunda, tetra-nükleotitlerde A12 ve A13 kromozomunda, penta-nükleotitlerde A07 kromozomunda, hekza-nükleotitlerde A03 kromozomunda bulunmaktadır (Çizelge 4.12).

Çizelge 4.12. Oryantal pamuk +/- sarmal için 5'-UTR dizilerinde SSR dağılımları

Krom	Di	Tri	Tetra	Penta	Hekza	Toplam
A01	11/2	4/0	1/0	1/0	3/0	20/2
A02	4/1	5/1	3 /0	3/0	3/0	18/2
A03	10/1	4/2	1/0	5/0	6 /4	26/7
A04	8/3	5/0	2/0	2/0	7/1	24/5
A05	23 /2	6/3	3 /0	6 /0	13 /1	51 /6
A06	17/1	4/1	1/0	2/0	4/0	28/2
A07	7/1	6/0	1/0	3 /1	10/1	27/3
A08	12/0	5/3	0/0	1/0	5/2	23/5
A09	6/1	7/1	2/0	2/0	7/0	24/2
A10	8/2	11 /2	2/0	2/0	4/0	27/4
A11	8 /4	8 /7	3 /0	2/0	10/0	31 / 11
A12	12/ 4	6/2	0/ 1	2/0	6/2	26/9
A13	10/2	3/1	1 /1	2/0	5/0	21/4

Çizelgede “/” işaretinin sol tarafı “+ sarmalı” sağ tarafı ise “– sarmalı” nitelemektedir.

Çizelge 4.13'te Oryantal pamukta +/- sarmal için 3'-UTR dizilerinde SSR dağılımları görülmektedir. Oryantal pamukta 3'-UTR bölgeleri toplam artı sarmal için 49, eksi sarmal için 305 SSR içermektedir. Çizelgede de görüldüğü üzere en fazla bulunan SSR artı sarmal üzerinde bulunan di-nükleotitler olup bunu sırasıyla tri-nükleotitler, penta-nükleotitler, 3 sayısı ile tetra- ve hekza-nükleotitler takip etmektedir. Eksi sarmal üzerinde en fazla bulunan SSR'lar ise sırasıyla di-nükleotitler, tri-nükleotitler, hekza-nükleotitler, tetra- ve penta-nükleotitler izlemektedir. Bu türün kromozomları incelendiğinde ise artı sarmal üzerinde en fazla SSR di-nükleotitlerde A11 kromozomunda yer alırken; tri-nükleotitlerde A05 ve A10 kromozomu, tetra-nükleotitlerde A01, A08 ve A10 kromozomu, penta-nükleotitlerde A11 kromozomunda, hekza-nükleotitlerde A08, A11 ve A13 kromozomunda yer almaktadır. Eksi sarmal üzerinde ise en fazla SSR di-nükleotitlerde A11 kromozomunda, tri-nükleotitlerde A05 kromozomunda, tetra-nükleotitlerde A10 kromozomunda, penta-nükleotitlerde A05 ve A11 kromozomunda, hekza-nükleotitlerde A12 kromozomunda bulunmaktadır (Çizelge 4.13).

Çizelge 4.13. Oryantal pamuk +/- sarmal için 3'-UTR dizilerinde SSR dağılımları

Krom	Di	Tri	Tetra	Penta	Hekza	Toplam
A01	2/6	0/2	1/0	0/1	0/0	3/9
A02	0/6	1/1	0/1	0/0	0/2	1/10
A03	1/9	1/8	0/1	1/2	0/5	3/25
A04	2/7	2/5	0/1	0/0	0/1	4/14
A05	2/14	3/11	0/1	0/3	0/2	5/31
A06	4/13	1/6	0/0	0/1	0/1	5/21
A07	0/11	0/4	0/1	0/0	0/4	0/20
A08	0/9	1/1	1/1	0/1	1/2	3/14
A09	2/9	0/4	0/0	0/0	0/3	2/16
A10	2/9	3/5	1/6	0/1	0/4	6/25
A11	6/24	2/9	0/3	3/3	1/2	12/41
A12	2/8	1/6	0/3	1/1	0/38	4/56
A13	0/11	0/7	0/1	0/1	1/3	1/23

Çizelgede “/” işaretinin sol tarafı “+ sarmalı” sağ tarafı ise “- sarmalı” nitelemektedir.

Çizelge 4.14'te Oryantal pamukta +/- sarmal için intergen SSR dağılımları görülmektedir. Oryantal pamukta intergen bölgeleri toplam artı sarmal için 33969, eksi sarmal için 33598 SSR içermektedir. Çizelgede de görüldüğü üzere en fazla bulunan SSR +/- sarmal üzerinde bulunan di-nükleotitler olup bunu sırasıyla tri-nükleotitler, hekza-nükleotitler, penta- ve tetra-nükleotitler izlemektedir. Bu türün kromozomları incelendiğinde ise artı ve eksi sarmal üzerinde en fazla SSR di-, tri- ve tetra-nükleotitlerde A11 kromozomunda, penta-nükleotitlerde A10 kromozomunda, hekza-nükleotitlerde ise A03 kromozomunda yer almaktadır (Çizelge 4.14).

Çizelge 4.14. Oryantal pamuk +/- sarmal için intergen dizilerinde SSR dağılımları

Krom	Di	Tri	Tetra	Penta	Hekza	Toplam
A01	1521/1494	624/624	106/107	156/154	251/252	2658/2631
A02	1266/1250	399/408	107/108	129/128	194/197	2095/2091
A03	1540/1518	627/609	120/117	181/182	399/379	2867/2805
A04	1347/1318	466/463	97/95	144/144	218/227	2272/2247
A05	1392/1359	589/582	109/106	190/186	233/246	2513/2479
A06	1687/1669	590/589	120/120	171/171	241/256	2809/2805
A07	1436/1412	568/569	115/113	166/161	240/244	2525/2499
A08	1614/1585	565/564	108/106	179/175	259/268	2725/2698
A09	1186/1169	516/519	91/90	137/138	238/247	2168/2163
A10	1668/1643	682/687	138/135	200/196	325/326	3013/2987
A11	1736/1689	710/704	143/141	186/182	282/305	3057/3021
A12	1528/1504	575/578	129/127	177/175	299/245	2708/2629
A13	1514/1496	552/547	102/100	139/140	252/260	2559/2543

Çizelgede “/” işaretinin sol tarafı “+ sarmalı” sağ tarafı ise “- sarmalı” nitelemektedir.

Çizelge 4.15’te Oryantal pamukta +/- sarmal için tüm gen SSR dağılımları görülmektedir. Oryantal pamukta tüm gen bölgeleri toplam artı sarmal için 1885, eksi sarmal için 2169 SSR içermektedir. Çizelgede de görüldüğü üzere en fazla bulunan SSR artı sarmal üzerinde bulunan di-nükleotitler olup bunu sırasıyla tri-nükleotitler, hekza-nükleotitler, tetra- ve penta-nükleotitler takip etmektedir. Eksi sarmal üzerinde en fazla bulunan SSR’lar ise sırasıyla di-nükleotitler, tri-nükleotitler, hekza-nükleotitler, penta- ve tetra-nükleotitler izlemektedir. Bu türün kromozomları incelendiğinde ise artı sarmal üzerinde en fazla SSR di-nükleotitlerde A06 kromozomunda yer alırken, tri-nükleotitlerde A11 kromozomu, tetra-nükleotitlerde A09 kromozomu, penta-nükleotitlerde A05 kromozomunda, hekza-nükleotitlerde A11 kromozomunda yer almaktadır. Eksi sarmal üzerinde ise en fazla SSR di-nükleotitlerde A11 kromozomunda, tri-nükleotitlerde A05 kromozomunda, tetra-nükleotitlerde A10 kromozomunda, penta-nükleotitlerde A11 kromozomunda ve hekza-nükleotitlerde A03 kromozomunda bulunmaktadır (Çizelge 4.15).

Çizelge 4.15. Oryantal pamuk +/- sarmal için tüm gen dizilerinde SSR dağılımları

Krom	Di	Tri	Tetra	Penta	Hekza	Toplam
A01	77/96	27/43	6/4	7/8	27/16	144/167
A02	36/42	26/20	10/7	10/7	14/8	96/84
A03	50/66	20/55	14/11	12/10	22/ 85	118/227
A04	51/80	18/26	10/7	9/8	18/14	106/135
A05	79/112	35/ 64	14/8	14/17	26/20	168/221
A06	87/73	33/23	6/4	6/6	36/12	168/118
A07	55/71	36/24	10/6	11/17	22/20	134/138
A08	74/86	35/43	9/9	5/10	27/21	150/169

Çizelge 4.15'in devamı

Krom	Di	Tri	Tetra	Penta	Hekza	Toplam
A09	67/71	36/35	16/6	7/3	24/19	150/134
A10	72/80	48/41	8/18	6/8	29/18	163/165
A11	72/132	57/62	13/8	8/18	45/22	195/242
A12	76/97	39/38	14/8	6/10	20/75	155/228
A13	73/69	29/46	5/7	9/4	22/15	138/141

Çizelgede “/” işaretinin sol tarafı “+ sarmalı” sağ tarafı ise “– sarmalı” nitelemektedir.

Pamuk dünya çapında en önemli lif ürünüdür. Lif kalitesi ve verimliliği, pamuk ovüllerinin epidermal hücrelerinin uzunluğu büyüme gelişim prosedüründen etkilenir. Tetraploit kültür pamuklarının ve kültürü yapılmayan diğer tetraploit pamukların, D genomu aralarında *G. raimondii*'nin atasını içeren eski bir hibridizasyondan kaynaklandığı görüşü yaygındır (Wendel ve Cronn 2003). Ayrıca, lif kalitesini kontrol eden kantitatif özellik lokuslarının çoğu, D-alt genomunda haritalanmıştır (Jiang vd. 1998). Bu nedenle, D-alt genomu üzerindeki bazı genlerin, özellikle transkripsiyon faktörlerinin, lif gelişiminde rol oynadığı düşünülmektedir.

Çizelge 4.16'da Peru pamukta +/- sarmal için upstream dizilerinde SSR dağılımları gösterilmektedir. Peru pamukta upstream bölgeleri toplam artı sarmal için 256, eksi sarmal için 232 SSR içermektedir. Çizelgede de görüldüğü üzere en fazla bulunan SSR artı sarmal üzerinde de bulunan di-nükleotitler olup bunu sırasıyla tri-, hekza-, tetra- ve penta-nükleotitler izlemektedir. Eksi sarmal üzerinde en fazla bulunan SSR'lar ise sırasıyla di-, tri-, hekza-, penta- ve tetra-nükleotitler izlemektedir. Bu türün kromozomları incelendiğinde ise artı sarmal üzerinde en fazla SSR di-, tri- ve tetra-nükleotitlerde D11 kromozomunda yer alırken, penta- ve hekza-nükleotitlerde D13 kromozomunda yer almaktadır. Eksi sarmal üzerinde ise en fazla SSR di-nükleotitlerde D05 kromozomunda, tri-nükleotitlerde D13 kromozomunda, tetra-nükleotitlerde D10 kromozomunda, penta-nükleotitlerde D09 kromozomunda, hekza-nükleotitlerde ise D13 kromozomunda bulunmaktadır (Çizelge 4.16).

Çizelge 4.16. Peru pamukta +/- sarmal için upstream dizilerinde SSR dağılımları

Krom	Di	Tri	Tetra	Penta	Hekza	Toplam
D01	10/5	5/1	0/1	2/0	4/3	21/10
D02	9/4	4/2	1/0	0/1	1/1	15/8
D03	2/6	3/3	2/1	1/1	0/2	8/13
D04	8/13	7/2	1/2	2/1	1/5	19/23
D05	15/19	5/7	1/2	2/0	4/2	27/30
D06	2/3	2/2	0/0	1/0	1/2	6/7
D07	6/6	4/3	1/1	0/2	3/3	14/15
D08	8/9	3/3	1/0	2/2	4/6	18/20
D09	13/5	3/6	1/2	1/4	1/4	19/21
D10	9/4	6/5	1/4	0/2	4/0	20/15

Çizelge 4.16'nın devamı

Krom	Di	Tri	Tetra	Penta	Hekza	Toplam
D11	16/13	12/6	6/1	1/2	9/1	44/23
D12	5/5	0/4	0/0	0/2	1/2	6/13
D13	13/10	9/15	3/1	4/0	10/8	39/34

Çizelgede “/” işaretinin sol tarafı “+ sarmalı” sağ tarafı ise “– sarmalı” nitelemektedir.

Çizelge 4.17’de Peru pamukta +/- sarmal için downstream dizilerinde SSR dağılımları gösterilmektedir. Peru pamukta downstream bölgeleri toplam artı sarmal için 14, eksi sarmal için 13 SSR içermektedir. Çizelgede de görüldüğü üzere en fazla bulunan SSR artı sarmal üzerinde de bulunan di-nükleotitler olup bunu sırasıyla tri-, hekza-, tetra-nükleotitler izlerken penta-nükleotitlerde SSR bulunmamaktadır. Eksi sarmal üzerinde en fazla bulunan SSR’lar di-nükleotitler olup tri- ve hekza-nükleotitlerde 2 SSR bulunurken, tetra- ve penta-nükleotitlerde 1 SSR bulunmaktadır. Bu türün kromozomları incelendiğinde ise artı sarmal üzerinde en fazla SSR di-, tri- ve tetra-nükleotitlerde D08 kromozomunda yer almakta olup, penta-nükleotitlerde artı sarmal üzerinde SSR bulunmayıp hekza-nükleotitlerde D13 kromozomunda yer almaktadır. Eksi sarmal üzerinde ise en fazla SSR di-nükleotitlerde D09 kromozomunda, tri-nükleotitlerde D08 ve D09 kromozomunda, tetra-nükleotitlerde D10 kromozomunda, penta-nükleotitlerde D09 kromozomunda, hekza-nükleotitlerde D09 ve D13 kromozomunda bulunmaktadır (Çizelge 4.17).

Çizelge 4.17. Peru pamukta +/- sarmal için downstream dizilerinde SSR dağılımları

Krom	Di	Tri	Tetra	Penta	Hekza	Toplam
D01	0/1	0/0	0/0	0/0	0/0	0/1
D02	0/1	1/0	0/0	0/0	0/0	1/1
D03	0/0	0/0	0/0	0/0	0/0	0/0
D04	1/0	1/0	0/0	0/0	0/0	2/0
D05	0/1	0/0	0/0	0/0	0/0	0/1
D06	0/0	0/0	0/0	0/0	0/0	0/0
D07	1/0	0/0	0/0	0/0	0/0	1/0
D08	3/0	2/1	1/0	0/0	0/0	6/1
D09	0/2	0/1	0/0	0/1	0/1	0/5
D10	1/1	0/0	0/1	0/0	0/0	½
D11	1/1	0/0	0/0	0/0	0/0	1/1
D12	0/0	0/0	0/0	0/0	1/0	1/0
D13	0/0	0/0	0/0	0/0	1/1	1/1

Çizelgede “/” işaretinin sol tarafı “+ sarmalı” sağ tarafı ise “– sarmalı” nitelemektedir.

Çizelge 4.18’de Peru pamukta +/- sarmal için ekson dizilerinde SSR dağılımları gösterilmektedir. Peru pamukta ekson bölgeleri toplam artı sarmal için 60, eksi sarmal için 102 SSR içermektedir. Çizelgede de görüleceği üzere en fazla bulunan SSR artı sarmal üzerinde de bulunan di-nükleotitler olup bunu sırasıyla hekza-, tri-, tetra- ve penta-nükleotitler yer almaktadır. Eksi sarmal üzerinde en fazla bulunan SSR’lar di-nükleotitler olurken; bunu tri-, hekza-, penta- ve tetra-nükleotitler izlemektedir. Bu türün kromozomları incelendiğinde ise artı sarmal üzerinde en fazla SSR di-nükleotitlerde D11 kromozomunda, tri-nükleotitlerde D02 kromozomunda, tetra-nükleotitlerde D02, D03, D04, D06, D08 ve D10 kromozomunda, penta-nükleotitlerde D01, D04, D05, D12 ve D13 kromozomunda, hekza-nükleotitlerde ise D05 ve D10 kromozomunda bulunmaktadır. Eksi sarmal üzerinde ise en fazla SSR di-nükleotitlerde D11 kromozomunda, tri- ve tetra-nükleotitlerde D13 kromozomunda, penta-nükleotitlerde D09 ve D11 kromozomunda iken; hekza-nükleotitlerde D13 kromozomunda bulunmaktadır (Çizelge 4.18).

Çizelge 4.18. Peru pamukta +/- sarmal için ekson dizilerinde SSR dağılımları

Krom	Di	Tri	Tetra	Penta	Hekza	Toplam
D01	4/3	3/1	0/0	1/0	1/0	9/4
D02	2/2	4/3	1/0	0/0	1/3	8/8
D03	0/2	0/1	1/0	0/1	1/1	2/5
D04	0/1	1/0	1/0	1/1	0/1	3/3
D05	2/4	0/4	0/0	1/1	3/0	6/9
D06	0/2	1/0	1/1	0/2	0/0	2/5
D07	1/4	1/3	0/2	0/0	2/0	4/9
D08	2/7	0/1	1/0	0/0	1/0	4/8
D09	0/6	0/1	0/1	0/3	1/1	1/12
D10	1/2	0/3	1/0	0/0	3/0	5/5
D11	11/7	2/1	0/0	0/3	2/2	5/13
D12	2/1	1/1	0/1	1/0	1/0	4/3
D13	4/6	3/5	0/3	1/0	1/4	7/18

Çizelgede “/” işaretinin sol tarafı “+ sarmalı” sağ tarafı ise “- sarmalı” nitelemektedir.

Çizelge 4.19’da Peru pamukta +/- sarmal için intron dizilerinde SSR dağılımları gösterilmektedir. Peru pamukta intron bölgeleri toplam artı sarmal için 233, eksi sarmal için 221 SSR içermektedir. Çizelgede de görüleceği üzere en fazla bulunan SSR artı sarmal üzerinde de bulunan di-nükleotitler olup bunu sırasıyla tri-, tetra-, hekza- ve penta-nükleotitler yer almaktadır. Eksi sarmal üzerinde ise en fazla bulunan SSR’lar di-nükleotitler olurken bunu tri-, hekza-, tetra- ve penta-nükleotitler izlemektedir. Bu türün kromozomları incelendiğinde ise artı sarmal üzerinde en fazla SSR di-nükleotitlerde D10 kromozomunda, tri-nükleotitlerde D09 kromozomunda, tetra-nükleotitlerde D12 kromozomunda, penta-nükleotitlerde D04, D09 ve D13 kromozomunda, hekza-nükleotitlerde D05 kromozomunda bulunmaktadır. Eksi sarmal üzerinde ise en fazla SSR di-nükleotitlerde D09 kromozomunda, tri-nükleotitlerde D11 kromozomunda, tetra-nükleotitlerde D03 ve D10 kromozomunda, penta-nükleotitlerde D12

kromozomunda, hekza-nükleotitlerde D04 kromozomunda bulunmaktadır (Çizelge 4.19).

Çizelge 4.19. Peru pamukta +/- sarmal için intron dizilerinde SSR dağılımları

Krom	Di	Tri	Tetra	Penta	Hekza	Toplam
D01	9/7	6/4	3/2	0/2	2/5	20/20
D02	6/5	0/3	1/0	1/2	0/2	8/12
D03	3/4	0/3	1/3	0/1	2/3	6/14
D04	8/12	1/2	1/2	2/1	0/7	12/24
D05	14/8	7/5	2/2	0/2	4/2	27/19
D06	4/1	2/0	3/1	0/0	0/1	9/3
D07	5/11	4/1	1/0	0/0	1/3	11/15
D08	11/7	6/1	3/1	1/0	2/5	23/14
D09	10/14	10/4	0/1	2/1	2/1	24/21
D10	17/11	5/5	4/3	1/2	3/3	30/24
D11	14/13	2/7	6/1	0/0	3/5	25/26
D12	4/3	1/3	10/0	1/3	0/2	16/11
D13	15/8	3/6	1/1	2/1	1/2	22/18

Çizelgede “/” işaretinin sol tarafı “+ sarmalı” sağ tarafı ise “- sarmalı” nitelemektedir.

Çizelge 4.20’de Peru pamukta +/- sarmal için 5'-UTR dizilerinde SSR dağılımları gösterilmektedir. Peru pamukta 5'-UTR bölgeleri toplam artı sarmal için 25, eksi sarmal için 11 SSR içermektedir. Çizelgede de görüleceği üzere en fazla bulunan SSR artı sarmal üzerinde de bulunan di-nükleotitler olup bunu sırasıyla hekza-, tri-, penta- ve tetra-nükleotitler izlemektedir. Eksi sarmal üzerinde en fazla bulunan SSR’lar di-nükleotitler olup tri-, tetra- ve hekza-nükleotitlerde 1 SSR bulunurken; penta-nükleotitlerde SSR bulunmamaktadır. Bu türün kromozomları incelendiğinde ise artı sarmal üzerinde en fazla SSR di-nükleotitlerde D08 kromozomunda, tri-nükleotitlerde D01 ve D13 kromozomunda, tetra-nükleotitlerde D02 ve D04 kromozomunda, penta-nükleotitlerde D01, D04, D05 ve D12 kromozomunda, hekza-nükleotitlerde ise D07 kromozomunda bulunmaktadır. Eksi sarmal üzerinde ise en fazla SSR di-nükleotitlerde D05 ve D11 kromozomunda, tri- ve tetra-nükleotitlerde D13 kromozomunda, hekza-nükleotitlerde D11 kromozomunda olup; penta-nükleotitlerde SSR bulunmamaktadır (Çizelge 4.20).

Çizelge 4.20. Peru pamukta +/- sarmal için 5'-UTR dizilerinde SSR dağılımları

Krom	Di	Tri	Tetra	Penta	Hekza	Toplam
D01	1/1	2/0	0/0	1/0	1/0	5/1
D02	1/0	0/0	1/0	0/0	0/0	2/0
D03	0/0	0/0	0/0	0/0	0/0	0/0
D04	0/0	0/0	1/0	1/0	0/0	2/0
D05	1/2	0/0	0/0	1/0	1/0	3/2
D06	0/0	0/0	0/0	0/0	0/0	0/0

Çizelge 4.20'nin devamı

Krom	Di	Tri	Tetra	Penta	Hekza	Toplam
D07	1/1	0/0	0/0	0/0	2/0	3/1
D08	2/1	0/0	0/0	0/0	1/0	3/1
D09	0/1	0/0	0/0	0/0	0/0	0/1
D10	0/0	0/0	0/0	0/0	0/0	0/0
D11	1/2	1/0	0/0	0/0	1/1	3/3
D12	0/0	0/0	0/0	1/0	0/0	1/0
D13	1/0	2/1	0/1	0/0	0/0	3/2

Çizelgede “/” işaretinin sol tarafı “+ sarmalı” sağ tarafı ise “– sarmalı” nitelemektedir.

Çizelge 4.21’de Peru pamukta +/- sarmal için 3'-UTR dizilerinde SSR dağılımları gösterilmektedir. Peru pamukta 3'-UTR bölgeleri toplam artı sarmal için 10, eksi sarmal için 26 SSR içermektedir. Çizelgede de görüleceği üzere en fazla bulunan SSR artı sarmal üzerinde de bulunan di-nükleotitler olup tri- ve hekza-nükleotitlerde 2 SSR bulunurken; tetra- ve penta-nükleotitlerde 1 SSR bulunmaktadır. Eksi sarmal üzerinde en fazla bulunan SSR’lar di-nükleotitler olmakla birlikte tri-nükleotitlerde 6 SSR, tetra- ve penta-nükleotitler de 3 SSR, hekza-nükleotitlerde ise 2 SSR yer almaktadır. Bu türün kromozomları incelendiğinde ise artı sarmal üzerinde en fazla SSR di-nükleotitlerde D01, D02, D10 ve D13 kromozomunda, tri-nükleotitlerde D01 ve D02 kromozomunda, tetra-nükleotitlerde D06 kromozomunda, penta-nükleotitlerde D13 kromozomunda, hekza-nükleotitlerde ise D10 ve D13 kromozomunda bulunmaktadır. Eksi sarmal üzerinde ise en fazla SSR di-nükleotitlerde D01, D05, D09, D11 ve D13 kromozomunda, tri-nükleotitlerde D02 kromozomunda, tetra-nükleotitlerde D07, D09 ve D12 kromozomunda, penta-nükleotitlerde D04, D06 ve D09 kromozomunda iken; hekza-nükleotitlerde D02 ve D04 kromozomunda bulunmaktadır (Çizelge 4.21).

Çizelge 4.21. Peru pamukta +/- sarmal için 3'-UTR dizilerinde SSR dağılımları

Krom	Di	Tri	Tetra	Penta	Hekza	Toplam
D01	1/2	1/0	0/0	0/0	0/0	2/2
D02	1/0	1/3	0/0	0/0	0/1	2/4
D03	0/1	0/0	0/0	0/0	0/0	0/1
D04	0/0	0/0	0/0	0/1	0/1	0/2
D05	0/2	0/0	0/0	0/0	0/0	0/2
D06	0/1	0/0	1/0	0/1	0/0	1/2
D07	0/0	0/0	0/1	0/0	0/0	0/1
D08	0/0	0/0	0/0	0/0	0/0	0/0
D09	0/2	0/0	0/1	0/1	0/0	0/4
D10	1/0	0/2	0/0	0/0	1/0	2/2
D11	0/2	0/0	0/0	0/0	0/0	0/2
D12	0/0	0/0	0/1	0/0	0/0	0/1
D13	1/2	0/1	0/0	1/0	1/0	3/3

Çizelgede “/” işaretinin sol tarafı “+ sarmalı” sağ tarafı ise “– sarmalı” nitelemektedir.

Çizelge 4.22’de Peru pamukta +/- sarmal için intergen dizilerinde SSR dağılımları gösterilmektedir. Peru pamukta intergen bölgeleri toplam artı sarmal için 17879, eksi sarmal için 18212 SSR içermektedir. Çizelgede de görüleceği üzere en fazla bulunan SSR artı sarmal üzerinde de bulunan di-nükleotitler olup bunu sırasıyla tri-, hekza-, penta- ve tetra-nükleotitler izlemektedir. Eksi sarmal üzerinde en fazla bulunan SSR’lar sırasıyla di-, tri-, hekza-, penta- ve tetra-nükleotitlerdir. Bu türün kromozomları incelendiğinde ise artı sarmal üzerinde en fazla SSR di-nükleotitlerde D05 kromozomunda, tri-nükleotitlerde D13 kromozomunda, tetra-nükleotitlerde D04 kromozomunda, penta-nükleotitlerde D11 kromozomunda, hekza-nükleotitlerde ise D10 kromozomunda bulunmaktadır. Eksi sarmal üzerinde ise en fazla SSR di-nükleotitlerde D05 kromozomunda, tri-nükleotitlerde D11 kromozomunda, tetra-nükleotitlerde D04 kromozomunda, penta-nükleotitlerde D11 kromozomunda iken; hekza-nükleotitlerde D10 kromozomunda yer almaktadır (Çizelge 4.22).

Çizelge 4.22. Peru pamukta +/- sarmal için intergen dizilerinde SSR dağılımları

Krom	Di	Tri	Tetra	Penta	Hekza	Toplam
D01	638/640	380/382	94/96	120/123	188/189	1420/1430
D02	542/537	285/276	100/102	118/114	190/192	1235/1221
D03	372/404	277/260	97/83	89/109	127/135	962/991
D04	620/618	331/333	174/182	119/114	188/184	1432/1431
D05	801/790	430/426	104/104	140/142	209/210	1684/1672
D06	465/502	321/334	127/138	98/99	200/205	1211/1278
D07	669/649	406/405	120/120	133/133	197/199	1525/1506
D08	604/633	374/383	103/99	109/111	204/208	1394/1434
D09	606/586	400/408	98/97	105/104	195/187	1404/1382
D10	697/686	410/408	137/132	125/127	224/227	1593/1580
D11	723/717	456/ 471	128/130	152/ 155	205/209	1664/ 1682
D12	329/446	193/275	65/102	84/81	121/147	792/1051
D13	635/642	466/462	110/110	153/142	199/198	1563/1554

Çizelgede “/” işaretinin sol tarafı “+ sarmalı” sağ tarafı ise “- sarmalı” nitelemektedir.

Çizelge 4.23’te Peru pamukta +/- sarmal için tüm gen dizilerinde SSR dağılımları gösterilmektedir. Peru pamukta tüm gen bölgeleri toplam artı sarmal için 452, eksi sarmal için 547 SSR içermektedir. Çizelgede de görüleceği üzere en fazla bulunan SSR artı sarmal üzerinde de bulunan di-nükleotitler olup bunu sırasıyla tri-, hekza-, tetra- ve penta-nükleotitler izlemektedir. Eksi sarmal üzerinde en fazla bulunan SSR’lar sırasıyla di-, tri-, hekza-, tetra- ve penta-nükleotitlerdir. Bu türün kromozomları incelendiğinde ise artı sarmal üzerinde en fazla SSR di-nükleotitlerde D05 kromozomunda, tri-nükleotitlerde D09 kromozomunda, tetra-nükleotitlerde D12 kromozomunda, penta-nükleotitlerde D04 kromozomunda, hekza-nükleotitlerde ise D01 ve D11 kromozomunda bulunmaktadır. Eksi sarmal üzerinde ise en fazla SSR di-nükleotitlerde D09 kromozomunda, tri-nükleotitlerde D11 ve D13 kromozomunda, tetra-nükleotitlerde D10 kromozomunda, penta-nükleotitlerde D09 kromozomunda iken; hekza-nükleotitlerde D13 kromozomunda yer almaktadır (Çizelge 4.23).

Çizelge 4.23. Peru pamukta +/- sarmal için tüm gen dizilerinde SSR dağılımları

Krom	Di	Tri	Tetra	Penta	Hekza	Toplam
D01	25/27	13/5	5/5	2/2	8/5	53/44
D02	14/15	7/8	4/3	2/4	1/5	28/35
D03	5/12	3/7	4/4	1/3	6/4	19/30
D04	13/23	5/10	3/2	4/4	0/9	25/48
D05	28/19	14/14	3/2	3/5	7/2	55/42
D06	10/7	6/2	4/3	0/2	0/2	20/16
D07	15/23	8/12	2/4	0/1	7/5	32/45
D08	16/24	11/6	5/1	1/1	7/8	40/40
D09	16/ 28	15/9	1/3	3/6	7/7	42/53
D10	24/22	6/15	5/7	1/5	6/5	42/54
D11	25/24	6/18	9/4	1/5	8/5	49/56
D12	5/9	6/5	11/2	1/3	2/2	25/21
D13	24/24	11/18	3/6	3/4	6/11	47/63

Çizelgede “/” işaretinin sol tarafı “+ sarmalı” sağ tarafı ise “– sarmalı” nitelemektedir.

Upland pamuğu dünya çapında yalnızca en önemli doğal lif üreten pamuk türü olmakla kalmaz, aynı zamanda biyotik ve abiyotik çevresel faktörlere karşı daha güçlü toleransı ve poliploidizasyonu incelemek için ideal bir araştırma sistemidir (Adams ve Wendel 2005; Paterson vd. 2012). *G. hirsutum* yıllık lif üretiminin %90'ını oluşturur ve *G. herbaceum* veya *G. arboreum*'a benzeyen A-genom benzeri ata türlerinin ve *G. raimondii*'ye benzeyen D-genom benzeri türlerin allopoliploidizasyon olayından kaynaklanır. Allotetraploit pamuklardaki intergenomik etkileşim, daha yüksek verim, daha iyi lif kalitesi, daha güçlü toleranslar ve mekanizasyona daha iyi uyum sağlamayı mümkün kılmıştır. Bu nedenle Upland pamuğun diploit pamuklarla karşılaştırılması son derece yararlıdır.

Çizelge 4.24'te Upland pamukta +/- sarmal için upstream dizilerinde SSR dağılımları gösterilmektedir. Upland pamukta upstream bölgeleri toplam artı sarmal için 623, eksi sarmal için 320 SSR içermektedir. Çizelgede de görüleceği üzere en fazla bulunan SSR artı sarmal üzerinde de bulunan tri-nükleotitler olup bunu sırasıyla hekza-, di-, penta- ve tetra-nükleotitler izlemektedir. Eksi sarmal üzerinde en fazla bulunan SSR'lar sırasıyla tri-, hekza-, tetra-, penta- ve di-nükleotitlerdir. Bu türün kromozomları incelendiğinde ise artı sarmal üzerinde en fazla SSR di-nükleotitlerde A02 ve A05 kromozomunda, tri- ve tetra-nükleotitlerde D10 kromozomunda, penta- ve hekza-nükleotitlerde D05 kromozomunda bulunmaktadır. Eksi sarmal üzerinde ise en fazla SSR di-nükleotitlerde D06 kromozomunda, tri-nükleotitlerde D10 kromozomunda, tetra-nükleotitlerde D01 kromozomunda, penta-nükleotitlerde D02 kromozomunda iken; hekza-nükleotitlerde A05 ve D12 kromozomunda yer almaktadır (Çizelge 4.24).

Çizelge 4.24. Upland pamukta +/- sarmal için upstream dizilerinde SSR dağılımları

Krom	Di	Tri	Tetra	Penta	Hekza	Toplam
A01	3/3	12/4	2/2	5/1	10/5	32/15
A02	7/2	5/2	3/1	3/0	6/2	24/7
A03	2/1	10/2	3/2	6/1	7/1	28/7
A04	2/0	4/1	1/3	1/0	4/1	12/5
A05	7/2	8/6	5/3	5/3	13/10	38/24
A06	3/1	3/4	2/1	0/2	5/0	13/8
A07	5/1	9/1	1/0	0/0	4/3	19/5
A08	3/3	4/5	2/1	2/0	6/3	17/12
A09	1/0	11/3	2/0	1/1	8/7	23/11
A10	4/2	14/2	1/3	2/2	7/2	28/11
A11	2/2	7/6	1/3	1/1	7/3	18/15
A12	6/0	10/4	3/0	3/2	6/5	28/11
A13	3/2	5/1	3/2	1/1	6/1	18/7
D01	5/0	7/2	4/5	3/2	9/2	28/11
D02	1/2	6/8	4/2	1/5	10/0	22/17
D03	4/0	3/2	3/2	3/0	5/5	18/9
D04	4/0	9/0	1/1	0/0	3/0	17/1
D05	4/1	10/7	5/2	7/3	14/3	40/16
D06	3/4	4/4	2/4	3/0	5/4	17/16
D07	3/1	11/6	1/3	3/1	7/0	25/11
D08	3/1	4/5	3/0	3/2	6/4	19/12
D09	1/1	5/8	3/2	5/2	9/4	23/17
D10	1/3	18/10	8/1	4/1	8/2	39/17
D11	4/1	7/8	3/4	4/3	5/2	23/18
D12	5/0	6/8	1/4	3/1	6/10	21/23
D13	6/2	9/5	3/2	6/3	9/2	33/14

Çizelgede “/” işaretinin sol tarafı “+ sarmalı” sağ tarafı ise “– sarmalı” nitelemektedir.

Çizelge 4.25’te Upland pamukta +/- sarmal için downstream dizilerinde SSR dağılımları gösterilmektedir. Upland pamukta downstream bölgeleri toplam artı sarmal için 8, eksi sarmal için 13 SSR içermektedir. Çizelgede de görüleceği üzere en fazla bulunan SSR artı sarmal üzerinde de bulunan hekza-nükleotitler olup di- ve tri-nükleotitlerde 2 SSR bulunurken; tetra- ve penta-nükleotitlerde SSR bulunmamaktadır. Eksi sarmal üzerinde ise 1. sırada bulunan 6 SSR ile hekza-nükleotitler, 2. sırada 5 SSR içeren tri-nükleotitler olup di- ve tetra-nükleotitlerde 1 SSR bulunup, penta-nükleotitlerde SSR bulunmamaktadır. Bu türün kromozomları incelendiğinde artı sarmal üzerinde en fazla SSR di-nükleotitlerde D01 ve D12 kromozomunda, tri-nükleotitlerde D11 ve D12 kromozomunda, tetra- ve penta-nükleotitlerde kromozomlarda SSR içermeyen hekza-nükleotitlerde en fazla A05, A11, D05 ve D09 kromozomunda bulunmaktadır. Eksi sarmal üzerinde ise en fazla SSR di-nükleotitlerde A11 kromozomunda, tri-nükleotitlerde D09 kromozomunda, tetra-nükleotitlerde D01 kromozomunda, penta-nükleotitlerde kromozomlarda SSR yer almazken; hekza-

nükleotitlerde A02, A04, D01, D06, D11 ve D12 kromozomunda yer almaktadır (Çizelge 4.25).

Çizelge 4.25. Upland pamukta +/- sarmal için downstream dizilerinde SSR dağılımları

Krom	Di	Tri	Tetra	Penta	Hekza	Toplam
A01	0/0	0/0	0/0	0/0	0/0	0/0
A02	0/0	0/0	0/0	0/0	0/1	0/1
A03	0/0	0/0	0/0	0/0	0/0	0/0
A04	0/0	0/0	0/0	0/0	0/1	0/1
A05	0/0	0/0	0/0	0/0	1/0	1/0
A06	0/0	0/0	0/0	0/0	0/0	0/0
A07	0/0	0/0	0/0	0/0	0/0	0/0
A08	0/0	0/0	0/0	0/0	0/0	0/0
A09	0/0	0/1	0/0	0/0	0/0	0/1
A10	0/0	0/0	0/0	0/0	0/0	0/0
A11	0/1	0/0	0/0	0/0	1/0	1/1
A12	0/0	0/0	0/0	0/0	0/0	0/0
A13	0/0	0/0	0/0	0/0	0/0	0/0
D01	1/0	0/1	0/1	0/0	0/1	1/3
D02	0/0	0/0	0/0	0/0	0/0	0/0
D03	0/0	0/0	0/0	0/0	0/0	0/0
D04	0/0	0/1	0/0	0/0	0/0	0/1
D05	0/0	0/0	0/0	0/0	1/0	1/0
D06	0/0	0/0	0/0	0/0	0/1	0/1
D07	0/0	0/0	0/0	0/0	0/0	0/0
D08	0/0	0/0	0/0	0/0	0/0	0/0
D09	0/0	0/2	0/0	0/0	1/0	1/2
D10	0/0	0/0	0/0	0/0	0/0	0/0
D11	0/0	1/0	0/0	0/0	0/1	1/1
D12	1/0	1/0	0/0	0/0	0/1	2/1
D13	0/0	0/0	0/0	0/0	0/0	0/0

Çizelgede “/” işaretinin sol tarafı “+ sarmalı” sağ tarafı ise “- sarmalı” nitelemektedir.

Çizelge 4.26’da Upland pamukta +/- sarmal için ekson dizilerinde SSR dağılımları gösterilmektedir. Upland pamukta ekson bölgeleri toplam artı sarmal için 1401, eksi sarmal için 1154 SSR içermektedir. Çizelgede de görüleceği üzere en fazla bulunan SSR artı sarmal üzerinde de bulunan tri-nükleotitler olup bunu sırasıyla hekza-, di-, penta- ve tetra-nükleotitler yer almaktadır. Eksi sarmal üzerinde ise en fazla bulunan tri-nükleotitler iken; bunu hekza-, tetra-, di- ve penta-nükleotitler izlemektedir. Bu türün kromozomları incelendiğinde artı sarmal üzerinde en fazla SSR di-nükleotitlerde A05 kromozomunda, tri-nükleotitlerde D11 kromozomunda, tetra-nükleotitlerde D10 kromozomunda, penta-nükleotitlerde A11 kromozomunda, hekza-nükleotitlerde D12 kromozomunda yer almakla birlikte, eksi sarmal üzerinde en fazla SSR di-, tri-, tetra- ve

penta-nükleotitlerde D05 kromozomunda iken; hekza-nükleotitlerde A11 kromozomunda yer almaktadır (Çizelge 4.26).

Çizelge 4.26 Upland pamukta +/- sarmal için ekson dizilerinde SSR dağılımları

Krom	Di	Tri	Tetra	Penta	Hekza	Toplam
A01	8/6	18/13	1/4	1/5	14/4	42/32
A02	4/4	15/8	2/10	5/4	14/11	40/37
A03	9/2	16/17	5/5	5/5	9/11	44/40
A04	3/2	5/6	0/6	2/2	5/3	15/19
A05	22/5	27/22	6/6	8/5	28/16	91/54
A06	10/6	15/11	4/4	1/8	7/12	37/41
A07	8/3	30/11	2/7	2/4	10/9	52/34
A08	8/6	17/12	5/6	2/9	11/9	43/42
A09	10/3	26/12	1/0	4/5	15/12	56/32
A10	9/3	17/10	1/2	4/2	14/10	45/27
A11	9/10	30/30	3/7	9/10	22/21	73/78
A12	10/5	28/12	1/9	2/5	10/13	51/44
A13	9/8	15/14	4/6	0/5	15/9	43/42
D01	11/6	26/16	3/6	3/4	15/8	58/40
D02	5/8	18/14	6/6	4/10	19/10	52/48
D03	8/3	16/10	4/8	5/6	15/10	48/37
D04	4/5	13/15	2/3	4/5	9/6	32/34
D05	19/22	29/40	4/11	7/12	23/14	82/99
D06	12/11	20/11	5/7	2/7	20/13	59/49
D07	8/4	28/17	4/8	5/7	17/7	62/43
D08	8/7	20/17	7/7	3/8	14/17	52/56
D09	5/7	23/10	4/6	6/2	11/12	49/37
D10	9/4	25/14	9/3	2/3	15/4	60/28
D11	16/11	33/28	4/7	6/4	20/19	79/69
D12	13/2	28/13	3/10	4/7	32/11	80/43
D13	10/6	23/23	4/10	5/5	14/5	56/49

Çizelgede “/” işaretinin sol tarafı “+ sarmalı” sağ tarafı ise “- sarmalı” nitelemektedir.

Çizelge 4.27’de Upland pamukta +/- sarmal için intron dizilerinde SSR dağılımları gösterilmektedir. Upland pamukta intron bölgeleri toplam artı sarmal için 279, eksi sarmal için 217 SSR içermektedir. Çizelgeden anlaşılabileceği üzere en fazla bulunan SSR artı sarmal üzerinde bulunan di-nükleotitler olup bunu sırasıyla hekza-, tri-, tetra- ve penta-nükleotitler yer almaktadır. Eksi sarmal üzerinde ise en fazla bulunan hekza-nükleotitler iken; bunu tri-, tetra-, penta- ve di-nükleotitler izlemektedir. Bu türün kromozomları incelendiğinde artı sarmal üzerinde en fazla SSR di-nükleotitlerde A01 ve D09 kromozomunda, tri-nükleotitlerde A01 kromozomunda, tetra-nükleotitlerde D05 kromozomunda, penta-nükleotitlerde D01 ve D13 kromozomunda, hekza-nükleotitlerde D10 kromozomunda yer alıp, eksi sarmal üzerinde en fazla SSR di-nükleotitlerde A01, A02, A05, A07, A11, D02, D03, D04 ve D13 kromozomları, tri- ve tetra-nükleotitlerde

D05 kromozomunda, penta-nükleotitlerde A08 ve D05 kromozomunda, hekza-nükleotitlerde D02, D05, D06 ve D11 kromozomlarında yer almaktadır (Çizelge 4.27).

Çizelge 4.27. Upland pamukta +/- sarmal için intron dizilerinde SSR dağılımları

Krom	Di	Tri	Tetra	Penta	Hekza	Toplam
A01	6/2	9/0	3/0	1/0	1/3	20/5
A02	2/2	4/0	1/0	1/0	1/2	9/4
A03	4/1	5/2	1/2	1/1	1/4	12/10
A04	3/0	1/2	2/1	0/1	1/1	7/5
A05	0/2	4/5	1/1	1/0	4/3	10/11
A06	0/1	1/3	0/0	1/1	1/1	3/6
A07	3/2	6/0	1/2	2/2	2/2	14/8
A08	0/0	3/1	0/1	1/3	2/3	6/8
A09	4/0	2/6	0/0	1/1	3/4	10/11
A10	4/0	3/2	0/0	0/0	2/0	9/2
A11	5/2	6/4	0/3	0/0	3/3	14/12
A12	0/1	4/1	1/4	0/1	1/1	6/8
A13	0/0	4/2	1/1	0/0	2/1	7/4
D01	5/0	4/2	1/1	3/1	3/0	16/4
D02	1/2	5/4	0/4	1/2	4/5	11/17
D03	1/2	4/2	1/4	0/2	3/2	9/12
D04	0/2	3/5	1/0	0/0	1/2	5/9
D05	3/0	3/8	5/5	1/3	3/5	15/21
D06	1/0	4/3	0/0	2/0	2/5	9/8
D07	3/1	3/2	0/0	0/0	0/2	6/5
D08	5/1	5/2	2/2	0/0	0/2	12/7
D09	6/1	0/2	0/0	2/2	2/1	10/6
D10	1/0	3/5	3/1	1/2	5/1	13/9
D11	5/1	6/6	3/2	1/1	3/5	18/15
D12	4/1	5/1	1/1	0/1	0/1	10/5
D13	5/2	4/1	4/1	3/0	2/1	18/5

Çizelgede “/” işaretinin sol tarafı “+ sarmalı” sağ tarafı ise “– sarmalı” nitelemektedir.

Çizelge 4.28’de Upland pamukta +/- sarmal için 5'-UTR dizilerinde SSR dağılımları gösterilmektedir. Upland pamukta 5'-UTR bölgeleri toplam artı sarmal için 719, eksi sarmal için 74 SSR içermektedir. Çizelgede de görüleceği üzere en fazla bulunan SSR artı sarmal üzerinde bulunan tri-nükleotitler olup bunu sırasıyla di-, hekza-, penta- ve tetra-nükleotitler yer almaktadır. Eksi sarmal üzerinde ise en fazla bulunan 28 SSR ile tri-nükleotitler iken; tetra- ve hekza-nükleotitlerde 13 SSR, di-nükleotitlerde 12 SSR, son olarak penta-nükleotitlerde 8 SSR bulunmaktadır. Bu türün kromozomları incelendiğinde artı sarmal üzerinde en fazla SSR di-nükleotitlerde D05 kromozomunda, tri-nükleotitlerde D11 kromozomunda, tetra-nükleotitlerde D08 kromozomunda, penta-nükleotitlerde A02, A05 ve A11 kromozomlarında, hekza-nükleotitlerde D12 kromozomunda yer alırken; eksi sarmal üzerinde en fazla SSR di- ve tri-nükleotitlerde

D05 kromozomunda, tetra-nükleotitlerde A02, A11, A12 ve A13 kromozomlarında, penta-nükleotitlerde A04, A06, A08, A10, D03, D05, D11 ve D12 kromozomlarında, hekza-nükleotitlerde A06 kromozomunda yer almaktadır (Çizelge 4.28).

Çizelge 4.28. Upland pamukta +/- sarmal için 5'-UTR dizilerinde SSR dağılımları

Krom	Di	Tri	Tetra	Penta	Hekza	Toplam
A01	7/0	7/0	0/0	1/0	9/0	24/0
A02	4/0	5/1	1/2	5/0	8/0	23/3
A03	9/1	6/2	1/1	4/0	3/0	23/4
A04	2/0	0/1	1/0	2/1	4/0	9/2
A05	17/0	12/1	4/0	5/0	11/1	49/2
A06	9/0	2/0	2/0	0/1	3/3	16/4
A07	6/0	13/0	1/0	2/0	6/0	28/0
A08	7/0	7/0	4/0	2/1	5/2	25/3
A09	7/0	10/1	1/0	3/0	9/2	30/3
A10	8/0	4/1	0/0	3/1	7/0	22/2
A11	9/1	8/2	2/2	5/0	10/1	34/6
A12	10/1	9/0	1/2	1/0	2/0	23/3
A13	5/0	9/1	2/2	0/0	6/0	22/3
D01	10/0	11/1	1/1	3/0	5/0	30/2
D02	4/0	7/1	5/0	3/0	11/0	30/1
D03	7/1	6/0	1/0	3/1	3/0	20/2
D04	4/1	9/4	2/0	2/0	7/0	24/5
D05	18/3	12/4	2/1	4/1	8/0	44/9
D06	10/2	10/1	2/0	2/0	6/0	30/3
D07	6/1	12/1	3/0	3/0	6/0	30/2
D08	6/0	9/1	6/0	3/0	4/1	28/2
D09	5/0	9/0	2/1	4/0	3/1	23/2
D10	9/0	7/1	3/0	1/0	5/0	25/1
D11	14/0	14/3	1/0	3/1	7/1	39/5
D12	11/1	9/0	2/1	3/1	12/1	37/4
D13	9/0	7/1	3/0	4/0	8/0	31/1

Çizelgede “/” işaretinin sol tarafı “+ sarmalı” sağ tarafı ise “– sarmalı” nitelemektedir.

Çizelge 4.29’da Upland pamukta +/- sarmal için 3'-UTR dizilerinde SSR dağılımları gösterilmektedir. Upland pamukta 3'-UTR bölgeleri toplam artı sarmal için 70, eksi sarmal için 320 SSR içermektedir. Çizelgede de görüleceği üzere en fazla bulunan SSR artı sarmal üzerinde bulunan tri-nükleotitler olup bunu sırasıyla hekza-, penta-, tetra- ve di-nükleotitler yer almaktadır. Eksi sarmal üzerinde ise en fazla bulunan tri-nükleotitler iken; bunu hekza-, tetra-, di- ve penta-nükleotitler izlemektedir. Bu türün kromozomları incelendiğinde artı sarmal üzerinde en fazla SSR di-nükleotitlerde A09 ve D11 kromozomunda, tri-nükleotitlerde D12 kromozomunda, tetra-nükleotitlerde A03 ve D10 kromozomunda, penta-nükleotitlerde A11 kromozomunda, hekza-nükleotitlerde D07 kromozomunda yer alırken; eksi sarmal üzerinde en fazla SSR di-nükleotitlerde

D05 kromozomunda, tri-nükleotitlerde D13 kromozomunda, tetra-nükleotitlerde A02 kromozomunda, penta-nükleotitlerde D02 ve D08 kromozomunda, hekza-nükleotitlerde ise A11 ve D02 kromozomlarında yer almaktadır (Çizelge 4.29).

Çizelge 4.29. Upland pamukta +/- sarmal için 3'-UTR dizilerinde SSR dağılımları

Krom	Di	Tri	Tetra	Penta	Hekza	Toplam
A01	0/2	0/3	0/2	0/1	0/2	0/10
A02	0/1	0/1	0/5	0/1	0/6	0/14
A03	0/1	0/3	2/2	1/2	1/3	4/11
A04	0/1	1/1	0/1	0/0	0/3	1/6
A05	1/1	2/7	0/2	0/1	1/3	4/14
A06	0/2	1/0	0/1	0/3	0/0	1/6
A07	0/1	0/2	1/0	0/1	0/5	1/9
A08	0/1	1/5	0/2	0/2	0/1	1/11
A09	2/1	1/2	0/0	0/0	0/2	3/5
A10	0/0	4/3	0/2	1/0	0/2	5/7
A11	0/5	0/3	0/3	3/3	0/7	3/21
A12	0/2	0/4	0/3	1/1	1/6	2/16
A13	0/2	1/3	1/1	0/1	1/3	3/10
D01	0/1	0/6	0/2	0/2	0/4	0/15
D02	1/4	1/3	0/1	1/4	0/7	3/19
D03	0/0	1/3	1/4	0/2	1/1	3/10
D04	0/2	1/1	0/0	0/3	0/2	1/8
D05	1/9	2/8	0/3	1/2	2/6	6/28
D06	1/0	1/2	0/1	0/3	1/3	3/9
D07	0/1	0/2	1/2	1/3	3/1	5/9
D08	0/1	0/1	0/2	0/4	1/6	1/14
D09	0/2	1/3	1/0	0/2	0/4	2/11
D10	0/3	0/2	2/1	0/1	2/1	4/8
D11	2/2	2/9	1/3	2/1	0/2	7/17
D12	0/0	5/4	0/4	1/1	0/5	6/14
D13	0/3	0/10	0/3	0/2	1/0	1/18

Çizelgede “/” işaretinin sol tarafı “+ sarmalı” sağ tarafı ise “– sarmalı” nitelemektedir.

Çizelge 4.30’da Upland pamukta +/- sarmal için intergen dizilerinde SSR dağılımları gösterilmektedir. Upland pamukta intergen bölgeleri toplam artı sarmal için 26492, eksi sarmal için 26716 SSR içermektedir. Çizelgede de görüleceği üzere en fazla bulunan SSR +/- sarmal üzerinde bulunan tri-nükleotitler olup bunu sırasıyla hekza-, tetra-, penta- ve di-nükleotitler yer almaktadır. Bu türün kromozomları incelendiğinde artı sarmal üzerinde en fazla SSR di-nükleotitlerde D05 kromozomunda, tri- ve tetra-nükleotitlerde A11 kromozomunda, penta- ve hekza-nükleotitlerde A05 kromozomunda yer alırken; eksi sarmal üzerinde en fazla SSR di-nükleotitlerde A05 kromozomunda, tri- ve tetra-nükleotitlerde A11 kromozomunda, penta- ve hekza-nükleotitlerde ise A05 kromozomunda yer almaktadır (Çizelge 4.30).

Çizelge 4.30. Upland pamukta +/- sarmal için intergen dizilerinde SSR dağılımları

Krom	Di	Tri	Tetra	Penta	Hekza	Toplam
A01	104/106	345/351	137/134	140/136	287/296	1013/1023
A02	96/99	287/292	135/127	132/134	271/275	921/927
A03	125/132	359/362	148/146	161/164	275/274	1068/1078
A04	71/72	282/283	107/103	118/119	208/209	786/786
A05	165/ 182	454/460	174/173	231/229	350/363	1374/1407
A06	129/135	368/371	145/144	196/188	300/299	1138/1137
A07	124/125	353/372	141/138	149/146	281/281	1048/1062
A08	158/160	376/383	131/131	172/166	342/340	1179/1180
A09	110/115	298/314	116/117	138/135	263/268	925/949
A10	117/123	452/455	174/172	184/187	319/323	1246/1260
A11	146/148	483/482	201/199	169/166	287/288	1286/1283
A12	135/140	370/379	172/166	163/160	294/293	1134/1138
A13	147/146	348/353	144/140	145/142	288/295	1072/1076
D01	129/133	366/373	158/155	130/131	237/240	1020/1032
D02	116/112	299/301	176/177	138/135	230/237	959/962
D03	93/98	241/246	118/113	99/99	186/189	737/745
D04	113/111	294/293	117/115	104/102	179/180	807/801
D05	179/180	417/403	165/157	141/137	235/241	1137/1118
D06	100/101	312/320	131/127	144/138	243/250	930/936
D07	111/114	324/336	145/142	123/122	202/215	905/929
D08	132/134	357/359	183/181	141/138	231/234	1044/1046
D09	119/117	292/305	102/101	105/107	225/231	843/861
D10	118/123	385/389	160/162	132/130	223/233	1018/1037
D11	128/130	384/386	153/148	137/135	256/255	1058/1054
D12	120/128	299/317	186/179	129/127	219/239	953/990
D13	101/105	342/344	148/140	124/123	176/187	891/899

Çizelgede “/” işaretinin sol tarafı “+ sarmalı” sağ tarafı ise “– sarmalı” nitelemektedir.

Çizelge 4.31’de Upland pamukta +/- sarmal için tüm gen dizilerinde SSR dağılımları gösterilmektedir. Upland pamukta tüm gen bölgeleri toplam artı sarmal için 2783, eksi sarmal için 2118 SSR içermektedir. Çizelgede de görüleceği üzere en fazla bulunan SSR artı sarmal üzerinde bulunan tri-nükleotitler olup bunu sırasıyla hekza-, di-, tetra- ve penta-nükleotitler yer almaktadır. Eksi sarmal üzerinde ise en fazla bulunan tri-nükleotitler iken; bunu sırasıyla hekza-, tetra-, penta- ve di-nükleotitler izlemektedir. Bu türün kromozomları incelendiğinde artı sarmal üzerinde en fazla SSR di- ve tri-nükleotitlerde D05 kromozomunda, tetra-nükleotitlerde D10 kromozomunda, penta-nükleotitlerde D05 kromozomunda, hekza-nükleotitlerde D13 kromozomunda bulunup, eksi sarmal üzerinde en fazla SSR di-, tri- ve tetra-nükleotitlerde D05 kromozomunda, penta-nükleotitlerde A05 kromozomunda, hekza-nükleotitlerde ise D11 kromozomunda yer almaktadır (Çizelge 4.31).

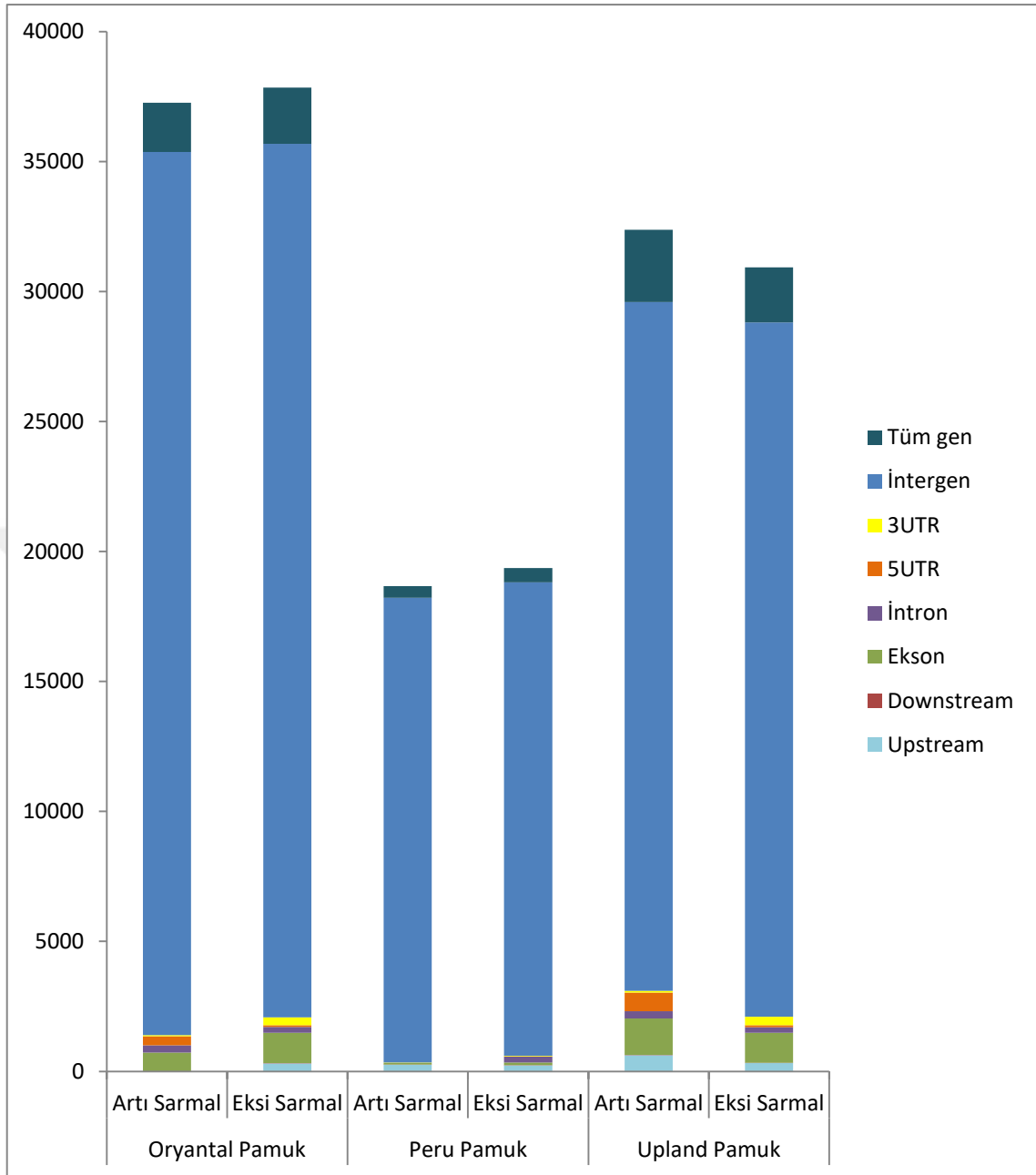
Çizelge 4.31. Upland pamukta +/- sarmal için tüm gen dizilerinde SSR dağılımları

Krom	Di	Tri	Tetra	Penta	Hekza	Toplam
A01	22/12	36/19	13/6	10/7	23/15	104/59
A02	11/5	40/13	10/14	6/7	19/14	86/53
A03	17/6	30/22	12/10	11/9	23/18	93/65
A04	8/3	8/19	10/13	5/2	12/5	43/42
A05	39/11	46/42	15/11	13/20	38/34	151/118
A06	17/11	22/22	13/5	3/13	15/24	70/75
A07	18/8	46/15	9/10	8/11	19/17	100/61
A08	24/13	34/26	11/8	6/17	19/15	94/79
A09	24/5	35/19	8/1	9/14	34/27	110/66
A10	22/5	27/14	11/3	7/3	24/13	91/38
A11	22/13	56/56	9/18	10/16	32/25	129/128
A12	21/9	40/30	14/17	4/10	18/22	97/88
A13	20/16	28/24	10/12	2/8	30/19	90/79
D01	30/9	48/39	19/9	9/12	21/15	127/84
D02	18/9	34/29	16/16	15/14	27/23	110/91
D03	15/11	27/26	14/14	8/10	21/16	85/77
D04	14/11	27/25	10/9	8/8	15/14	74/67
D05	47/28	61/68	13/20	20/18	34/27	175/161
D06	30/13	34/18	8/11	7/9	26/20	105/71
D07	21/4	37/35	14/12	8/9	28/15	108/75
D08	30/11	34/34	20/13	4/16	26/25	114/99
D09	26/11	38/27	12/9	13/5	20/19	109/71
D10	20/8	38/28	28/12	5/10	23/8	114/66
D11	44/17	56/64	10/14	8/9	35/ 38	153/142
D12	29/11	52/20	15/17	7/9	40/22	143/79
D13	20/13	35/36	17/15	15/9	21/11	108/84

Çizelgede “/” işaretinin sol tarafı “+ sarmalı” sağ tarafı ise “– sarmalı” nitelemektedir.

Şekil 4.14’de görüldüğü üzere 3 pamuk türünün gen elementlerine ait diziler bir arada incelendiğinde artı ve eksi sarmalda upstream bölgesi Upland pamukta 623/320 SSR içermesi ile 1. sırada yer alırken; Oryantal pamuk 445/307 SSR ile 2. sırada, Peru pamuğu 256/232 SSR ile 3. sırada yer almaktadır. Downstream bölgesi artı sarmalda 14 SSR ile Peru pamuğu ilk sırada olup 8 SSR ile Upland pamuk 2. sırada ve 5 SSR ile Oryantal pamuğu 3. sırada bulunmakla birlikte, eksi sarmalda Peru pamuğu ve Upland pamukta 13 SSR, Oryantal pamukta ise 12 SSR bulunmaktadır.

Ekson bölgesi artı sarmalda 1401 SSR ile Upland pamuk 1. sırada, 725 SSR ile Oryantal pamuk 2. sırada, 60 SSR ile Peru pamuk 3. sırada yer alırken; eksi sarmalda ise 1162 SSR ile Oryantal pamuk 1. sırada, 1154 SSR ile Upland pamuk 2. sırada, 102 SSR ile Peru pamuğu 3. sırada yer almaktadır. İntron bölgesi artı sarmalda 279 SSR ile Upland pamuk 1. sırada, Oryantal pamuk 277 SSR ile 2. sırada, 233 SSR ile Peru pamuğu 3. sırada bulunup eksi sarmalda sırasıyla Oryantal pamuk 233 SSR, Peru pamuk 221 ve Upland pamuk 217 SSR içermektedir. 5'-UTR bölgesi artı ve eksi sarmalda Upland pamukta 719/74 SSR ile 1. sırada bulunurken; Oryantal pamuk 346/62 SSR ile 2. sırada, Peru pamuğu ise 25/11 SSR ile 3. sırada bulunmaktadır. 3'UTR bölgesinde artı ve eksi sarmalda Upland pamuk 70/320 SSR ile ilk sırada olup bunu sırasıyla 49/305 SSR ile Oryantal pamuk ve 10/26 SSR ile Peru pamuğu takip etmektedir. İntergen bölgesinde artı ve eksi sarmalda Oryantal pamuk 33969/33598 SSR ile 1. sırada, Upland pamuk 26492/26716 SSR ile 2. sırada, Peru pamuğu ise 17879/18212 SSR ile 3. sırada yer almaktadır. Tüm gen incelendiğinde artı sarmalda 2783 SSR ile Upland pamuk ilk sırada, 1885 SSR ile Oryantal pamuk 2. sırada, 452 SSR ile Peru pamuk 3. sırada bulunurken; eksi sarmalda 2169 SSR ile Oryantal pamuk 1. sırada, 2118 SSR ile Upland pamuk 2. sırada 547 SSR ile Peru pamuk 3. sırada yer almaktadır (Şekil 4.14).



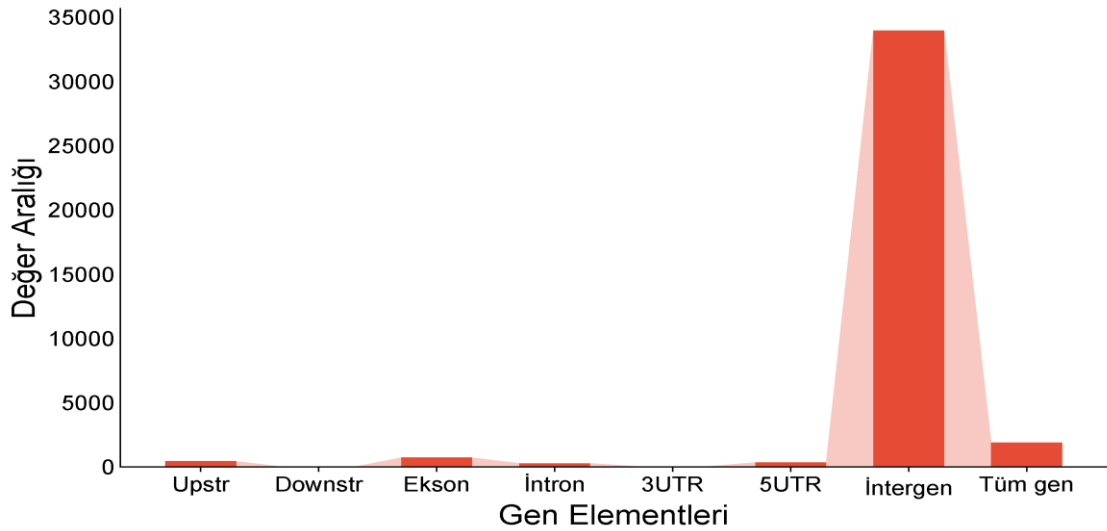
Şekil 4.14. Oryantal pamuk, Peru pamuk ve Upland pamukta gen element dizilerinin SSR yönünden karşılaştırılması

Üç pamuk türünün genom büyüklükleri ve gen elementleri bir arada ayrı ayrı incelendiğinde 1593,75 Mb genom büyüklüğü içeren A genomuna sahip Oryantal pamukta artı sarmalda ve eksi sarmalda en fazla bulunan dizilerin beklendiği gibi intergen dizileri olup %86'lık bir kısım içerirken; en az artı ve eksi sarmalda downstream'dir. 739,54 Mb genom büyüklüğü içeren D genomuna sahip Peru pamukta artı ve eksi sarmalda %91'lik kısım içeren intergen olup; en az artı sarmalda 3'-UTR, eksi sarmalda ise downstream bulunmaktadır. A alt genomuna sahip 1446,29 Mb genom büyüklüğü ve D alt genomu içeren 836,33 Mb genom büyüklüğü bulunan Upland

pamukta artı ve eksi sarmalda da en fazla intergen yer alırken; en az ise artı ve eksi sarmalda downstream gen elementi bulunmaktadır. Downstream’de az bulunmasının nedeni bu çalışmada 200 bp alındığı için kaynaklanmış olabilir.

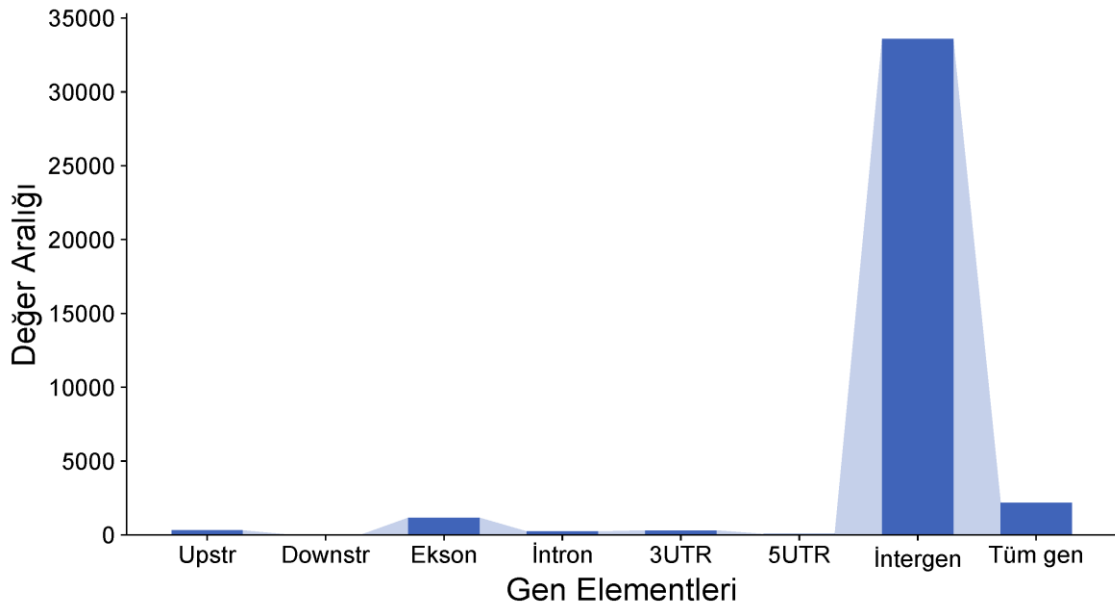
D genomu örülebilir pamuk tipleri olup intergenik SSR’ların A genomu pamuk tiplerine kıyasla daha fazla olması SSR’ların kültür ve pamuk lif özellikleriyle alakalı olduğunu gösterebilir. A genomu ile AD genomu karşılaştırıldığında artı sarmalda downstream ve intronlar benzerlik gösterirken; D genomu ile AD genomu karşılaştırıldığında benzerlik görülmemektedir. Eksi sarmal incelendiğinde A ve AD genomu arasında upstream, downstream, ekson, 3'-UTR ve 5'-UTR benzerlik gösterirken; D ve AD genomu incelendiğinde downstream ve intron sayıca benzerlik göstermektedir. A alt genomu ile D alt genomu kıyaslandığı zaman artı sarmalda upstream, downstream, intron ve 3'-UTR benzerlik görülmekte olup eksi sarmalda downstream ve 5'-UTR benzerlik göstermektedir. A alt genomu D alt genomuna göre gen elementlerinde SSR’lar da sayıca artış yaşanırken intergende azalış yaşanmıştır. Upland pamukta A genomunda benzerlerin olması tetraploit pamuğun atalarından birisidir. Donör olması da SSR karakterizasyon yönünden desteklemiştir. Gerçek D genomu eğrilebilir özelliği olmayan kısa lifli pamuk iken; tetraploitteki D genomu eğrilebilir, yüksek kaliteli pamuk üretir. Evrimleşme ile türleşmenin olduğunun bir kanıtıdır. Bu nedenle de diploit gen sayısı farklıdır.

Şekil 4.15 incelendiğinde Oryantal pamuğa ait (A genom) artı sarmal için gen elementlerinin karşılaştırılması gösterilmektedir. Bu grafik göz önüne alındığında intergenik bölgenin en fazla değere sahip olduğu bunu tüm gen, ekson, upstream, 5'-UTR, intron, 3'-UTR ve downstream dizileri izlemektedir. Bu gen ekspresyonunun ile SSR dağılımları arasında bir pozitif ilişki olabileceğini anlamına gelmektedir. Ekson ve upstream gen elementlerinin kodlama bölgelerinin ve gen düzenleme bölgelerinde önemli olduğunu göstermektedir. 5'-UTR ve 3'-UTR bölgeleri daha düşük sıralamada yer alıyor, bu da bu bölgelerin artı sarmallarda daha az temsil edildiğini gösterir. Ancak; bu bölgelerin gen ekspresyonu üzerindeki etkileri yine de önemli olabilir. İntron ve downstream bölgeleri en düşük sıralamada yer alıyor, bu da bu bölgelerin pozitif etkiler açısından daha az temsil edildiğini gösterir (Şekil 4.15). Gen elementlerinin çoğunluğunun artı sarmallarda yüksek temsil edilmesi, bu bölgelerin gen ekspresyonunu destekleyici veya teşvik edici işlevler gördüğünü gösterir. Eksonlar ve upstream bölgeler gibi kodlama ve düzenleyici bölgelerin yüksek temsil edilmesi, bu bölgelerin genlerin pozitif düzenlenmesinde önemli olduğunu göstermektedir.



Şekil 4.15. Oryantal pamukta (A genom) artı sarmal için gen elementlerinin karşılaştırılması

Şekil 4.16’da Oryantal pamuğa ait (A genom) eksi sarmal için gen elementlerinin karşılaştırılması gösterilmektedir. Bu karşılaştırmaya göre gen elementleri içerisinde en yüksek değer aralığına sahip olan intergenik bölge olup, bunu sırasıyla tüm gen, ekson, upstream, 3'-UTR, intron, 5'-UTR ve downstream takip etmektedir. İntergen ve tüm gen bölgeleri yine yüksek değerlerde temsil edilmektedir. Ancak; bu bölgelerdeki yüksek yoğunluk, negatif etkilerin bu bölgelerde yoğun olduğunu gösterebilir. Ekson ve upstream bölgeleri, eksi sarmalda da önemli bir temsil oranına sahiptir. Bu, gen kodlama bölgelerinin ve gen düzenleme bölgelerinin negatif etkiler açısından belirgin olduğunu gösterebilir. 3'-UTR bölgeleri eksi sarmalda daha yüksek sıralamada yer alıyor, bu da bu bölgelerin gen ekspresyonunu baskılayıcı bir rol oynayabileceğini öne sürebilir. 5'-UTR ve downstream bölgeleri ise eksi sarmalda düşük sıralamada bulunduğundan bu bölgelerin negatif etkiler açısından daha az temsil edildiğini göstermektedir (Şekil 4.16). Eksi sarmallarda gen elementlerinin temsil sıralamasındaki değişiklikler, bu bölgelerin gen ekspresyonunu baskılayan veya düzenleyen etkiler yarattığını öne sürebilir. Özellikle 3'-UTR bölgelerinin yüksek sıralamada olması, bu bölgelerin negatif düzenleyici etkiler taşıdığını gösterebilir.

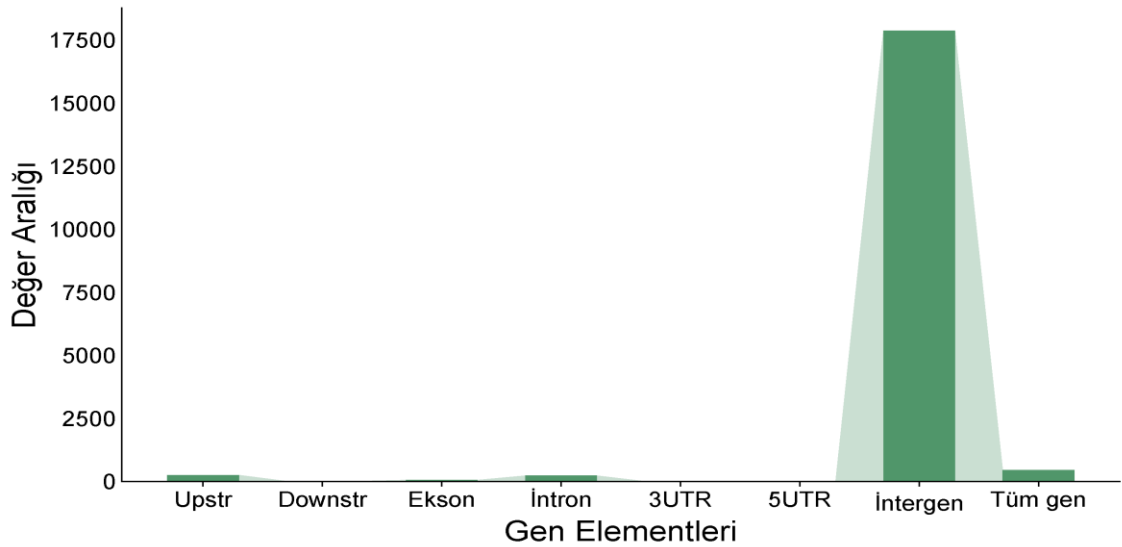


Şekil 4.16. Oryantal pamukta (A genom) eksi sarmal için gen elementlerinin karşılaştırılması

Şekil 4.17’de Peru pamukta (D genom) artı sarmal için gen elementlerinin karşılaştırılması gösterilmektedir. Bu grafiğe göre en yüksek değere sahip olan intergenik bölge olmakla beraber bunu tüm gen, upstream, intron, ekson, 5'-UTR, downstream ve 3'-UTR izlemektedir. İntergen ve tüm gen bölgelerinin yüksek sıralamada olması, bu bölgelerin artı sarmalda pozitif etkiler açısından kritik olduğunu ve gen ekspresyonunu teşvik eden düzenleyici elemanlar olarak işlev gördüğünü göstermektedir.

Upstream bölgelerinin yüksek sıralamada olması, gen düzenleyici elemanların gen ekspresyonunu pozitif yönde etkileyebileceğini ve genlerin bu bölgelerdeki düzenleyici etkilerden faydalandığını öne sürer. İntronların yüksek sıralamada bulunması, bu bölgelerin artı sarmalda gen ekspresyonunun pozitif düzenlemesinde önemli rol oynadığını gösterebilir.

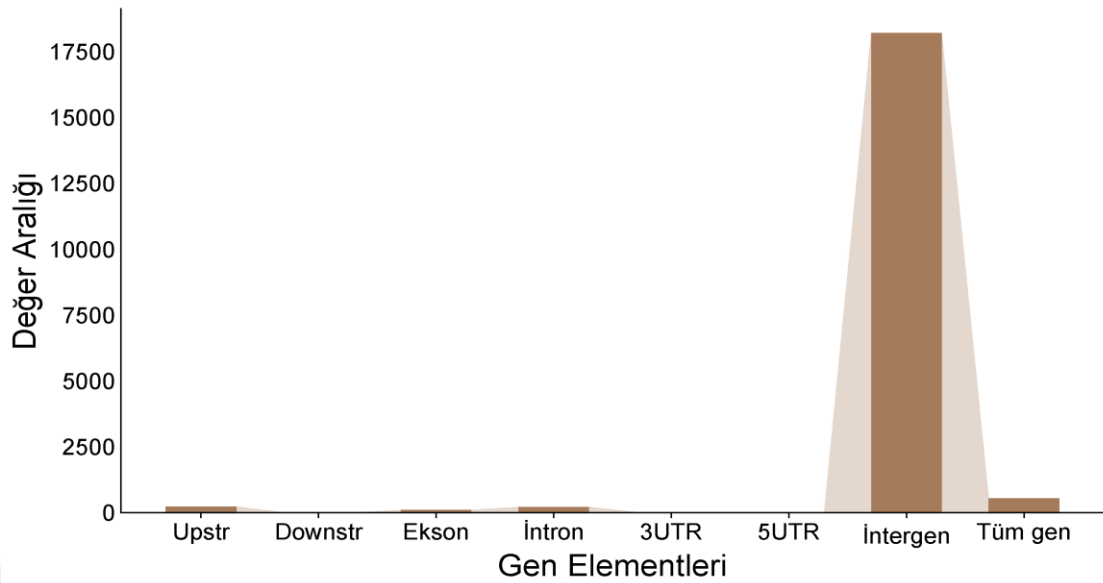
İntronlar, genlerin işlevselliği ve düzenlenmesinde önemli etkiler sağlayabilir. Eksonların daha düşük sıralamada olması, bu bölgelerin pozitif etkiler açısından daha az belirgin olduğunu gösterebilir. Ancak, eksonlar yine de önemli gen kodlama bölgeleridir. 5'-UTR ve 3'-UTR bölgelerinin artı sarmalda daha düşük sıralamada olması, bu bölgelerin pozitif düzenlemede daha az etkili olduğunu öne sürebilir. Ancak, 5'-UTR ve 3'-UTR bölgeleri, gen ekspresyonunun düzenlenmesinde önemli roller oynayabilir. Downstream bölgelerinin en düşük sıralamada bulunması, bu bölgelerin artı sarmalda gen ekspresyonunu destekleyici etkiler açısından daha az temsil edildiğini gösterebilir (Şekil 4.17).



Şekil 4.17. Peru pamukta (D genom) artı sarmal için gen elementlerinin karşılaştırılması

Şekil 4.18’de Peru pamukta (D genom) eksi sarmal için gen elementlerinin karşılaştırılması gösterilmektedir. Bu karşılaştırmaya göre en fazla intergenik bölge yer alırken; bunu tüm gen, upstream, intron, ekson, 3'-UTR, downstream ve 5'-UTR izlemektedir. Eksi sarmalda ise intergen, tüm gen ve upstream bölgeleri eksi sarmalda da yüksek sıralamada yer almakta, bu bölgelerin negatif etkiler açısından da önemli olduğunu ve gen ekspresyonunun baskılandığını öne sürebilir. İntronlar eksi sarmalda da yüksek sıralamada, bu da bu bölgelerin gen ekspresyonunu negatif olarak düzenlemede önemli rol oynayabileceğini öne sürer.

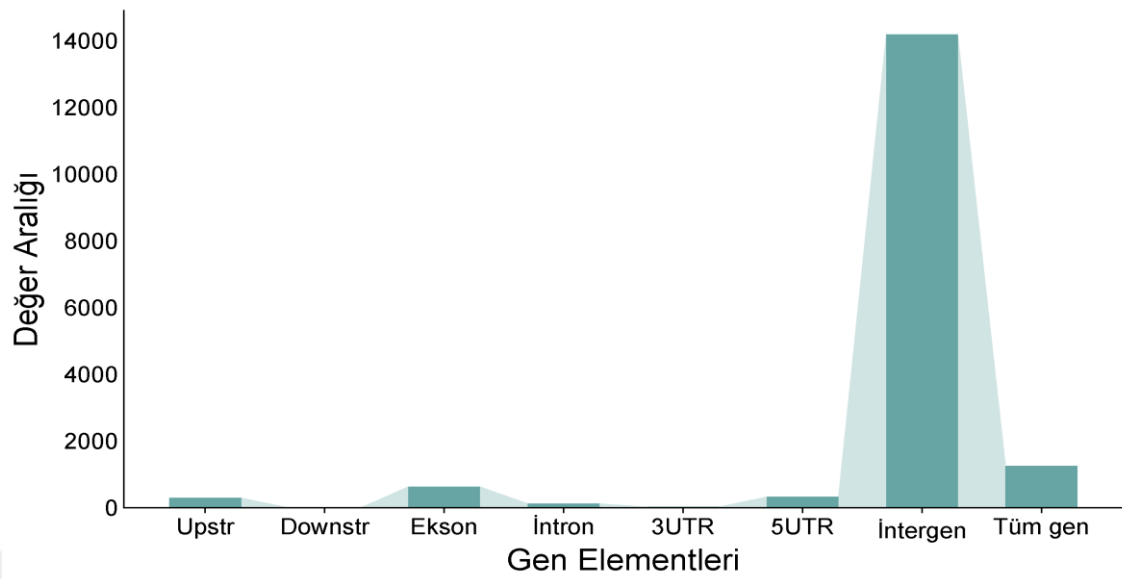
Eksonların eksi sarmalda daha düşük sıralamada bulunması, eksonların negatif etkiler açısından daha az belirgin olduğunu gösterebilir. Ancak, eksonlar gen kodlama işlevi açısından hâlâ önemli rol oynamaktadır. 3'-UTR bölgeleri eksi sarmalda daha yüksek sıralamada, bu bölgelerin gen ekspresyonunu negatif olarak düzenleyici bir rol oynayabileceğini gösterir. Downstream bölgeleri de düşük sıralamada, bu da negatif düzenlemede daha az etkili olduğunu öne sürer. 5'-UTR bölgesi eksi sarmalda en düşük sıralamada yer almaktadır. Bu, bu bölgenin negatif etkiler açısından en az temsil edildiğini ve dolayısıyla gen ekspresyonunun baskılanmasında daha az etkili olduğunu gösterebilir (Şekil 4.18).



Şekil 4.18. Peru pamukta (D genom) eksi sarmal için gen elementlerinin karşılaştırılması

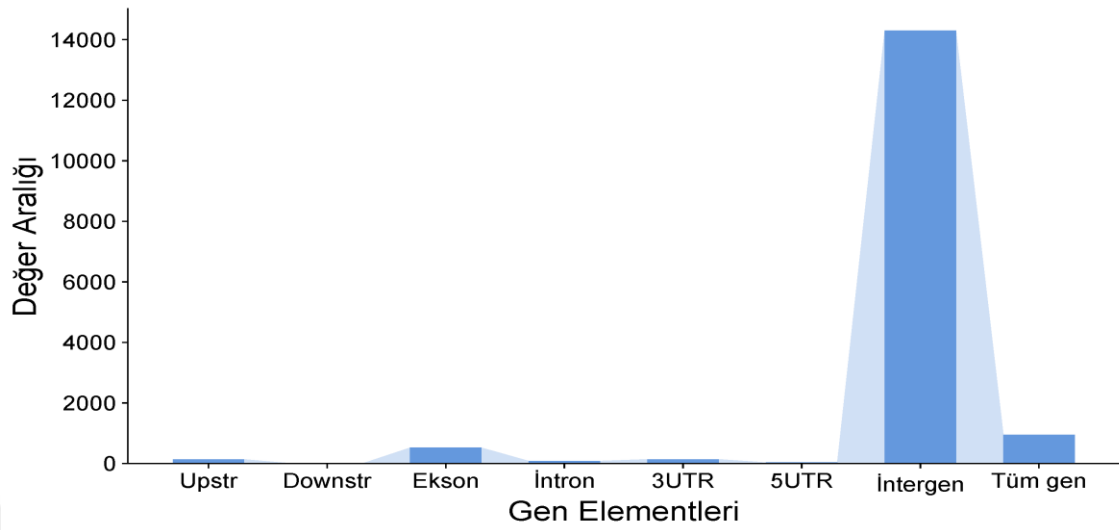
Şekil 4.19’da Upland pamuğa ait (A alt genom) artı sarmal için gen elementleri karşılaştırıldığında ilk sırada intergenik bölge olmakla beraber bunu tüm gen, ekson, 5'-UTR, upstream, intron, 3'-UTR ve downstream yer almaktadır. Bu grafik incelendiğinde intergen bölgelerinin genlerin pozitif düzenlemesinde merkezi bir rol oynadığını, tüm genin ise gen ekspresyonunun desteklenmesi, bu bölgelerin genetik düzenleme açısından etkili olduğunu göstermektedir. Eksonların yüksek sıralamada bulunması, bu bölgelerin pozitif düzenlemede önemli rol oynadığını ve gen kodlama işlevinde etkili olduğunu gösterir. Eksonlar, protein sentezinde kritik rol oynar ve gen ekspresyonunun artmasını destekleyebilir. 5'-UTR bölgeleri, artı sarmalda eksonlardan sonra gelen bir sıralamada yer almaktadır. Bu, bu bölgelerin pozitif düzenlemede önemli rol oynadığını ancak eksonlar kadar belirgin olmadığını öne sürebilir. 5'-UTR bölgeleri translasyonun başlama aşamasında kritik rol oynar.

Upstream bölgeleri, genlerin başlangıç bölgelerindeki düzenleyici elemanları içerir ve pozitif düzenlemede önemli rol oynayabilir. Ancak, artı sarmalda eksonlar ve 5'-UTR bölgelerinin altında kalmaktadır, bu da upstream bölgelerin pozitif etkilerde daha az belirgin olduğunu öne sürebilir. İntronların düşük sıralamada olması, bu bölgelerin artı sarmalda pozitif düzenlemede daha az etkili olduğunu öne sürebilir. İntronlar, genlerin işlevselliğini etkileyebilir ancak artı sarmalda daha az belirgin olabilirler. 3'-UTR bölgelerinin en düşük sıralamada bulunması, bu bölgelerin pozitif düzenlemede daha az etkili olduğunu öne sürebilir. 3'-UTR bölgeleri gen ekspresyonunun sonlandırılması ve düzenlenmesinde rol oynar. Downstream bölgeleri en düşük sıralamada, bu da bu bölgelerin artı sarmalda pozitif düzenlemede daha az etkili olduğunu gösterir. Gen ekspresyonunun düzenlenmesinde downstream bölgelerin rolü minimal olabilir (Şekil 4.19).



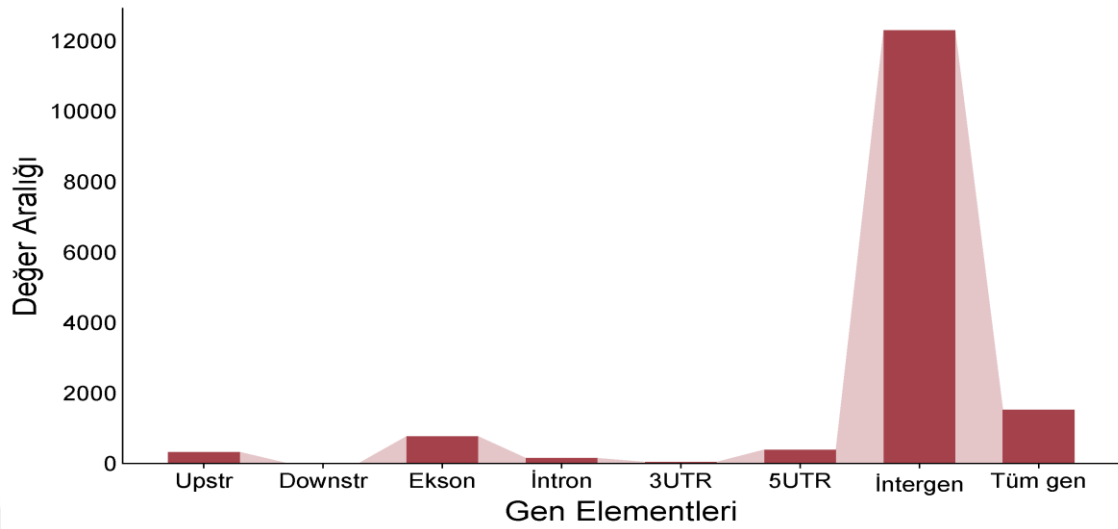
Şekil 4.19. Upland pamukta (A alt genom) artı sarmal için gen elementlerinin karşılaştırılması

Şekil 4.20’de Upland pamuğa ait (A alt genom) eksi sarmal da gen elementlerinin karşılaştırılması gösterilmektedir. Bu karşılaştırılmaya göre ilk sırada intergenik bölge olup bunu tüm gen, ekson, 3'-UTR, upstream, intron, 5'-UTR ve downstream izlemektedir. Eksi sarmalda da intergen bölgelerinin yüksek sıralamada bulunması, bu bölgelerin gen ekspresyonunu negatif yönde etkileyen düzenleyici elemanlar içerdiğini gösterir. Bu, intergen bölgelerinin negatif düzenlemede merkezi bir rol oynadığını ve gen ekspresyonunun baskılanmasında önemli olduğunu öne sürer. Tüm genlerin yüksek sıralamada olması, gen ekspresyonunun negatif düzenlemede önemli rol oynadığını ve genetik materyalin negatif düzenleme açısından etkili olduğunu göstermektedir. Eksonlar, eksi sarmalda da yüksek sıralamada bulunur, bu da eksonların negatif düzenlemede etkili olduğunu ve gen ekspresyonunu baskılayan bir rol oynayabileceğini gösterir. Eksonlar gen kodlama bölgeleri olup, negatif etkiler açısından belirgin olabilirler. 3'-UTR bölgelerinin eksi sarmalda orta-yüksek sıralamada yer alması, bu bölgelerin negatif düzenlemede önemli rol oynadığını öne sürer. 3'-UTR bölgeleri, gen ekspresyonunun sonlandırılması ve düzenlenmesinde kritik olabilir. Upstream bölgeleri, negatif düzenlemede de önemli rol oynar ancak eksonlar ve 3'-UTR bölgelerinin altında kalır. Bu, upstream bölgelerin negatif düzenlemede daha az belirgin olduğunu öne sürebilir. İntronların düşük sıralamada bulunması, bu bölgelerin negatif düzenlemede daha az etkili olduğunu öne sürebilir. İntronlar gen ekspresyonunu etkileyebilir, ancak eksi sarmalda daha az belirgin olabilirler. 5'-UTR bölgelerinin en düşük sıralamada bulunması, bu bölgelerin negatif düzenlemede daha az etkili olduğunu öne sürebilir. 5'-UTR bölgeleri gen ekspresyonunun başlama aşamasında önemli olabilir ancak; negatif düzenlemede etkileri minimal olabilir. Downstream bölgeleri eksi sarmalda en düşük sıralamada olup bu da bu bölgelerin negatif düzenlemede daha az etkili olduğunu gösterir. Gen ekspresyonunun düzenlenmesinde downstream bölgelerinin rolü minimal olabilir (Şekil 4.20).



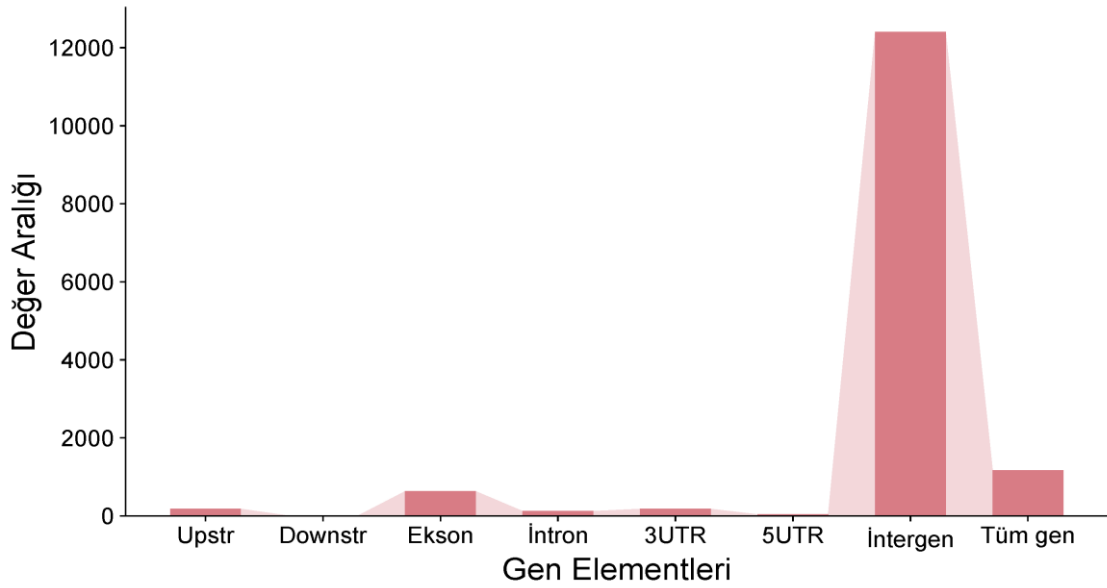
Şekil 4.20. Upland pamukta (A alt genom) eksi sarmal için gen elementlerinin karşılaştırılması

Şekil 4.21’de görüldüğü üzere Upland pamuğa ait (D alt genom) artı sarmal için gen elementleri karşılaştırıldığında intergen en fazla değere sahip iken; bunu tüm gen, ekson, 5'-UTR, upstream, intron, 3'-UTR ve downstream izlemektedir. Artı sarmalda intergen bölgelerinin en yüksek sıralamada bulunması, bu bölgelerin gen ekspresyonunun pozitif düzenlemesinde kritik bir rol oynadığını ve genler arasındaki düzenleyici elemanların bu bölgelerde yoğun olduğunu gösterir. Gen ekspresyonunu teşvik eden cis-düzenleyici elemanların intergen bölgelerinde bol bulunabilir. Tüm genler artı sarmalda yüksek sıralamada yer almaktadır. Bu durum, genlerin pozitif düzenlemede etkili olduğunu ve gen ekspresyonunu artırmaya yönelik önemli bir rol oynadığını gösterir. Eksonların orta sıralamada bulunması, bu bölgelerin pozitif düzenlemede belirgin bir rol oynadığını ancak; diğer bölgeler kadar etkili olmadığını öne sürer. Eksonlar, protein kodlama bölgeleri olarak önemli olmasına rağmen, artı sarmalda etkileri daha az belirgin olabilir. 5'-UTR bölgeleri artı sarmalda orta-yüksek sıralamada yer alır. Bu, bu bölgelerin pozitif düzenlemede etkili olduğunu ve transkripsiyonun başlama aşamasında önemli rol oynayabileceğini gösterir. Upstream bölgeleri, genlerin başlangıç bölgelerindeki düzenleyici elemanları içerip bu bölgelerin pozitif düzenlemede etkili olduğunu ancak; etkisinin bazı bölgeler kadar belirgin olmadığı görülmektedir. İntronlar orta-düşük sıralamada yer almaktadır. Bu, intronların pozitif düzenlemede etkisinin sınırlı olabileceğini gösterir. İntronlar, gen ekspresyonu üzerinde düzenleyici etkilere sahip olabilir ancak; bu etkiler artı sarmalda daha az belirgin olabilir. 3'-UTR bölgeleri artı sarmalda düşük sıralamada bulunur. Bu, bu bölgelerin pozitif düzenlemede etkisinin sınırlı olduğunu gösterir. 3'-UTR bölgeleri gen ekspresyonunun son aşamalarında rol oynar. Downstream bölgeleri en düşük sıralamada bulunmaktadır. Bu durum bu bölgelerin artı sarmalda pozitif düzenlemede en az etkili olduğunu öne sürer. Gen ekspresyonunun düzenlenmesinde downstream bölgelerinin rolü minimal olabilir (Şekil 4.21).



Şekil 4.21. Upland pamukta (D alt genom) artı sarmal için gen elementlerinin karşılaştırılması

Şekil 4.22’de Upland pamukta (D alt genom) eksi sarmal için gen elementlerinin karşılaştırılması yapıldığında, intergenik bölgenin en yüksek değer ile 1. sırada bulunup, bunu sırasıyla tüm gen, ekson, upstream, 3'-UTR, intron, 5'-UTR ve downstream yer almaktadır. Eksi sarmalda intergen bölgelerinin en yüksek sıralamada olması, bu bölgelerin negatif düzenlemede kritik bir rol oynadığını ve genler arasındaki düzenleyici elemanların bu bölgelerde yoğun olduğunu gösterir. Gen ekspresyonunu baskılayan cis-düzenleyici elemanlar burada bol bulunabilir. Tüm genlerin yüksek sıralamada yer alması, gen ekspresyonunun negatif düzenlemesinde önemli bir rol oynadığını gösterir. Genlerin negatif düzenlemede etkili olduğunu ve gen ekspresyonunun baskılandığını öne sürer. Eksonlar orta sıralamada yer alır. Bu, eksonların negatif düzenlemede belirgin bir rol oynayabileceğini ancak; diğer bölgeler kadar etkili olmayabileceğini gösterir. Eksonlar, protein kodlama bölgeleri olarak önemli olabilir ancak; negatif düzenlemede etkileri sınırlı olabilir. Upstream bölgeleri eksi sarmalda orta-yüksek sıralamada bulunur. Bu, bu bölgelerin negatif düzenlemede etkili olabileceğini ve gen ekspresyonunu baskılayan düzenleyici elemanlar içerebileceğini gösterir. 3'-UTR bölgeleri eksi sarmalda orta-düşük sıralamada bulunur. Bu, bu bölgelerin negatif düzenlemede sınırlı etkisi olduğunu öne sürebilir. 3'-UTR bölgeleri gen ekspresyonunun son aşamalarında rol oynar. İntronlar düşük sıralamada yer almaktadır. Bu, intronların negatif düzenlemede etkisinin sınırlı olduğunu ve diğer bölgeler kadar belirgin olmadığını gösterir. 5'-UTR bölgeleri eksi sarmalda daha düşük sıralamada bulunur. Bu, bu bölgelerin negatif düzenlemede en az etkili olduğunu öne sürer. 5'-UTR bölgeleri gen ekspresyonunun başlama aşamasında önemli olabilir ancak; negatif düzenlemedeki etkileri minimal olabilir. Downstream bölgeleri eksi sarmalda en düşük sıralamada bulunmaktadır. Bu, bu bölgelerin negatif düzenlemede en az etkili olduğunu ve gen ekspresyonunun baskılanmasında minimal rol oynadığını gösterir (Şekil 4.22).



Şekil 4.22. Upland pamukta (D alt genom) eksi sarmal için gen elementlerinin karşılaştırılması

Gossypium türlerinde, SSR'lar gen elementlerinde görülmekte ve kromozom boyunca dağılmaktadır. Bu da genomun geniş bir şekilde kullanıldığını göstermektedir. Diğer markırlarla birleştirildiğinde SSR'lar genomu incelemek için ideal araçlardır. SSR'ların alt genomlar ve kromozomlar üzerindeki eşit olmayan dağılımı, SSR motiflerinin dağılımını belirlememize yardımcı olabilir ve SSR dizilerinin fonksiyonu, pamuk genomundaki diğer hedef genlere yardımcı olabilir. Basit dizi tekrarları (SSR'lar), ökaryotik genomlarda hem kodlayan hem de kodlamayan bölgelerde bulunan genom alanlarıdır. Genom alanlarından fazla SSR'lar bir genom özelliği, genom karakterizasyonudur. Her ne kadar kodlayan (örnek EST-SSR'lar) genellikle düşük polimorfizm ile karakterize edilse de kodlayan bölgedeki genik SSR'ların polimorfizm düzeyi düşük ama intergenik bölgelerde polimorfizm düzeyi yüksektir. Genomik SSR'lar yüksek polimorfizm düzeyine sahiptir. Bunların DNA-yan dizileri, ilgilenilen özellikleri kontrol eden genlerle ilişkili oldukları için genetik araştırmalara ve üremeye yardımcı olabilecek yararlı bir DNA markır kaynağı olabilir.

Çizelge 4.32'de Oryantal pamukta +/- sarmal için 2000up ve 200down gen elementlerinin SSR ve sıklık dağılımları gösterilmektedir. Oryantal pamukta 2000up gen elementinde toplam 435/307 SSR içermektedir. Kromozom düzeyinde en çok SSR içeren artı sarmalda A11, eksi sarmalda A03 kromozomudur. 2000up gen elementinin kromozomlar bazında sıklığı incelendiğinde en fazla artı sarmalda yer alan A05 kromozomu olurken; eksi sarmalda A03 ve A05 kromozomlarıdır. 200down gen elementi incelendiğinde ise toplam 5/12 SSR bulunmaktadır. Kromozom düzeyinde en fazla SSR artı sarmalda A02, A03, A05, A08 ve A10 kromozomları olup eksi sarmalda A12 kromozomudur. 200down gen bölgesinin sıklık dağılımları ele alındığında artı sarmalda A05, eksi sarmalda ise A12 kromozomu yer almaktadır.

Çizelge 4.32. Oryantal pamukta + ve - sarmal için kromozomlara göre 2000up ve 200down gen elementlerinin SSR ve sıklık dağılımları

Krom	Büyüklik (Mb)	UPR +/-	UPR sıklık	DWR +/-	DWR sıklık
A01	123,41	29/21	0,23/0,17	0/0	0/0
A02	115,15	22/13	0,19/0,11	1/0	0,009/0
A03	139,24	28/45	0,20/0,32	1/0	0,007/0
A04	122,51	28/18	0,23/0,15	0/0	0/0
A05	93,16	38/30	0,41/0,32	1/2	0,010/0,021
A06	143,37	32/16	0,22/0,11	0/1	0/0,007
A07	106,03	31/17	0,29/0,16	0/0	0/0
A08	142,3	38/27	0,27/0,19	1/1	0,007/0,007
A09	89,4	36/16	0,40/0,18	0/2	0/0,022
A10	131,76	41/17	0,31/0,13	1/2	0,008/0,016
A11	137,85	46/38	0,33/0,28	0/1	0/0,007
A12	119,98	36/27	0,30/0,23	0/3	0/0,025
A13	129,59	30/22	0,23/0,17	0/0	0/0
Toplam	1593,75	435/307	0,27/0,19	5/12	0,003/0,007

*UPR: upstream, *DWR: downstream, Çizelgede “/” işaretinin sol tarafı “+ sarmalı” sağ tarafı ise “- sarmalı” nitelemektedir.

Çizelge 4.33’te Oryantal pamukta +/- sarmal için ekson ve intron gen elementlerinin SSR ve sıklık dağılımları gösterilmektedir. Oryantal pamukta ekson bölgesi için toplam 725/1169 SSR içermektedir Kromozom düzeyinde en fazla SSR artı sarmalda A05 ve A11, eksi sarmalda A11 kromozomundadır Ekson gen elementinin sıklığı incelendiğinde en yüksek sayı ile artı ve eksi sarmalda A05 kromozomudur. İtron gen bölgesi için ise toplam 277/233 SSR içermektedir. Kromozom düzeyinde en fazla SSR artı sarmalda A11, eksi sarmalda A12 kromozomlarında yer almaktadır. İtron gen elementinin sıklığı incelendiğinde en yüksek sayı ile artı sarmalda A09 olurken; eksi sarmalda A12 kromozomlarıdır.

Çizelge 4.33. Oryantal pamukta + ve - sarmal için kromozomlara göre ekson ve intron gen elementlerinin SSR ve sıklık dağılımları

Krom	Büyüklik (Mb)	Ekson +/-	Ekson sıklık	İtron +/-	İtron sıklık
A01	123,41	46/78	0,37/0,63	11/22	0,09/0,18
A02	115,15	38/51	0,33/0,44	24/7	0,21/0,07
A03	139,24	45/128	0,32/0,92	19/10	0,14/0,07
A04	122,51	38/69	0,31/0,56	12/14	0,09/0,11
A05	93,16	83/130	0,89/1,40	17/18	0,18/0,19
A06	143,37	66/74	0,46/0,52	19/5	0,13/0,03
A07	106,03	53/69	0,50/0,65	23/21	0,22/0,20
A08	142,3	52/86	0,37/0,60	30/17	0,21/0,12
A09	89,4	54/60	0,60/0,67	27/7	0,30/0,08

Çizelge 4.33'ün devamı

Krom	Büyüklik (Mb)	Ekson +/-	Ekson sıklık	İntron +/-	İntron sıklık
A10	131,76	65/90	0,49/0,68	22/22	0,17/0,17
A11	137,85	83/137	0,60/0,99	31/21	0,22/0,15
A12	119,98	60/127	0,50/1,05	18/47	0,15/0,39
A13	129,59	42/70	0,32/0,54	24/22	0,19/0,17
Toplam	1593,75	725/1169	0,45/0,73	277/233	0,17/0,15

Çizelgede “/” işaretinin sol tarafı “+ sarmalı” sağ tarafı ise “- sarmalı” nitelemektedir.

Çizelge 4.34'te Oryantal pamukta +/- sarmal için 5'-UTR ve 3'-UTR gen elementlerinin SSR ve sıklık dağılımları gösterilmektedir. Oryantal pamukta 5'-UTR gen elementinde toplam 346/62 SSR bulunmaktadır Kromozomlara göre SSR sayısı ile sıklık dağılımları ele alındığında en fazla artı sarmalda A05, eksi sarmalda ise A11 ve A12 kromozomları olmaktadır. 3'-UTR gen elementine göre toplam 49/305 SSR içermektedir. Kromozom düzeyinde en fazla SSR sayısı ve sıklığı incelendiğinde artı sarmalda A11; eksi sarmalda A12 kromozomunda yer almaktadır.

Çizelge 4.34. Oryantal pamukta + ve - sarmal için kromozomlara göre 5'-UTR ve 3'-UTR gen elementlerinin SSR ve sıklık dağılımları

Krom	Büyüklik (Mb)	5'-UTR +/-	5'-UTR sıklık	3'-UTR +/-	3'-UTR sıklık
A01	123,41	20/2	0,162/0,016	3/9	0,024/0,073
A02	115,15	18/2	0,156/0,017	1/10	0,008/0,086
A03	139,24	26/7	0,186/0,050	3/25	0,021/0,179
A04	122,51	24/5	0,195/0,040	4/14	0,032/0,114
A05	93,16	51/6	0,547/0,064	5/31	0,053/0,332
A06	143,37	28/2	0,195/0,013	5/21	0,034/0,146
A07	106,03	27/3	0,254/0,028	0/20	0/0,188
A08	142,3	23/5	0,161/0,035	3/14	0,021/0,098
A09	89,4	24/2	0,268/0,022	2/16	0,022/0,179
A10	131,76	27/4	0,204/0,030	6/25	0,046/0,190
A11	137,85	31/11	0,224/0,079	12/41	0,088/0,297
A12	119,98	26/9	0,216/0,075	4/56	0,033/0,466
A13	129,59	21/4	0,162/0,030	1/23	0,008/0,177
Toplam	1593,75	346/62	0,217/0,038	49/305	0,030/0,191

Çizelgede “/” işaretinin sol tarafı “+ sarmalı” sağ tarafı ise “- sarmalı” nitelemektedir.

Çizelge 4.35'te Peru pamukta +/- sarmal için kromozomlara göre 2000up ve 200down gen elementlerinin SSR ve sıklık dağılımları gösterilmektedir. Peru pamukta 2000up gen elementi için toplam 256/242 SSR içermektedir. Kromozom düzeyinde en

fazla SSR artı sarmalda D11, eksi sarmalda D13 kromozomundadır. 2000up gen elementinin sıklığı incelendiğinde en yüksek sayı ile artı ve eksi sarmalda D13 kromozomudur. 200down gen elementi incelendiğinde toplam 14/13 bulunmaktadır. Kromozom düzeyinde en fazla SSR artı sarmalda D08 olurken; eksi sarmalda D09 kromozomudur. 200down gen bölgesinin sıklık dağılımları ele alındığında en yüksek sayı ile artı sarmalda D08, eksi sarmalda D10 kromozomlarıdır.

Çizelge 4.35. Peru pamukta artı ve eksi sarmal için kromozomlara göre 2000up ve 200down gen elementlerinin SSR ve sıklık dağılımları

Krom	Büyüklik (Mb)	UPR +/-	UPR sıklık	DWR +/-	DWR sıklık
D01	52,42	21/10	0,400/0,190	0/1	0/0,019
D02	48,27	15/8	0,310/0,165	1/1	0,020/0,020
D03	63,02	8/13	0,126/0,206	0/0	0/0
D04	50,65	19/23	0,375/0,454	2/0	0,039/0
D05	57,14	27/30	0,472/0,525	0/1	0/0,017
D06	60,4	6/7	0,099/0,115	0/0	0/0
D07	53,5	14/15	0,261/0,280	1/0	0,018/0
D08	62,61	18/20	0,287/0,319	6/1	0,095/0,015
D09	49,93	19/21	0,380/0,420	0/5	0/0,100
D10	61,94	20/15	0,322/0,242	1/2	0,015/ 0,031
D11	63,19	44/23	0,696/0,363	1/1	0,015/0,015
D12	57,72	6/13	0,103/0,225	1/0	0,017/0
D13	58,75	39/34	0,663/0,578	1/1	0,017/0,017
Toplam	739,54	256/242	0,346/0,302	14/13	0,018/0,017

*UPR: upstream, *DWR: downstream, Çizelgede “/” işaretinin sol tarafı “+ sarmalı” sağ tarafı ise “- sarmalı” nitelemektedir.

Çizelge 4.36’da Peru pamukta +/- sarmal için kromozomlara göre ekson ve intron gen elementlerinin SSR ve sıklık dağılımları gösterilmektedir. Peru pamukta ekson gen elementinde toplam 60/102 SSR içermektedir. Kromozom düzeyinde en fazla SSR sayısı ve sıklık yönünden artı sarmalda D01, eksi sarmalda D13 kromozomunda yer almaktadır. İtron gen elementinde ise toplam 233/221 SSR bulunmaktadır. Kromozom düzeyinde en fazla SSR artı sarmalda D10, eksi sarmalda D11 kromozomlarında yer almaktadır İtron gen bölgesinin kromozomlara göre sıklık dağılımları ele alındığında en yüksek sayı ile artı sarmalda D10 olup; eksi sarmalda D04 kromozomu olmaktadır.

Çizelge 4.36. Peru pamukta + ve - sarmal için kromozomlara göre ekson ve intron gen elementlerinin SSR ve sıklık dağılımları

Krom	Büyüklik (Mb)	Ekson +/-	Ekson sıklık	İtron +/-	İtron sıklık
D01	52,42	9/4	0,171/0,076	20/20	0,381/0,381
D02	48,27	8/8	0,165/0,165	8/12	0,165/0,248
D03	63,02	2/5	0,031/0,079	6/14	0,095/0,222

Çizelge 4.36'nın devamı

Krom	Büyüklik (Mb)	Ekson +/-	Ekson sıklık	İntron +/-	İntron sıklık
D04	50,65	3/3	0,059/0,059	12/24	0,236/ 0,473
D05	57,14	6/9	0,105/0,157	27/19	0,472/0,332
D06	60,4	2/5	0,033/0,082	9/3	0,149/0,049
D07	53,5	4/9	0,074/0,168	11/15	0,205/0,280
D08	62,61	4/8	0,063/0,127	23/14	0,367/0,223
D09	49,93	1/12	0,020/0,240	24/21	0,480/0,420
D10	61,94	5/5	0,080/0,080	30/24	0,484/0,387
D11	63,19	5/13	0,079/0,205	25/ 26	0,395/0,411
D12	57,72	4/3	0,069/0,051	16/11	0,277/0,190
D13	58,75	7/18	0,119/0,306	22/18	0,374/0,306
Toplam	739,54	60/102	0,081/0,137	233/221	0,315/0,298

Çizelgede “/” işaretinin sol tarafı “+ sarmalı” sağ tarafı ise “– sarmalı” nitelemektedir.

Çizelge 4.37’de Peru pamukta +/- sarmal için kromozomlara göre 5'-UTR ve 3'-UTR gen elementlerinin SSR ve sıklık dağılımları gösterilmektedir. Peru pamukta 5'-UTR gen elementi için toplam 25/11 SSR bulunmaktadır. Kromozom düzeyinde en fazla SSR sayısı ve sıklık yönünden artı sarmalda D01, eksi sarmalda D11 kromozomunda yer almaktadır. 3'-UTR gen bölgesi için ise toplam 10/26 SSR içermektedir. Kromozom düzeyinde en fazla SSR sayısı ve sıklık açısından artı sarmalda D13, eksi sarmalda D02 kromozomlarıdır.

Çizelge 4.37. Peru pamukta artı ve eksi sarmal için kromozomlara göre 5'-UTR ve 3'-UTR gen elementlerinin SSR ve sıklık dağılımları

Krom	Büyüklik (Mb)	5'-UTR +/-	5'-UTR sıklık	3'-UTR +/-	3'-UTR sıklık
D01	52,42	5/1	0,095/0,019	2/2	0,038/0,038
D02	48,27	2/0	0,041/0	2/4	0,041/ 0,082
D03	63,02	0/0	0/0	0/1	0/0,015
D04	50,65	2/0	0,039/0	0/2	0/0,039
D05	57,14	3/2	0,052/0,035	0/2	0/0,035
D06	60,4	0/0	0/0	1/2	0,016/0,033
D07	53,5	3/1	0,056/0,018	0/1	0/0,018
D08	62,61	3/1	0,047/0,015	0/0	0/0
D09	49,93	0/1	0/0,020	0/4	0/0,080
D10	61,94	0/0	0/0	2/2	0,032/0,032
D11	63,19	3/3	0,047/0,047	0/2	0/0,031
D12	57,72	1/0	0,017/0	0/1	0/0,017
D13	58,75	3/2	0,051/0,034	3/3	0,051/0,051
Toplam	739,54	25/11	0,033/0,014	10/26	0,013/0,035

Çizelgede “/” işaretinin sol tarafı “+ sarmalı” sağ tarafı ise “– sarmalı” nitelemektedir.

Çizelge 4.38’de Upland pamukta +/- sarmal için kromozomlara göre 2000up ve 200down gen elementlerinin SSR ve sıklık dağılımları gösterilmektedir. Upland pamukta 2000up gen elementinde toplam 623/320 SSR bulunmaktadır. Kromozom düzeyinde en fazla SSR artı sarmalda D05, eksi sarmalda A05 kromozomundadır. 2000up gen elementinin sıklığı incelendiğinde artı sarmalda en yüksek sayı D05 kromozomunda olurken; eksi sarmalda D12 kromozomunda yer almaktadır. 200down gen elementi incelendiğinde toplam 8/13 SSR içermektedir. Kromozom düzeyinde en fazla SSR sayısı ve sıklığı incelendiğinde artı sarmalda D12 olurken; eksi sarmalda D01 kromozomudur.

Çizelge 4.38. Upland pamukta artı ve eksi sarmal için kromozomlara göre 2000up ve 200down gen elementlerinin SSR ve sıklık dağılımları

Krom	Büyüklik (Mb)	UPR +/-	UPR sıklık	DWR +/-	DWR sıklık
A01	119,76	32/15	0,267/0,125	0/0	0/0
A02	108,14	24/7	0,221/0,064	0/1	0/0,009
A03	113,69	28/7	0,246/0,061	0/0	0/0
A04	89,18	12/5	0,134/0,056	0/1	0/0,011
A05	111,10	38/24	0,342/0,216	1/0	0,009/0
A06	129,20	13/8	0,100/0,061	0/0	0/0
A07	98,90	19/5	0,192/0,050	0/0	0/0
A08	127,50	17/12	0,133/0,094	0/0	0/0
A09	85,34	23/11	0,269/0,128	0/1	0/0,011
A10	118,18	28/11	0,236/0,093	0/0	0/0
A11	124,18	18/15	0,144/0,120	1/1	0,008/0,008
A12	109,47	28/11	0,255/0,100	0/0	0/0
A13	111,65	18/7	0,161/0,062	0/0	0/0
Toplam	1446,29	298/138	0,259/0,207	2/4	0,001/0,003
D01	65,21	28/11	0,429/0,168	1/3	0,015/0,046
D02	72,19	22/17	0,304/0,235	0/0	0/0
D03	54,96	18/9	0,327/0,163	0/0	0/0
D04	58,23	17/1	0,291/0,017	0/1	0/0,017
D05	66,48	40/16	0,601/0,240	1/0	0,015/0
D06	66,68	17/16	0,254/0,239	0/1	0/0,014
D07	59,44	25/11	0,420/0,185	0/0	0/0
D08	69,43	19/12	0,273/0,172	0/0	0/0
D09	54,44	23/17	0,422/0,312	1/2	0,018/0,036
D10	68,09	39/17	0,572/0,249	0/0	0/0
D11	72,82	23/18	0,315/0,247	1/1	0,013/0,013
D12	63,26	21/23	0,331/0,363	2/1	0,031/0,015
D13	65,10	33/14	0,506/0,215	0/0	0/0
Toplam	836,33	325/182	0,388/0,217	6/9	0,007/0,010
GenelΣ	2282,62	623/320	0,272/0,140	8/13	0,003/0,006

*UPR: upstream, *DWR: downstream, Çizelgede “/” işaretinin sol tarafı “+ sarmalı” sağ tarafı ise “- sarmalı” nitelemektedir.

Çizelge 4.39’da Upland pamukta +/- sarmal için kromozomlara göre ekson ve intron gen elementlerinin SSR ve sıklık dağılımları gösterilmektedir. Upland pamukta ekson gen bölgesi için toplam 1401/1154 SSR içermektedir. Kromozom düzeyinde en fazla SSR artı sarmalda A05, eksi sarmalda D05 kromozomundadır. Ekson gen elementinin sıklığı incelendiğinde artı sarmalda en yüksek sayı D12 kromozomunda olup, eksi sarmalda D05 kromozomunda yer almaktadır. İtron gen elementi için ise toplam 279/217 SSR bulunmaktadır. Kromozom düzeyinde en fazla SSR sayısı artı sarmalda A01 iken; eksi sarmalda D05 kromozomudur. İtron gen bölgesinin sıklığı ele alındığında en fazla görüldüğü kromozomlar artı sarmalda D13, eksi sarmalda ise D05 kromozomlarıdır.

Çizelge 4.39. Upland pamukta artı ve eksi sarmal için kromozomlara göre ekson ve intron gen elementlerinin SSR ve sıklık dağılımları

Krom	Büyüklik (Mb)	Ekson +/-	Ekson sıklık	İtron +/-	İtron sıklık
A01	119,76	42/32	0,350/0,267	20/5	0,167/0,041
A02	108,14	40/37	0,369/0,342	9/4	0,083/0,036
A03	113,69	44/40	0,387/0,351	12/10	0,105/0,087
A04	89,18	15/19	0,168/0,231	7/5	0,078/0,056
A05	111,10	91/54	0,819/0,486	10/11	0,090/0,099
A06	129,20	37/41	0,286/0,317	3/6	0,023/0,046
A07	98,90	52/34	0,525/0,343	14/8	0,141/0,080
A08	127,50	43/42	0,337/0,329	6/8	0,047/0,062
A09	85,34	56/32	0,656/0,374	10/11	0,117/0,128
A10	118,18	45/27	0,380/0,228	9/2	0,076/0,016
A11	124,18	73/78	0,587/0,628	14/12	0,112/0,096
A12	109,47	51/44	0,465/0,401	6/8	0,054/0,073
A13	111,65	43/42	0,385/0,376	7/4	0,062/0,035
Toplam	1446,29	632/522	0,436/0,360	127/94	0,087/0,064
D01	65,21	58/40	0,889/0,613	16/4	0,245/0,061
D02	72,19	52/48	0,720/0,664	11/17	0,152/0,235
D03	54,96	48/37	0,873/0,673	9/12	0,163/0,218
D04	58,23	32/34	0,549/0,583	5/9	0,085/0,154
D05	66,48	82/99	1,233/1,489	15/21	0,225/0,315
D06	66,68	59/49	0,884/0,734	9/8	0,134/0,119
D07	59,44	62/43	1,043/0,723	6/5	0,100/0,084
D08	69,43	52/56	0,748/0,806	12/7	0,172/0,100
D09	54,44	49/37	0,900/0,679	10/6	0,183/0,110
D10	68,09	60/28	0,881/0,411	13/9	0,190/0,132
D11	72,82	79/69	1,084/0,947	18/15	0,247/0,205
D12	63,26	80/43	1,264/0,679	10/5	0,158/0,079
D13	65,10	56/49	0,860/0,752	18/5	0,276/0,076
Toplam	836,33	769/632	0,919/0,755	152/123	0,181/0,147
GenelΣ	2282,62	1401/1154	0,613/0,505	279/217	0,122/0,095

Çizelgede “/” işaretinin sol tarafı “+ sarmalı” sağ tarafı ise “- sarmalı” nitelemektedir.

Çizelge 4.40'ta Upland pamukta +/- sarmal için kromozomlara göre 5'-UTR ve 3'-UTR gen elementlerinin SSR ve sıklık dağılımları gösterilmektedir. Upland pamukta 5'-UTR gen elementinde toplam 719/74 bulunmaktadır. Kromozom düzeyinde en fazla SSR sayısı ve sıklığı artı sarmalda A05 kromozomunda iken; D05 kromozomunda ise SSR sayısı olarak eksi sarmalda en fazla olup, sıklık olarak hem artı sarmal hem de eksi sarmalda ilk sırada yer almaktadır. 3'-UTR gen bölgesi için ise toplam 70/320 SSR içermektedir. Kromozom düzeyinde en fazla SSR sayısı ve sıklığı artı sarmalda D11, eksi sarmalda D05 kromozomundadır.

Çizelge 4.40. Upland pamukta +/- sarmal için kromozomlara göre 5'-UTR ve 3'-UTR gen elementlerinin SSR ve sıklık dağılımları

Krom	Büyüklik (Mb)	5'-UTR +/-	5'-UTR sıklık	3'-UTR +/-	3'-UTR sıklık
A01	119,76	24/0	0,200/0	0/10	0/0,083
A02	108,14	23/3	0,212/0,027	0/14	0/0,129
A03	113,69	23/4	0,202/0,035	4/11	0,035/0,096
A04	89,18	9/2	0,100/0,022	1/6	0,011/0,067
A05	111,10	49/2	0,441/0,018	4/14	0,036/0,126
A06	129,20	16/4	0,123/0,030	1/6	0,007/0,046
A07	98,90	28/0	0,283/0	1/9	0,010/0,091
A08	127,50	25/3	0,196/0,023	1/11	0,007/0,086
A09	85,34	30/3	0,351/0,035	3/5	0,035/0,058
A10	118,18	22/2	0,186/0,016	5/7	0,042/0,059
A11	124,18	34/6	0,273/0,048	3/21	0,024/0,169
A12	109,47	23/3	0,210/0,027	2/16	0,018/0,146
A13	111,65	22/3	0,197/0,026	3/10	0,026/0,089
Toplam	1446,29	328/35	0,226/0,024	28/140	0,019/0,096
D01	65,21	30/2	0,460/0,030	0/15	0/0,230
D02	72,19	30/1	0,415/0,013	3/19	0,041/0,263
D03	54,96	20/2	0,363/0,036	3/10	0,054/0,181
D04	58,23	24/5	0,412/0,085	1/8	0,017/0,137
D05	66,48	44/9	0,661/0,135	6/28	0,090/0,421
D06	66,68	30/3	0,449/0,044	3/9	0,044/0,134
D07	59,44	30/2	0,504/0,033	5/9	0,084/0,151
D08	69,43	28/2	0,403/0,028	1/14	0,014/0,201
D09	54,44	23/2	0,422/0,036	2/11	0,036/0,202
D10	68,09	25/1	0,367/0,014	4/8	0,058/0,117
D11	72,82	39/5	0,535/0,068	7/17	0,096/0,233
D12	63,26	37/4	0,584/0,063	6/14	0,094/0,221
D13	65,10	31/1	0,476/0,015	1/18	0,015/0,276
Toplam	836,33	391/39	0,467/0,046	42/180	0,050/0,215
GenelΣ	2282,62	719/74	0,314/0,032	70/320	0,030/0,140

Çizelgede “/” işaretinin sol tarafı “+ sarmalı” sağ tarafı ise “- sarmalı” nitelemektedir.

4.2. Mikrosatellit İçeren Selüloz Sentaz ve Selüloz Sentaz Benzeri Genler ve Filogenetik İlişkiler

Gossypium cinsi, tekstil endüstrisinin ana hammaddesi olan pamuk türlerini içermesi nedeniyle bu türler arasındaki genetik çeşitliliği ve adaptasyon stratejilerini anlamak için genetik analizler büyük önem taşımaktadır. Bu çalışmada, *Gossypium arboreum*, *Gossypium raimondii* ve *Gossypium hirsutum* türlerinde mikrosatellit (SSR) içeren selüloz sentaz (*CesA*) ve selüloz sentaz benzeri (*Csl*) genlerin barkod geni olarak filogenetik çalışmalarda kullanımı ve pamuk bitkilerinin genetik kaynaklarının korunması ve iyileştirilmesi konularına ilişkin temel bilgi sağlamaktadır.

Çizelge 4.41’de listelenen Oryantal pamuğa ait 38 gen arasında 5’i *CesA* ve *CesA-like* genlerinden oluşmakta iken 6 tanesi selüloz sentaz benzeri (*Csl*) genlerinden oluşmaktadır. Mikrosatellit tanımlama kriterlerine göre bu genlerde bulunan toplam 16 mikrosatellit tanımlanmıştır. Ancak 16 mikrosatellit 11 gende mevcut olup bazı genlerde daha fazla sayıda SSR bulunmaktadır. Çizelge 4.41’de görülebileceği gibi bu genlerin 6 tanesi artı sarmalda, 5 tanesi ise eksi sarmalda yer almaktadır.

Çizelge 4.41. *Gossypium arboreum*’a ait *CesA* ve *Csl* genlerini, motif içeriklerini ve sarmal bilgilerini içeren mikrosatellitlerin listesi

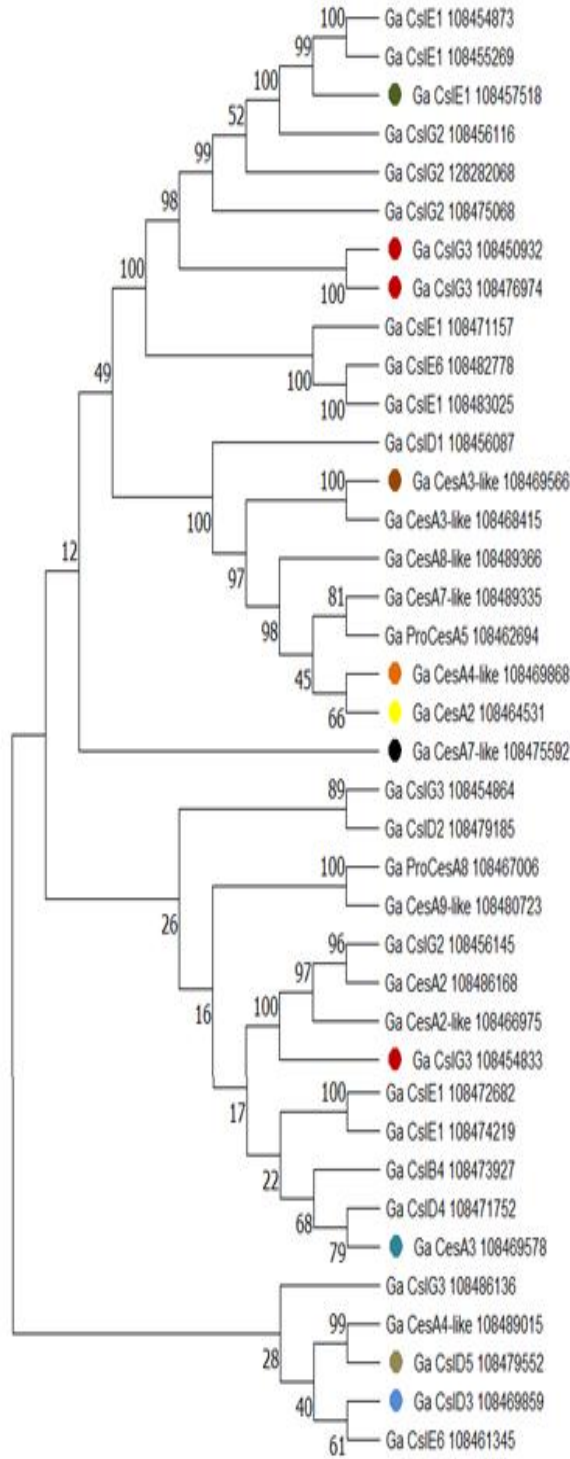
UID No	Gen	Kromozom	Motif	Sarmal
108469566	<i>CesA3-like</i>	A08	[GA] ₂₂	(+)
108457518	<i>CslE1</i>	A09	[AT] ₁₀	(+)
108464531	<i>CesA2</i>	A12	[AG] ₁₂ [TA] ₉	(+)
108476974	<i>CslG3</i>	A12	[TA] ₁₁	(+)
108475592	<i>CesA7-like</i>	A03	[AT] ₈ [AG] ₁₉ [CT] ₉	(-)
108469578	<i>CesA3</i>	A08	[AT] ₂₈	(-)
108469868	<i>CesA4-like</i>	A08	[TTC] ₈	(+)
108454833	<i>CslG3</i>	A09	[AAT] ₅	(-)
108450932	<i>CslG3</i>	A11	[TACA] ₁₈	(+)
108479552	<i>CslD5</i>	A12	[TAT] ₅ [ATTT] ₅ [TTTAA] ₄	(-)
108469859	<i>CslD3</i>	A08	[AGAAAG] ₅	(-)

Kromozomlar arasında en fazla *CesA* ve *Csl* geni A08 kromozomunda olup 2’si *CesA-like*, 1’i *CesA* ve *Csl* geninden oluşmaktadır. A08 kromozomu 4 tane SSR içermekte olup en fazla tekrarlanan eksi sarmaldaki *CesA3* genine ait AT motifi 28 kez tekrarlanmıştır. Ancak; SSR yönünden incelendiğinde 6 SSR içermesi ile en fazla A12 kromozomu olmaktadır. Tespit edilmiş olan 16 motif arasında en yüksek sayı di-

nükleotitlerdir. En düşük sayı ise en az tekrar sayısına da sahip olan penta-nükleotitler [TTTTA]₄ ve hekza-nükleotitlerdir [AGAAAG]₅.

řekil 4.23’de Oryantal pamuđa ait mikrosatellit içeren selüloz sentaz ve selüloz sentaz benzeri genlerin kullanılmasıyla elde edilen filogenetik analizi gösterilmiştir. Bu pamuk türüne ait 38 *CesA* ve *Csl* nükleotit dizisi Maksimum olasılık yöntemi ve Tamura-Nei modelini kullanılarak çalışılmıştır. Yapılan analizde 8929 pozisyon (nükleotit ekleme-çıkarma-değişimi) bulunmakta olup MEGA11 yazılım programı kullanılarak elde edilmiştir (Tamura vd. 2021). İlgili filogenetik analiz incelendiğinde 11 SSR içeren *CesA* ve *Csl* geninin filogenetik pozisyonlarının göreceli olarak benzer pozisyonlarda bulunduğu görülmektedir (řekil 4.23).





Şekil 4.23. Oryantal pamukta mikrosatellit içeren selüloz sentaz ve selüloz sentaz benzeri genlerin filogenetik analizi

SSR içeren ID numarası 108457518 olan *CsIE1* geni ile ID numaraları sırasıyla 108454873 ve 108455269 olan *CsIE1* genleri ve ID numarası 108456116 *CsIG2* gen ile benzerlik göstermektedir. ID numaraları 108450932 ile 108476974 SSR içerip ID numaraları 128282068 *CsIG2* ve 108475068 *CsIG2*, 108471157 *CsIE1*, 108482778 *CsIE6*, 108483025 *CsIE1* genleriyle yakından ilişkilidir. SSR içeren ID'si 108469566 *CesA3-like* geni, 108468415 *CesA3-like*, 108456087 *CsID1*, 108489366 *CesA8-like*, 108489335 *CesA7-like*, 108462694 *ProCesA5* ile SSR içeren 108469868 *CesA4-like* ve 108464531 *CesA2* geniyle benzerlik gösterir. ID numarası 108475592 *CesA7-like* geni SSR içerirken diğer genlere göre daha mesafeli konumdadır. ID numarası 108454833 *CsIG3* geni SSR içerip 108466975 *CesA2-like*, 108486168 *CesA2*, 108456145 *CsIG2* genleri ile benzerlik göstermektedir. SSR içeren 108469578 *CesA3* geni, ID numaraları 108471752 *CsID4* ve 108473927 *CsIB4* genleriyle yakından ilişkilidir. 108479552 *CsID5* geni SSR içermekte 108489015 *CesA4-like* geni ile benzerlik gösterirken, SSR bulunduran ID numarası 108469859 *CsID3* geni, 108461345 *CsIE6* ile yakından ilişkilidir.

Çizelge 4.42'de listelenen Peru pamuğa ait 46 gen arasında 6 tanesi *CesA*, *CesA-like* ve olası *CesA* genleri iken; 8'i selüloz sentaz benzeri (*Csl*) genlerinden oluşmaktadır. Mikrosatellit (SSR) tanımlama kriterlerine göre gen içeren toplam 19 mikrosatellit tanımlanmıştır. Çizelge 4.42'de görüldüğü üzere 8 tanesi artı sarmalda, 6 tanesi ise eksi sarmalda bulunmaktadır.

Çizelge 4.42. *Gossypium raimondii*'ye ait *CesA* ve *Csl* genlerini, motif içeriklerini ve sarmal bilgilerini içeren mikrosatellitlerin listesi

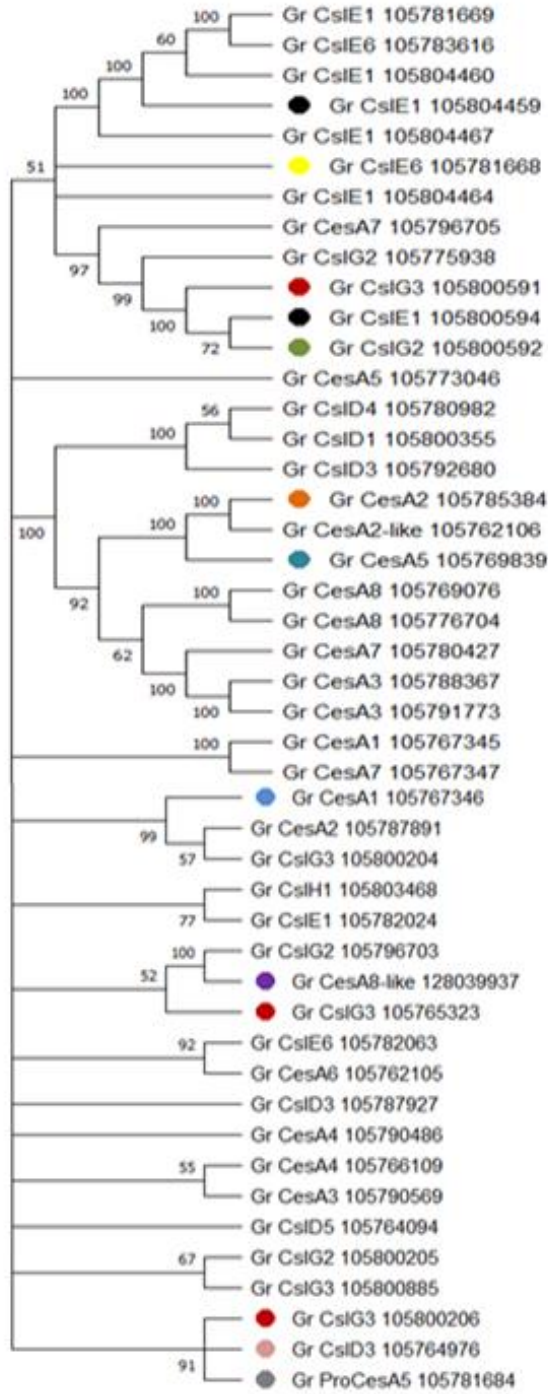
UID No	Gen	Kromozom	Motif	Sarmal
105800592	<i>CsIG2</i>	D09	[AT] ₁₄	(+)
105804459	<i>CsIE1</i>	D11	[TA] ₉	(+)
105765323	<i>CsIG3</i>	D12	[TA] ₉	(+)
105781668	<i>CsIE6</i>	D13	[TA] ₁₅	(+)
128039937	<i>CesA8-like</i>	D03	[AT] ₁₈ [TA] ₁₁ [TGG] ₇	(-)
105767346	<i>CesA1</i>	D05	[AT] ₁₃ [TTA] ₅ [TTTA] ₄ [CGACC] ₅	(-)
105769839	<i>CesA5</i>	D05	[TGC] ₅	(+)
105800594	<i>CsIE1</i>	D09	[ATT] ₅	(+)
105785384	<i>CesA2</i>	D11	[TTA] ₇	(+)
105800206	<i>CsIG3</i>	D09	[TAA] ₁₇	(-)
105764976	<i>CsID3</i>	D13	[AAT] ₅	(-)
105781684	<i>Prob-CesA5</i>	D13	[AAAT] ₄	(-)
105800591	<i>CsIG3</i>	D09	[GGTGC] ₄ [TGCAA] ₄	(+)

D09 ve D05 kromozomları en fazla *CesA* ve *CsI* geni içerdiği belirlenmiştir. En fazla tekrarlanan ise D03 kromozomunda bulunan eksi sarmaldaki *CesA8-like* genine ait AT motifi olup 18 kez tekrarlanmıştır. En az tekrarlanan ise D09 kromozomunda bulunan artı sarmaldaki *CsIG3* genine ait penta-nükleotitler olan [GGTGC]₄, [TGCAA]₄ motifleri ve D13 kromozomundaki eksi sarmalda bulunan *Prob-CesA5* genine ait olan [AAAT]₄ motifleridir. 19 motif arasında en yüksek sayı di- ve tri-nükleotitlerdir. En düşük sayı ise tetra-nükleotitlerdir.

Şekil 4.24'te Peru pamuğa ait mikrosatellit içeren selüloz sentaz ve selüloz sentaz benzeri genlerin filogenetik analizi gösterilmiştir. Bu pamuk türüne ait 46 *CesA* ve *CsI* nükleotit dizisi Maksimum olasılık yöntemi ve Tamura-Nei modelini kullanılarak çalışılmıştır. Yapılan analizde 22881 pozisyon bulunmakta olup MEGA11 yazılım programı kullanılarak gerçekleştirilmiştir (Tamura vd. 2021).

İlgili filogenetik analizde 13 gende SSR bulunmakta olup bu SSR içeren genlerin filogenetik pozisyonlarının benzer olduğu görülmektedir. ID numarası 105804459 olan *CsIE1* geni SSR içermekte olup ID numaraları sırasıyla 105781669, 105804460 ve 105804467 olan *CsIE1* genleri ve ID numarası 105783616 *CsIE6* ile benzerlik göstermektedir. SSR içeren ID numarası 105781668 *CsIE6* geni, 105804464 *CsIE1*, 105796705 *CesA7*, 105775938 *CsIG2*, SSR içeren ID numarası 105800591 *CsIG3*, ID numarası 105800594 *CsIE1* ve ID numarası 105800592 *CsIG2* genleri ile benzerlik gösterir. SSR içeren ID numarası 105785384 *CesA2* ve 105769839 *CesA5* genleri 105762106 *CesA2-like* geni ile benzerlik göstermektedir. ID numarası 105767346 *CesA1* SSR içermekle beraber ID numaraları sırasıyla 105787891 *CesA2* ile ID numarası 105800204 olan *CsIG3* geni ile yakından ilişkilidir. ID numarası 128039937 *CesA8-like* geni ve 105765323 *CsIG3* genleri SSR içerip ID numarası 105796703 olan *CsIG2* geni ile benzerlik gösterirken; ID numaraları sırasıyla 105800206 *CsIG3*, 105764976 *CsID3* ve 105781684 *ProCesA5* genlerinde SSR içermekle birlikte birbirleriyle yakından ilişkilidir.

İlgili filogenetik analizde en çok SSR içeren *CsI* ailesi olan *CsIE* genlerin filogenetik pozisyonları birbirine yakın olmakla birlikte diğer *CsI* ailesinden de belirgin olarak ayrıldığı görülmektedir. Diğer bir gen grubu olan *CesA* ailesinin de mikrosatellit içerikleri nedeniyle filogenetik pozisyonları ilişkilendirilmiştir. Verilerin birlikte incelenmesine de mikrosatellit içerikleri yönünden genlerin evrimsel dönüşümlerinin de benzer olabileceği, mikrosatellitlerin dizilerinde bulundukları genlerin sadece fonksiyonları değil aynı zamanda evrimsel dönüşümlerinde de önemli olabileceği şeklinde yorumlanmıştır. Ancak; konuyla ilgili detaylı çalışmaların yapılmasında yararlar olacaktır.



Şekil 4.24. Peru pamukta için mikrosatellit içeren selüloz sentaz ve selüloz sentaz benzeri genlerin filogenetik analizi

Çizelge 4.43'te listelenen Upland pamuğa ait 82 *CesA* ve *Csl* geni arasında 12 tanesi *CesA*, *CesA-like* ve olası *CesA* genleri iken; 10'u selüloz sentaz benzeri (*Csl*) genlerinden oluşmaktadır. Mikrosatellit tanımlama kriterlerine göre genleri içeren toplam 27 mikrosatellit tanımlanmıştır. Çizelge 4.43'te görüldüğü üzere 9 tanesi artı sarmalda olup 13 tanesi eksi sarmalda bulunmaktadır. Diploit türlerdeki mikrosatellit sayıları tetraploit türün mikrosatellit sayısından daha az olduğu görülmekle birlikte, iki diploit türdeki toplam sayı olan 35 sayısından daha az olduğu görülmektedir. Bu bilgi ışığında poliploit oluşumu matematiksel toplam olarak genom katlamasının olmadığı mikrosatellit sayısında da görülmüştür.

Çizelge 4.43. *Gossypium hirsutum*'a ait *CesA* ve *Csl* genlerini, motif içeriklerini ve sarmal bilgilerini içeren mikrosatellitlerin listesi

UID No	Gen	Kromozom	Motif	Sarmal
107928590	<i>CslG2</i>	A09	[AT] ₈	(+)
107914099	<i>CesA5</i>	D05	[TC] ₈ [TGC] ₅	(+)
107900907	<i>CslI1</i>	D08	[AG] ₁₁	(+)
107896452	<i>CslG3</i>	A09	[TA] ₉	(-)
107907451	<i>CesA7</i>	D05	[TA] ₁₂ [TAA] ₆ [TTTA] ₄	(-)
107953681	<i>CesA7-like</i>	D07	[AC] ₈	(-)
107935226	<i>CslE1</i>	D11	[TA] ₈	(-)
121229536	<i>CesA5-like</i>	A05	[TGC] ₅	(+)
107955410	<i>CesA4</i>	A07	[AAG] ₅	(+)
107898377	<i>CslI1</i>	A11	[ATT] ₅	(+)
107937655	<i>CesA4-like</i>	D07	[AAG] ₅	(+)
107911480	<i>CesA2</i>	D11	[TTA] ₇	(+)
107949226	<i>CslD4</i>	A04	[ATT] ₅	(-)
107958687	<i>CesA1</i>	A05	[TAT] ₅	(-)
107907453	<i>CesA1</i>	D05	[TAA] ₉	(-)
107901762	<i>CesA5</i>	D06	[TTA] ₈	(-)
107935218	<i>CslE1</i>	D11	[AAT] ₆ [AAAT] ₄	(-)
107896451	<i>CslG2</i>	A09	[TTTC] ₄	(-)
107935224	<i>CslE1</i>	D11	[ACAT] ₄	(-)
107945371	<i>Prob-CesA8</i>	D03	[ATTTT] ₄	(+)
107950592	<i>CesA2</i>	D03	[AGAAA] ₄ [AAAAG] ₅	(-)
107912646	<i>CslH1</i>	D11	[CCATGC] ₄	(-)

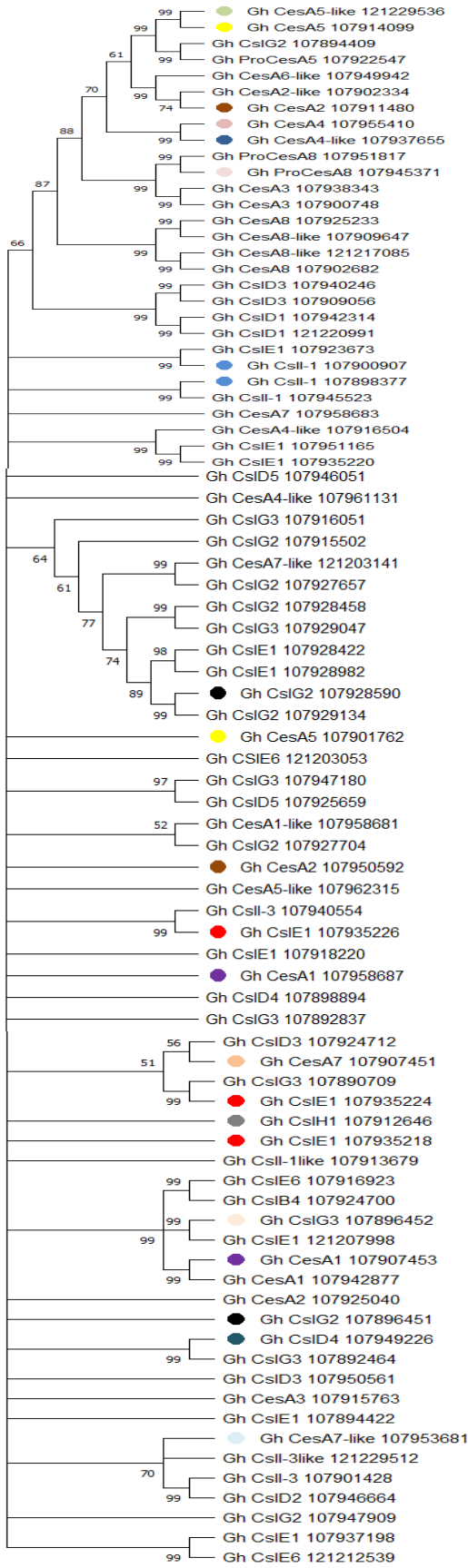
Kromozomlar SSR içerikleri yönünden karşılaştırıldığında, en fazla bulunan D11 kromozomu olmasına rağmen mikrosatellit olarak ilk sırayı 6 SSR içermesi yönünden D05 kromozomu ile paylaşmaktadır. En fazla tekrarlanan motif D05 kromozomunda bulunan *CesA7* genine ait eksi sarmalda yer alan [TA] motifi 12 kez tekrarlanmıştır. En az tekrarlanan motifler ise A09 kromozomunda bulunan *CslG2* genine ait eksi sarmalda yer alan [TTTC]₄, D03 kromozomunda bulunan *CesA2* genine ait eksi sarmalda yer alan [AGAAA]₄ ve *Prob-CesA8* genine ait artı sarmaldaki [ATTTT]₄; D05 kromozomunda bulunan *CesA7* genine ait eksi sarmaldaki [TTTA]₄, D11 kromozomundaki eksi sarmalda olan tetra-nükleotitler [AAAT]₄, [ACAT]₄ ve *CslH1* genine ait hekza-nükleotit motifi olan [CCATGC]₄ motifleridir. 27 SSR arasında en yüksek sayı tri-nükleotitler olup en düşük sayıya sahip olan ise hekza-nükleotitlerdir.

Şekil 4.25'te Upland pamuğa ait mikrosatellit içeren selüloz sentaz ve selüloz sentaz benzeri genlerin filogenetik analizi gösterilmiştir. Bu pamuk türüne ait 82 *CesA* ve *Csl* nükleotit dizisi Maksimum olasılık yöntemi ve Tamura-Nei modelini kullanılarak çalışılmıştır. Yapılan analizde 20044 pozisyon bulunmakta olup MEGA11 yazılım programı kullanılarak gerçekleştirilmiştir (Tamura vd. 2021). Upland pamuğun filogenetik analizi incelendiğinde ID numarası 121229536 *CesA5-like* ile 107914099 *CesA5* genleri SSR bulundurmakla beraber; SSR içermeyen ID numaraları 107894409 *CslG2* ve 107922547 *ProCesA5* genleri ile benzerlik göstermektedir. SSR içeren ID numarası 107911480 *CesA2* geni, SSR bulundurmayan ID numarası 107902334 *CesA2-like* geni ile 107949942 *CesA6-like* geni ile yakından ilişkilidir. ID numarası 107955410 *CesA4* ile 107937655 *CesA4-like* genleri hem SSR içermekte hem de birbirleriyle benzerlik göstermektedir. ID numarası *ProCesA8* geni SSR içermekte olup sırasıyla ID numaraları 107951817 *ProCesA8* geni, 107938343 *CesA3* ve 107900748 *CesA3* genleriyle yakın mesafededir. SSR içeren ID numarası 107900907 *CslI1* geni, 107923673 ID numarasına sahip *CslE1* ile; SSR bulunduran 107898377 *CslI1* geni, 107898377 ID numaralı *CslI1* ile benzerlik göstermektedir.

SSR içeren ID numarası 107928590 *CslG2* geni, SSR içermeyen ID numarası 107929134 *CslG2*, 107928982 *CslE1*, 107928422 *CslE1*, 107929047 *CslG3*, 107928458 *CslG2*, 107927657 *CslG2* ve 121203141 *CesA7-like* genleri ile fazla benzerlik gösterirken; 107915502 *CslG2* ve 107916051 *CslG3* ile daha az benzerlik göstermesinin yanında, ID numarasına sahip 107901762 *CesA5*, 107950592 *CesA2*, 107958687 *CesA1*, 107912646 *CslH1*, 107935218 *CslE1* ve 107896451 *CslG2* genleri hem SSR içerip hem de diğer genlere göre daha mesafeli konumdadır. 107935226 ID numarasına sahip *CslE1* geni SSR içermekle birlikte 107940554 ID numaralı *CslI3* gen ile; SSR içeren 107907451 *CesA7* ve 107935224 *CslE1* genleri SSR içermeyen 107924712 *CslD3* ve 107890709 *CslG3* genleri ile benzerlik göstermektedir. SSR bulunduran 107896452 *CslG3* ve 107907453 *CesA1* genleri, SSR bulundurmayan 107916923 *CslE6*, 107924700 *CslB4*, 121207998 *CslE1* ve 107942877 *CesA1* genleri ile yakından ilişkilidir. SSR içeren 107949226 *CslD4* geni, SSR içermeyen 107892464 ID numarasına sahip *CslG3* geni benzerlik göstermekle birlikte son kümeleme grubunda bulunup SSR içeren 107953681 ID numaralı *CesA7-like* geni, SSR bulundurmayan 121229512 *CslI-3like*, 107901428 *CslI-3* ve 107946664 *CslD2* genleriyle yakından ilişkilidir.

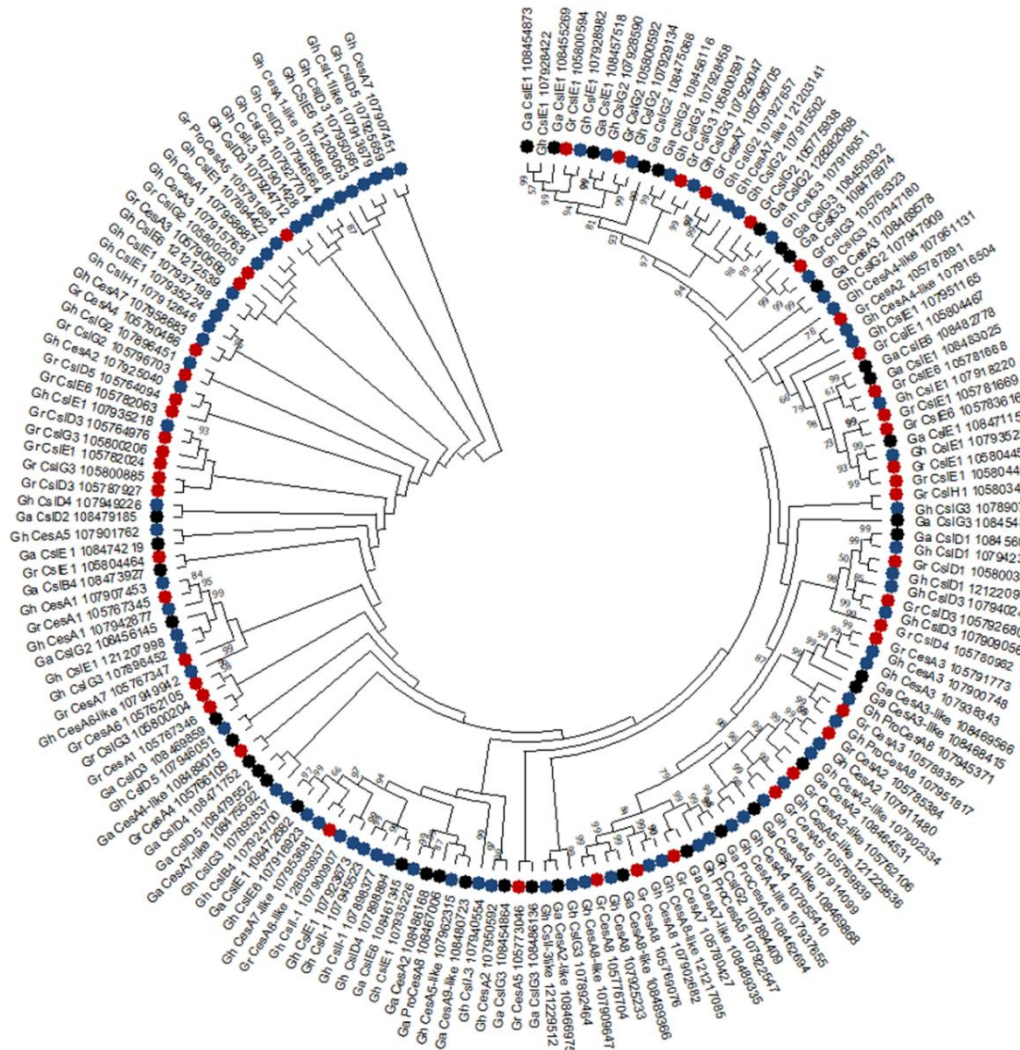
CesA ve *Csl* nükleotit dizilerinin Upland pamuktaki filogenetik analizi bu genlerin diploit türlerindeki filogenetik analizleri ile karşılaştırıldığında gen aileleri aralarındaki ilişkilerin diploit türlerde daha belirgin olduğu ancak; Upland pamuk türünde gerek daha düşük bootstrap değerleri ve gerekse de ilişkiler daha açık olarak ortaya çıkmamaktadır. Bu durum poliploit türdeki *CesA* ve *Csl* nükleotit dizilerinin diploitlerdeki atasal dizilerden farklı oluşu ve hâlâ değişimin yaşandığını göstermesi yönünden ilginç bulunmuştur (Şekil 4.26). Diğer taraftan poliploit pamuk türlerinde *CesA* ve *Csl* nükleotit dizilerin iyi birer barkod gen olmadıklarını bu nedenle de tür çeşit ayrımı ile filogenetik çalışmalar için etkin gen aileleri olmadığını göstermiştir.





Şekil 4.25. Upland pamukta mikrosatellit içeren selüloz sentaz ve selüloz sentaz benzeri genlerin filogenetik analizi

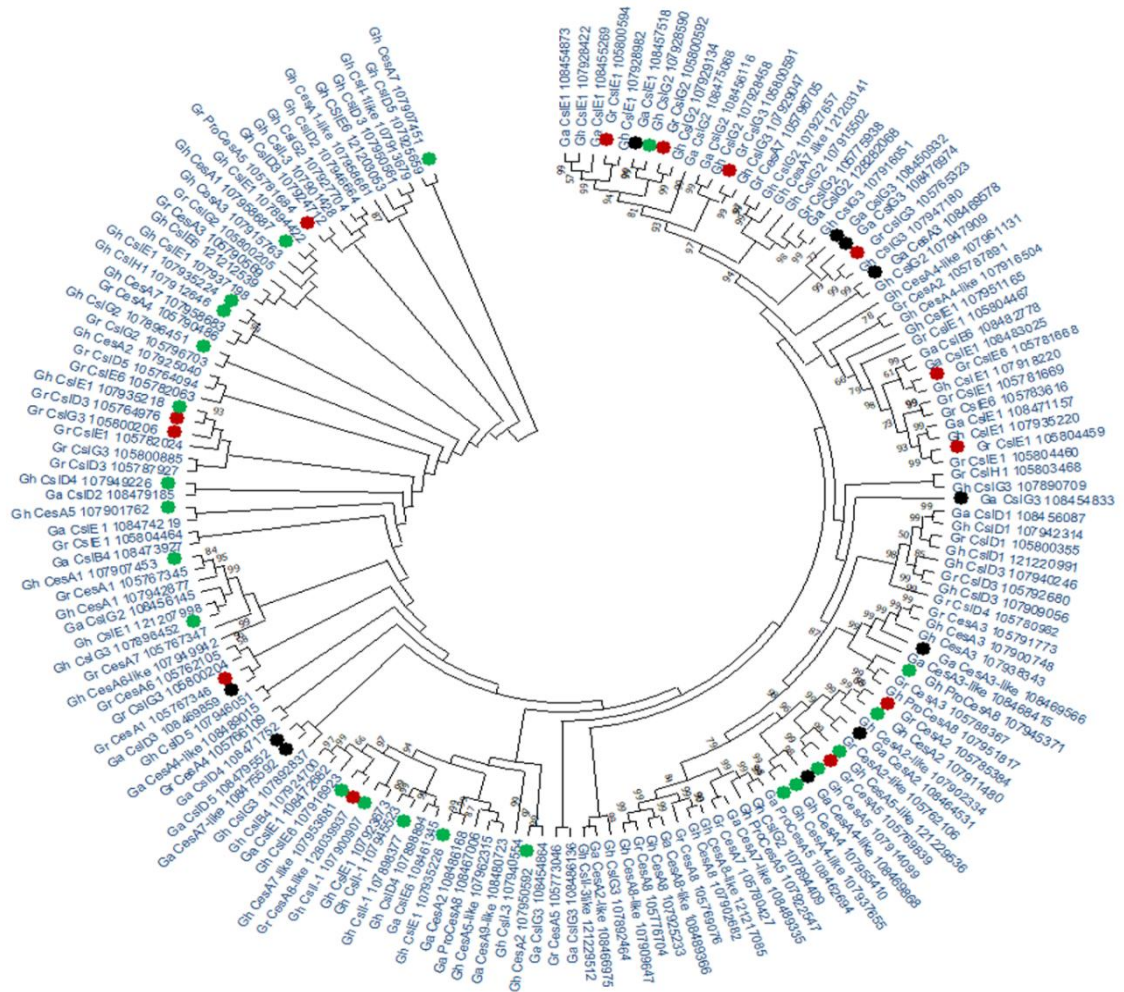
Tez çalışmasında üç pamuk türüne ait 166 selüloz sentaz ve selüloz sentaz benzeri gen dizileri maksimum olasılık yöntemi ve Tamura-Nei modelini kullanılarak çalışıldı ve tahmin edilen evrimsel analiz ağaç şeklinde Şekil 4.26'da sunulmuştur. Yapılan analizde 166 gen dizisinde toplam 38530 pozisyon analizi MEGA11 yazılım program kullanılarak gerçekleştirilmiştir (Tamura vd. 2021). İlgili filogenetik analiz incelendiğinde 166 genin 4 ana grupta filogenetik ilişki gösterdiği ağacın uç kısmında Upland pamuğunun genlerinin daha yoğun olarak bulunduğu görülmektedir. Üç pamuk türünün birlikte incelendiğinde Upland pamuğun diğer iki türe göre daha yeni olduğu, oryantal pamuğun ise daha eski pamuk türü olduğu Peru pamuğunun ise bu iki türün arasında yer aldığı ilgili filogeni analizinde görülmektedir.



Şekil 4.26. Maksimum olasılık yöntemi ve Tamura-Nei modelini kullanılarak tahmin edilen evrimsel analiz.

Şekilde bağlantılı taksa'nın bir araya toplandığı ağaçların Bootstrap değerleri dalların yanında gösterilmektedir. Heuristik arama için başlangıç ağaçları, Tamura-Nei modeli kullanılarak tahmin edilen çift yönlü mesafelerin bir matrisine Komşu-Birlik ve BioNJ algoritmaları uygulayarak ve daha sonra üst düzey log olasılık değeri olan topolojiyi seçerek otomatik olarak elde edildi. Üç pamuk türüne ait taksonlar 3 farklı markır ile gösterilmektedir. Mavi markır Upland, kırmızı Peru ve siyah Oryantal pamuğu göstermektedir.

Bu tez çalışmasında 3 pamuk türünün genomlarında bulunan ve mikrosatellit içeren selüloz sentaz ve selüloz sentaz benzeri gen dizileri ile SSR içermeyen genlerin maksimum olasılık yöntemi ve Tamura-Nei modelini kullanılarak çalışıldı ve tahmin edilen evrimsel analiz ağaç şeklinde Şekil 4.27'de sunulmuştur. Analizde örneğin Peru ve Upland pamuklarındaki *CesA2* genlerinde SSR içermekte ve bu iki türün *CesA* genleri filogenetik benzerlik göstermekte diğer taraftan Oryantal pamuğun *CesA2* geni SSR içermekle birlikte filogenetik açıdan Peru ve Upland pamuklarına biraz daha mesafeli konumdadır. *CesA2* genlerinin ilgili çizelgelerde (Çizelge 4.42, Çizelge 4.43) gösterildiği gibi SSR içerikleri farklıdır. Genlerin filogenetik olarak yakın olmalarına rağmen SSR içeriklerinin farklı olması ilgili genlerin SSR kısımlarının diğer gen kısımlarına oranla daha fazla değişim göstermiş olduğu anlaşılmaktadır.



Şekil 4.27. Maksimum olasılık yöntemi ve Tamura-Nei modeli kullanılarak SSR içeren ve SSR içermeyen *CesA* ve *Csl* genlerinin filogenetik analizi

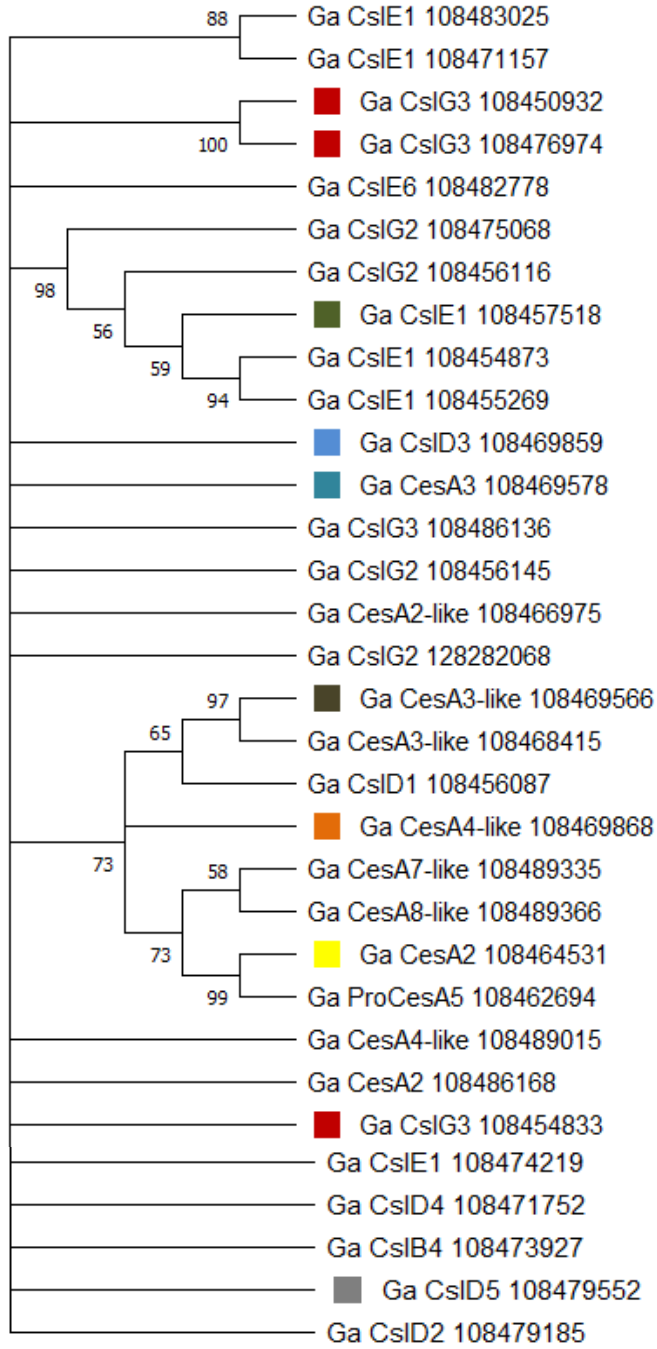
Filogenetik çalışmalar DNA dizileri ile yapılabildiği gibi protein dizileri (aminoasit) kullanılarak da gerçekleştirilmiştir. Genel bir kural olarak eğer evrimsel değişim tam doğal şartlarda gerçekleşiyor ise nükleik asit dizileri daha yüksek çözünürlüklü sonuç üretebilirken protein dizileri daha az çözünürlük göstermekle birlikte belirgin morfolojik verilere daha uygun filogenetik ilişkiler ortaya koyabilmektedir.

Şekil 4.28’de Oryantal pamuğa ait mikrosatellit içeren selüloz sentaz ve selüloz sentaz benzeri genlerin protein tabanlı filogenetik analizi gösterilmiştir. Evrimsel tarih, Maksimum Olasılık yöntemi ve JTT matris tabanlı model kullanılarak çıkarılmıştır. 100 tekrardan elde edilen bootstrap konsensüs (uzlaş) ağacı, analiz edilen taksonların evrimsel tarihini temsil eder. Bootstrap tekrarlarının %50'sinden daha azını yeniden üreten dallar birleştirilmiştir. İlgili taksonların bootstrap testindeki tekrarların (100 tekrar) hangi yüzdeyle kümelendiği dalların yanında gösterilmiştir.

Heüristik arama için başlangıç ağaçları, JTT modelini kullanarak tahmin edilen çiftler arası mesafe matrisine uygulanan Komşu-Birleşim (Neighbor-Joining) ve BioNJ algoritmalarıyla otomatik olarak elde edilmiştir ve ardından en iyi log-olasılık değerine sahip topoloji seçilmiştir. Bu analiz 38 amino asit dizisini içermiştir ve nihai veri setinde toplam 1393 pozisyon bulunmuştur. Evrimsel analizler MEGA11 yazılımı kullanılarak gerçekleştirilmiştir (Tamura vd. 2021).

İlgili analizde ID numaraları 108450932 *CsIG3*, 108476974 *CsIG3* benzerlik gösterirken, 108454833 ID numarasına sahip *CsIG3* geni diğer genlere göre mesafeli konumdadır. Ayrıca diğer SSR içeren genlerden daha mesafeli olan genler ID numaralarına göre sırasıyla 108469859 *CsID3*, 108469578 *CesA3*, 108475592 *CesA7-like* ve 108479552 *CsID5* genleridir. SSR içeren ID numarası 108457518 *CsIE1* geni SSR içermeyen ID numaraları 108454873 ve 108455269 *CsIE1* genleri ile 108475068 ve 108456116 *CsIG2* genleri ile benzerlik göstermektedir. Diğer bir kümelene içinde bulunan ID numarası 108469566 *CesA3-like* SSR bulundurmayan 108468415 *CesA3-like* ve 108456087 *CsID1* ile benzerlik göstermekle beraber SSR içeren 108464531 ID numarasına sahip *CesA2* geni SSR’ı olmayan 108462694 *ProCesA5*, 108489335 *CesA7-like*, 108489366 *CesA8-like* ile SSR’a sahip 108469868 *CesA4-like* geni ile yakından ilişkilidir.

Tez çalışmasının sonraki kısmında tekrar değinileceği üzere mikrosatellitin genlerin kodlama bölgelerinde bulunmaları translasyon ürünlerinde de aminoasit tekrarların oluşumu ile sonuçlanması beklenmektedir. Nükleik asitlerdeki dizi tekrarların amino asitlerde oluşturduğu dizi tekrarları her ne kadar birbirleri ile ilişkili olmalarına rağmen tamamen farklı tekrar biçimlerine dönüşebilmektedir. Örneğin DNA üzerinde bulunan di-nükleotit tekrarı translasyon sonucu heteromer özellikte dipeptit tekrarlarına dönüşebildiği gibi, tri-nükleotit ve hekza-nükleotit tekrarları homomer ve dimer özellikte aminoasit tekrarlarına dönüşebilmektedir. Diğer taraftan tetra-nükleotit ve penta-nükleotit tekrarları tetrapeptit ve pentapeptit tekrarlarına translasyon sonucu dönüşebilmektedir. Bu nedenle mikrosatellit ürünleri farklı filogenetik sonuçlar ortaya çıkarabilme özelliğine sahip DNA dizileridir.



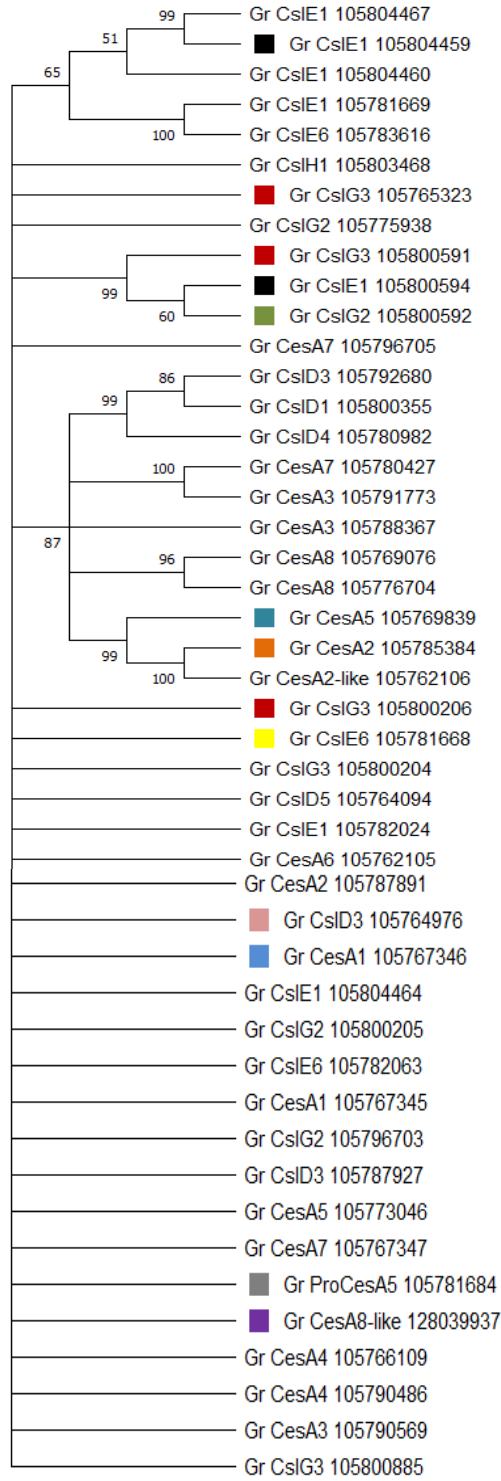
Şekil 4.28. Oryantal pamukta mikrosatellit içeren selüloz sentaz ve selüloz sentaz benzeri genlerin protein için filogenetik analizi

Şekil 4.29’da Peru pamuğa ait mikrosatellit içeren selüloz sentaz ve selüloz sentaz benzeri genlerin protein tabanlı filogenetik analizi gösterilmiştir. Evrimsel tarih, Maksimum Olasılık yöntemi ve JTT matris tabanlı model kullanılarak çıkarılmıştır. 100 tekrardan elde edilen bootstrap konsensüs ağacı, analiz edilen taksonların evrimsel tarihini temsil eder. Bootstrap tekrarlarının %50'sinden daha azını yeniden üreten dallar birleştirilmiştir. İlgili taksonların bootstrap testindeki tekrarların (100 tekrar üzerinden değerlendirildiğinde) hangi yüzdeyle kümelendiği dalların yanında gösterilmiştir.

Heüristik arama için başlangıç ağaçları, JTT modelini kullanarak tahmin edilen çiftler arası mesafe matrisine uygulanan Komşu-Birleşim (Neighbor-Joining) ve BioNJ algoritmalarıyla otomatik olarak elde edilmiştir ve ardından en iyi log-olasılık değerine sahip topoloji seçilmiştir. Bu analiz 46 amino asit dizisini içermiştir ve nihai veri setinde toplam 4822 pozisyon bulunmuştur. Evrimsel analizler MEGA11 yazılımı kullanılarak gerçekleştirilmiştir (Tamura vd. 2021).

İlgili analizde SSR içerenler 105804459 *CsIE1*, 105800592 *CsIG2*, 105800591 *CsIG3* genleri birbirleriyle yüksek ölçüde benzerlik göstermekle beraber ID numaraları 105765323 ve 105800206 *CsIG3* genleri biraz daha mesafeli konumdadır. 105769839 *CesA5* ve 105785384 *CesA2* genleri SSR içerip 105762106 *CesA2-like* ile benzerlik göstermektedir. Ayrıca diğer SSR içeren genlerin ID numaraları sırasıyla 105781668 *CsIE6*, 105764976 *CsID3*, 105767346 *CesA1*, 105781684 *ProCesA5* ve 128039937 *CesA8-like* genleridir.

Şekil 4.29’da görüleceği üzere amino asit tabanlı filogenetik analizlerde mikrosatellitin varlığı ve tekrar sayısının uzunluğu *CesA* ve *CsI* genlerinin filogenetik pozisyonlarını etkileyebilmektedir. Örneğin ilgili Şekil 4.29’a bakıldığında *CsIE* ve *CsIG* gen ailesinden mikrosatellit içeren genler filogenetik olarak daha ilişkili pozisyonda gözükmemektedir. Benzer şekilde *CesA5* ve *CesA2* genleri mikrosatellit içerikleri nedeniyle filogenetik analizde ilişkili görülmektedir.



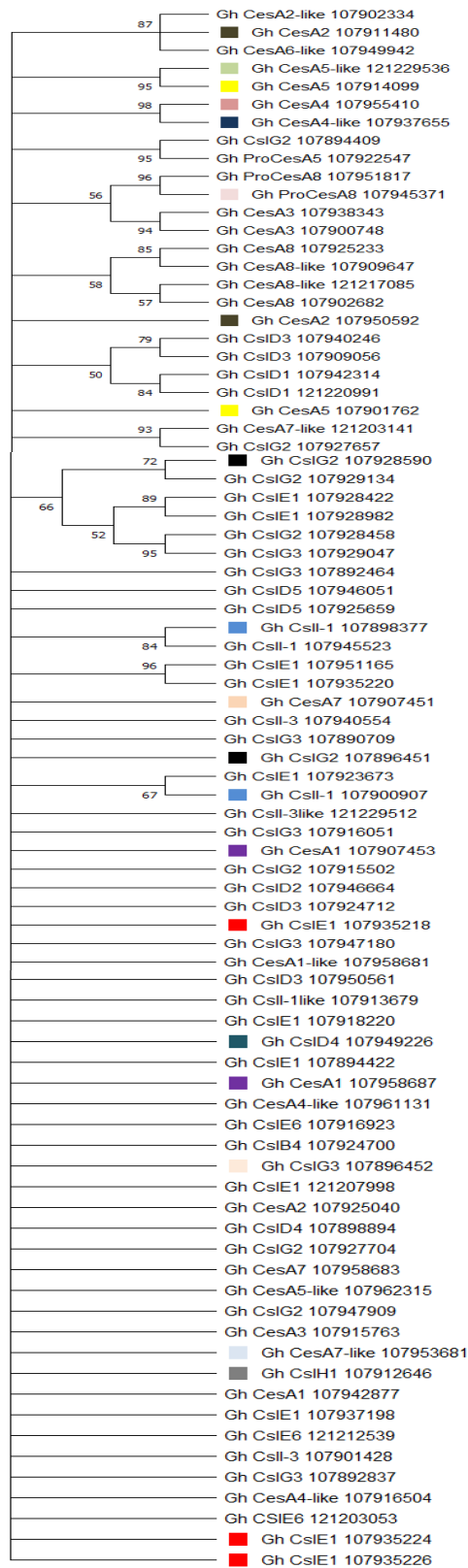
Şekil 4.29. Peru pamukta mikrosatellit içeren selüloz sentaz ve selüloz sentaz benzeri genlerin protein için filogenetik analizi

Şekil 4.30'da Upland pamuğa ait mikrosatellit içeren selüloz sentaz ve selüloz sentaz benzeri genlerin protein tabanlı filogenetik analizi gösterilmiştir. Evrimsel tarih, Maksimum Olasılık yöntemi ve JTT matris tabanlı model kullanılarak çıkarılmıştır. 100 tekrardan elde edilen bootstrap konsensüs ağacı, analiz edilen taksonların evrimsel tarihini temsil eder. %50'den daha az bootstrap tekrarı tarafından yeniden üretilen dallar birleştirilmiştir. Bootstrap testindeki (100 tekrar) ilgili taksonların hangi yüzdeyle kümелendiği dalların yanında gösterilmiştir.

Heüristik arama için başlangıç ağaçları, JTT modelini kullanarak tahmin edilen çiftler arası mesafe matrisine Komşu-Birleşim (Neighbor-Joining) ve BioNJ algoritmaları uygulayarak otomatik olarak elde edilmiştir ve ardından en yüksek log-olasılık değerine sahip topoloji seçilmiştir. Bu analiz 82 amino asit dizisini içermiştir ve nihai veri setinde toplam 3572 pozisyon bulunmuştur. Evrimsel analizler MEGA11 kullanılarak gerçekleştirilmiştir (Tamura vd. 2021).

Bu analizde SSR içeren 107911480 *CesA2* geni, 107902334 *CesA2-like* ve 107949942 *CesA6-like* ile yüksek ölçüde benzerlik gösterip 107950592 *CesA2* geni mesafeli konumdadır. SSR'a sahip ID numaraları 121229536 *CesA5-like*, 107914099 *CesA5* geni benzerlik gösterip 107901762 *CesA5* geni biraz daha mesafelidir. ID numaraları 107955410 *CesA4* ve 107937655 *CesA4-like* genleri SSR içermekle beraber benzerlik göstermektedir.

SSR bulunan 107945371 *ProCesA8* geni ile 107951817 *ProCesA8* geni, SSR içeren 107928590 *CslG2* ile 107929134 *CslG2* geni, diğer bir SSR içeren 107898377 *CsII-1* gen ile 107945523 *CsII-1* geni ve son olarak SSR içeren 107900907 ID numarasına sahip *CsII-1* geni SSR bulunmayan 107923673 *CsIE1* geni ile yakından ilişkilidir. Ayrıca diğer SSR içeren genler filogenetik analiz konumuna göre sırasıyla 107907451 *CesA7*, 107907453 *CesA1*, 107935218 *CsIE1*, 107949226 *CsID4*, 107958687 *CesA1*, 107896452 *CslG3*, 107953681 *CesA7-like*, 107912646 *CsIH1*, 107935224 ve 107935226 *CsIE1* genleridir.

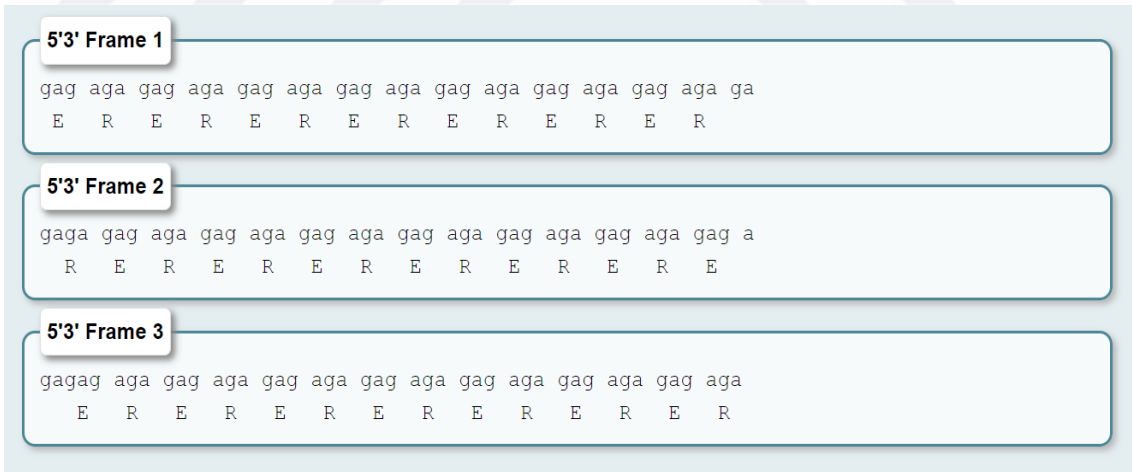


Şekil 4.30. Upland pamukta mikrosatellit içeren selüloz sentaz ve selüloz sentaz benzeri genlerin protein için filogenetik analizi

4.3. Selüloz Sentaz ve Selüloz Sentaz Benzeri Genlerdeki Mikrosatellit Dizilerin Kodladığı Aminoasitler

Oryantal pamuk türünde mikrosatellit içeren selüloz sentaz (*CesA*) ve selüloz sentaz benzeri (*Csl*) genlerindeki mikrosatellit dizileri ExPASy programı kullanılarak in silico translasyona tabi tutulmuştur. Böylelikle mikrosatellitlerin filogenetik ilişkilerde ortaya koyduğu etki düzeyinin aminoasite dönüşme potansiyelleri de incelenebilmiş ve SSR tekrar dizilerin genlerin yapısal ve fonksiyonel özelliklerine katkıları daha detaylı bir şekilde karakterize edilmesini sağlamıştır. Oryantal pamukta 38 gen arasında 5'i *CesA* ve *CesA-like* genlerinden oluşmakta iken; 6 tanesi selüloz sentaz benzeri (*Csl*) genlerinden oluşmaktadır. Mikrosatellit tanımlama kriterlerine göre bu genlerde toplam 16 mikrosatellit tanımlanmış olup bu mikrosatellitlerin 3 olası okuma çerçevesinde kodladıkları amino asitler belirlenmiştir.

Şekil 4.31'de A08 kromozomunda bulunan, ID numarası 108469566 olan *CesA3-like* genine ait oryantal pamuğu 22 kez tekrarlanan [GA] motifi 5' => 3' Frame 1'de 7'li E (Glutamik asit) ve R (Arjinin) aminoasitlerini kodlamıştır. 5' => 3' Frame 2'de 7'li R (Arjinin) ve E (Glutamik asit) aminoasitlerine dönüşmüştür. 5' => 3' Frame 3'te ise 7'li E (Glutamik asit) ve R (Arjinin) aminoasitlerini kodlamaktadır. Burada mikrosatellit tanımlama kriterine bağlı olarak GA tekrarı olarak nitelenen mikrosatellitin ribozomlarda translasyon sırasında AGA veya GAG genetik kodonları olarak tanımlanabileceği açıkça görülmektedir. Bu mikrosatellit tipinin arttırılması veya azaltılmasıyla üretilen proteinin pozitif veya negatif yükünün değiştirilebileceği, enzimin aktif bölgesinde bulunması durumunda ise aktivitesi üzerine önemli etkiler olabileceği açıktır.



Şekil 4.31. Oryantal pamuğa ait ID numarası 108469566 olan *CesA3-like* genine ait [GA]₂₂ motifinin kodladığı aminoasitler

Şekil 4.32’de A09 kromozomunda bulunan, ID numarası 108457518 olan *CsIE1* genine ait Oryantal pamuğu 10 kez tekrarlanan [AT] motifi 5' => 3' Frame 1’de 3’lü I (İzolösın) ve Y (Tırosın) aminoasitlerini kodlamıştır. 5' => 3' Frame 2’de 3’lü Y (Tırosın) ve I (İzolösın) aminoasitlerine dönüşmüştür. 5' => 3' Frame 3’te ise tekrar 3’lü I (İzolösın) ve Y (Tırosın) aminoasitlerini kodlamıştır.

5'3' Frame 1
ata tat ata tat ata tat at
I Y I Y I Y
5'3' Frame 2
atat ata tat ata tat ata t
Y I Y I Y I
5'3' Frame 3
atata tat ata tat ata tat
I Y I Y I Y

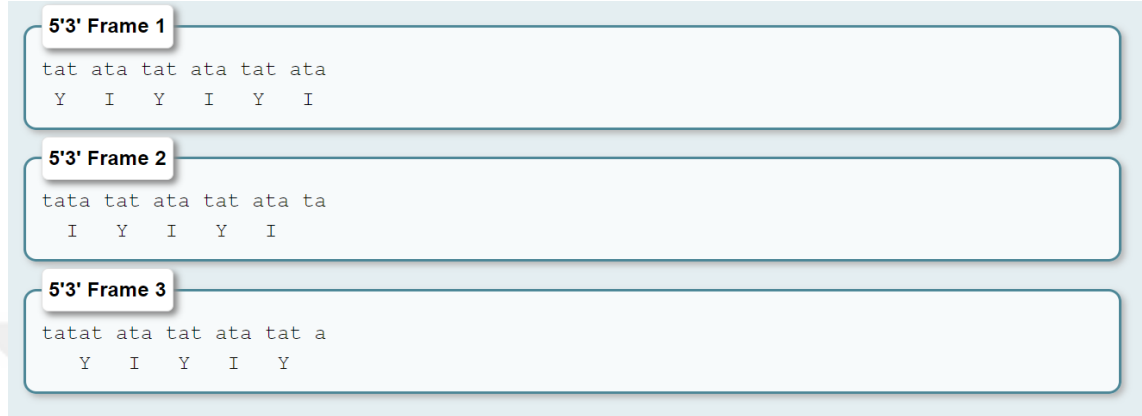
Şekil 4.32. Oryantal pamuğa ait ID numarası 108457518 olan *CsIE1* genine ait [AT]₁₀ motifinin kodladığı aminoasitler

Şekil 4.33’te A12 kromozomunda bulunan, ID numarası 108464531 olan *CesA2* genine ait Oryantal pamuğu 12 kez tekrarlanan [AG] motifi 5' => 3' Frame 1’de 4’lü R (Arjinin) ve E (Glutamik asit) aminoasitlerini kodlamıştır. 5' => 3' Frame 2’de GAG-AGA şeklinde olurken E (Glutamik asit) ve R (Arjinin) aminoasitlerinde olmaktadır. Frame 3’te ise GAGA-GAG şeklini alıp tekrar R (Arjinin) ve E (Glutamik asit) aminoasitlerine denk gelmektedir.

5'3' Frame 1
aga gag aga gag aga gag aga gag
R E R E R E R E
5'3' Frame 2
agag aga gag aga gag aga gag ag
E R E R E R E
5'3' Frame 3
agaga gag aga gag aga gag aga g
R E R E R E R

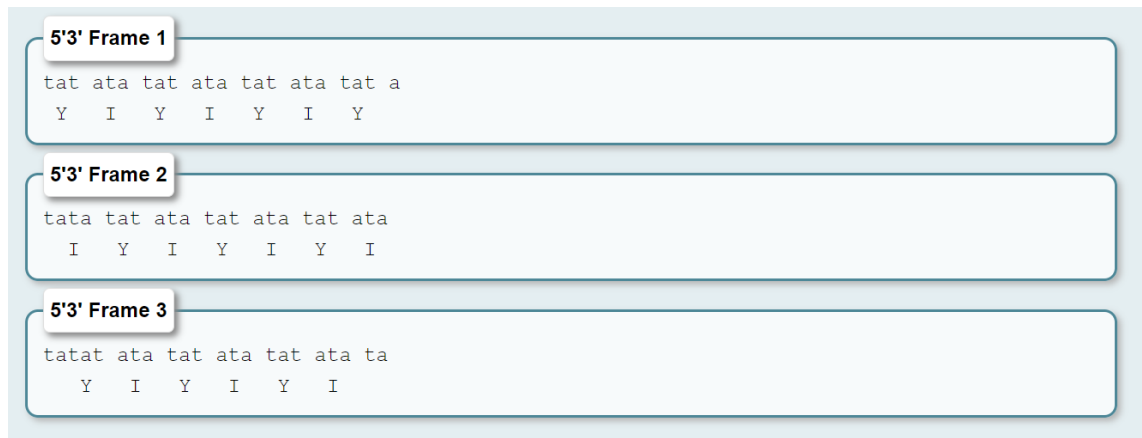
Şekil 4.33. Oryantal pamuğa ait ID numarası 108464531 olan *CsIE1* genine ait [AG]₁₂ motifinin kodladığı aminoasitler

Şekil 4.34'te A12 kromozomunda bulunan, ID numarası 108464531 olan *CesA2* genine ait Oryantal pamuğu 9 kez tekrarlanan [TA] motifi 5' => 3' Frame 1'de 3'lü Y (Tirozin) ve I (İzolösün) aminoasitlerini kodlamış olup 5' => 3' Frame 2'de TATA-TAT şeklinde olurken I (İzolösün) ve Y (Tirozin) aminoasitlerinde olmaktadır. Frame 3'te ise TATAT-ATA şeklini alıp tekrar Y (Tirozin) ve I (İzolösün) aminoasitlerine denk gelmektedir.



Şekil 4.34. Oryantal pamuğa ait ID numarası 108464531 olan *CesA2* genine ait [TA]₉ motifinin kodladığı aminoasitler

Şekil 4.35'te A12 kromozomunda bulunan, ID numarası 108476974 olan *CsIG3* genine ait Oryantal pamuğu 11 kez tekrarlanan [TA] motifi 5' => 3' Frame 1'de TAT-ATA motifine dönüşüp Y (Tirozin) ve I (İzolösün) aminoasitlerini kodlarken; 5' => 3' Frame 2'de TATA-TAT şeklini alıp I (İzolösün) ve Y (Tirozin) aminoasitlerinde olmaktadır. Frame 3'te ise TATAT-ATA motifinde tekrar Y (Tirozin) ve I (İzolösün) aminoasitlerini kodlamıştır.



Şekil 4.35. Oryantal pamuğa ait ID numarası 108476974 olan *CsIG3* genine ait [TA]₁₁ motifinin kodladığı aminoasitler

Şekil 4.36'da A12 kromozomunda bulunan, ID numarası 108475592 olan *CesA7-like* genine ait Oryantal pamuğu 8 kez tekrarlanan [AT] motifi 5' => 3' Frame 1'de ATA-TAT motifine dönüşüp I (İzolösin) ve Y (Tirozin) aminoasitlerini kodlamıştır. 5' => 3' Frame 2'de ATAT-ATA motifini alıp Y (Tirozin) ve I (İzolösin) aminoasitlerinde olmaktadır. Frame 3'te ise ATATA-TAT motifinde tekrar I (İzolösin) ve Y (Tirozin) aminoasitlerinde olmaktadır.

5'3' Frame 1
ata tat ata tat ata t
I Y I Y I
5'3' Frame 2
atat ata tat ata tat
Y I Y I Y
5'3' Frame 3
atata tat ata tat at
I Y I Y

Şekil 4.36. Oryantal pamuğa ait ID numarası 108475592 olan *CesA7-like* genine ait [AT]₈ motifinin kodladığı aminoasitler

Şekil 4.37'de A03 kromozomunda bulunan, ID numarası 108475592 olan *CesA7-like* genine ait Oryantal pamuğu 19 kez tekrarlanan [AG] motifi 5' => 3' Frame 1'de AGA-GAG motifine dönüşüp R (Arjinin) ve E (Glutamik asit) aminoasitlerini, 5' => 3' Frame 2'de ATAG-AGA motifinde E (Glutamik asit) ve R (Arjinin) aminoasitlerini, Frame 3'te ise ATAGA-GAG motifinde tekrar R (Arjinin) ve E (Glutamik asit) aminoasitlerini kodlamıştır.

5'3' Frame 1
aga gag aga gag aga gag aga gag aga gag aga gag ag
R E R E R E R E R E R E R E
5'3' Frame 2
agag aga gag aga gag aga gag aga gag aga gag aga g
E R E R E R E R E R E R E R
5'3' Frame 3
agaga gag aga gag aga gag aga gag aga gag aga gag
R E R E R E R E R E R E R E

Şekil 4.37. Oryantal pamuğa ait ID numarası 108475592 olan *CesA7-like* genine ait [AG]₁₉ motifinin kodladığı aminoasitler

Şekil 4.38’de A03 kromozomunda bulunan, ID numarası 108475592 olan *CesA7-like* genine ait Oryantal pamuğu 19 kez tekrarlanan [AG] motifini 5' => 3' Frame 1’de CTC-TCT motifine dönüşüp L (Lösin) ve S (Serin) aminoasitlerini, 5' => 3' Frame 2’de CTCT-CTC motifinde S (Serin) ve L (Lösin) aminoasitlerini, Frame 3’te ise CTCTC-TCT motifinde tekrar L (Lösin) ve S (Serin) aminoasitlerini kodlamıştır.

5'3' Frame 1
ctc tct ctc tct ctc tct
L S L S L S
5'3' Frame 2
ctct ctc tct ctc tct ct
S L S L S
5'3' Frame 3
ctctc tct ctc tct ctc t
L S L S L

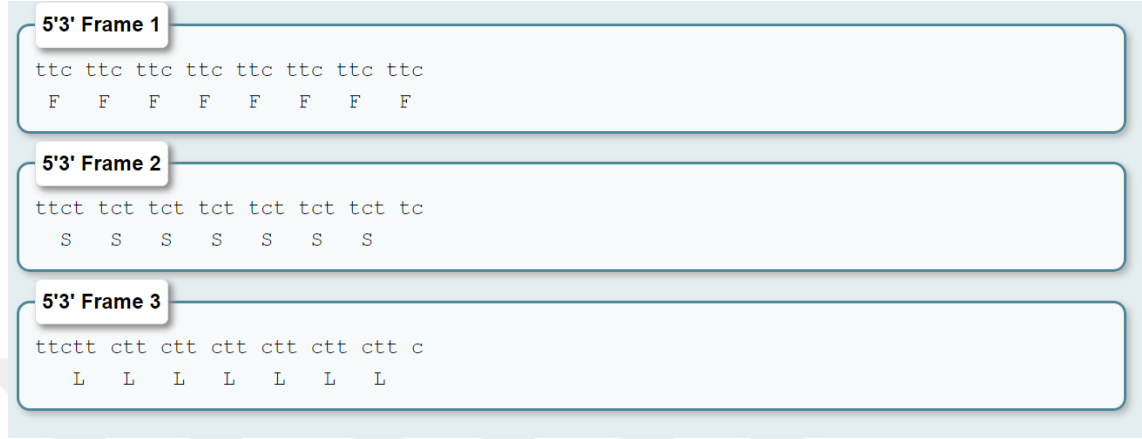
Şekil 4.38. Oryantal pamuğa ait ID numarası 108475592 olan *CesA7-like* genine ait [AG]₁₉ motifinin kodladığı aminoasitler

Şekil 4.39’da A08 kromozomunda bulunan, ID numarası 108469578 olan *CesA3* genine ait Oryantal pamuğu 28 kez tekrarlanan [AT] motifini 5' => 3' Frame 1’de ATA-TAT motifinde I (İzolösin) ve Y (Tirozin) aminoasitlerini; 5' => 3' Frame 2’de ATAT-ATA motifinde Y (Tirozin) ve I (İzolösin) aminoasitlerinde olmaktadır. Frame 3’te ise ATATA-TAT motifinde tekrar I (İzolösin) ve Y (Tirozin) aminoasitlerini kodlamıştır.

5'3' Frame 1
ata tat ata tat ata tat ata tat ata tat ata tat ata tat ata tat ata tat at
I Y I Y I Y I Y I Y I Y I Y I Y I Y
5'3' Frame 2
atat ata tat ata tat ata tat ata tat ata tat ata tat ata tat ata tat ata t
Y I Y I Y I Y I Y I Y I Y I Y I Y I
5'3' Frame 3
atata tat ata tat ata tat ata tat ata tat ata tat ata tat ata tat ata tat
I Y I Y I Y I Y I Y I Y I Y I Y I Y

Şekil 4.39. Oryantal pamuğa ait ID numarası 108469578 olan *CesA3* genine ait [AT]₂₈ motifinin kodladığı aminoasitler

Şekil 4.40'ta A08 kromozomunda bulunan, ID numarası 108469868 olan *CesA4-like* genine ait Oryantal pamuğu 8 kez tekrarlanan [TTC] motifi 5' => 3' Frame 1'de TTC-TTC motifine dönüşüp F (Fenilalanin) aminoasitlerini; 5' => 3' Frame 2'de TTCT-TCT motifini alıp S (Serin) aminoasitlerini, Frame 3'te ise TTCTT-CTT motifinde L (Lösin) aminoasitlerini kodlamaktadır.



Şekil 4.40. Oryantal pamuğa ait ID numarası 108469868 olan *CesA4-like* genine ait [TTC]₈ motifinin kodladığı aminoasitler

Şekil 4.41’de A09 kromozomunda bulunan, ID numarası 108454833 olan *CsIG3* genine ait Oryantal pamukta 5 kez tekrarlanan [AAT] motifi 5' => 3' Frame 1’de AAT-AAT motifine dönüşüp N (Asparajin) aminoasitlerini; 5' => 3' Frame 2’de AATA-ATA motifini alıp I (İzolösin) aminoasitlerini kodlamaktadır. Frame 3’te ise AATAA-TAA (TAA: durdurma kodonu) motifinde olup herhangi bir aminoasiti kodlamamaktadır.

5'3' Frame 1
aat aat aat aat aat
N N N N N
5'3' Frame 2
aata ata ata ata at
I I I I
5'3' Frame 3
aataa taa taa taa t
- - - -

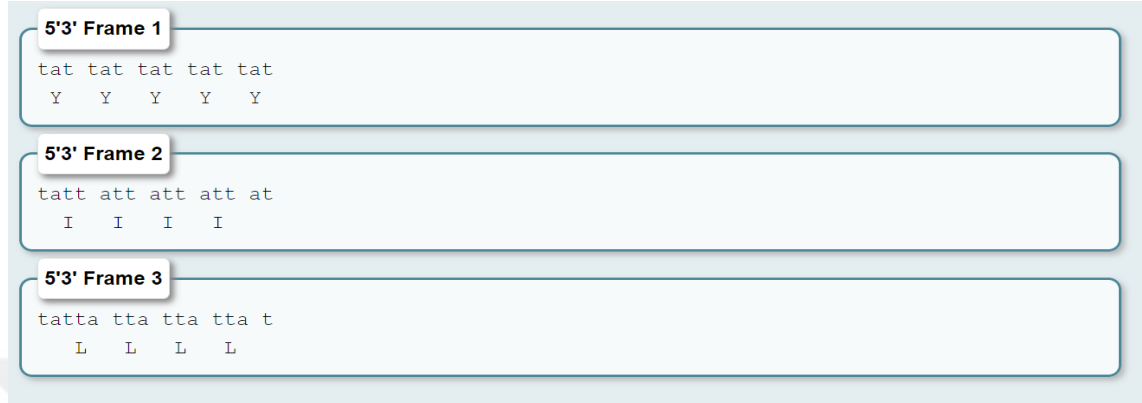
Şekil 4.41. Oryantal pamuğa ait ID numarası 108454833 olan *CsIG3* genine ait [AAT]₅ motifinin kodladığı aminoasitler

Şekil 4.42’de A09 kromozomunda bulunan, ID numarası 108450932 olan *CsIG3* genine ait Oryantal pamukta 18 kez tekrarlanan [TACA] motifi 5' => 3' Frame 1’de TAC-ATA motifine dönüşüp Y (Tirozin), I (İzolösin), H (Histidin) ve T (Treonin) aminoasitlerini, 5' => 3' Frame 2’de TACA-TAC motifini alıp T (Treonin), Y (Tirozin), I (İzolösin), H (Histidin) aminoasitlerini, Frame 3’te ise TACAT-ACA motifinde olup H (Histidin), T (Treonin), Y (Tirozin), I (İzolösin) aminoasitlerini kodlamaktadır.

5'3' Frame 1
tac ata cat aca tac ata cat aca tac ata cat aca tac ata cat aca tac ata cat aca
Y I H T Y I H T Y I H T Y I H T Y I H T
tac ata cat aca
Y I H T
5'3' Frame 2
taca tac ata cat aca tac ata cat aca tac ata cat aca tac ata cat aca tac ata cat
T Y I H T Y I H T Y I H T Y I H T Y I H
aca tac ata ca
T Y I
5'3' Frame 3
tacat aca tac ata cat aca tac ata cat aca tac ata cat aca tac ata cat aca tac ata
H T Y I H T Y I H T Y I H T Y I H T Y I
cat aca tac a
H T Y

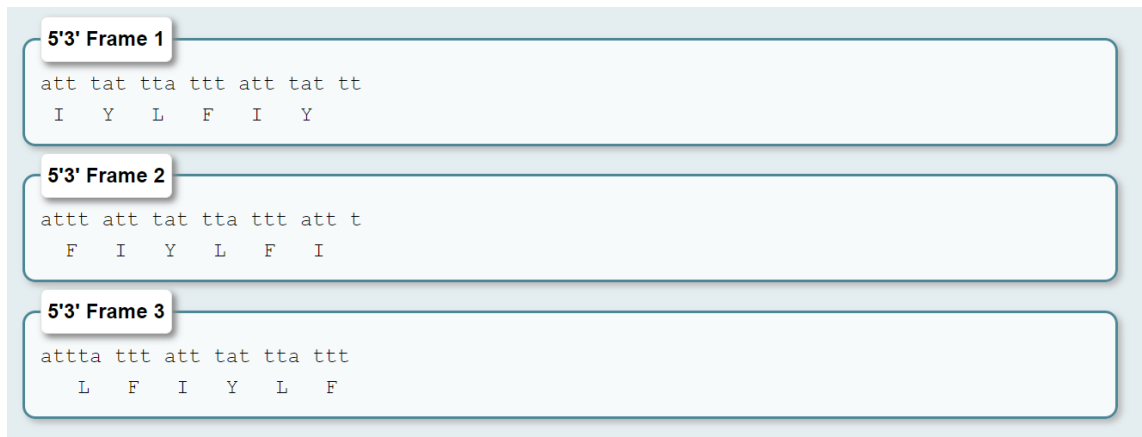
Şekil 4.42. Oryantal pamuğa ait ID numarası 108450932 olan *CsIG3* genine ait [TACA]₁₈ motifinin kodladığı aminoasitler

Şekil 4.43'te A12 kromozomunda bulunan, ID numarası 108479552 olan *CsID5* genine ait Oryantal pamuğu 5 kez tekrarlanan [TAT] motifi 5' => 3' Frame 1'de TAT-TAT motifine dönüşüp Y (Tirozin) aminoasitini; 5' => 3' Frame 2'de TATT-TTT motifini alıp I (İzolösin) aminoasitini, Frame 3'te ise TATTA-TTA motifinde olup, L (Lösin) aminoasitini kodlamaktadır.



Şekil 4.43. Oryantal pamuğa ait ID numarası 108479552 olan *CsID5* genine ait [TAT]₅ motifinin kodladığı aminoasitler

Şekil 4.44'te A12 kromozomunda bulunan, ID numarası 108479552 olan *CsID5* genine ait Oryantal pamuğu 5 kez tekrarlanan [ATTT] motifi 5' => 3' Frame 1'de ATT-TAT motifine dönüşüp I (İzolösin), Y (Tirozin), L (Lösin) F (Fenilalanin) aminoasitlerini, 5' => 3' Frame 2'de ATTT-ATT motifini alıp F (Fenilalanin), I (İzolösin) Y (Tirozin) ve L (Lösin) aminoasitlerini, Frame 3'te ise ATTTA-TTT motifinde olup, L (Lösin) F (Fenilalanin), I (İzolösin), Y (Tirozin) aminoasitlerini kodlamaktadır.



Şekil 4.44. Oryantal pamuğa ait ID numarası 108479552 olan *CsID5* genine ait [ATTT]₅ motifinin kodladığı aminoasitler

Şekil 4.45'te A12 kromozomunda bulunan, ID numarası 108479552 olan *CsID5* genine ait Oryantal pamuğu 4 kez tekrarlanan [TTTTA] motifi 5' => 3' Frame 1'de TTT-TAT motifine dönüşüp F (Fenilalanin), Y (Tirozin), I (İzolösin) ve L (Lösin) aminoasitlerini, 5' => 3' Frame 2'de TTTT-ATT motifini alıp F (Fenilalanin), I (İzolösin), L (Lösin), Y (Tirozin) aminoasitlerini, Frame 3'te ise TTTTA-TTT motifinde olup, L (Lösin) F (Fenilalanin), Y (Tirozin), I (İzolösin) aminoasitlerini kodlamaktadır.

5'3' Frame 1
ttt tat ttt att tta ttt a F Y F I L F
5'3' Frame 2
tttt att tta ttt tat tta F I L F Y L
5'3' Frame 3
tttta ttt tat ttt att ta L F Y F I

Şekil 4.45. Oryantal pamuğa ait ID numarası 108479552 olan *CsID5* genine ait [TTTTA]₄ motifinin kodladığı aminoasitler

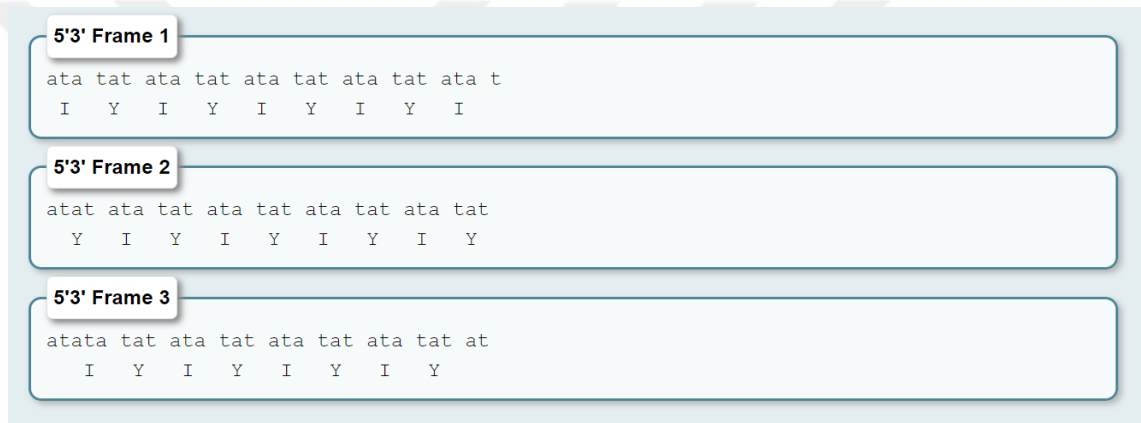
Şekil 4.46'da A12 kromozomunda bulunan, ID numarası 108469859 olan *CsID3* genine ait Oryantal pamuğu 5 kez tekrarlanan [AGAAAG] motifi 5' => 3' Frame 1'de AGA-AAG motifine dönüşüp R (Arjinin), K (Lisin) aminoasitlerini, 5' => 3' Frame 2'de AGAA-AGA motifini alıp E (Glutamik asit) ve R (Arjinin) aminoasitlerini, Frame 3'te ise AGAAA-GAG motifinde olup, K (Lisin) ve E (Glutamik asit) aminoasitlerini kodlamaktadır.

5'3' Frame 1
aga aag aga aag aga aag aga aag aga aag R K R K R K R K R K
5'3' Frame 2
agaa aga gaa aga gaa aga gaa aga gaa ag E R E R E R E R E
5'3' Frame 3
agaaa gag aaa gag aaa gag aaa gag aaa g K E K E K E K E K

Şekil 4.46. Oryantal pamuğa ait ID numarası 108469859 olan *CsID3* genine ait [AGAAAG]₅ motifinin kodladığı aminoasitler

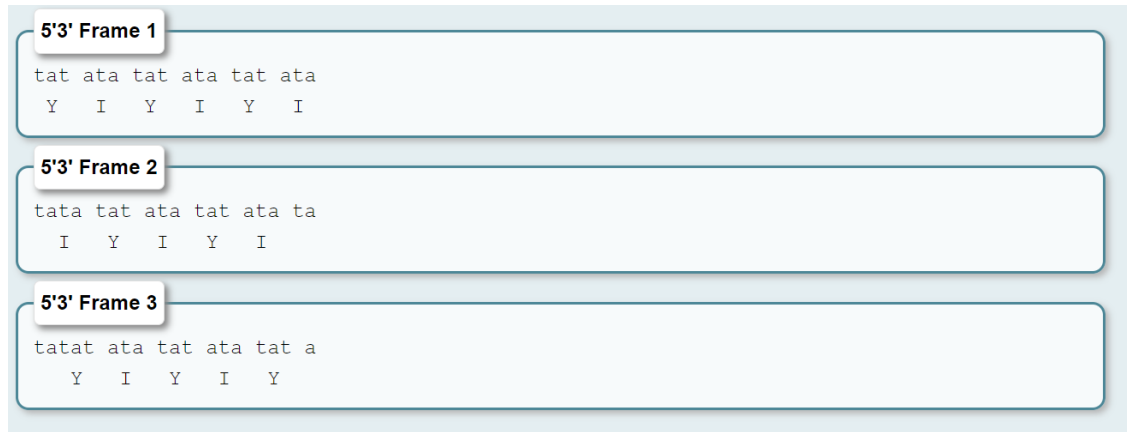
Peru pamuk türüne ait genetik verilerin analizi için ExPASy (Expert Protein Analysis System) biyoinformatik platformu kullanılmıştır. ExPASy aracılığıyla elde edilen protein dizilim verileri ve yapı bilgileri, *G. raimondii*'nin genetik ve fonksiyonel özelliklerinin derinlemesine incelenmesine olanak sağlamış, bu türdeki spesifik genlerin yapısal ve biyokimyasal analizlerini gerçekleştirerek, pamuk bitkilerinin evrimsel ve genetik çeşitliliğinin anlaşılmasına önemli katkılarda bulunmuştur. Peru pamuğuna ait 46 gen arasında 6 tanesi *CesA*, *CesA-like* ve olası *CesA* genlerinden oluşmakta iken; 8'i selüloz sentaz benzeri (*Csl*) genlerinden oluşmaktadır. Mikrosatellit tanımlama kriterlerine göre genleri içeren toplam 19 mikrosatellit tanımlanmıştır.

Şekil 4.47'de D09 kromozomunda bulunan, ID numarası 105800592 olan *CslG2* genine ait Peru pamuğu 14 kez tekrarlanan [AT]₁₄ motifine dönüşüp I (İzolösin) ve Y (Tirozin) aminoasitlerini, 5'=>3' Frame 1'de ATA-TAT motifine dönüşüp I (İzolösin) ve Y (Tirozin) aminoasitlerini, 5'=>3' Frame 2'de ATAT-ATA motifini alıp Y (Tirozin) ve I (İzolösin) aminoasitlerini, Frame 3'te ise ATATA-TAT motifinde olup, I (İzolösin) ve Y (Tirozin) aminoasitlerini kodlamaktadır.



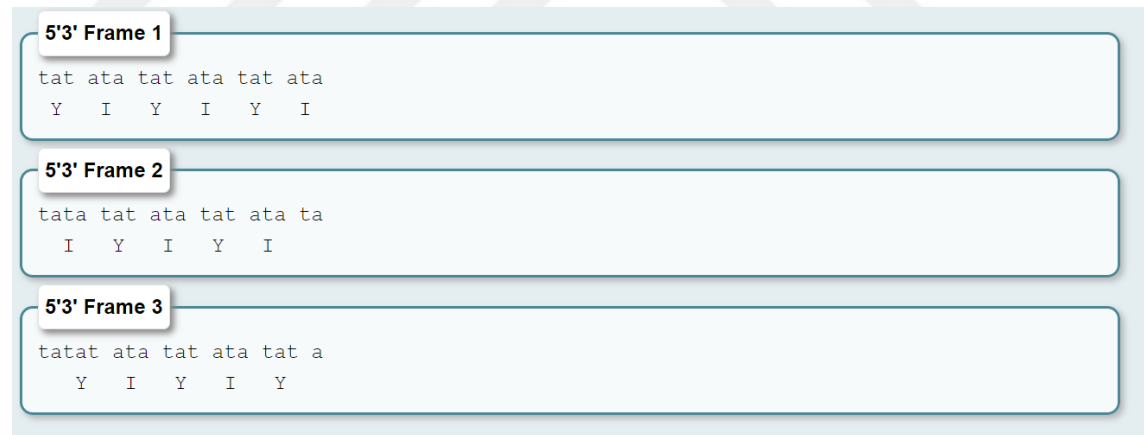
Şekil 4.47. Peru pamuğuna ait ID numarası 105800592 olan *CslG2* genine ait [AT]₁₄ motifinin kodladığı aminoasitler

Şekil 4.48'de D11 kromozomunda bulunan, ID numarası 105804459 olan *CslE1* genine ait Peru pamukta 9 kez tekrarlanan [TA]₉ motifine dönüşüp Y (Tirozin) ve I (İzolösin) aminoasitlerini, 5' => 3' Frame 1'de TAT-ATA motifine dönüşüp Y (Tirozin) ve I (İzolösin) aminoasitlerini, 5' => 3' Frame 2'de TATAT-TAT motifini alıp I (İzolösin) ve Y (Tirozin) aminoasitlerini, Frame 3'te ise TATAT-ATA motifinde olup, Y (Tirozin) ve I (İzolösin) aminoasitlerini kodlamaktadır.



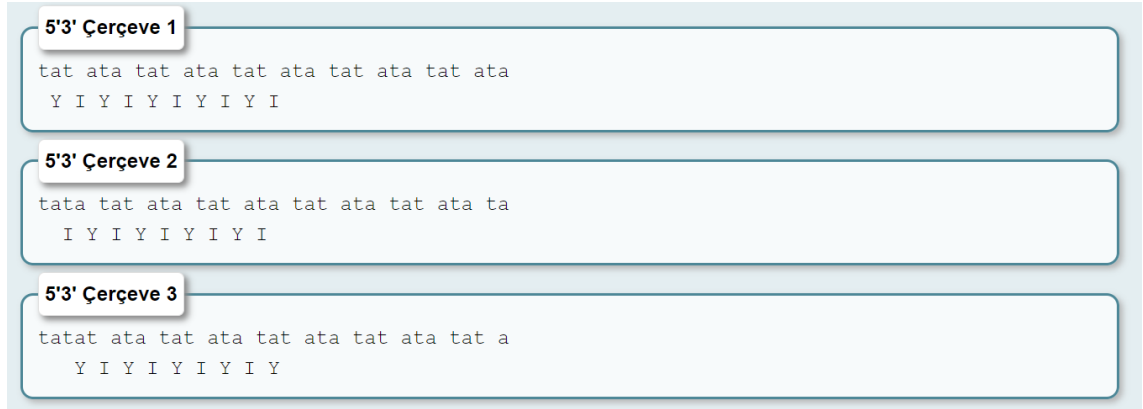
Şekil 4.48. Peru pamuğuna ait ID numarası 105804459 olan *CsIE1* genine ait [TA]₉ motifinin kodladığı aminoasitler

Şekil 4.49'da D12 kromozomunda bulunan, ID numarası 105765323 olan *CsIG3* genine ait Peru pamuğu 9 kez tekrarlanan [TA] motifi 5' => 3' Frame 1'de TAT-ATA motifine dönüşüp Y (Tirozin) ve I (İzolösün) aminoasitlerini, 5' => 3' Frame 2'de TATA-TAT motifini alıp I (İzolösün) ve Y (Tirozin) aminoasitlerini, Frame 3'te ise TATAT-ATA motifinde olup, Y (Tirozin) ve I (İzolösün) aminoasitlerini kodlamaktadır.



Şekil 4.49. Peru pamuğuna ait ID numarası 105765323 olan *CsIG3* genine ait [TA]₉ motifinin kodladığı aminoasitler

Şekil 4.50'de D13 kromozomunda bulunan, ID numarası 105781668 olan *CsIE6* genine ait Peru pamuğu 15 kez tekrarlanan [TA] motifi 5' => 3' Frame 1'de TAT-ATA motifine dönüşüp Y (Tirozin) ve I (İzolösün) aminoasitlerini, 5' => 3' Frame 2'de TATA-TAT motifini alıp I (İzolösün) ve Y (Tirozin) aminoasitlerini, Frame 3'te ise TATAT-ATA motifinde olup, Y (Tirozin) ve I (İzolösün) aminoasitlerini kodlamaktadır.



Şekil 4.50. Peru pamuğuna ait ID numarası 105781668 olan *CsIE6* genine ait [TA]₁₅ motifinin kodladığı aminoasitler

Şekil 4.51'de D13 kromozomunda bulunan, ID numarası 128039937 olan *CesA8-like* genine ait Peru pamuğu 18 kez tekrarlanan [AT] motifi 5' => 3' Frame 1'de ATA-TAT motifine dönüşüp I (İzolösin) ve Y (Tirozin) aminoasitlerini, 5' => 3' Frame 2'de ATAT-ATA motifini alıp Y (Tirozin) ve I (İzolösin) aminoasitlerini, Frame 3'te ise ATATA-TAT motifinde olup, I (İzolösin) ve Y (Tirozin) aminoasitlerini kodlamaktadır.



Şekil 4.51. Peru pamuğuna ait ID numarası 128039937 olan *CesA8-like* genine ait [AT]₁₈ motifinin kodladığı aminoasitler

Şekil 4.52’de D03 kromozomunda bulunan, ID numarası 128039937 olan *CesA8-like* genine ait Peru pamuğu 11 kez tekrarlanan [TA] motifini 5' => 3' Frame 1’de TAT-ATA motifine dönüşüp Y (Tirozin) ve I (İzolösin) aminoasitlerini, 5' => 3' Frame 2’de TATA-TAT motifini alıp I (İzolösin) ve Y (Tirozin) aminoasitlerini, Frame 3’te TATAT-ATA motifinde olup, Y (Tirozin) ve I (İzolösin) aminoasitlerini kodlamaktadır.

5'3' Frame 1
tat ata tat ata tat ata tat a
Y I Y I Y I Y
5'3' Frame 2
tata tat ata tat ata tat ata
I Y I Y I Y I
5'3' Frame 3
tatat ata tat ata tat ata ta
Y I Y I Y I

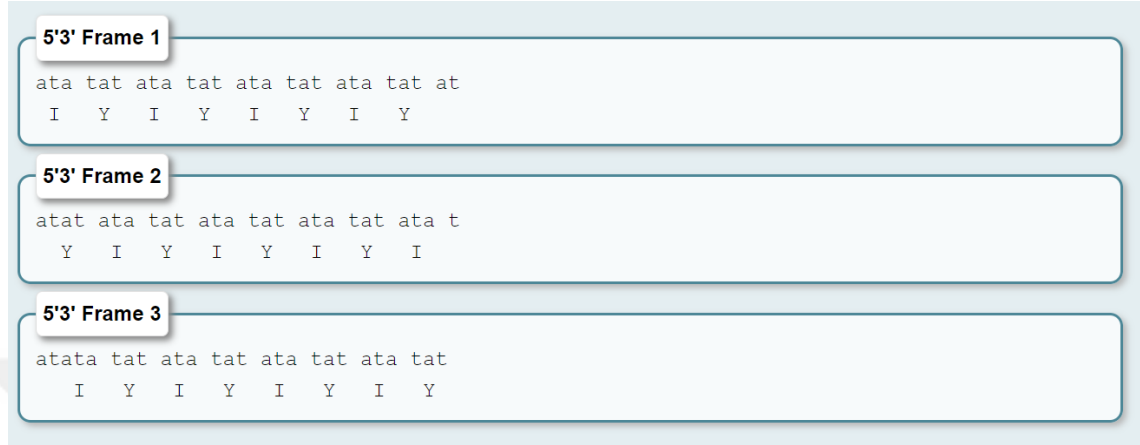
Şekil 4.52. Peru pamuğuna ait ID numarası 128039937 olan *CesA8-like* genine ait [TA]₁₁ motifinin kodladığı aminoasitler

Şekil 4.53’te D03 kromozomunda bulunan, ID numarası 128039937 olan *CesA8-like* genine ait Peru pamuğu 7 kez tekrarlanan [TGG] motifini 5' => 3' Frame 1’de TGG-TGG motifine dönüşüp W (Tryptofan) aminoasitini, 5' => 3' Frame 2’de TGGT-GGT motifini alıp G (Glisin) aminoasitini, Frame 3’te ise TGGTG-GTG motifinde olup, V (Valin) aminoasitlerini kodlamaktadır.

5'3' Frame 1
tgg tgg tgg tgg tgg tgg tgg
W W W W W W W
5'3' Frame 2
tggT ggt ggt ggt ggt ggt gg
G G G G G G G
5'3' Frame 3
tggTg gtg gtg gtg gtg gtg g
V V V V V V V

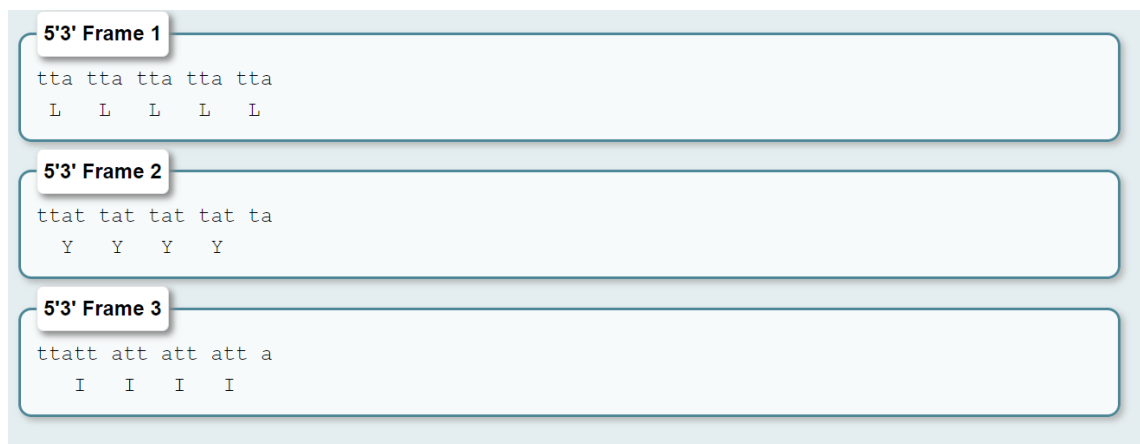
Şekil 4.53. Peru pamuğuna ait ID numarası 128039937 olan *CesA8-like* genine ait [TGG]₇ motifinin kodladığı aminoasitler

Şekil 4.54'te D05 kromozomunda bulunan, ID numarası 105767346 olan *CesA1* genine ait Peru pamuğu 13 kez tekrarlanan [AT] motifi 5' => 3' Frame 1'de ATA-TAT motifine dönüşüp I (İzolösin) ve Y (Tirozin) aminoasitlerini, 5' => 3' Frame 2'de ATAT-ATA motifini alıp Y (Tirozin) ve I (İzolösin) aminoasitini, Frame 3'te ise ATATA-TAT motifinde olup, I (İzolösin) ve Y (Tirozin) aminoasitlerini kodlamaktadır.



Şekil 4.54. Peru pamuğuna ait ID numarası 105767346 olan *CesA1* genine ait [AT]₁₃ motifinin kodladığı aminoasitler

Şekil 4.55'te D05 kromozomunda bulunan, ID numarası 105767346 olan *CesA1* genine ait Peru pamuğu 5 kez tekrarlanan [TTA] motifi 5' => 3' Frame 1'de TTA-TTA motifine dönüşüp L (Lösin) aminoasitini, 5' => 3' Frame 2'de TTAT-TAT motifini alıp Y (Tirozin) aminoasitinde olmaktadır. Frame 3'te ise TTATT-ATT motifinde olup, I (İzolösin) aminoasitlerini kodlamaktadır.



Şekil 4.55. Peru pamuğuna ait ID numarası 105767346 olan *CesA1* genine ait [TTA]₅ motifinin kodladığı aminoasitler

Şekil 4.56’da D05 kromozomunda bulunan, ID numarası 105767346 olan *CesA1* genine ait Peru pamuğu 4 kez tekrarlanan [TTTA] motifi 5' => 3' Frame 1’de TTT-ATT motifine dönüşüp F (Fenilalanin), I (İzolösin), Y (Tirozin) ve L (Lösin) aminoasitlerini, 5' => 3' Frame 2’de TTTA-TTT motifini alıp L (Lösin), F (Fenilalanin), I (İzolösin), Y (Tirozin) aminoasitinde olmaktadır. Frame 3’te ise TTTAT-TTA motifinde olup, Y (Tirozin), L (Lösin), F (Fenilalanin) ve I (İzolösin) aminoasitlerini kodlamaktadır.

5'3' Frame 1
ttt att tat tta ttt a
F I Y L F
5'3' Frame 2
ttta ttt att tat tta
L F I Y L
5'3' Frame 3
tttat tta ttt att ta
Y L F I

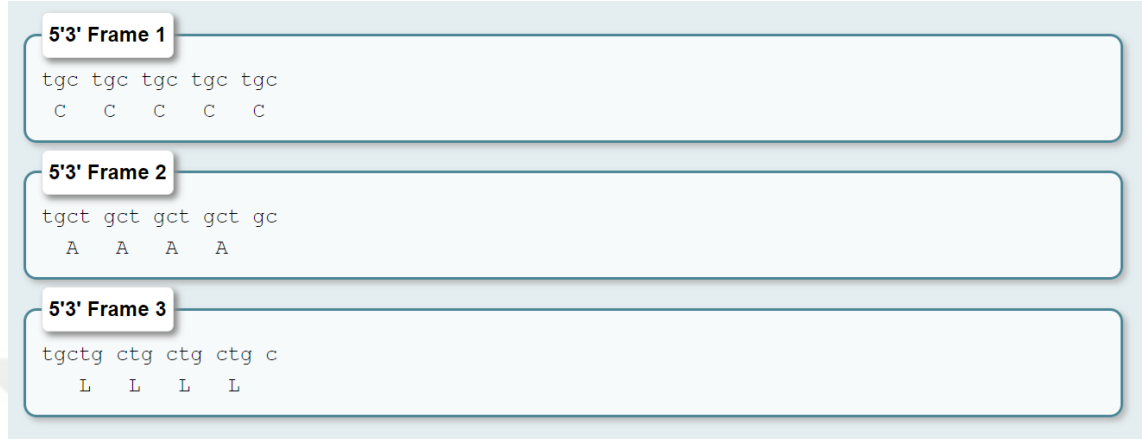
Şekil 4.56. Peru pamuğuna ait ID numarası 105767346 olan *CesA1* genine ait [TTTA]₄ motifinin kodladığı aminoasitler

Şekil 4.57’de D05 kromozomunda bulunan, ID numarası 105767346 olan *CesA1* genine ait Peru pamuğu 5 kez tekrarlanan [CGACC] motifi 5' => 3' Frame 1’de CGA-CCC motifine dönüşüp R (Arjinin), P (Prolin), D (Aspartik asit), T (Treonin) aminoasitlerini, 5' => 3' Frame 2’de CGAC-CCG motifini alıp D (Aspartik asit), P (Prolin), T (Treonin), R (Arjinin) aminoasitinde olmaktadır. Frame 3’te ise CGACC-CGA motifinde olup, T (Treonin), R (Arjinin), P (Prolin) ve D (Aspartik asit) aminoasitlerini kodlamaktadır.

5'3' Frame 1
cga ccc gac ccg acc cga ccc gac c
R P D P T R P D
5'3' Frame 2
cgac ccg acc cga ccc gac ccg acc
D P T R P D P T
5'3' Frame 3
cgacc cga ccc gac ccg acc cga cc
T R P D P T R

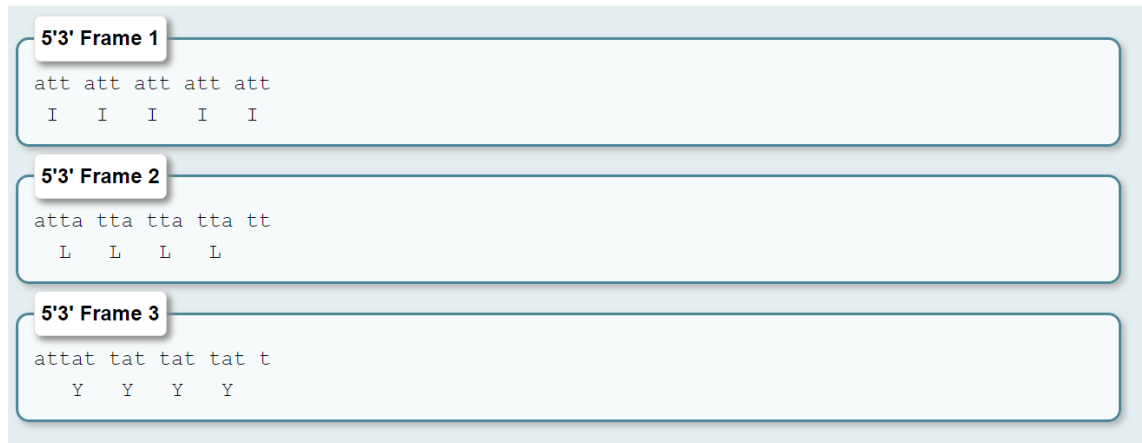
Şekil 4.57. Peru pamuğuna ait ID numarası 105767346 olan *CesA1* genine ait [CGACC]₅ motifinin kodladığı aminoasitler

Şekil 4.58’de D05 kromozomunda bulunan, ID numarası 105769839 olan *CesA1* genine ait Peru pamuğu 5 kez tekrarlanan [TGC] motifi 5' => 3' Frame 1’de TGC-TGC motifine dönüşüp C (Sistein) aminoasitini, 5' => 3' Frame 2’de TGCT-GCT motifini alıp A (Alanin) aminoasitinde olmaktadır. Frame 3’te ise TGCTG-CTG motifinde olup, L (Lösin) aminoasitlerini kodlamaktadır.



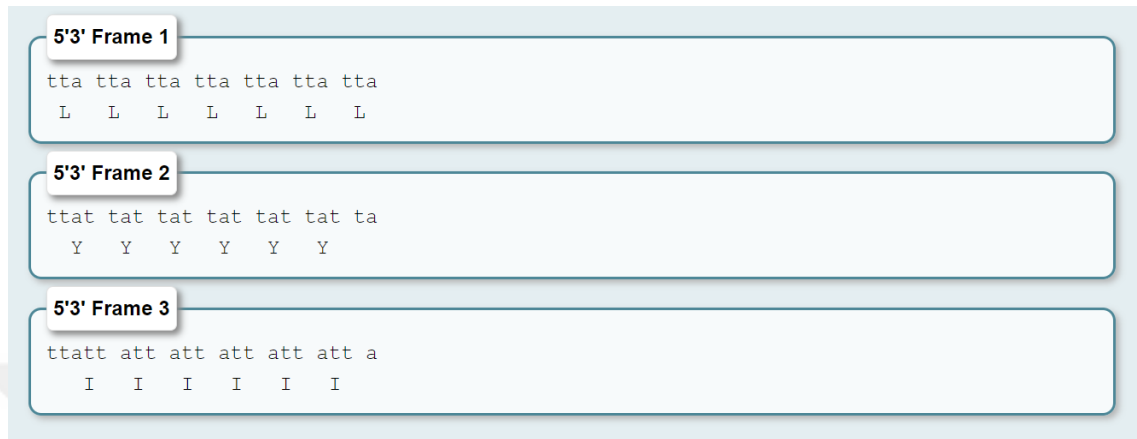
Şekil 4.58. Peru pamuğuna ait ID numarası 105769839 olan *CesA1* genine ait [TGC]₅ motifinin kodladığı aminoasitler

Şekil 4.59’da D09 kromozomunda bulunan, ID numarası 105800594 olan *CsIE1* genine ait Peru pamuğu 5 kez tekrarlanan [ATT] motifi 5' => 3' Frame 1’de ATT-ATT motifine dönüşüp I (İzolösin) aminoasitini, 5' => 3' Frame 2’de ATTA-TTA motifini alıp L (Lösin) aminoasitini, Frame 3’te ise ATTAT-TAT motifinde olup, Y (Tirozin) aminoasitlerini kodlamaktadır.



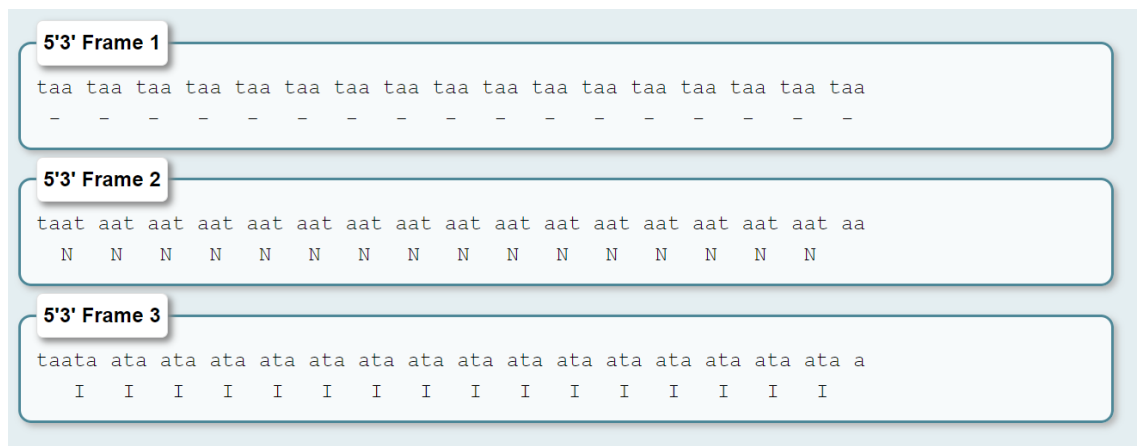
Şekil 4.59. Peru pamuğuna ait ID numarası 105800594 olan *CsIE1* genine ait [ATT]₅ motifinin kodladığı aminoasitler

Şekil 4.60'da D09 kromozomunda bulunan, ID numarası 105785384 olan *CesA2* genine ait Peru pamuğu 7 kez tekrarlanan [TTA] motifi 5' => 3' Frame 1'de TTA-TTA motifine dönüşüp L (Lösin) aminoasitini, 5' => 3' Frame 2'de TTAT-TAT motifini alıp Y (Tirozin) aminoasitini, Frame 3'te ise TTATT-ATT motifinde olup, I (İzolösin) aminoasitlerini kodlamaktadır.



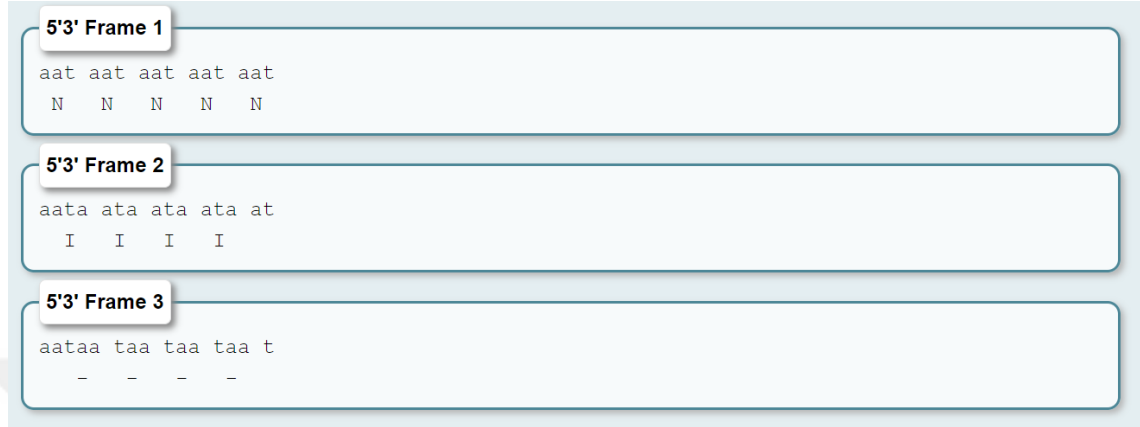
Şekil 4.60. Peru pamuğuna ait ID numarası 105785384 olan *CesA2* genine ait [TTA]₇ motifinin kodladığı aminoasitler

Şekil 4.61’de D09 kromozomunda bulunan, ID numarası 105800206 olan *Cs/LG3* genine ait Peru pamuğu 17 kez tekrarlanan [TAA] motifi 5' => 3' Frame 1’de TAA-TAA motifine dönüşüp durdurma kodonuna karşılık gelmektedir. 5' => 3' Frame 2’de TAAT-AAT motifini alıp N (Asparajin) aminoasitinde olmaktadır. Frame 3’te ise TAATA-ATA motifinde olup, I (İzolösin) aminoasitlerini kodlamaktadır.



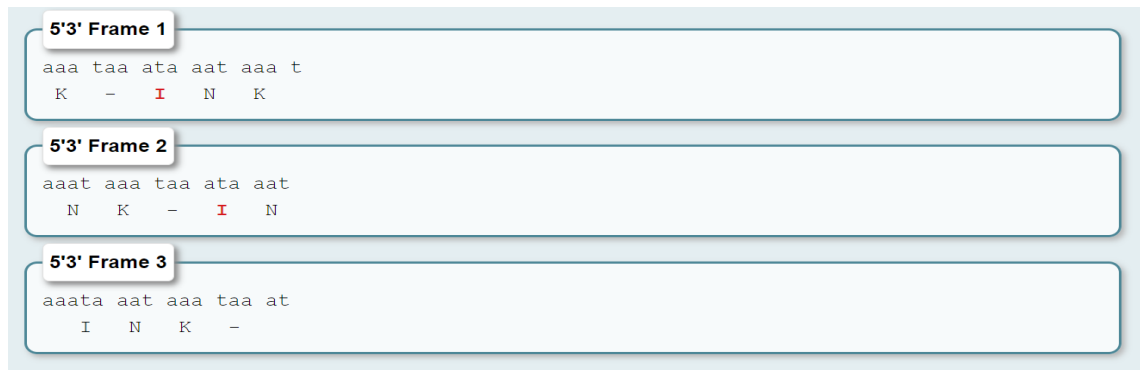
Şekil 4.61. Peru pamuğuna ait ID numarası 105800206 olan Cs/G3 genine ait [TAA]₁₇ motifinin kodladığı aminoasitler

Şekil 4.62’de D13 kromozomunda bulunan, ID numarası 105764976 olan *CsID3* genine ait Peru pamuğu 5 kez tekrarlanan [AAT] motifi 5' => 3' Frame 1’de AAT-AAT motifine dönüşüp N (Asparajin) aminoasitini, 5' => 3' Frame 2’de AATA-ATA motifini alıp I (İzolösin) aminoasitinde olmaktadır. Frame 3’te ise AATAA-TAA motifinde olup durdurma kodonlarını kodlamaktadır.



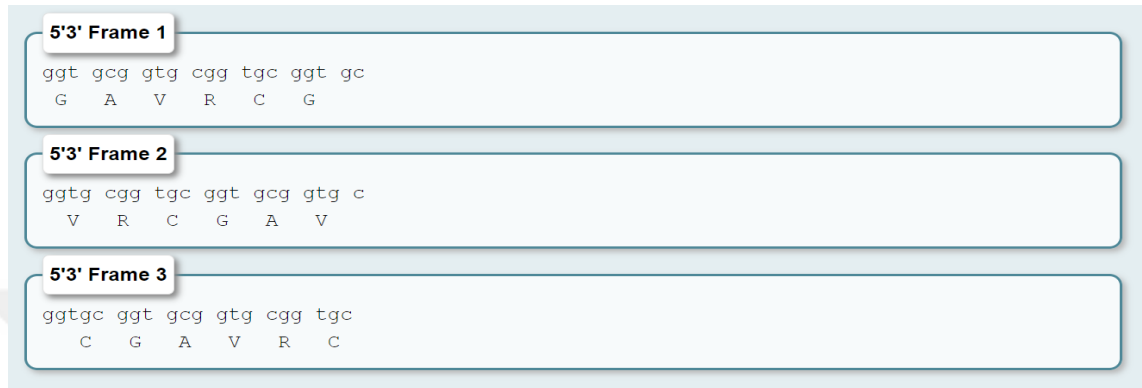
Şekil 4.62. Peru pamuğuna ait ID numarası 105764976 olan *CsID3* genine ait [AAT]₅ motifinin kodladığı aminoasitler

Şekil 4.63’te D13 kromozomunda bulunan, ID numarası 105781684 olan *Prob-CesA5* genine ait Peru pamuğu 4 kez tekrarlanan [AAAT] motifi 5' => 3' Frame 1’de AAA-TTA motifine dönüşüp K (Lisin), TAA (durdurma kodonu), I (İzolösin), N (Asparajin)’i, 5' => 3' Frame 2’de AAAT-AAA motifini alıp N (Asparajin), K (Lisin), TAA (durdurma kodonu), I (İzolösin) aminositlerinden, Frame 3’te ise AAATA-AAT motifinde olup I (İzolösin), N (Asparajin), K (Lisin) ve TAA (durdurma kodonu)’dan oluşmaktadır.



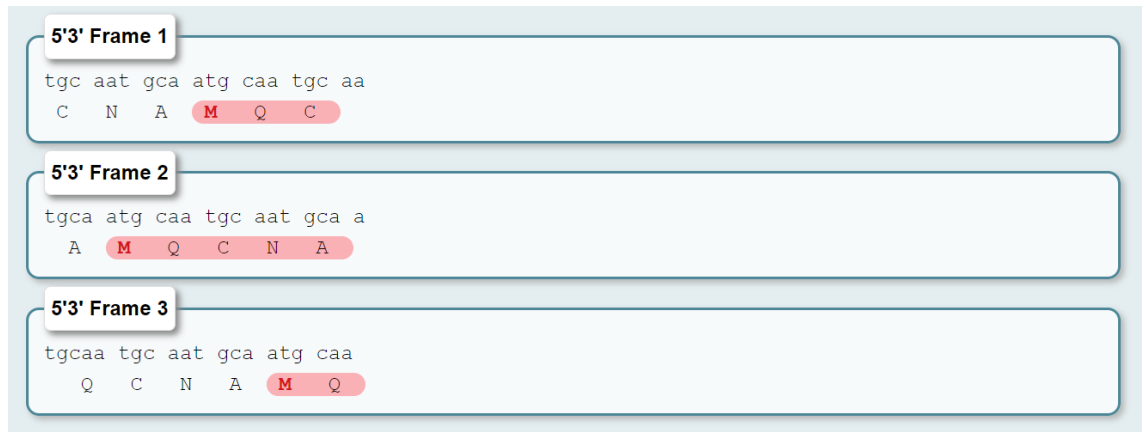
Şekil 4.63. Peru pamuğuna ait ID numarası 105781684 olan *Prob-CesA5* genine ait [AAAT]₄ motifinin kodladığı aminoasitler

Şekil 4.64'te D09 kromozomunda bulunan, ID numarası 105800591 olan *CsIG3* genine ait Peru pamuğu 4 kez tekrarlanan [GGTGC] motifi 5' => 3' Frame 1'de GGT-GCG motifine dönüşüp G (Glisin), A (Alanin), V (Valin), R (Arjinin) ve C (Sistein) aminoasitlerinden oluşmaktadır. 5' => 3' Frame 2'de GGTG-CGG motifini alıp V (Valin), R (Arjinin), C (Sistein), G (Glisin) ve A (Alanin)'den oluşmaktadır. Frame 3'te ise GGTGC-GGT motifinde olup C (Sistein), G (Glisin), A (Alanin), V (Valin), R (Arjinin) aminoasitlerinden oluşmaktadır.



Şekil 4.64. Peru pamuğuna ait ID numarası 105800591 olan *CsIG3* genine ait [GGTGC]₄ motifinin kodladığı aminoasitler

Şekil 4.65'te D09 kromozomunda bulunan, ID numarası 105800591 olan *CsIG3* genine ait Peru pamuğu 4 kez tekrarlanan [TGCAA] motifi 5' => 3' Frame 1'de TGC-AAT motifine dönüşüp C (Sistein), N (Asparajin), A (Alanin), M (Metiyonin), Q (Glutamin) aminoasitlerinden oluşmaktadır. 5' => 3' Frame 2'de TGCA-ATG motifini alıp A (Alanin), M (Metiyonin), Q (Glutamin), C (Sistein) ve N (Asparajin)'den oluşmaktadır. Frame 3'te ise TGCAA-TGC motifinde olup Q (Glutamin), C (Sistein), N (Asparajin), A (Alanin), M (Metiyonin) aminoasitlerinden oluşmaktadır.



Şekil 4.65. Peru pamuğuna ait ID numarası 105800591 olan *CsIG3* genine ait [TGCAA]₄ motifinin kodladığı aminoasitler

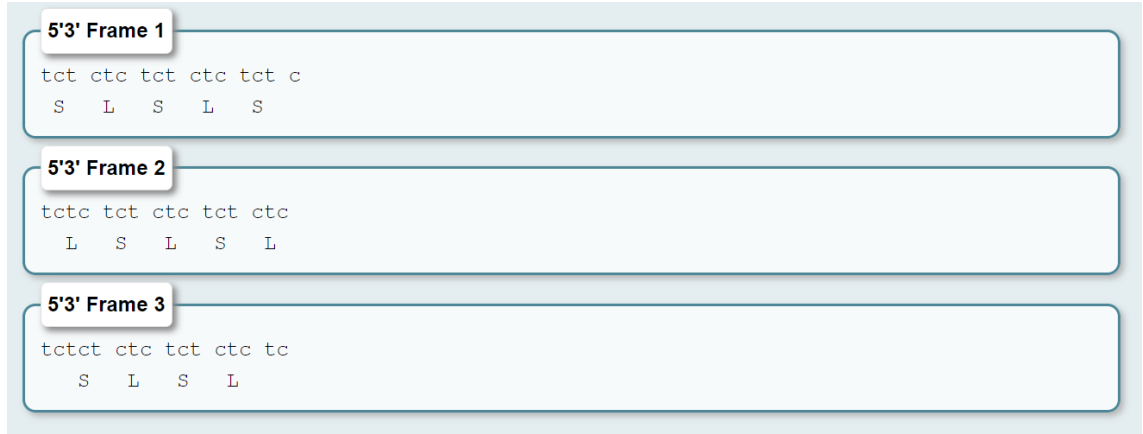
Upland pamuk (*Gossypium hirsutum*) türüne yönelik yapılan bu çalışmada, ExPASy (Expert Protein Analysis System) biyoinformatik aracı kullanılarak, genetik verilerin protein düzeyindeki analizleri gerçekleştirilmiştir. ExPASy platformu, Upland pamukta yer alan kritik genlerin protein dizilimlerini ve fonksiyonel özelliklerini ayrıntılı bir şekilde incelemeye olanak tanıyarak, bu türün biyokimyasal profillerinin ve genetik çeşitliliğinin daha kapsamlı bir şekilde değerlendirilmesini sağlamıştır. Upland pamukta 82 gen arasında 12'si *CesA*, *CesA-like* ve olası *CesA* genlerinden oluşmakta iken; 10 tanesi selüloz sentaz benzeri (*Csl*) genlerinden oluşmaktadır. Mikrosatellit tanımlama kriterlerine göre genleri içeren toplam 27 mikrosatellit tanımlanmıştır.

Şekil 4.66'da A09 kromozomunda bulunan, ID numarası 107928590 olan *CslG2* genine ait Upland pamuğu 8 kez tekrarlanan [AT]₈ motifine dönüşüp I (İzolösin) ve Y (Tirozin) aminoasitlerinden oluşmaktadır. 5' => 3' Frame 1'de ATA-TAT motifine dönüşüp I (İzolösin) ve Y (Tirozin) aminoasitlerinden oluşmaktadır. 5' => 3' Frame 2'de ATAT-ATA motifini alıp Y (Tirozin) ve I (İzolösin) aminoasitlerinden oluşmaktadır. Frame 3'te ise ATATA-TAT motifinde olup I (İzolösin) ve Y (Tirozin) aminoasitlerini kodlamıştır.

5'3' Frame 1
ata tat ata tat ata t
I Y I Y I
5'3' Frame 2
atat ata tat ata tat
Y I Y I Y
5'3' Frame 3
atata tat ata tat at
I Y I Y

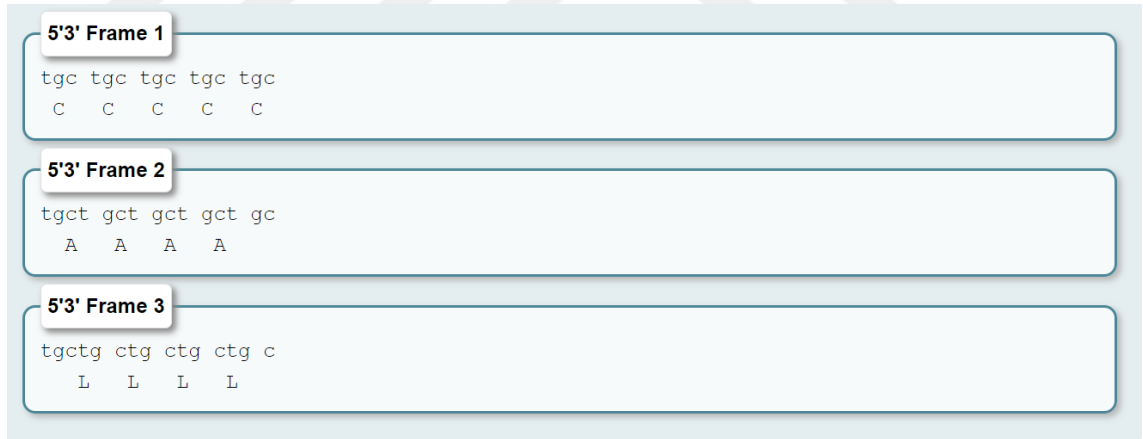
Şekil 4.66. Upland pamuğuna ait ID numarası 107928590 olan *CslG2* genine ait [AT]₈ motifinin kodladığı aminoasitler

Şekil 4.67'de D05 kromozomunda bulunan, ID numarası 107914099 olan *CesA5* genine ait Upland pamuğu 8 kez tekrarlanan [TC]₈ motifine dönüşüp S (Serin) ve L (Lösin) aminoasitlerinden oluşmaktadır. 5' => 3' Frame 1'de TCT-CTC motifine dönüşüp S (Serin) ve L (Lösin) aminoasitlerinden oluşmaktadır. 5' => 3' Frame 2'de TCTC-TCT motifini alıp L (Lösin) ve S (Serin)'den oluşmaktadır. Frame 3'te ise TCTCT-CTC motifinde olup S (Serin) ve L (Lösin) aminoasitlerinden oluşmaktadır.



Şekil 4.67. Upland pamuğuna ait ID numarası 107914099 olan *CesA5* genine ait [TC]₈ motifinin kodladığı aminoasitler

Şekil 4.68’de D05 kromozomunda bulunan, ID numarası 107914099 olan *CesA5* genine ait Upland pamuğu 5 kez tekrarlanan [TGC] motifi 5' => 3' Frame 1’de TGC-TGC motifine dönüşüp C (Sistein) aminoasitlerinden oluşmaktadır. 5' => 3' Frame 2’de TGCT-GCT motifini alıp A (Alanin)’den oluşmaktadır. Frame 3’te ise TGCTG-CTG motifinde olup L (Lösin) aminoasitlerinden oluşmaktadır.



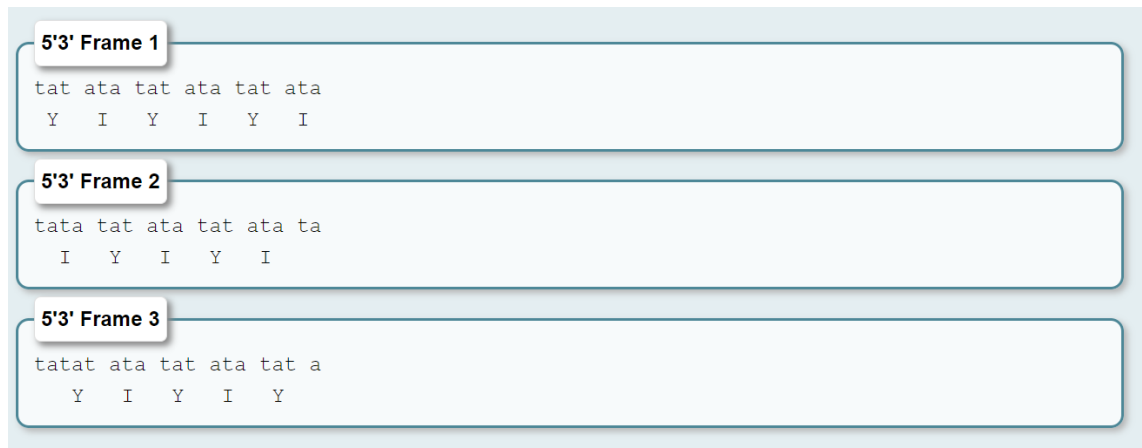
Şekil 4.68. Upland pamuğuna ait ID numarası 107914099 olan *CesA5* genine ait [TGC]₅ motifinin kodladığı aminoasitler

Şekil 4.69’da D08 kromozomunda bulunan, ID numarası 107900907 olan *CsII1* genine ait Upland pamuğu 11 kez tekrarlanan [AG] motifi 5' => 3' Frame 1’de AGA-GAG motifine dönüşüp R (Arjinin) ve E (Glutamik asit) aminoasitlerinden oluşmaktadır. 5' => 3' Frame 2’de AGAG-AGA motifini alıp E (Glutamik asit) ve R (Arjinin)’den oluşmaktadır. Frame 3’te ise AGAGA-GAG motifinde olup R (Arjinin) ve E (Glutamik asit) aminoasitlerinden oluşmaktadır.



Şekil 4.69. Upland pamuğuna ait ID numarası 107900907 olan *CsII1* genine ait [AG]₁₁ motifinin kodladığı aminoasitler

Şekil 4.70’te A09 kromozomunda bulunan, ID numarası 107896452 olan *CsIG3* genine ait Upland pamuğu 9 kez tekrarlanan [TA] motifi 5' => 3' Frame 1’de TAT-ATA motifine dönüşüp Y (Tirozin) ve I (İzolösin) aminoasitlerinden oluşmaktadır. 5' => 3' Frame 2’de TATA-TAT motifini alıp I (İzolösin) ve Y (Tirozin)’den oluşmaktadır. Frame 3’te ise TATAT-ATA motifinde olup Y (Tirozin) ve I (İzolösin) aminoasitlerinden oluşmaktadır.



Şekil 4.70. Upland pamuğuna ait ID numarası 107896452 olan *CsIG3* genine ait [TA]₉ motifinin kodladığı aminoasitler

Şekil 4.71’de D05 kromozomunda bulunan, ID numarası 107907451 olan *CesA7* genine ait Upland pamuğu 12 kez tekrarlanan [TA] motifi 5' => 3' Frame 1’de TAT-ATA motifine dönüşüp Y (Tirozin) ve I (İzolösin) aminoasitlerinden oluşmaktadır. 5' => 3' Frame 2’de TATA-TAT motifini alıp I (İzolösin) ve Y (Tirozin)’den oluşmaktadır. Frame 3’te ise TATAT-ATA motifinde olup Y (Tirozin) ve I (İzolösin) aminoasitlerinden oluşmaktadır.

5'3' Frame 1
tat ata tat ata tat ata tat ata
Y I Y I Y I Y I

5'3' Frame 2
tata tat ata tat ata tat ata ta
I Y I Y I Y I

5'3' Frame 3
tatat ata tat ata tat ata tat a
Y I Y I Y I Y

Şekil 4.71. Upland pamuğuna ait ID numarası 107907451 olan *CesA7* genine ait [TA]₁₂ motifinin kodladığı aminoasitler

Şekil 4.72’de D05 kromozomunda bulunan, ID numarası 107907451 olan *CesA7* genine ait Upland pamuğu 6 kez tekrarlanan [TAA] motifi 5' => 3' Frame 1’de TAA-TAA motifine dönüşüp stop kodonlarından oluşmaktadır. 5' => 3' Frame 2’de TAAT-AAT motifini alıp N (Asparajin) aminoasitinden oluşmaktadır. Frame 3’te ise TAATA-ATA motifinde olup I (İzolösin) aminoasitlerinden oluşmaktadır.

5'3' Frame 1
taa taa taa taa taa taa
- - - - -

5'3' Frame 2
taat aat aat aat aat aa
N N N N N

5'3' Frame 3
taata ata ata ata ata a
I I I I I

Şekil 4.72. Upland pamuğuna ait ID numarası 107907451 olan *CesA7* genine ait [TAA]₆ motifinin kodladığı aminoasitler

Şekil 4.73'te D05 kromozomunda bulunan, ID numarası 107907451 olan *CesA7* genine ait Upland pamuğu 4 kez tekrarlanan [TTTA] motif 5' => 3' Frame 1'de TTT-ATT motifine dönüşüp F (Fenilalanin), I (İzolösin), Y (Tirozin) ve L (Lösin) aminoasitlerinden oluşmaktadır. 5' => 3' Frame 2'de TTTA-TTT motifini alıp L (Lösin), F (Fenilalanin), I (İzolösin) ve Y (Tirozin) aminoasitinden oluşmaktadır. Frame 3'te ise TTTAT-TTA motifinde olup Y (Tirozin), L (Lösin), F (Fenilalanin), I (İzolösin) aminoasitlerinden oluşmaktadır.

5'3' Frame 1
ttt att tat tta ttt a
F I Y L F
5'3' Frame 2
ttta ttt att tat tta
L F I Y L
5'3' Frame 3
tttat tta ttt att ta
Y L F I

Şekil 4.73. Upland pamuğuna ait ID numarası 107907451 olan *CesA7* genine ait [TTTA]₄ motifinin kodladığı aminoasitler

Şekil 4.74'te D05 kromozomunda bulunan, ID numarası 107953681 olan *CesA7-like* genine ait Upland pamuğu 8 kez tekrarlanan [AC] motif 5' => 3' Frame 1'de ACA-CAC motifine dönüşüp T (Treonin), H (Histidin) aminoasitlerinden oluşmaktadır. 5' => 3' Frame 2'de ACAC-ACA motifini alıp H (Histidin), T (Treonin) aminoasitlerinden oluşmaktadır. Frame 3'te ise ACACA-CAC motifinde olup T (Treonin) ve H (Histidin) aminoasitlerinden oluşmaktadır.

5'3' Frame 1
aca cac aca cac aca c
T H T H T
5'3' Frame 2
acac aca cac aca cac
H T H T H
5'3' Frame 3
acaca cac aca cac ac
T H T H

Şekil 4.74. Upland pamuğuna ait ID numarası 107953681 olan *CesA7-like* genine ait [AC]₈ motifinin kodladığı aminoasitler

Şekil 4.75'te D11 kromozomunda bulunan, ID numarası 107935226 olan *CsIE1* genine ait Upland pamuğu 8 kez tekrarlanan [TA]₈ motifine dönüşüp Y (Tirozin), I (İzolösin) aminoasitlerinden oluşmaktadır. 5' => 3' Frame 2'de TATA-TAT motifini alıp I (İzolösin), Y (Tirozin) aminoasitlerinden oluşmaktadır. Frame 3'te ise TATAT-ATA motifinde olup Y (Tirozin), I (İzolösin) aminoasitlerinden oluşmaktadır.

5'3' Frame 1
tat ata tat ata tat a
Y I Y I Y
5'3' Frame 2
tata tat ata tat ata
I Y I Y I
5'3' Frame 3
tatat ata tat ata ta
Y I Y I

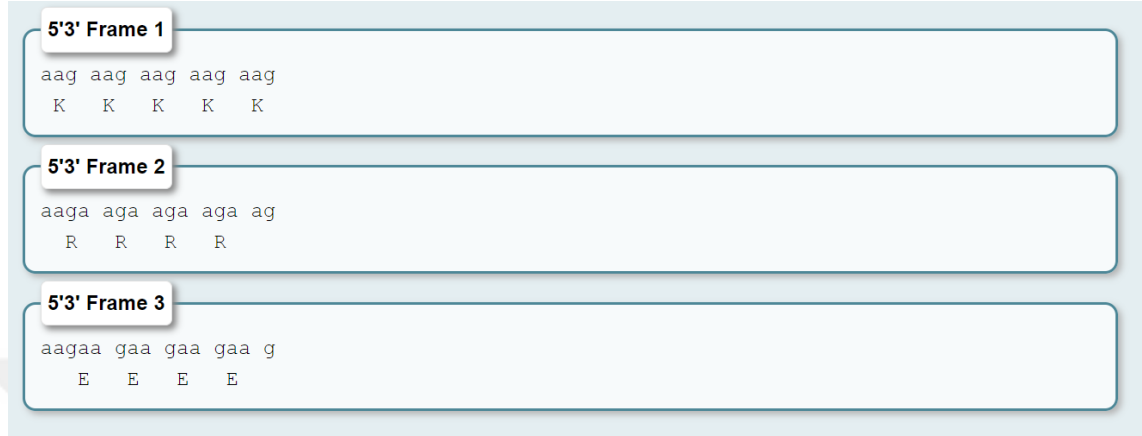
Şekil 4.75. Upland pamuğuna ait ID numarası 107935226 olan *CsIE1* genine ait [TA]₈ motifinin kodladığı aminoasitler

Şekil 4.76'da A05 kromozomunda bulunan, ID numarası 121229536 olan *CesA5-like* genine ait Upland pamuğu 5 kez tekrarlanan [TGC] motifine dönüşüp C (Sistein) aminoasitinden oluşmaktadır. 5' => 3' Frame 2'de TGCT-GCT motifini alıp A (Alanin) aminoasitinden oluşmaktadır. Frame 3'te ise TGCTG-CTG motifinde olup L (Lösin) aminoasitlerinden oluşmaktadır.

5'3' Frame 1
tgc tgc tgc tgc tgc
C C C C C
5'3' Frame 2
tgct gct gct gct gc
A A A A
5'3' Frame 3
tgctg ctg ctg ctg c
L L L L

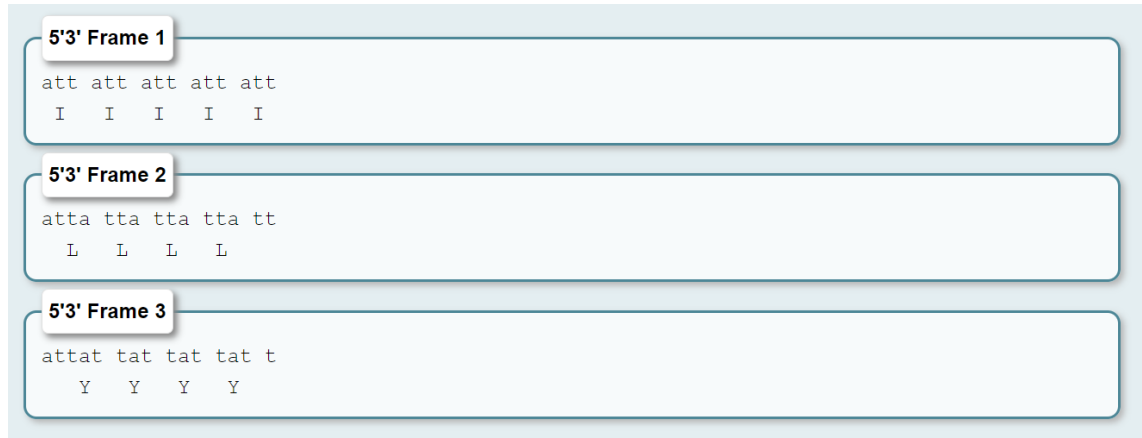
Şekil 4.76. Upland pamuğuna ait ID numarası 121229536 olan *CesA5-like* genine ait [TGC]₅ motifinin kodladığı aminoasitler

Şekil 4.77’de A07 kromozomunda bulunan, ID numarası 107955410 olan *CesA4* genine ait Upland pamuğu 5 kez tekrarlanan [AAG] motifi 5' => 3' Frame 1’de AAG-AAG motifine dönüşüp K (Lisin) aminoasitinden oluşmaktadır. 5' => 3' Frame 2’de AGA-AGA motifini alıp R (Arjinin) aminoasitinden oluşmaktadır. Frame 3’te ise AAGAA-GAA motifinde olup E (Glutamik asit) aminoasitlerinden oluşmaktadır.



Şekil 4.77. Upland pamuğuna ait ID numarası 107955410 olan *CesA4* genine ait [AAG]₅ motifinin kodladığı aminoasitler

Şekil 4.78’de A11 kromozomunda bulunan, ID numarası 107898377 olan *CsII1* genine ait Upland pamuğu 5 kez tekrarlanan [ATT] motifi 5' => 3' Frame 1’de ATT-ATT motifine dönüşüp I (İzolösin) aminoasitinden oluşmaktadır. 5' => 3' Frame 2’de ATTA-TTA motifini alıp L (Lösin) aminoasitinden oluşmaktadır. Frame 3’te ise ATTAT-TAT motifinde olup Y (Tirozin) aminoasitlerinden oluşmaktadır.



Şekil 4.78. Upland pamuğuna ait ID numarası 107898377 olan *CsII1* genine ait [ATT]₅ motifinin kodladığı aminoasitler

Şekil 4.79’da D07 kromozomunda bulunan, ID numarası 107937655 olan *CesA4-like* genine ait Upland pamuğu 5 kez tekrarlanan [AAG] motifi 5' => 3' Frame 1’de AAG-AAG motifine dönüşüp K (Lisin) aminoasitinden oluşmaktadır. 5' => 3' Frame 2’de AAGA-AGA motifini alıp R (Arjinin) aminoasitinden oluşmaktadır. Frame 3’te ise AAGAA-GAA motifinde olup E (Glutamik asit) aminoasitlerinden oluşmaktadır.

5'3' Frame 1
aag aag aag aag aag
K K K K K
5'3' Frame 2
aaga aga aga aga ag
R R R R
5'3' Frame 3
aagaa gaa gaa gaa g
E E E E

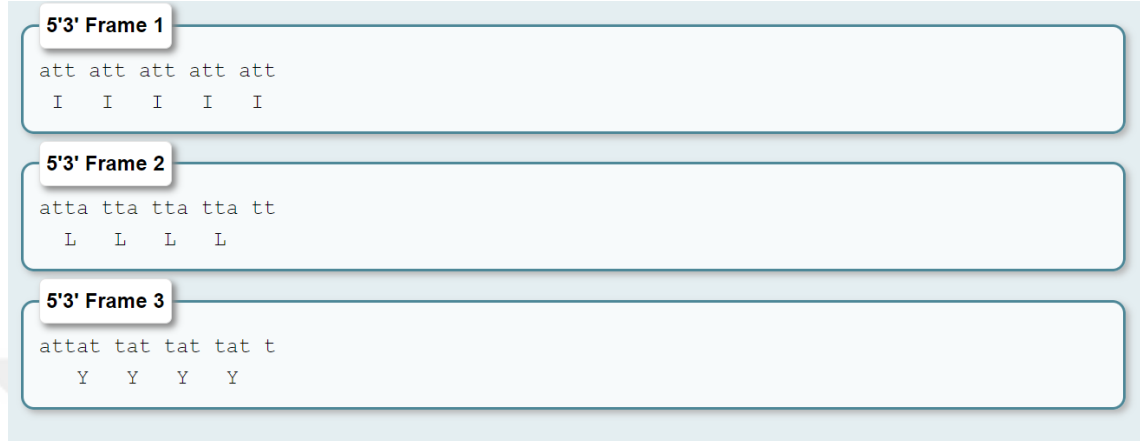
Şekil 4.79. Upland pamuğuna ait ID numarası 107937655 olan *CesA4-like* genine ait [AAG]₅ motifinin kodladığı aminoasitler

Şekil 4.80’de D11 kromozomunda bulunan, ID numarası 107911480 olan *CesA2* genine ait Upland pamuğu 7 kez tekrarlanan [TTA] motifi 5' => 3' Frame 1’de TTA-TTA motifine dönüşüp L (Lösin) aminoasitinden oluşmaktadır. 5' => 3' Frame 2’de TTAT-TAT motifini alıp Y (Tirozin) aminoasitinden oluşmaktadır. Frame 3’te ise TTATT-ATT motifinde olup I (İzolösin) aminoasitlerinden oluşmaktadır.

5'3' Frame 1
tta tta tta tta tta tta tta
L L L L L L L
5'3' Frame 2
ttat tat tat tat tat tat ta
Y Y Y Y Y Y Y
5'3' Frame 3
ttatt att att att att att a
I I I I I I I

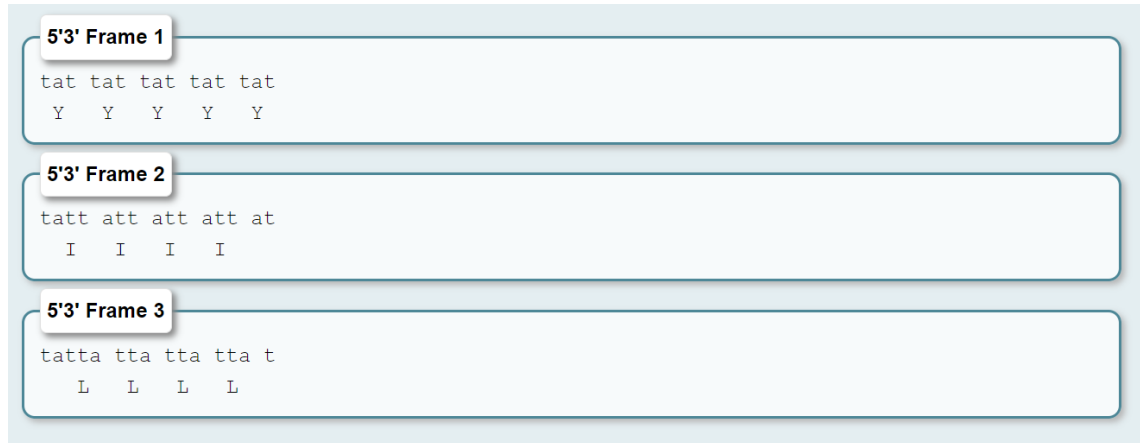
Şekil 4.80. Upland pamuğuna ait ID numarası 107911480 olan *CesA2* genine ait [TTA]₇ motifinin kodladığı aminoasitler

Şekil 4.81’de A04 kromozomunda bulunan, ID numarası 107949226 olan *CsID4* genine ait Upland pamuğu 5 kez tekrarlanan [ATT] motifini 5' => 3' Frame 1’de ATT-ATT motifine dönüşüp I (İzolösin) aminoasitinden oluşmaktadır. 5' => 3' Frame 2’de ATTA-TTA motifini alıp L (Lösin) aminoasitinden oluşmaktadır. Frame 3’te ise ATTAT-TAT motifinde olup Y (Tirozin) aminoasitlerinden oluşmaktadır.



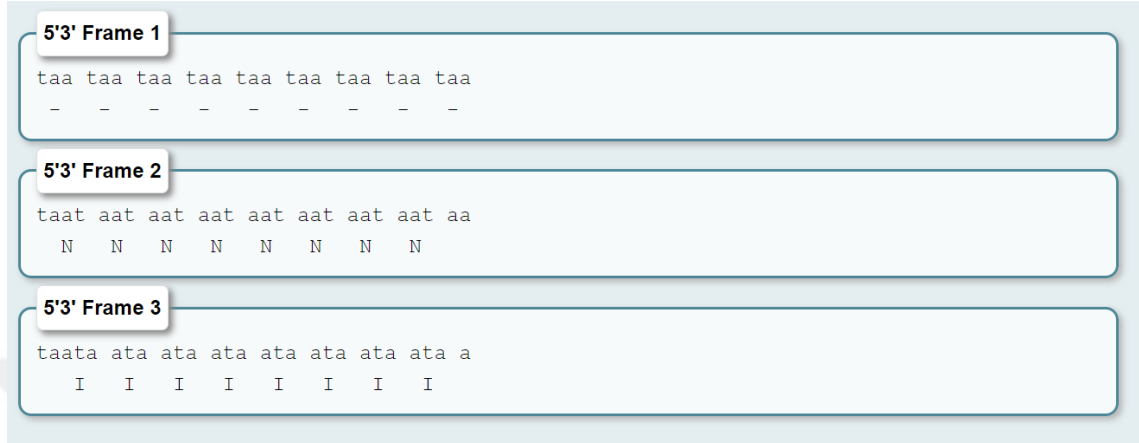
Şekil 4.81. Upland pamuğuna ait ID numarası 107949226 olan *CsID4* genine ait [ATT]₅ motifinin kodladığı aminoasitler

Şekil 4.82’de A05 kromozomunda bulunan, ID numarası 107958687 olan *CesA1* genine ait Upland pamuğu 5 kez tekrarlanan [TAT] motifini 5' => 3' Frame 1’de TAT-TAT motifine dönüşüp Y (Tirozin) aminoasitinden oluşmaktadır. 5' => 3' Frame 2’de TATT-ATT motifini alıp I (Tirozin) aminoasitinden oluşmaktadır. Frame 3’te ise TATTA-TTA motifinde olup L (Lösin) aminoasitlerinden oluşmaktadır.



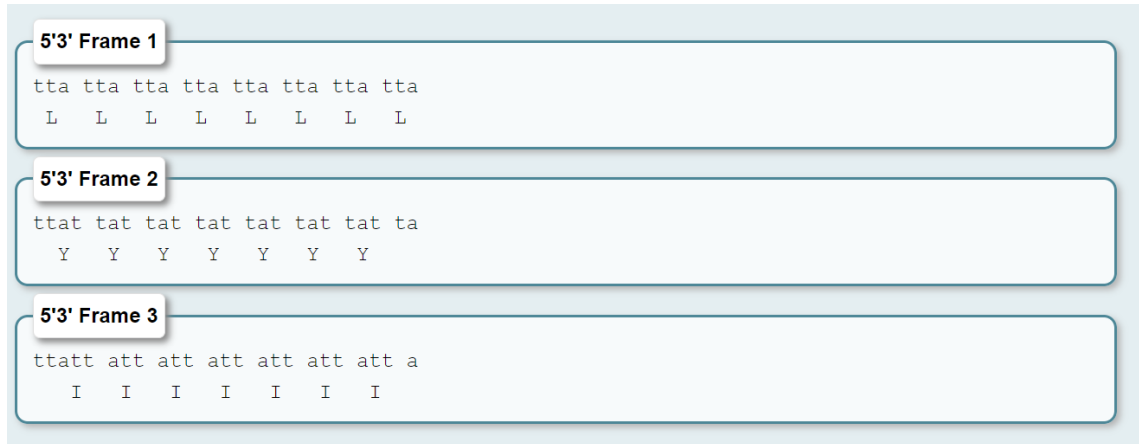
Şekil 4.82. Upland pamuğuna ait ID numarası 107958687 olan *CesA1* genine ait [TAT]₅ motifinin kodladığı aminoasitler

Şekil 4.83'te D05 kromozomunda bulunan, ID numarası 107907453 olan *CesA1* genine ait Upland pamuğu 9 kez tekrarlanan [TAA] motifi 5' => 3' Frame 1'de TAA-TAA motifine dönüşüp stop kodonlarından oluşmaktadır. 5' => 3' Frame 2'de TAAT-AAT motifini alıp N (Asparjin) aminoasitinden oluşmaktadır. Frame 3'te ise TAATA-ATA motifinde olup I (İzolösin) aminoasitlerinden oluşmaktadır.



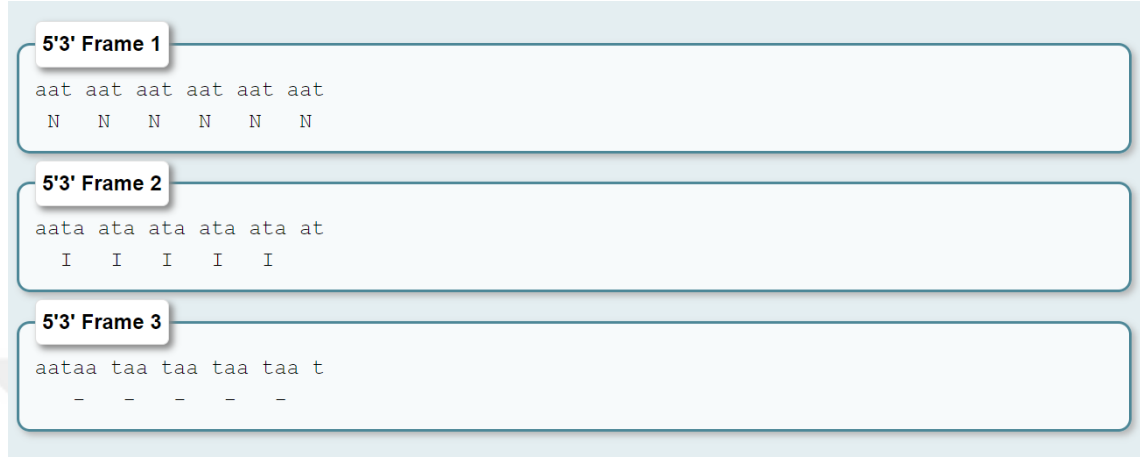
Şekil 4.83. Upland pamuğuna ait ID numarası 107907453 olan *CesA1* genine ait [TAA]₉ motifinin kodladığı aminoasitler

Şekil 4.84'te D06 kromozomunda bulunan, ID numarası 107901762 olan *CesA5* genine ait Upland pamuğu 8 kez tekrarlanan [TAA] motifi 5' => 3' Frame 1'de TTA-TTA motifine dönüşüp L (Lösin) aminoasitinden oluşmaktadır. 5' => 3' Frame 2'de TTAT-TAT motifini alıp Y (Tirozin) aminoasitinden oluşmaktadır. Frame 3'te ise TTATT-ATT motifinde olup I (İzolösin) aminoasitlerinden oluşmaktadır.



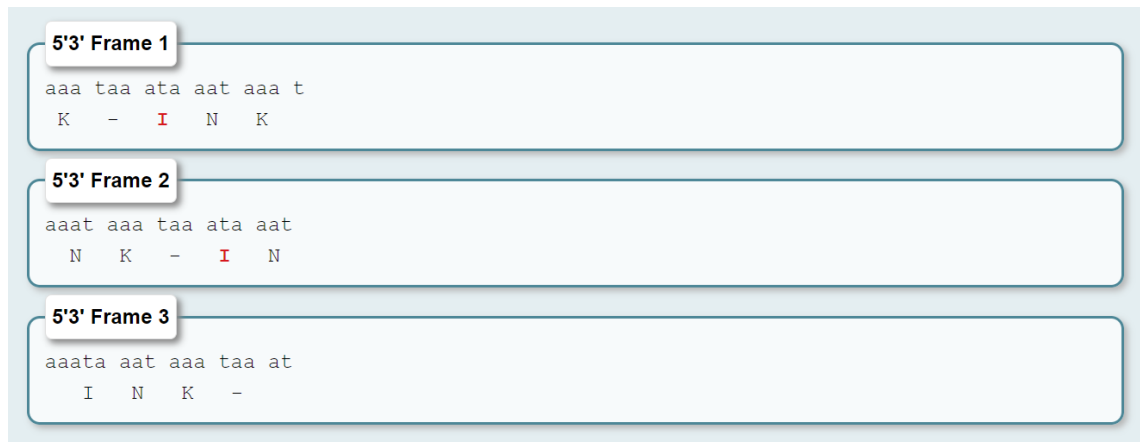
Şekil 4.84. Upland pamuğuna ait ID numarası 107901762 olan *CesA5* genine ait [TAA]₈ motifinin kodladığı aminoasitler

Şekil 4.85'te D11 kromozomunda bulunan, ID numarası 107935218 olan *CsIE1* genine ait Upland pamuğu 6 kez tekrarlanan [AAT] motifi 5' => 3' Frame 1'de AAT-AAT motifine dönüşüp N (Asparjin) aminoasitinden oluşmaktadır. 5' => 3' Frame 2'de AATA-ATA motifini alıp I (İzolösin) aminoasitinden oluşmaktadır. Frame 3'te ise AATAA-TAA motifinde olup stop kodonlarından oluşmaktadır.



Şekil 4.85. Upland pamuğuna ait ID numarası 107935218 olan *CsIE1* genine ait [AAT]₆ motifinin kodladığı aminoasitler

Şekil 4.86'da D11 kromozomunda bulunan, ID numarası 107935218 olan *CsIE1* genine ait Upland pamuğu 4 kez tekrarlanan [AAAT] motifi 5' => 3' Frame 1'de AAA-TAA motifine dönüşüp K (Lisin), TAA (durdurma kodonu), I (İzolösin) ve N (Asparjin)'den oluşmaktadır. 5' => 3' Frame 2'de AAAT-AAA motifini alıp N (Asparjin), K (Lisin), TAA (durdurma kodonu) ve I (İzolösin)'den oluşmaktadır. Frame 3'te ise AAATA-AAT motifinde olup I (İzolösin), N (Asparjin), K (Lisin) ve TAA (stop kodonu)'dan oluşmaktadır.



Şekil 4.86. Upland pamuğuna ait ID numarası 107935218 olan *CsIE1* genine ait [AAAT]₄ motifinin kodladığı aminoasitler

Şekil 4.87’de A09 kromozomunda bulunan, ID numarası 107896451 olan *CsIG2* genine ait Upland pamuğu 4 kez tekrarlanan [TTTC] motifi 5' => 3' Frame 1’de TTC-CTT motifine dönüşüp F (Fenilalanin), L (Lösin), S (Serin)’den oluşmaktadır. 5' => 3' Frame 2’de TTTC-TTT motifini alıp F (Fenilalanin), L (Lösin), S (Serin)’den oluşmaktadır. Frame 3’te ise TTTCT-TTC motifinde olup S (Serin) ve F (Fenilalanin), L (Lösin) aminoasitlerini kodlamaktadır.

5'3' Frame 1
ttt ctt tct ttc ttt c
F L S F F
5'3' Frame 2
tttc ttt ctt tct ttc
F F L S F
5'3' Frame 3
tttct ttc ttt ctt tc
S F F L

Şekil 4.87. Upland pamuğuna ait ID numarası 107896451 olan *CsIG2* genine ait [TTTC]₄ motifinin kodladığı aminoasitler

Şekil 4.88’de D11 kromozomunda bulunan, ID numarası 107935224 olan *CsIE1* genine ait Upland pamuğu 4 kez tekrarlanan [ACAT] motifi 5' => 3' Frame 1’de ACA-TAC motifine dönüşüp T (Treonin), Y (Tirozin), I (İzolösin) ve H (Histidin)’den oluşmaktadır. 5' => 3' Frame 2’de ACAT-ACA motifini alıp H (Histidin), T (Treonin), Y (Tirozin), I (İzolösin)’den oluşmaktadır. Frame 3’te ise ACATA-CAT motifinde olup I (İzolösin), H (Histidin), T (Treonin) ve Y (Tirozin) aminoasitlerinden oluşmaktadır.

5'3' Frame 1
aca tac ata cat aca t
T Y I H T
5'3' Frame 2
acat aca tac ata cat
H T Y I H
5'3' Frame 3
acata cat aca tac at
I H T Y

Şekil 4.88. Upland pamuğuna ait ID numarası 107935224 olan *CsIE1* genine ait [ACAT]₄ motifinin kodladığı aminoasitler

Şekil 4.89’da D11 kromozomunda bulunan, ID numarası 107945371 olan *Prob-CesA8* genine ait Upland pamuğu 4 kez tekrarlanan [ATTTT] motifi 5' => 3' Frame 1’de ATT-TTA motifine dönüşüp I (İzolösin), L (Lösin), F (Fenilalanin) ve Y (Tirozin)’den oluşmaktadır. 5' => 3' Frame 2’de ATTT-TAT motifini alıp F (Fenilalanin), Y (Tirozin), I (İzolösin) ve L (Lösin)’den oluşmaktadır. Frame 3’te ise ATTTT-ATT motifinde olup F (Fenilalanin), I (İzolösin), L (Lösin), Y (Tirozin) aminoasitlerinden oluşmaktadır.

5'3' Frame 1
att tta ttt tat ttt att tt
I L F Y F I
5'3' Frame 2
atTTt tat ttt att tta ttt t
F Y F I L F
5'3' Frame 3
atTTTt att tta ttt tat ttt
F I L F Y F

Şekil 4.89. Upland pamuğuna ait ID numarası 107945371 olan *Prob-CesA8* genine ait [ATTTT]₄ motifinin kodladığı aminoasitler

Şekil 4.90’da D03 kromozomunda bulunan, ID numarası 107950592 olan *CesA2* genine ait Upland pamuğu 4 kez tekrarlanan [AGAAA] motifi 5' => 3' Frame 1’de AGA-AAA motifine dönüşüp R (Arjinin), K (Lisin), E (Glutamik asit)’den oluşmaktadır. 5' => 3' Frame 2’de AGAA-AAG motifini alıp E (Glutamik asit), K (Lisin), R (Arjinin)’den oluşmaktadır. Frame 3’te ise AGAAA-AGA motifinde olup K (Lisin), R (Arjinin), E (Glutamik asit) aminoasitlerinden oluşmaktadır.

5'3' Frame 1
aga aaa gaa aag aaa aga aa
R K E K K R
5'3' Frame 2
agaa aag aaa aga aaa gaa a
E K K R K E
5'3' Frame 3
agaaa aga aaa gaa aag aaa
K R K E K K

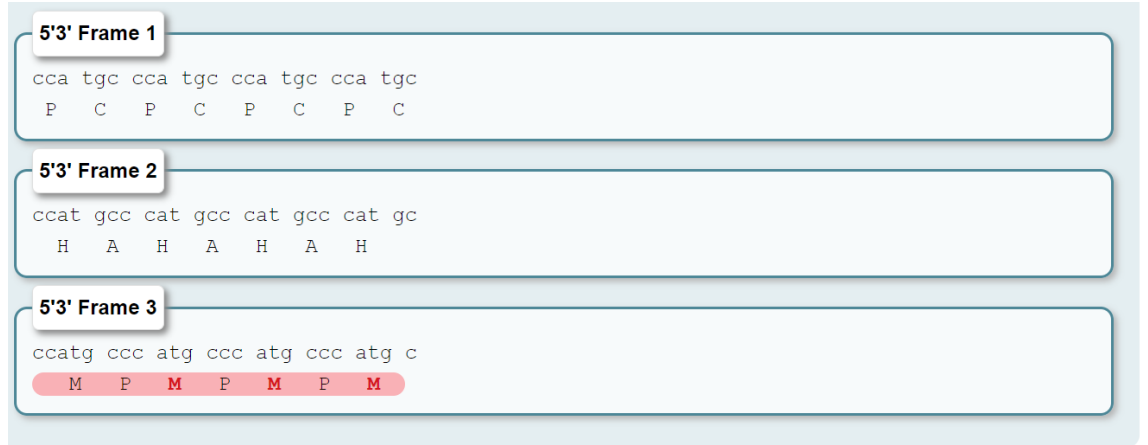
Şekil 4.90. Upland pamuğuna ait ID numarası 107950592 olan *CesA2* genine ait [AGAAA]₄ motifinin kodladığı aminoasitler

Şekil 4.91’de D03 kromozomunda bulunan, ID numarası 107950592 olan *CesA2* genine ait Upland pamuğu 5 kez tekrarlanan [AAAAG] motifi 5' => 3' Frame 1’de AAA-AGA motifine dönüşüp K (Lisin), R (Arjinin), E (Glutamik asit)’den oluşmaktadır. 5' => 3' Frame 2’de AAAA-GAA motifini alıp K (Lisin), E (Glutamik asit), R (Arjinin) oluşmaktadır. Frame 3’te ise AGAAA-AGA motifinde olup K (Lisin), R (Arjinin), E (Glutamik asit) aminoasitlerinden oluşmaktadır. Her üç olasılıkta da lisin ve glutamik asit kodlandığı bu nedenle de ilgili SSR motifinin lisin ve glutamik asit sayısını artırdığı görülmektedir.

5'3' Frame 1
aaa aga aaa gaa aag aaa aga aaa g
K R K E K K R K
5'3' Frame 2
aaaa gaa aag aaa aga aaa gaa aag
K E K K R K E K
5'3' Frame 3
aaaag aaa aga aaa gaa aag aaa ag
K K R K E K K

Şekil 4.91. Upland pamuğuna ait ID numarası 107950592 olan *CesA2* genine ait [AAAAG]₅ motifinin kodladığı aminoasitler

Şekil 4.92’de D11 kromozomunda bulunan, ID numarası 107912646 olan *CsIH1* genine ait mikrosatellitin 4 kez tekrarlanan [CCATGC] motifinin translasyon sonucu oluşturulabilen amino asit dizileri üç okunabilir çerçeve (frame) olarak gösterilmektedir. Bunlar 5' => 3' Frame 1’de CCA-TGC motifine dönüşüp, P (Prolin), C (Sistein)’den oluşmaktadır. 5' => 3' Frame 2’de CCAT-GCC motifini alıp H (Histidin), A (Alanin) oluşmaktadır. Frame 3’te ise CCATG-CCC motifinde olup M (Metionin) ve P (Prolin) aminoasitlerinden oluşmaktadır. Mikrosatellit *CsIH1* geninde Metionin-Prolin, Histidin-Alanin veya Prolin-Sistein amino asitlerin ikili tekrarlarının kodlanmasına neden olmaktadır. Prolin amino asitinin protein zincirinde arttırılması protein katlanmasına katkı sunabilecek iken; Histidin amino asitinin arttırılması, proteinin pozitif yükünün artmasına neden olabilecektir. Ayrıca histidin aminoasiti de histon proteinleri için önemlidir.



Şekil 4.92. Upland pamuğuna ait ID numarası 107912646 olan *CslH1* genine ait [CCATGC]₄ motifinin kodladığı aminoasitler

5. SONUÇLAR

Bu tez çalışması kapsamında 3 pamuk türüne ait bütün genom referans gen dizileri kullanılmıştır. Pamuğun ekonomik öneminin ötesinde allotetraploit özellikte bir poliploit olması, genomik ve evrimsel çalışmalar için model bir bitki olarak öne çıkmasına da neden olmaktadır. Ayrıca allotetraploit oluşumunda atasal genom sunucuları olan iki farklı diploit genomun da bütün genom dizilerinin tamamlanmış olması ve dizi bilgilerinin araştırmacılar için kullanılabilir niteliği olması da pamuğun bu tez çalışmasında kullanılmasının nedenleri arasında bulunmaktadır. Bu özelliklerden dolayı pamuk bitkisinde mikrosatellitlerin motif ve içerikleri yönünden genom içi ve genomlar arası dağılımını, gen elementlerindeki durumu, genlerdeki durumunun ise selüloz sentaz ve selüloz sentaz benzeri genler çalışılarak ortaya konulması için ideal kaynak olmasından da kaynaklanmıştır.

Veriler ışığında çalışmada kullanılan üç pamuk türüne ait diploit genomları, tetraploit genomun alt genomları ve türlerin kromozomları mikrosatellit sayıları ve içeriklerinin ile mikrosatellit tekrar sayılarının da önemli farklar olduğu belirlenmiştir. Tri-, tetra-, penta- ve hekza-nükleotitlerin tetraploit pamuk türü olan Upland (*G. hirsutum*) da daha yüksek, di-nükleotit tekrarları mikrosatellitlerin diploit pamuk türleri olan Asya pamuğu *G. arboreum* ve Peru pamuğu *G. raimondii* de daha yüksek olduğu belirlenmiştir. Di-nükleotit türler kendi aralarında karşılaştırıldığında ise Asya pamuğu *G. arboreum*'un mikrosatellitleri Peru pamuğu *G. raimondii*'den daha fazla olduğu sonucuna varılmıştır.

Üç türün her kromozomu mikrosatellit sayıları yönünden karşılaştırıldığında her bir türün değişiklik gösterdiği tespit edilmiş olup en yüksek değerler Oryantal pamukta bulunmuştur. Diğer taraftan *G. arboreum* kromozomları ve *G. hirsutum*'un A alt genomu ile *G. raimondii* kromozomları, *G. hirsutum*'un D alt genomuna kıyasla daha yüksek sayıda mikrosatellit içerdiği sonucuna ulaşılmıştır. Sonuçlar her genom içindeki kromozomların çeşitli mikrosatellit sayıları içerdiğini ve mikrosatellit içeriklerini pamuk genomlarına özgü özellikler olduğunu ortaya koymuştur.

Üç pamuk türünün SSR yoğunluklarının farklı olduğu ortaya konmuştur. *G. hirsutum* da genik-SSR ve intergenik SSR daha fazla iken *G. raimondii* de genik ve intergenik SSR'ın daha az olduğu sonucuna varılmıştır. Bu çalışma mikrosatellitlerin, gen elementlerinin önemli bir bileşeni olduğunu doğrulamıştır. Analiz edilen gen elementleri içinde mikrosatellit bolluğunun farklılıkları mevcuttur. Oryantal pamuk (A genomu) ve Upland pamukta (AD genomu) tri-nükleotit tekrarları kodlama bölgelerinde daha yaygın olurken; Peru pamukta di-nükleotit tekrarlarının daha fazla olduğu sonucuna ulaşılmıştır.

Üç pamuk türünün genik ve genler arası SSR sıklığı karşılaştırıldığında genik-SSR yönünden en fazla sıklığa sahip olan Upland pamuğu (1,22 SSR/Mb) iken bunu Peru (1,18 SSR/Mb) ve Oryantal pamuğu (1,18 SSR/Mb) takip etmiştir. Genler arası-SSR yönünden ise en yüksek sıklığa sahip olan Oryantal pamuğu iken (21,00 SSR/Mb) bunu Peru pamuğu (14,71 SSR/Mb) ve Upland pamuğu (11,61 SSR/Mb) takip etmiştir. Her bir pamuk genomunda tanımlanan her SSR tipi için tekrar motifleri ayrı ayrı incelenmiş olup Peru pamuğu ve Upland pamuğunun her bir motifi aynı sonuç vermesine rağmen Ağaç pamuğunu diğerlerinden ayıran farklılığın ise tri-nükleotitler olduğu sonucuna varılmıştır. Ayrıca her bir pamuk kromozomunda farklı SSR tiplerinin sıklığı ve dağılımı incelendiğinde SSR lokuslarının sıklığı kromozom boyutuyla ilişkili olmadığı belirlenmiştir. Örneğin Oryantal pamukta A05 kromozomunun yoğunluğu en yüksek olmasına rağmen ağaç pamuğun en küçük kromozomlarından biridir. En fazla SSR sayısı Oryantal pamukta A11 kromozomunda yer alırken Peru pamukta D05 kromozomunda, Upland pamukta ise A05 kromozomunda olduğu sonucuna ulaşılmıştır.

Bu çalışma mikrosatellit DNA'nın gen elementlerinin önemli bir bileşeni olduğunu doğrulamıştır. Oryantal pamuk (A genomu) ve Upland pamukta (AD genomu) tri-nükleotit tekrarları kodlama bölgelerinde daha yaygın olurken; Peru pamukta di-nükleotit tekrarlarının daha fazla olduğu sonucuna ulaşılmıştır. Mikrosatellit dizilerin üç farklı okuma çerçevesinde kodladıkları amino asitler belirlenmiş ve SSR dizilerinden tri-nükleotit ve penta-nükleotitlerin mono-peptit veya di-peptit tekrarlı amino asitleri kodladığı, di-nükleotit tekrarların ise di-peptit amino asitleri kodladığı belirlenirken, tetra- ve penta-nükleotit tekrarların tetra- ve penta-peptit tekrarlı amino asitleri sentezleyebildikleri belirlenmiştir.

Canlılar aleminin en önemli polisakkariti olan selülozu sentezleyen selüloz ve selüloz sentaz benzeri genler mikrosatellit içerenler ve içermeyenler olarak gruplandırıldı ve mikrosatellit içeren selüloz ve selüloz sentaz benzeri genleri içerdikleri mikrosatellitlerin fonksiyonları ve filogenetik pozisyonları yönünden etkili oldukları sonucuna varılmıştır. Mikrosatellitlerin verici diploit genomlardaki sayıları ile tetraploit genomdaki sayıları karşılaştırıldığında evrim sürecinde bir kısım mikrosatellitlerin kayıp olduğu ancak bazı yeni mikrosatellitlerin oluştuğu tespit edilmiştir.

NGS referans genom bilgileri kullanılarak herhangi bir organizmanın tüm mikrosatellit türlerini tanımlamak ve genik ile intergenik mikrosatellitleri karşılaştırmak mümkündür. Bu sayede gen, gen bölgelerine özgü fonksiyonel mikrosatellit (SSR) markırlarının elde edilebileceği bu tez çalışmasıyla teyit edilebilmiştir.

6. KAYNAKLAR

- Abdullaev, A.A., Salakhutdinov, I.B., Egamberdiev, S.S., Khurshut, E.E., Rizaeva, S.M., Ulloa, M. and Abdurakhmonov, I.Y. 2017. Genetic diversity, linkage disequilibrium, and association mapping analyses of *Gossypium barbadense* L. germplasm. *PLoS One*, 12, e0188125.
- Adams, K.L. and Wendel, J.F. 2005. Polyploidy and genome evolution in plants. *Curr Opin Plant Biol*, 8(2):135–41.
- Aho, A.V., Kernighan, B.W. and Weinberger P.J. 1988. The AWK Programming Language. New York. *Scientific Research an Academic Publisher*.
- Akkaya, M.S., Bhagwat, A.A. and Cregan, P.B. 1992. Length polymorphisms of simple sequence repeat DNA in soybean. *Genetics*, 132:1131–1139.
- Anonim 1: <https://learn.genetics.utah.edu/cotton> [Son erişim tarihi: 16.10.2023].
- Appel, R.D., Bairoch, A. and Hochstrasser, D.F. 1994. A new generation of information retrieval tools for biologists: the example of the ExPASy WWW server. *Trends Biochem Science*, 19, 258-260.
- Appenzeller, L., Doblin, M., Barreiro, R., Wang, H., Niu, X., Kollipara, K., Carrigan, L., Tomes, D., Chapman, M. and Dhugga, K. 2004. Cellulose synthesis in maize isolation and expression analysis of the cellulose synthase (*CesA*) gene family. *Cellulose*, 11(3–4):287–99.
- Arioli, T., Peng, L., Betzner, A.S., Burn, J., Wittke, W., Herth, W., Camilleri, C., Höfte, H., Plazinski, J., Birch, R., Cork, A., Glover, J., Redmond, J. and Williamson, R.E. 1998. Molecular analysis of cellulose biosynthesis in Arabidopsis. *Science*, 279:717–20.
- Artimo, P., Jonnalagedda, M., Arnold, K., Baratin, D., Csardi, G., de Castro, E., Duvaud, S., Flegel, V., Fortier, A., Gasteiger, E. et al. 2012. ExPASy: SIB bioinformatics resource portal. *Nucleic Acids Research*, 40 W597–W603. 10.1093/nar/gks400.
- Ashley, M.V. 2010. Plant Parentage, Pollination, and Dispersal: How DNA Microsatellites have Altered the Landscape. *Critical Reviews in Plant Sciences*, 29, 148-161.
- Azpilicueta, M.M., Gallo, L.A., Van Zonneveld, M., Thomas, E., Moreno, C. and Marchelli, P. 2013. Management of Nothofagus Genetic Resources: Definition of Genetic Zones Based on a Combination of Nuclear and Chloroplast Marker Data. *Forest Ecology and Management*, 302, 414-424.
- Bagshaw, A.T. 2017. Functional mechanisms of microsatellite DNA in eukaryotic genomes. *Genome Biol Evol*, 9:2428–2443.
- Bairoch, S. 1999. *ExPASy: The Expert Protein Analysis System*. Online. <https://www.expasy.org> [Son erişim tarihi: 08.09.2024].
- Bao, E., Jiang, T. and Girke, T. 2014. AlignGraph: algorithm for secondary de novo genome assembly guided by closely related references. *Bioinformatics*, 30(12):319–28.

- Bardak, A. and Bölek, Y. 2012. Genetic Diversity of Diploid and Tetraploid Cottons Determined by Ssr and Issr Markers. *Turkish Journal Of Field Crops*, 17(2), 139-144.
- Benson, G. 1999. Tandem Repeats Finder: a program to analyze DNA sequences. *Nucleic Acids Research*, 27: 573-580.
- Bertini, C.H., Schuster, I., Sediya, T., Barros, E.G.D. and Moreira, M.A. 2006. Characterization and genetic diversity analysis of cotton cultivars using microsatellites. *Genet Mol Biol*, 29: 321-329.
- Bhandari, S., Tyagi, K., Singh, B. and Goutam, U. 2021. Role of molecular markers to study genetic diversity in Bamboo: A Review. *Plant cell biotechnology and molecular biology*, 86-97.
- Bilgen, M., Karaca, M., Onus, A. N. and Ince, A. G. 2004. A software program combining sequence motif searches with keywords for finding repeats containing DNA sequences. *Bioinformatics*, 20: 3379-3386.
- Bradnam, K.R., Fass, J.N., Alexandrov, A., Baranay, P., Bechner, M., Birol, I., Boisvert, S., Chapman, J.A., Chapuis, G., Chikhi, R. et al. 2013. Assemblathon 2: evaluating de novo methods of genome assembly in three vertebrate species. *Gigascience*, 2013;2:1-31.
- Burton, R.A., Shirley, N.J., King, B.J., Harvey, A.J. and Fincher, G.B. 2004. The CESA gene family of barley. Quantitative analysis of transcripts reveals two groups of co-expressed genes. *Plant Physiol*, 134:224-36.
- Cai, C., Zhu, G., Zhang, T. and Guo, W. 2017. High-density 80 K SNP array is a powerful tool for genotyping *G. hirsutum* accessions and genome analysis. *BMC Genomics*, 18, 654.
- Cantu, D., Vicente, A.R., Labavitch, J.M., Bennett, A.B. and Powell, A.L. 2008. Strangers in the matrix: plant cell walls and pathogen susceptibility. *Trends Plant Science*, 13, 610-617.
- Cao, S., Cheng, H., Zhang, J., Aslam, M., Yan, M., Hu, A., Lin, L., Ojolo, S.P., Zhao, H., Priyadarshani, S., Yu, Y., Cao, G. and Qin, Y. 2019. Genome-wide identification, expression pattern analysis and evolution of the *Ces/Csl* Gene superfamily in Pineapple (*Ananas comosus*). *Plants (Basel)* 8:275.
- Card, D.C., Schield, D.R., Reyes-Velasco, J., Fujita, M.K., Andrew, A.L., Oyler-McCance, S.J., Fike, J.A., Tomback, D.F., Ruggiero, R.P. and Castoe, T.A. 2014. Two low coverage bird genomes and a comparison of reference-guided versus de novo genome assemblies. *PLoS One*, 9(9):e106649.
- Chakraborty, M., Baldwin-Brown, J.G., Long, A.D. and Emerson, J.J. 2016. Contiguous and accurate de novo assembly of metazoan genomes with modest long read coverage. *Nucleic Acids Research*, 44(19):e147.
- Chen, G. and Du, X.M. 2006. Genetic diversity of source germplasm of Upland cotton in China as determined by SSR marker analysis. *Acta Genetica Sinica*. 33(8):733-45.

- Chen, M., Tan, Z.Y., Jiang, J.H., Li, M.F., Chen, H.J., Shen, G.L. and Yu, R.Q. 2009. Similar distribution of simple sequence repeats in diverse completed Human Immunodeficiency Virus Type 1 genomes. *FEBS Lett*, 583:2959–2963.
- Chen, Z., Feng, K., Grover, C.E., Li, P., Liu, F., Wang, Y., Xu, Q., Shang, M., Zhou, Z., Cai, X. et al. 2016. Chloroplast DNA structural variation, phylogeny, and age of divergence among diploid cotton species. *PLoS One*, 11:e0157183.
- Chen, Z., Grover, C.E., Li, P., Wang, Y., Nie, H., Zhao, Y., Wang, M., Liu, F., Zhou, Z., Wang, X. et al. 2017a. Molecular evolution of the plastid genome during diversification of the cotton genus. *Mol Phylogenet Evol*, 112, 268–276.
- Chen, Z., Nie, H., Grover, C.E., Wang, Y., Li, P., Wang, M., Pei, H., Zhao, Y., Li, S., Wendel, J.F. and Hua, J. 2017b. Entire nucleotide sequences of *Gossypium raimondii* and *G. arboreum* mitochondrial genomes revealed A-genome species as cytoplasmic donor of the allotetraploid species. *Plant Biol*, 19, 484–493.
- Chen, Z.J., Sreedasyam, A., Ando, A., Song, Q., De-Santiago, L.M., Hulse-Kemp, A.M., Ding, M., Ye, W., Kirkbride, R.C., Jenkins, J. et al. 2020. Genomic diversifications of five *Gossypium* allopolyploid species and their impact on cotton improvement. *Nature Genetics*, 52, 525–533.
- Cronn, R.C., Small, R.L., Haselkorn, T. and Wendel, J.F. 2002. Rapid diversification of the cotton genus (*Gossypium*: Malvaceae) revealed by analysis of sixteen nuclear and chloroplast genes. *American Journal of Botany*, 89, 707–725.
- Cronn, R., Small, R.L., Haselkorn, T. and Wendel, J.F. 2003. Cryptic repeated genomic recombination during speciation in *Gossypium Gossypioides*. *Evolution* 57, 2475–2489.
- Delgrange, O. and Rivals, E. 2004. STAR: An Algorithm to Search for Tandem Approximate Repeats. *Bioinformatic*, 16, 2812–2820.
- Delhom, C.D., Wanjura, J.D. and Hequet, E.F. 2022. Cotton fibre elongation: a review. *The Journal of The Textile Institute*, 1, 1–12.
- Dong, C., Wang, J., Yu, Y., Ju, L., Zhou, X., Ma, X., Mei, G., Han, Z., Si, Z., Li, B., Chen, H. and Zhang, T. 2019. Identifying functional genes influencing *Gossypium hirsutum* fiber quality. *Front Plant Science*, 9, 1968.
- Du, X., Huang, G., He, S., Yang, Z., Sun, G., Ma, X., Li, N., Zhang, X., Sun, J., Liu, M. et al. 2018. Resequencing of 243 diploid cotton accessions based on an updated genome identifies the genetic basis of key agronomic traits. *Nature Genetics*, 50, 796–802.
- Dufresnes, C., Brelsford, A., Beziers, P. and Perrin, N. 2014. Stronger Transferability but Lower Variability in Transcriptomic- than in Anonymous Microsatellites: Evidence From Hylid Frogs. *Molecular Ecology Resources*, 14, 716–725.
- Ekblom, R. and Wolf, J.B.W. 2014 A field guide to whole-genome sequencing, assembly and annotation. *Evol Appl*, 7(9):1026–42.
- Erkılınç, A. and Karaca, M. 2005. Assessment of Genetic Variation in Some Cotton Varieties (*Gossypium hirsutum* L.) Grown in Turkey Using Microsatellite. *Akdeniz Üniversitesi Ziraat Fakültesi Dergisi*, 18, 201–206.

- Fang, L., Gong, H., Hu, Y., Liu, C., Zhou, B., Huang, T., Wang, Y., Chen, S., Fang, D.D., Du, X. et al. 2017a. Genomic insights into divergence and dual domestication of cultivated allotetraploid cottons. *Genome Biol*, 18, 33.
- Fang, L., Wang, Q., Hu, Y., Jia, Y., Chen, J., Liu, B., Zhang, Z., Guan, X., Chen, S., Zhou, B. et al. 2017. Genomic analyses in cotton identify signatures of selection and loci associated with fiber quality and yield traits. *Nature Genetics*, 49, 1089–1098.
- Fang, L., Zhao, T., Hu, Y., Si, Z., Zhu, X., Han, Z., Liu, G., Wang, S., Ju, L., Guo, M. et al. 2021. Divergent improvement of two cultivated allotetraploid cotton species. *Plant Biotechnol. Journal*, 19, 1325–1336.
- Fryxell, P. A. .1969. A classification of *Gossypium* L. (Malvaceae). *Taxon* 18, 585–591.
- Fryxell, P. A. 1978. The Natural History of the Cotton Tribe (Malvaceae tribe, *Gossypieae*). College Station, TX: 18-245.
- Gao, H., Cai, S., Yan, B., Chen, B. and Yu, F. 2009. Discrepancy variation of dinucleotide microsatellite repeats in eukaryotic genomes. *Biol Research*, 42: 365-375, 2009.
- Goodwin, S., McPherson, J.D. and McCombie, W.R. 2016. Coming of Age: Ten Years of Next-Generation Sequencing Technologies. *Nature Reviews Genetics*, 17, 333-351.
- Gordon, D., Huddleston, J., Chaisson, M.J.P., Hill, C.M., Kronenberg, Z.N., Munson, K.M., Malig, M., Raja, A., Fiddes, I., Hillier, L.W. et al. 2016. Long-read sequence assembly of the gorilla genome. *Science*, 352(6281).
- Govindaraj, M., Vetriventhan, M. and Srinivasan, M. 2015. Importance of genetic diversity assessment in crop plants and its recent advances: an overview of its analytical perspectives. *Genetics Research International*, doi: 10.1155/2015/431487.
- Grimson, M.J., Haigler, C.H. and Blanton, R.L. 1996. Cellulose microfibrils, cell motility, and plasma membrane protein organization change in parallel during culmination in *Dictyostelium discoideum*. *J Cell Science*, 109(Pt 13), 3079–3087.
- Grover, C.E., Kim, H.R., Wing, R.A., Paterson, A.H. and Wendel, J.F. 2007. Microcolinearity and genome evolution in the *AdhA* region of diploid and polyploid cotton (*Gossypium*). *Plant J*, 50, 995–1006.
- Grover, C.E., Arick, M.A., Thrash, A., Conover, J.L., Sanders, W.S., Peterson, D.G., Frelichowski, J.E., Scheffler, J.A., Scheffler, B.E. and Wendel, J.F. 2019. Insights into the evolution of New World diploid cottons (*Gossypium*, subgenus *houzingenia*) based on genome sequencing. *Genome Biol. Evolution*, 11, 53–71.
- Gu, Y., Kaplinsky, N., Bringmann, M., Cobb, A., Carroll, A., Sampathkumar, A., Baskin, T., Persson, S. and Somerville, C. 2010. Identification of a novel *CesA*-associated protein required for cellulose biosynthesis. *Proc Natl Acad Sci USA*, 107:12866–12871.

- Hatakeyama, M., Aluri, S., Balachadran, M.T., Sivarajan, S.R., Patrignani, A., Gr ter, S., Poveda, L., Shimizu-Inatsugi, R., Baeten, J., Francoijs, K. et al. 2018. Multiple hybrid de novo genome assembly of finger millet, an orphan allotetraploid crop. *DNA Res*, Feb 1;25(1):39-47.
- Hendrix, B. and Stewart, J.M. 2005. Estimation of the nuclear DNA content of *Gossypium* species. *Ann Bot (Lond)*, 95: 789-797.
- Hu, Y., Chen, J., Fang, L., Zhang, Z., Ma, W., Niu, Y., Ju, L., Deng, J., Zhao, T., Lian, J. et al. 2019. *Gossypium barbadense* and *Gossypium hirsutum* genomes provide insights into the origin and evolution of allotetraploid cotton. *Nature Genetics*, 51, 739–748.
- Huang, C., Nie, X., Shen, C., You, C., Li, W., Zhao, W., Zhang, X. and Lin, Z. 2017. Population structure and genetic basis of the agronomic traits of upland cotton in China revealed by a genome-wide association study using high-density SNPs. *Plant Biotechnol J*, 15, 1374–1386.
- Huang, G., Wu, Z., Percy, R.G., Bai, M., Li, Y., Frelichowski, J.E. Hu, J., Wang, K., Zu, J.Z. and Zhu, Y. 2020. Genome sequence of *Gossypium herbaceum* and genome updates of *Gossypium arboreum* and *Gossypium hirsutum* provide insights into cotton a-genome evolution. *Nature Genetics*, 52, 516–524.
- Huang, X. and Han, B. 2014. Natural variations and genome-wide association studies in crop plants. *Annu. Rev. Plant Bio*, 65, 531–551.
- Huang, X., Kurata, N., Wei, X., Wang, Z-X., Wang, A., Zhao, Q., Zhao, Y., Liu, K., Lu, H., Li, W. et al. 2012. A map of rice genome variation reveals the origin of cultivated rice. *Nature*, 490, 497–501.
- Huang, X., Wei, X., Sang, T., Zhao, Q., Feng, Q.I., Zhao, Y., Li, C., Zhu, C., Lu, T., Zhang, Z. et al. 2010. Genome-wide association studies of 14 agronomic traits in rice landraces. *Nature Genetics*, 42, 961–967.
- Ince, A.G. 2010. Doku/Organ Spesifik Mikrosatellit DNA Gen İ eriklerinin *Capsicum* cDNA K t phanelerinde *In silico* ve *In vitro* Yaklařımlarla Belirlenmesi. Doktora Tezi, Akdeniz  niversitesi, Antalya, 300 s.
- Ince, A.G. and Karaca, M. 2023. Genome-Wide Comparisons of Microsatellites in *Gossypium* L. Asia Pasific 8th International Modern Sciences Congress, September 11-12, Delhi, India.
- Ince, A.G., Karaca, M. and Onus, A.N. 2010. Differential Expression Patterns of Genes Containing Microsatellites in *Capsicum annuum* L. *Molecular Breeding*, 25, 645- 658.
- Ince, A.G., Karaca, M., Bilgen, M. and Onus, A.N. 2008. Digital Differential Display Tools for Mining Microsatellite Containing Organism, Organ and Tissue. *Plant Cell Tissue and Organ Culture*, 94, 281-290.
- Ince, A.G., Karaca, M., Uygur G  er, E. and Aydın, A. 2017. Descriptive Statistics and PIC Values of Genomic- and Trascriptomic-Microsatellites in Several Plant Species. *Jornal of Scientific and Engineering Research*, 4, 236-246.

- Islam, S., Thyssen, G.N., Jenkins, J.N., Zeng, L., Delhom, C.D., McCarty, J.C., Deng, D.D., Hinchliffe, D.J., Jones, D.C. and Fang, D.D. 2016. A MAGIC population-based genome-wide association study reveals functional association of GhRBB1_A07 gene with superior fiber quality in cotton. *BMC Genomics*, 17, 903.
- Jia, G., Huang, X., Zhi, H., Zhao, Y., Zhao, Q., Li, W., Chai, Y., Yang, L., Liu, K., Lu, H. et al. 2013. A haplotype map of genomic variations and genome-wide association studies of agronomic traits in foxtail millet (*Setaria italica*). *Nature Genetics*, 45, 957–961.
- Jiang, C. X., Wright, R. J., El-Zik, K. M. and Paterson, A. H. 1998. Polyploid formation created unique avenues for response to selection in *Gossypium* (cotton). *Proc. Natl. Acad. Sci. U. S. A*, 95, 4419–4424.
- Joshi, S.P., Ranjekar P.K. and Gupta V.S. 1999. Molecular markers in plant genome analysis. *Current Science*: 230–240.
- Jurka, J. and Pethiyagoda, C. 1995. Simple repetitive DNA sequences from primates: compilation and analysis. *J Mol Evol*, 40:120–126.
- Kantety, R. V., Rota, M. L., Matthews, D. E. and Sorrells, M. E. 2002. Data mining for simple sequence repeats in expressed sequence tags from barley, maize, rice, sorghum and wheat. *Plant Molecular Biology*, 48: 501–510.
- Karaca, M. 2001. Characterization of *Cynodon* spp. and *Gossypium* spp. genomes using molecular and cytological techniques. Doktora Tezi, Mississippi State University, Mississippi State.
- Karaca, M., and Ince, A.G. 2023. Cellulose Synthases Specific Microsatellite Primers in *Gossypium hirsutum* L. 2. International Paris Congress on Agriculture & Animal Husbandry October 24-26.
- Karaca, M., and Ince, A.G. 2024. Grafting based DNA methylation alteration of snoRNAs in upland cotton (*Gossypium* L.) *Physiology and Molecular Biology of Plants*, 30, 893-907.
- Karaca, M., Bilgen, M., Onus, A.N., Ince, A.G. and Elmasulu, S.Y. 2005. Exact tandem repeats analyzer (E-TRA) for DNA sequence mining. *Journal of Genetics*, 84: 49-54.
- Karaca, M., Ince, A.G. and Aydın, A. 2019. Identification of a Putative Epiallele Expressed in 25 DPA Cotton Fiber. 8-10 April, International Congress on Agriculture and Forestry Research, ss.493-500. Marmaris, Turkey.
- Karaca, M., Ince, A.G. and Reddy, U.K. 2020. Interspecific grafting between *Gossypium hirsutum*, *G. barbadense* and *G. herbaceum* lines. *Scientific Reports*, 10, 18649.
- Karaca, M., Ince, A.G., Aydın, A., Elmasulu, S.Y. and Turgut, K. 2015. Microsatellites for Genetic and Taxonomic Research on Thyme (*Thymus* L.). *Turkish Journal of Biology*, doi:10.3906/biy-1406-20.
- Keegstra, K. 2010. Plant cell walls. *Plant Physiol*, 154, 483–486.

- Keskinkılıç, K. 2017. Türkiye Pamuk Durumundaki Gelişmeler. <https://www.researchgate.net/publication/320471439>. 45s. [Son erişim tarihi: 08.09.2024].
- Kofler, R.S., Schlötterer, C. and Lelley, T. 2007. SciRoKo: A New Tool for Whole Genome Microsatellite Search and Investigation. *Bioinformatics*, (13), 1683-1685.
- Koren, S. and Phillippy, A.M. 2015. One chromosome, one contig: complete microbial genomes from long-read sequencing and assembly. *Curr Opin Microbiol*, 23:110–20.
- Kundu, S. and Ghosh, S.K. 2015. Trend of different molecular markers in the last decades for studying human migrations. *Gene*, 556:81–90.
- Kurtz, S., Jomuna, V.C., Ohlebusch, E., Schleiermacher, C., Stoye, J. and Giegerich, R. 2001. REPuter: the manifold applications of repeat analysis on a genomic scale. *Nucleic Acids Research*, 29: 4633-4642.
- Langercrantz, U., H, Ellegren. and L, Andersson. 1993. The abundance of various polymorphic microsatellite motifs differs between plants and vertebrates. *Nucleic Acids Research*, 21:1111–1115.
- Lee, J.J., Woodward, A.W. and Chen, Z.J. 2007. Gene expression changes and early events in cotton fibre development. *Ann Bot*, 100, 1391–1401.
- Li, C., Fu, Y., Sun, R., Wang, Y. and Wang, Q. 2018. Single-locus and multi-locus genome-wide association studies in the genetic dissection of fiber quality traits in Upland cotton (*Gossypium hirsutum* L.). *Front. Plant Science*, 9, 1083.
- Li, F., Fan, G., Lu, C., Xiao, G., Zou, C., Kohel, R.J., Ma, Z. Shang, H., Ma, X., Wu, J. et al. 2015. Genome sequence of cultivated upland cotton (*Gossypium hirsutum* TM-1) provides insights into genome evolution. *Nature Biotechnol*, 33, 524.
- Li, F., Fan, G., Wang, K., Sun, F., Yuan, Y., Song, G., Li, Q. Ma, Z., Lu, C., Zou, C. et al. 2014. Genome sequence of the cultivated cotton *Gossypium arboreum*. *Nature Genetics*, 46, 567–572.
- Li, H. 2015. *SeqTK: A toolkit for processing sequences in FASTA/Q formats*. Online. <https://github.com/lh3/seqtk> [Son erişim tarihi: 08.09.2024].
- Li, H., Peng, Z., Yang, X., Wang, W., Fu, J., Wang, J., Han, Y., Chai, Y., Guo, T., Yang, N. et al. 2013. Genome-wide association study dissects the genetic architecture of oil biosynthesis in maize kernels. *Nature Genetics*, 45, 43–50.
- Li, Y.C., Korol, A.B., Fahima, T., Beiles, A. and Nevo, E. 2002. Microsatellites: genomic distribution, putative functions and mutational mechanisms: a review. *Mol Ecol*, 11:2453–2465.
- Lin, Z., Wang, Y., Zhang, X. and Zhang, J. 2012. Functional markers for cellulose synthase and their comparison to SSRs in cotton. *Plant Mol Biol Reporter*, 30:1270–1275.
- Litt, M. and Luty, J.A. 1989. A Hypervariable Microsatellite Revealed by *In Vitro* Amplification of a Dinucleotide Repeat within the Cardiac Muscle Actin Gene. *The American Journal of Human Genetics*, 44, 397-401.

- Liu, R., Gong, J., Xiao, X., Zhang, Z., Li, J., Liu, A., Lu, Q., Shang, H., Shi, Y., Ge, Q. et al. 2018. GWAS analysis and QTL identification of fiber quality traits and yield components in Upland cotton using enriched high-density SNP markers. *Front Plant Science*, 9, 1067.
- Liu, X., Zhao, B., Zheng, H.J., Hu, Y., Lu, G., Yang, C.Q., Chen, J.D., Chen, J.J., Chen, D.Y., Zhang, L. 2015. *Gossypium barbadense* genome sequence provides insight into the evolution of extra-long fibers and specialized metabolites. *Science Rep*, 5:14139.
- Lu, C., Zou, C., Zhang, Y., Yu, D., Cheng, H., Jiang, P., Yang, W., Wang, Q., Feng, X., Prosper, M.A. et al. 2015. Development of chromosome-specific markers with high polymorphism for allotetraploid cotton based on genome-wide characterization of simple sequence repeats in diploid cottons (*Gossypium arboreum* L. and *Gossypium raimondii* Ulbrich). *BMC Genomics*, 16-55.
- Ma, J., Geng, Y., Pei, W., Wu, M., Li, X., Liu, G., Li, D., Ma, Q., Zang, X., Yu, S. et al. 2018a. Genetic variation of dynamic fiber elongation and developmental quantitative trait locus mapping of fiber length in upland cotton (*Gossypium hirsutum* L.). *BMC Genomics*, 19, 882.
- Ma, Z., He, S., Wang, X., Sun, J., Zhang, Y., Zhang, G., Wu, L., Li, Z., Liu, Z., Sun, G. et al. 2018b. Resequencing a core collection of upland cotton identifies genomic variation and loci influencing fiber quality and yield. *Nature Genetics*, 50, 803–813.
- Mace, E.S., Tai, S., Gilding, E.K., Li, Y., Prentis, P.J., Bian, L., Campbell, B.C., Hu, W., Innes, D.J., Han, X. et al. 2013. Whole-genome sequencing reveals untapped genetic potential in Africa's indigenous cereal crop sorghum. *Nature Communications*, 4, 2320.
- Mason, A.S. 2015. SSR genotyping. *Methods Mol Biol*, 1245:77-89.
- Matthysse, A.G. and Thomas, D.I. 1995. Mechanism of Cellulose Synthesis in *Agrobacterium tumefaciens*. *J. Bacteriol*, 177, 1076–1081.
- McFarlane, H., Döring, A., Persson, S. 2014. The cell biology of cellulose synthesis. *Annu Rev Plant Biol*, 65(1):69–94.
- Mert, M. 2007. Pamuk Tarımının Temelleri. TMMOB Ziraat Mühendisleri Odası Teknik Yayınlar Dizisi: 7, Hatay, 282 s.
- Mudunuri, S.B., Nagarajaram, H.A. and Bateman, A. 2007. IMEx: Imperfect Microsatellite Extractor. *Bioinformatics*, (10), 1181-1187.
- Nadeem, M.A., Nawaz, M.A., Shahid, M.Q., Doğan, Y., Comertpay, G., Yıldız, M., Hatipoğlu, R., Ahmad, F., Alsaleh, A., Labhane, N. 2018. DNA molecular markers in plant breeding: current status and recent advancements in genomic selection and genome editing. *Biotechnol Biotechnol Equip*, 32:261–285.
- Ning, Z., Chen, H., Mei, H. and Zhang, T. 2014. Molecular tagging of QTLs for fiber quality and yield in the upland cotton cultivar Acala-Prema. *Euphytica*. 195:143–156.

- Paape, T., Hatakeyama, M., Shimizu-Inatsugi, R., Cereghetti, T., Onda, Y., Kenta, T., Sese, J. and Shimizu, K.K. 2016. Conserved but attenuated parental gene expression in allopolyploids: constitutive zinc Hyperaccumulation in the Allotetraploid *Arabidopsis kamchatica*. *Mol Biol Evol*, 33(11):2781–800.
- Parisi, V., Fonzo, V. D. and Aluf-Pentini, F. 2003. STRING: finding tandem repeats in DNA sequences. *Bioinformatics*, 19: 1733-1738.
- Paterson, A. H., Wendel, J.F., Gundlach, H., Guo, H., Jenkins, J., Jin, D., Llewellyn, D., Showmaker, K.C., Shu, S., Udall, J. et al. 2012. Repeated polyploidization of *Gossypium* genomes and the evolution of spinnable cotton fibres. *Nature*, 492, 423–427.
- Pear, J.R., Kawagoe, Y., Schreckengost, W.E., Delmer, D.P., Stalker, D.M. 1996. Higher plants contain homologs of the bacterial *CslA* genes encoding the catalytic subunit of cellulose synthase. *Proc Natl Acad Science*, 93(22):12637–42.
- Pedersen, G.B., Blaschek, L., Frandsen, K.E.H., Noack, L.C. and Persson, S. 2023. Cellulose synthesis in land plants. *Molecular Plant*, 16, 206-231.
- Pop, M., Phillippy, A., Delcher, A.L. and Salzberg, S.L. 2004. Comparative genome assembly. *Brief Bioinform*, 5(3):237–48.
- Powell, W., Machray, G.C. and Provan, J. 1996. Polymorphism revealed by simple sequence repeats. *Trends Plant Science*, 1(7):215–222.
- Pryszcz, L.P. and Gabaldón, T. 2016. Redundans: an assembly pipeline for highly heterozygous genomes. *Nucleic Acids Research*, 44(12):e113.
- Quinlan, A.R. and Hall, I.M. 2010. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics*, Volume 26, Issue 6, March. Pages 841–842.
- Rehman, A. and Iqbal, M. 2022. Genome –wide Allelic Variations and Population Structure Analysis of Cotton Genotypes using SSR Markers. *Pakistan Journal Of Biochemistry and biotechnology*, Vol.3 No.1.
- 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. et al. 2012. GAGE: a critical evaluation of genome assemblies and assembly algorithms. *Genome Research*, 22(3):557–67.
- Saxena, I.M. and Brown, R.M.Jr. 2005. Cellulose biosynthesis: current views and evolving concepts. *Ann Bot*, 96, 9–21.
- Schneeberger, K., Ossowski, S., Ott, F., Klein, J.D., Wang, X., Lanz, C., Smith, L.M., Cao, J., Fitz, J., Warthmann, N. et al. 2011. Reference-guided assembly of four diverse *Arabidopsis thaliana* genomes. *Proc Natl Acad Sci USA*, 108(25):10249–54.
- Shao, Q., Zhang, F., Tang, S., Liu, Y., Fang, X., Liu, D., Liu, D., Zhang, J., Teng, Z., Paterson, A.H and Zhang, Z. 2014. Identifying QTL for fiber quality traits with three upland cotton (*Gossypium hirsutum* L.) populations. *Euphytica*, 198:43–58.

- Shi, J., Zhang, Y., Wang, N., Xu, Q., Huo, F., Liu, X. and Yan, G. 2023. Genetic diversity and population structure analysis of upland cotton (*Gossypium hirsutum* L.) germplasm in China based on SSR markers. *Genet Resour Crop Evol*, <https://doi.org/10.1007/s10722-023-01594-9>.
- Shrestha, N., Zhang, K., Gowda, S. A., Abdelraheem, A., Jones, D. J. and Kuraparthi, V. 2023. Identification of quantitative trait loci for fiber quality, yield, and plant height traits in Upland cotton. *Crop Science*. 63, 1421-1441.
- Simon, R. B., Justin, T. P., Joshua, A. U., William, S. S., Daniel, G. P., Mark, A. A. Grover, C.E. and Wendel, J.F. Independent domestication of two old world cotton species. *Genome Biol. Evol*, 8 1940–1947.
- Smith, C. 2022. Guide to NGS Platforms. Life Science Articles. Online. <https://www.biocompare.com> [Son erişim tarihi: 21.11.2024].
- Soranzo, N., Provan, J. and Powell, W. 1999. An example of microsatellite length variation in the mitochondrial genome of conifers. *Genome*, 42, 158–161.
- Sreenu, V. B., Vishwanath, A., Nagaraju, J. and Nagarajaram, H. A. 2003. MICdb: database of prokaryotic microsatellites. *Nucleic Acids Research*, 31: 106-108.
- Su, J., Li, L., Pang, C., Wei, H., Wang, C., Song, M., Wang, H., Zhao, S., Zhang, C., Mao, G. et al. 2016. Two genomic regions associated with fiber quality traits in Chinese upland cotton under apparent breeding selection. *Science Reporter*, 6, 38496.
- Sun, Z., Wang, X., Liu, Z., Gu, Q., Zhang, Y., Li, Z., Ke, H., Yang, J., Wu, J., Wu, L. et al. 2017. Genome-wide association study discovered genetic variation and candidate genes of fibre quality traits in *Gossypium hirsutum* L. *Plant Biotechnol J*, 15, 982–996.
- Tamura, K., Stecher, G. and Kumar, S. 2021. MEGA11: Molecular Evolutionary Genetics Analysis Version 11. *Mol Biol Evol*, Jun 25; 38(7):3022-3027.
- Tang, D., Chen, M., Huang, X., Zhang, G., Zeng, L., Zhang, G., Wu, S. and Wang, Y. 2023. SRplot: A free online platform for data visualization and graphing. *PLoS One*, Nov 9;18(11):e0294236 PMID: 37943830.
- Tautz, D. 1989. Hypervariability of simple sequences as a general source for polymorphic DNA markers. *Nucleic Acids Research*, 17:6463–6471.
- Tautz, D. and Renz, M. 1984. "Simple sequences are ubiquitous among eukaryotes." *Nucleic Acids Research*, 12(10), 4127-4138.
- Thiel, T., Michalek, W., Varshney, R. K. and Graner, A. 2003. Exploiting EST databases for the development of cDNA derived microsatellite markers in barley (*Hordeum vulgare* L.). *Journal of Genetics*, 106: 411-422.
- Thompson, K. 1974. *grep: A program for searching text*. Unix operating system.
- TÜİK 2019. Tahıllar ve Diğer Bitkisel Ürünler (Pamuk Üretim Miktarı). <https://biruni.tuik.gov.tr/medas/?kn=92&locale=tr> [Son erişim tarihi: 27.08.2024].

- TÜİK 2021a. Tahıllar ve Diğer Bitkisel Ürünler (Pamuk Ekim Alanı). <https://biruni.tuik.gov.tr/medas/?kn=92&locale=tr> [Son erişim tarihi: 27.08.2024].
- TÜİK 2021b. Tahıllar ve Diğer Bitkisel Ürünler (Kütlü Pamuk Üretim Miktarı). <https://biruni.tuik.gov.tr/medas/?kn=92&locale=tr> [Son erişim tarihi: 27.08.2024].
- TÜİK 2021c. Tahıllar ve Diğer Bitkisel Ürünler (Pamuk Verim). <https://biruni.tuik.gov.tr/medas/?kn=92&locale=tr> [Son erişim tarihi: 27.08.2024].
- Tyagi, P., Gore, M.A., Bowman, D.T., Campbell, B.T., Udall, J.A. and Kuraparthi, V. 2014. Genetic diversity and population structure in the US Upland cotton (*Gossypium hirsutum* L.). *Theor Appl Genet*, 127, 283–295.
- Udall, J.A., Long, E., Hanson, C., Yuan, D., Ramaraj, T., Conover, J.L., Gong, L., Arick, M.A., Grover, C.E., Peterson, D.G. and Wendel, J.F. 2019. De novo genome sequence assemblies of *Gossypium raimondii* and *Gossypium turneri*. *Genes Genomes Genetics*, 9, 3079–3085.
- Ugurlu, M. 2020. Pamuk Üretiminin Ekonomik Analizi: Manisa Örneği. *Ziraat Fakültesi Dergisi*, 206-217.
- USDA 2020. United States Department of Agriculture Foreign Agricultural Service, December 2020 Report, Cotton: World Markets and Trade, 1-28.
- Varshney, R.K., Graner, A. and Sorrells, M.E. 2005. Genic microsatellite markers in plants: features and applications. *Trends Biotechnol*, 23(1):48–55.
- Vašek, J., Cílová, D., Melounová, M. Svoboda, P., Vejl, P., Štikarová, R., Vostrý, L., Kuchtová, P. and Ovesná, J. 2020. New EST-SSR markers for individual genotyping of opium poppy cultivars (*Papaver somniferum* L.) *Plants*, 9:10.
- Vezzi, F., Cattonaro, F. and Policriti, A. 2011. E-RGA: enhanced reference guided assembly of complex genomes. *EMBnet J*, 17(1):46–54.
- Wang, K., Wang, Z., Li, F., Ye, W., Wang, J., Song, G., Yue, Z., Gong, L., Shang, H., Zhu, S. et al. 2012. The draft genome of a diploid cotton *Gossypium raimondii*. *Nature Genetics*, 44, 1098–1103.
- Wang, M., Tu, L., Lin, M., Lin, Z., Wang, P., Yang, Q. Ye, Z., Shen, C., Li, J., Zhang, L. et al. 2017. Asymmetric subgenome selection and cis-regulatory divergence during cotton domestication. *Nature Genetics*, 49 579–587.
- Wang, M., Tu, L., Yuan, D., Zhu, D., Shen, C., Li, J., Liu, F., Pei, L., Wang, P., Zhao, G. et al. 2019. Reference genome sequences of two cultivated allotetraploid cottons, *Gossypium hirsutum* and *Gossypium barbadense* . *Nature Genetics*, 51, 224–229.
- Wang, M.L., Barkley, N.A., Jenkins, T.M. 2009. Microsatellite markers in plants and insects. Part I. Applications of biotechnology. *Genes Genomes Genomics*, 3:54–67.
- Wang, T., McFarlane, H.E. and Persson, S. 2016. The impact of abiotic factors on cellulose synthesis. *J. Exp. Bot*, 67, 543–552.

- Wang, X. and Wang, L. 2016. GMATA: An Integrated Software Package for Genome-Scale SSR Mining, Marker Development and Viewing. *Frontiers in Plant Science*, 7:1350.
- Wang, X. Y., Yang, X. M., Lu, B., Zhou, L. H. and Wu, K. M. 2017. Genetic variation and phylogeographic structure of the cotton aphid, *Aphis gossypii*, based on mitochondrial DNA and microsatellite markers. *Scientific Reports*, 7, 1920.
- Wang, X., Lu, P. and Luo, Z. 2013. GMATo: A Novel Tool for the Identification and Analysis of Microsatellites in Large Genomes, 9(10), 541-4.
- Wang, Z., Weber, J.L., Zhang, G. and Tanksley, S.D. 1994. Survey of plant short tandem DNA repeats. *Theor Appl Genet*, 88:1–6.
- Weising, K. and Gardner, R.C. 1999. A set of conserved PCR primers for the analysis of simple sequence repeat polymorphisms in chloroplast genomes of dicotyledonous angiosperms. *Genome*, 42, 9–19.
- Wen, X., Zhai, Y., Zhang, L., Chen, Y., Zhu, Z., Chen, G., Wang, K. and Zhu, Y. 2022. Molecular studies of cellulose synthase super complex from cotton fiber reveal its unique biochemical properties. *Science China Life Sciences*, 65, 1776-1793.
- Wendel, J. F. 1989. New world tetraploid cottons contain old world cytoplasm. *Proc. Natl. Acad. Sci. U.S.A.*, 86, 4132–4136.
- Wendel, J.F. and Grover, E.C. 2015. Taxonomy and evolution of the Cotton genus, *Gossypium*. 7: 25-44.
- Wendel, J.F., and Cronn, R.C. 2003. Polyploidy and the evolutionary history of cotton. *Adv. Agron*, 78, 139–186.
- Wendel, J.F., Brubaker, C.L. and Seelanan, T. 2010. “The origin and evolution of *Gossypium*” *Physiology of Cotton*. (Dordrecht: Springer Netherlands). 1–18.
- Wray, G.A., Hahn, M.W., Abouheif, E., Balhoff, J.P., Pizer, M., Rockman, M.V., Romano, L.A. 2003. The evolution of transcriptional regulation in eukaryotes. *Mol Biol Evol*, 20: 1377–1419.
- Yang, J.H., Liu, D.Y., Wang, X.W., Ji, C.M., Cheng, F., Liu, B.N., Hu, Z.Y., Chen, S., Pental, D., Ju, Y.H. et al. 2016. The genome sequence of allopolyploid Brassica Juncea and analysis of differential homoeolog gene expression influencing selection. *Nature Genetics*, 48(10):1225–32.
- Yang, Z., Ge, X., Yang, Z., Qin, W., Sun, G., Wang, Z., Li, Z. Liu, J., Wu, J., Wang, Y. et al. 2019. Extensive intraspecific gene order and gene structural variations in upland cotton cultivars. *Nature Communications*, 10, 2989.
- Yano, K., Yamamoto, E., Aya, K., Takeuchi, H., Lo, P.C., Hu, L., Yamasaki, M. Yoshida, S., Kitano, H., Hirano, K. and Matsuoka, M. et al. 2016. Genome-wide association study using whole-genome sequencing rapidly identifies new genes influencing agronomic traits in rice. *Nature Genetics*, 48, 927–934.
- Yin, Y., Huang, J. and Xu, Y. 2009. The cellulose synthase superfamily in fully sequenced plants and algae. *BMC Plant Biol*, 9:99–113.

- You, M.S., Yue, Z., He, W.Y., Yang, X.H., Yang, G., Xie, M., Zhan, D.L., Baxter, S.W., Vasseur, L., Gurr, G.M., et al. 2013. A heterozygous moth genome provides insights into herbivory and detoxification. *Nature Genetics*, 45(2):220–5.
- Yuan, D., Grover, C.E., Hu, G., Pan, M., Miller, E.R., Conover, J.L., Hunt, S.P. Udall, J.A. and Wendel, J.F. 2021. Parallel and intertwining threads of domestication in allopolyploid cotton. *Adv. Science*, 8, 2003634.
- Zhang, H., Li, D., Zhao, X., Pan, S., Wu, X., Peng, S., Huang, H., Shi, R., Tan, Z. 2020. Relatively semi-conservative replication and a folded slippage model for short tandem repeats. *BMC Genomics*, 21:1–14.
- Zhang, T.T., Zhang, N.Y., Li, W., Zhou, X.J., Pei, X.Y., Liu, Y.G., Ren, Z.Y., He, K.L., Zhang, W.S., Zhou, K.H. et al. 2020. Genetic structure, gene flow pattern and association analysis of superior germplasm resources in domesticated upland cotton (*Gossypium hirsutum* L.). *Plant Diversity*, 42, 189-197.
- Zhang, T., Hu, Y., Jiang, W., Fang, L., Guan, X., Chen, J., Zhang, J., Saski, C.A., Scheffler, B.E., Stelly, D.M. et al. 2015. Sequencing of allotetraploid cotton (*Gossypium hirsutum* L. acc. TM-1) provides a resource for fiber improvement. *Nature biotechnology*, 33, 531–537.
- Zhang, X., Xue, Y., Guan, Z., Zhou, C., Nie, Y., Men, S., Wang, Q., Shen, C., Zhang, D., Jin, S., Tu, L., Yin, P. and Zhang, X. 2021. Structural insights into homotrimeric assembly of cellulose synthase CesA7 from *Gossypium hirsutum*. *Plant Biotechnology Journal*, 19, 1579-587.
- Zhao, Y., Wang, H., Chen, W. and Li, Y. 2014. Genetic structure, linkage disequilibrium and association mapping of *Verticillium* wilt resistance in elite cotton (*Gossypium hirsutum* L.) germplasm population. *PLoS One*, 9, e86308.
- Zou, C., Lu, C., Zhang, Y., Song, G. 2012. Distribution and characterization of simple sequence repeats in *Gossypium raimondii* genome. *Bioinformatics*, 8(17): 801–806.
- Zou, X., Zhen, Z., Ge, Q., Fan, S., Liu, A., Gong, W., Li, J., Gong, J., Shi, Y., Wang, Y. et al. 2018. Genome-wide identification and analysis of the evolution and expression patterns of the cellulose synthase gene superfamily in *Gossypium* species. *Gene*, 646, 28-38.

ÖZGEÇMİŞ

DAMLA AKMAN

ÖĞRENİM BİLGİLERİ

Yüksek Lisans 2021-2025	Akdeniz Üniversitesi Fen Bilimleri Enstitüsü, Tarla Bitkileri AnaBilim Dalı, Antalya
Lisans 2017-2021	Akdeniz Üniversitesi Ziraat Fakültesi, Tarla Bitkileri Bölümü, Antalya