



**BİYOİNFORMATİK; GENSEL VERİ ANALİZİ
TABANLI YAZILIMLARIN KARŞILAŞTIRILMASI**

Tahsin ERTAŞ

BİYOFİZİK ANABİLİM DALI

**Tez Danışmanı
Yrd. Doç. Dr. Onur ÖZTÜRK**

Yüksek Lisans Tezi-2017

**T.C.
İNÖNÜ ÜNİVERSİTESİ
SAĞLIK BİLİMLERİ ENSTİTÜSÜ**

**BIYOİNFORMATİK; GENSEL VERİ ANALİZİ TABANLI YAZILIMLARIN
KARŞILAŞTIRILMASI**

Tahsin ERTAŞ

**Biyofizik Anabilim Dalı
Yüksek Lisans Tezi**

**Tez Danışmanı
Yrd. Doç. Dr. Onur ÖZTÜRK**


**MALATYA
2017**

KABUL VE ONAY SAYFASI

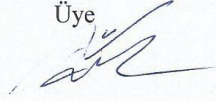
İnönü Üniversitesi Sağlık Bilimleri Enstitüsü Biyofizik Anabilim Dalı Yüksek Lisans Programı çerçevesinde yürütülmüş olan; **Tahsin ERTAŞ'ın “Biyoinformatik; Gensel Veri Analizi Tabanlı Yazılımların Karşılaştırılması”** konulu bu çalışması, aşağıdaki jüri tarafından Yüksek Lisans tezi olarak kabul edilmiştir.

Tez Savunma Tarihi: 29/06/2017

Prof. Dr. Yunus KARAKOÇ
Sağlık Bilimleri Üniversitesi
Jüri Başkanı



Yrd. Doç. Dr. Onur ÖZTÜRK
İnönü Üniversitesi
Tez Danışmanı
Üye



Yrd. Doç. Dr. Harika G.GÖZÜKARA BAĞ
İnönü Üniversitesi
Üye



ONAY

Bu tez, İnönü Üniversitesi Lisansüstü Eğitim-Öğretim Yönetmeliği'nin ilgili maddeleri uyarınca yukarıdaki jüri üyeleri tarafından kabul edilmiş ve Enstitü Yönetim Kurulu'nun/...../2017 tarih ve 2017/..... sayılı Kararıyla da uygun görülmüştür.

Prof. Dr. Yusuf TÜRKÖZ
Enstitü Müdürü

İÇİNDEKİLER

ÖZET.....	v1
ABSTRACT.....	v11
SİMGELER VE KISALTMALAR DİZİNİ.....	v111
ŞEKİLLER DİZİNİ.....	ix
TABLOLAR DİZİNİ.....	x
1. GİRİŞ.....	1
2. GENEL BİLGİLER.....	3
2.1. Haplotip analizi.....	3
2.2. Hb D-Los Angeles.....	3
2.3. İstatistiksel Kavramlar.....	4
2.3.1. Hardy- Weinberg Eşitliği.....	4
2.3.2. Moleküle Özgü Varyans Analizi (AMOVA).....	4
2.3.3. Bağlantı Dengesizliği; Linkage Disequilibrium (LD).....	6
2.3.4. Uyumsuzluk (Mismatch) Dağılım Analizi.....	6
2.4.5. Tarafsızlık (Neutrality) Testleri.....	8
3. MATERYAL VE METOT.....	9
3.1. Arlequin; İstatistik Tabanlı Yazılım.....	9
3.2. Power Marker; İstatistik Tabanlı Yazılım.....	11
4. BULGULAR	14
5. TARTIŞMA	28
6. SONUÇ VE ÖNERİLER	32
KAYNAKLAR.....	34
EKLER.....	38
EK.1. ÖZGEÇMİŞ.....	38
EK.2. ETİK KURUL ONAYINA GEREK OLMADIĞINA DAİR BELGE.....	39

TEŐEKKÜR

Biyofizik yüksek lisans öğrenimim ve tez çalışmam süresince her türlü yardım ve değerli katkıları için tez danışmanım Yrd. Doç. Dr. Onur ÖZTÜRK 'e TEŐEKKÜRLERİMİ SUNARIM.



ÖZET

Biyoinformatik; Gensel Veri Analizi Tabanlı Yazılımların Karşılaştırılması

Amaç: Beta geni üzerinde yer alan, yayınlanmış polimorfik odaklardan yararlanılarak, bu odakların Hb D-Los Angeles mutasyonu ile olan olası ilişkilerinin istatistiksel yazılım programları (Arlequin 3.1) (10) ve Power Marker 3.25 (11) ile karşılaştırmalı analizini yapmaktır. Gensel verilerin farklı yazılımlar tarafından işlenmesi ile elde edilecek sonuçların, gensel veri analizi güvenilirliği, programların uyumu, varsa farklılıkları ve bu farklılıkların nedenlerinin tartışılması, araştırmacıların çalışma amaçlarına yönelik program seçimi ve kıyaslamalarında yol göstermesi bakımından kaynak veri sağlayacağı öngörülmektedir. Bu bakış açısı altında; tez çalışmasının temel amacı, Hb D-Los Angeles modelinde beta globin gen ailesi ve beta globin geni üzerindeki gensel verilerin analizindeki olası farklılıkların ya da ortak sonuçların iki istatistiksel yazılım ile karşılaştırılabilmesidir.

Materyal ve Metot: Tez çalışmamız da Hb D-Los Angeles mutasyonu taşıyan bireyler (40 adet) ile normal bireylere (59 adet) ait yayınlanmış veriler kullanılarak istatistik tabanlı Arlequin 3.1 ve Power Marker 3.25 yazılımlarıyla karşılaştırmalı test değerlendirme aşamaları gerçekleştirilmiştir.

Bulgular: Arlequin ve Power Marker yazılımları, her iki popülasyonda da Hb D-Los Angeles mutasyonu ile ilişkili haplotip sıklıklarını eşit frekans oranlarına sahip biçimde hesaplamıştır. Moleküler çeşitlilik ve Mismatch dağılım parametreleri dikkate alındığında Arlequin yazılımı popülasyonların tarihsel gelişim süreçlerinin belirlemede önemli avantaj sağlayabilmektedir. Her bir lokus için her iki popülasyona ait allel frekans hesaplamaları, allel çifti frekans hesaplamaları, haplotip eğilim regresyonu, ve farklılık testi gibi parametreler Power Marker yazılımına özgü testler olarak sunulmaktadır.

Sonuç: Arlequin yazılımı kullanılarak elde edilen Tau mutasyon yaşı, LD (bağlantı dengesizliği) ve popülasyonların tarihsel süreçlerini yansıtan parametreler gibi sonuçlar, Power Marker yazılımına göre daha ayrıntılı ve kapsamlı olarak hesaplanmaktadır.

Anahtar Kelimeler: Biyoinformatik, gensel veri analizi, Arlequin, Power Marker, popülasyon genetiği

ABSTRACT

Bioinformatics; The Comparison of Softwares based on Genetics Data Analysis

Aim: The comparative analysis with relational statistics programs (Arlequin 3.1 (10) and Power Marker 3.25 (11) is performed, using the published polymorphic loci on the Beta gene to investigate possible associations of these loci with the Hb D-Los Angeles mutation. It is envisaged that the results obtained by the processing of genetic data by different software will provide source data in terms of reliability of the analysis of the genetic data, compatibility of the programs, discrepancies, if any, and reasons for these differences, and the researchers guide the program selection and benchmarking for the study purposes. Under this point of view; The main purpose of the thesis study was to compare the possible differences or common results of analysis of gene data on beta globin gene family and beta globin gene in Hb D-Los Angeles model with two statistical software.

Material and Method: In our thesis study, comparative test evaluation steps were performed with statistical-based Arlequin 3.1 and Power Marker 3.25 software using the published data of Hb D-Los Angeles mutated individuals and (40) normal individuals (59).

Results: Arlequin and Power Marker software calculated the haplotype frequencies associated with the Hb D-Los Angeles mutation in both populations with equal frequency values. Considering the molecular diversity and mismatch distribution parameters, Arlequin software can provide important advantages in determining the historical development processes of populations. For each locus, parameters such as allele frequency calculations, allele pair frequency calculations, haplotype tendency regression, and difference test for both populations were presented as specific tests of Power Marker software.

Conclusion: Results obtained using Arlequin software such as Tau mutation age, LD (linkage disequilibrium) and parameters reflecting historical processes of populations are calculated in more detail and comprehensively than Power Marker software.

Key Words: Bioinformatics, genetic data analysis, Arlequin, Power Marker, Population genetics

SİMGELER VE KISALTMALAR DİZİNİ

DNA	: Deoksiribonükleik asit
Hb	: Hemoglobin
mt DNA	: Mitokondriyal DNA
PCR	: Polimeraz zincir reaksiyonu
RFLP	: Restriksiyon fragmenti uzunluk polimorfizmi
SNP	: Tek nükleotid polimorfizmi
OTU	: Gözlenen sınıflandırma birimi
LRT	: Olasılık oranı testi
EM	: Beklenti maksimizasyonu (<i>Expectation-maximization</i>)
LD	: Bağlantı Dengesizliği (<i>Linkage Disequilibrium</i>)
AMOVA	: Moleküle Özgü Varyans Analizi
HWE	: Hardy-Weinberg Eşitliği

ŞEKİLLER DİZİNİ

<u>Şekil No</u>	<u>Sayfa No</u>
Şekil 2.1. Beta globin gen ailesi ve haplotiplemede kullanılan 7 odak. Enzim kesim bölgesi (↑)	3
Şekil 2.2. Arlequin istatistiksel yazılım programı AMOVA arayüzü.....	5
Şekil 2.3. Power Marker istatistiksel yazılım programı AMOVA arayüzü.....	6
Şekil 4.1. Normal popülasyonun ani genişleme modelinde gözlenen çift farklılığı ve beklenen uyumsuzluk dağılımları.....	14
Şekil 4.2. Hb D-Los Angeles popülasyonunun ani genişleme modelinde gözlenen çift farklılığı ve beklenen uyumsuzluk dağılımları.....	14

TABLULAR DİZİNİ

Tablo No	Sayfa No
Tablo 2.1. Hb D Ailesi Çeşitliliği.....	4
Tablo 3.1. Arlequin yazılımı ile gerçekleştirilebilen popülasyon içi istatistiksel analiz yöntemleri.....	10
Tablo 3.2. Arlequin yazılımı ile gerçekleştirilebilen popülasyonlar arası istatistiksel analiz yöntemleri.....	11
Tablo 3.3. Power Marker yazılımı ile gerçekleştirilebilen istatistiksel analizler.....	13
Tablo 4.1. Hb D-Los Angeles popülasyonu için İkili Bağlantı Dengesizliği (<i>Pairwise Linkage Disequilibrium</i>); Arlequin sonuçları.....	15
Tablo 4.2. Hb D-Los Angeles popülasyonu için Linkage disequilibrium (<i>LD</i>) anlamlılık diyagramı; Arlequin sonuçları.....	15
Tablo 4.3. Hb D-Los Angeles popülasyonu için Linkage disequilibrium (<i>LD</i>) anlamlılık diyagramı; Power Marker sonuçları.....	16
Tablo 4.4. Normal popülasyonu için İkili Bağlantı Dengesizliği (<i>Pairwise Linkage Disequilibrium</i>); Arlequin sonuçları.....	16
Tablo 4.5. Normal popülasyonu için Linkage disequilibrium (<i>LD</i>) anlamlılık diyagramı; Arlequin sonuçları.....	17
Tablo 4.6. Normal popülasyon için Linkage disequilibrium (<i>LD</i>) anlamlılık diyagramı; Power Marker sonuçları.....	17
Tablo 4.7. Hb D-Los Angeles popülasyonu [$\beta 121$ (<i>GH4</i>) <i>Glu</i> → <i>Gln</i> , <i>GAA</i> → <i>CAA</i>] haplotip çeşitliliği; Arlequin sonuçları.....	17
Tablo 4.8. Hb D-Los Angeles popülasyonu [$\beta 121$ (<i>GH4</i>) <i>Glu</i> → <i>Gln</i> , <i>GAA</i> → <i>CAA</i>] haplotip çeşitliliği; Power Marker sonuçları.....	18

Tablo 4.9. Normal popülasyon haplotip çeşitliliği; Arlequin sonuçları.....	18
Tablo 4.10. Normal popülasyon haplotip çeşitliliği; Power Marker sonuçları.....	19
Tablo 4.11. Normal ve Hb D-Los Angeles popülasyonları arasında hesaplanan (<i>AMOVA</i>) F-istatistiği; Arlequin sonuçları.....	20
Tablo 4.12. Normal ve Hb D-Los Angeles popülasyonları arasında hesaplanan (<i>AMOVA</i>) F-istatistiği; Power Marker sonuçları.....	20
Tablo 4.13. Normal ve Hb D-Los Angeles popülasyonu için Özel F-İstatistiği; Power Marker sonuçları.....	21
Tablo 4.14. İki popülasyon için molekül sel çeşitliliğin özeti; Arlequin sonuçları.....	21
Tablo 4.15. Mismatch dağılımı test istatistiklerinin SSD ve r_g değerleri; Arlequin sonuçları.....	21
Tablo 4.16. Normal popülasyondaki tüm lokuslar için Hardy-Weinberg Eşitliği (<i>HWE</i>) testi; Arlequin sonuçları.....	22
Tablo 4.17. Normal popülasyondaki tüm lokuslar için Hardy-Weinberg Eşitliği (<i>HWE</i>) testi; Power Marker sonuçları.....	22
Tablo 4.18. Hb-D Los Angeles popülasyonundaki tüm lokuslar için Hardy-Weinberg Eşitliği (<i>HWE</i>) testi; Arlequin sonuçları.....	23
Tablo 4.19. Hb-D Los Angeles popülasyonundaki tüm lokuslar için Hardy-Weinberg Eşitliği (<i>HWE</i>) testi; Power marker sonuçları.....	23
Tablo 4.20. Her bir lokus için normal popülasyona ait allel frekans hesaplamaları; Power Marker sonuçları.....	24
Tablo 4.21. Her bir lokus için Hb D-Los Angeles popülasyonuna ait allel frekans hesaplamaları; Power Marker sonuçları.....	24
Tablo 4.22. Her bir lokus için normal popülasyona ait allel çifti frekans hesaplamaları; Power Marker sonuçları.....	25

Tablo 4.23. Her bir lokus için Hb D-Los Angeles popülasyonuna ait allel çifti frekans hesaplamaları; Power Marker sonuçları.....	26
Tablo 4.24. Normal ve Hb D-Los Angeles popülasyonları için Haplotip Eğilim Regresyonu; Power Marker sonuçları.....	27
Tablo 4.25. Normal ve Hb D-Los Angeles popülasyonları için Farklılık Testi; Power Marker sonuçları.....	27



1. GİRİŞ

Günümüzde genetik teknolojilerin ulaştığı noktada kalıtsal klinik sorunlara neden olan veya klinik olarak belirti göstermeyen ancak nesilden nesile kalıtılabilen olgular üzerinde yapılan araştırmalar, literatür veri bankalarının gelişmesine önemli katkılarda bulunmaktadır (1). Hastalıkların önemli bir bölümünün temeli genetik (gene özgü) olduğundan, mümkün olan tedavi yöntemi yine gen kaynaklı olmalıdır. Günümüzde bu yönde çalışmalar bulunmakla birlikte henüz gen tedavisi uygulaması olanaklı değildir. Genler ve gen ürünü olan proteinleri konu alan çalışmaların ve bu çalışmalarda üretilen verilerin hızla artması genetik ve sağlık sorunu oluşturan hastalıkların anlaşılmasında ve olası tedavi yöntemlerinin geliştirilmesinde umut vadetmektedir. Doğal olarak genetik çalışmaların sayısının hızla artması ve karmaşık veri kümelerinin üretilmesi ile genetik teknoloji ile bilgisayar teknolojisinin ortak çalışma alanı bulmasına neden olmaktadır. Bu amaçla üretilen verilerin saklanması, analizi ve paylaşımı gibi temel konularda genetik veri tabanlı yazılımlar oldukça yararlı olmaktadır. Bu bağlamda, moleküler biyolojinin içinde, canlı sistem modelinde bilgi toplamak, işlemek, depolamak ve kullanmak gibi bir temel araştırma temasının olması gerektiği kabul edilmektedir (2). Son 30 yılda araştırmacılar çok sayıda organizmanın genom dizisini belirleyebilmek için çaba göstermektedir (3). Bunların en büyüğü, 1980'li yıllarda ortaya atılan ve 1987'lerde oluşturulmaya başlanan ve 1990 Ekim ayında resmen başlayan 'insan genom projesi' olmuştur. Bu büyük proje ve diğer genom projelerinden elde edilen DNA, RNA ve protein dizi verilerinin saklanması, analizinin yapılması, işlenmesi ve sonuçlarının kullanıma sunulabilmesi bilgisayar teknolojilerinin ve tıp biliminin ortak çalışmasına olanak sağlayan 'Biyoinformatik' bilim dalı ile mümkün olmaktadır. Gerek dünya da ve gerekse ülkemizde bu yeni bilim dalının önemi giderek artmaktadır. Biyoinformatik bilim dalının temel amacı; hızla gelişen ve artan bilgiyi toplamak, yönetmek, dağıtmak, bilgiye en hızlı ve kolay biçimde ulaşılmasını sağlamak, birçok alt sistemden oluşan yapıdaki biyolojik sistemlerin işlevlerini tanımlamaya çalışmaktır (4). Çalışmamızdaki sonuçların yazılım tercihleri açısından tartışılmasının, geçmişe dönük genetik süreçlerin araştırılmasında ve antropolojik yaklaşımların geliştirilmesinde değerli veriler üretebileceği öngörülmektedir (5). Bu kapsamda, tez çalışmamızda model olarak seçtiğimiz ve daha önce yayınlanmış olan (6-9) anormal hemoglobin (*Hb D-Los*

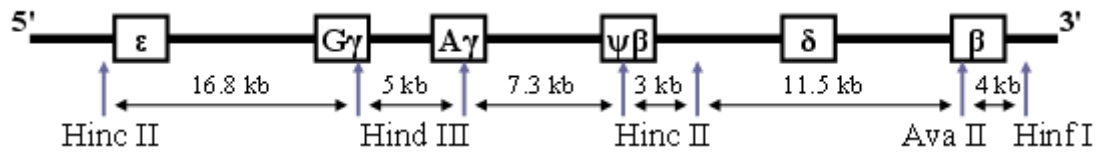
Angeles) popülasyonu ve normal popülasyon haplotip analizi verileri referans alınarak, literatürde yer alan ve sıklıkla tercih edilen biyoinformatik amaçlı iki farklı yazılımdan (*Arlequin ver. 3.5 ve Power Marker ver. 3.25*) elde edilen istatistiksel sonuçlar değerlendirilmektedir (10,11). Bu sonuçların değerlendirilmesi ile elde edilen bilgilerin, arařtırcıların alıřma amalarına yönelik program seimi ve uygun yöntemlerin seilmesi konusunda yararlı olacağı öngörülmektedir.



2. GENEL BİLGİLER

2.1. Haplotip analizi

Haplotip analizi, tek nükleotid değişimine sahip bölgelerin PCR ve RFLP tekniklerinin uygulanması, sonuçların ise uygun gensel yazılımlarla test edilmesi işlemidir. Polimorfizm, herhangi bir sorun yaratmayan, bölgesel nükleotid değişimleri olarak tanımlanır. Polimorfizmler, nesilden nesile kalıtılma özelliklerinden dolayı geçmişe dönük gensel analiz köken araştırmaları için oldukça değerli veriler sunmaktadır (7). Çalışmamızda, Hb D-Los Angeles [$\beta 121(GH4)Glu \rightarrow Gln$] mutasyonlu olgularda, β -globin gen ailesindeki toplam yedi bölgenin ($G\gamma/A\gamma$ -globin, $\psi\beta$ -globin, δ -globin ve β -globin) yayınlanmış enzim kesim sonuçları kullanılarak haplotip analizleri yapılmıştır (7, (Şekil 2.1)).



Şekil 2.1. Haplotip analizinde yer alan 7 odak. Enzim kesim bölgesi (↑) (12, 13).

2.2. Hb D-Los Angeles

Çalışmamızda model olarak belirlediğimiz Hb D-Los Angeles mutasyonu, β -zinciri 121. kodonda G>C farklılaşması ile glutamik asit \rightarrow glutamin değişimi ile oluşan amino asit farklılığından kaynaklanmış bir proteindir (6- 9). İlk olarak, Itano tarafından 1951 yılında yayınlanmış ve Hb-D Los Angeles adını almıştır (8). Bu anormal hemoglobinin molekül özellikleri 1964 yılında tanımlanmış ve 21 farklı üyesi literatüre kazandırılmıştır (Tablo 2.1). Bu üyelerden bazıları, benzer gensel farklılığa sahiptir, ancak farklı bölgelerde tanımlanmaları nedeni ile gerçekte 9 farklı molekül çeşidi bulunmaktadır (15).

Tablo 2.1. Hb D Çeşitliliği (15)

Amino asit Çeşitliliği	Hb Adı	Amino asit Çeşitliliği	Hb Adı	
β 16(A13) Gly→Arg	D-Bushman	α 68 (E17) Asn→Lys	D-Baltimore	
β 19(B1) Asn→Lys	D-Ouled Rabah		D-St. Louis	
β 22(B4) Glu→Gln	D-Iran		D-Washington	
Glu→Val	D-Granada		G-Philadelphia	
β 87(F3) Thr→Lys	D-Ibadan		Stanleyville-T	
β 121(GH4) Glu→Val	D-Camperdown		G-Bristol	
	D-Beograd		G-Knoxville-T	
β 121(GH4) Glu→Ala	D-Neath		G-Azakuoli	
Glu→Gln	D-Los Angeles			
	D-Punjab			
	D-Chicago			
	D-Nort Carolina			
	D-Portugal			
	D-Oak Ridge			

2.3. İstatistiksel Kavramlar

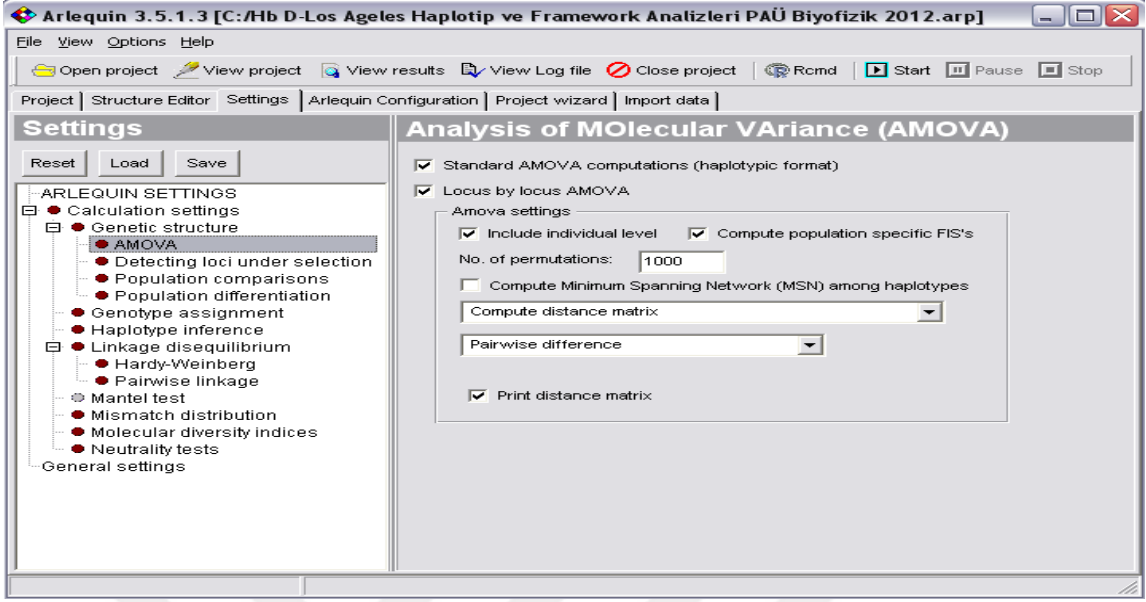
2.3.1. Hardy-Weinberg Eşitliği

Genetik ilişkilendirme çalışmalarında önemli alleller için Hardy-Weinberg Dengesi (HWE) testi önemli bir erken adım olarak önerildi. Örneğin, iki aşamalı test prosedürleri sıklıkla önerilmektedir (16), burada HWE testi, allel ve hastalık arasındaki ilişkiyi test etmeden önce ön adımdır (17, 18). Buna ek olarak, bazı G-E etkileşim test yaklaşımları HWE'yi üstlendiği için, bir hastalıkla bağlantılı olarak gen ortamı (G-E) etkileşimleri için daha fazla test yapılmadan önce HWE için test yapılmaktadır (19). Hardy-Weinberg kanununu genotip frekanslarını " $p^2+2pq+q^2=1$ " denkleminde göre test eder. Test sonuçlarına göre varsayımların geçerlilik düzeylerine karar verilir (20).

2.3.2. Moleküle Özgü Varyans Analizi (AMOVA)

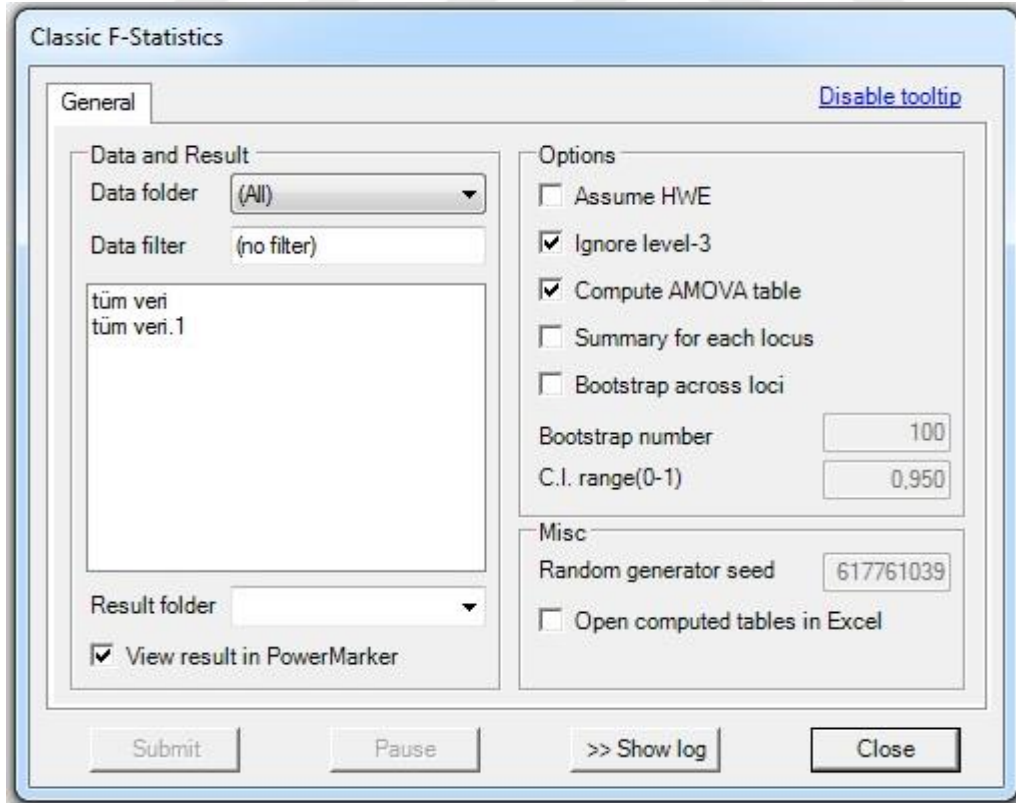
Bu analiz yönteminde haplotip çeşitliliği yüzde sıklık ile allel olasılık verileri kullanılarak popülasyonlara özgü gensel parametreler hesaplanmaktadır (21). Popülasyonların olası farklı alt grupları (*Popülasyonun bireyleri, popülasyonlar, popülasyonun alt grupları içi ve popülasyon alt grupları arasında*) ile ilişkili bileşenlerin testi parametrik olmayan istatistiksel yöntemler ile gerçekleştirilmektedir (21).

Şekil 2.2' de Arlequin yazılımının haplotip çeşitliliği verilerini referans alan varyans analizi arayüzü gösterilmektedir.



Şekil 2.2. Arlequin istatistiksel yazılım programı AMOVA arayüzü

Power Marker istatistiksel yazılımıyla da AMOVA analizi yapılmaktadır (bkz. Şekil 2.3).



Şekil 2.3. Power Marker istatistiksel yazılım programı AMOVA arayüzü

2.3.3. Bağlantı Dengesizliği; Linkage Disequilibrium (LD)

Aynı kromozom üzerindeki polimorfik odakların birbirinden bağımsız olarak kalıtılmalarındaki zorluk nedeni ile yeni haplotip çeşitlerinin (*rekombinant ürünler*) ortaya çıkması olasılığı düşmektedir (*linkage-bağlantı*). LD aynı veya farklı kromozom üzerindeki allellerin yüksek olasılıklı bağlantılarıdır (22). Doğal seçim yoluyla görülme olasılığı artan bir gen bölgesi, içerdiği ve bağlantılı olduğu polimorfik odakların ortaya çıkma olasılığını arttırmaktadır. Bu bağlantı sayesinde gene özgü evrimsel süreç geçmişe dönük olarak araştırılabilmektedir. Söz konusu araştırmanın gerçekleştirilebilmesi için gen bölgesine özgü polimorfik odak haritalandırılmasının tespitine gereksinim vardır. Yüksek sayılı popülasyon çalışmalarında LD testleri gerçekleştirilirken hangi allelin hangi atadan kalıtıldığını belirlemek zor olduğundan tahmini haplotipler arasında olasılığı en yüksek haplotip çeşidi tespit edilmeye çalışılır ve anlamlılık testleri uygulanır (23, 24). Söz konusu bu istatistiksel bağlantı analizleri “Maksimum olabilirliğe dayalı algoritmalar (*Expectation Maximization (EM) algoritması*) ile ML (*Maximum-Likelihood*)” testlerini içeren Power Marker ve Arlequin yazılımları ile mümkün olmaktadır. İki yazılımda ortak olarak, LD sonuçlarının belirlenmesi için “Ki-Kare testi” uygulanmaktadır. Ki-Kare testi; H_0 hipotezinin ($H_0=iki\ lokus\ arasında\ LD\ vardır$) onaylanıp onaylanmayacağını test etmektedir. Eğer, $p < 0.05$ ise H_0 hipotezi onaylanır ve değişkenler arasında bağlantı olduğu yani gen bölgelerinin, üzerlerindeki polimorfik odaklarla beraber bir sonraki nesile aktarılacağı sonucuna varılmaktadır.

2.3.4. Uyumsuzluk (*Mismatch*) Dağılım Analizi

Mismatch analizi popülasyon içi haplotip çiftleri arasındaki gözlenen farklılıkların dağılımını hesaplamaktadır. Popülasyon birey sayısı olarak sabit olduğunda bu dağılım çok modlu (*multimodal*) bir grafik olarak kendini göstermektedir. Nüfus yoğunluğu üstel olarak artan popülasyonların mismatch analizi verileri, tek modlu (*unimodal*) grafik olarak gözlenmektedir. Mismatch dağılımı toplulukların birey sayısı değişimleri ile ilgili geçmişe dönük araştırmalar için değerli bilgiler sağlamaktadır (25, 26). Söz konusu analiz sonuçları; toplum birey sayısındaki değişim parametrelerinden (τ , θ_0 , θ_1 , SSD, RI gibi) yararlanılarak hesaplanmaktadır. θ_0 ve θ_1

popülasyonun ilk ve son nüfus büyüklük değerlerini tanımlamaktadır. Raggedness indeks (RI), p değerine bağlı olarak mismatch dağılım analizinin anlamlılığını test etmektedir. Sum of square deviations (SSD), beklenen ve gözlenen değerler üzerinden ani popülasyon yayılma modeli ile mismatch dağılımını karşılaştırmaktadır (27). Analizler sonucu, rg indekslerinin düşük değerlerde ve SSD p-değerlerinin > 0.05 olması bu popülasyonların nüfussal olarak benzer genişleme oranlarına sahip olduklarını ifade etmektedir (28). SSD ve rg indekslerinin p değerleri, gözlenen verilerin tahmin edilen ya da beklenene ne kadar iyi uyduğunu-uyumadığını belirlemeye yönelik matematiksel bir kanıt sunmaktadır. Bu test; uyum-iyiliği (*Goodness of fit*) olarak ifade edilmektedir. Arlequin yazılımında sıfır hipotezinin uyum-iyiliğine karar vermek için geliştirilen istatistiksel testlerden biri olan Ki-kare (χ^2) kullanılmaktadır. Eğer, $p > 0,05$ ise H_0 hipotezi ($H_0 = \text{dağılım tek modludur}$) kabul edilir. Başka bir ifade ile, sıfır hipotezinin kabul edilmesi, gözlenen ikili çeşitlilik dağılımının, ani popülasyon yayılması modeli altındaki beklenen çeşitlilikten istatistiksel olarak önemli derecede farklı olmadığını işaret etmektedir. Tek modlu grafik, nüfusun herhangi bir kısıtlama veya göç olayı etkisi altında kalmadan üstel olarak arttığını ifade etmektedir (25-27, 29).

Mutasyon parametreleri modeline göre, $\theta_1 = 2uN_0$, $\theta_2 = 2uN_1$ ve $\tau = 2ut$ (u; mutasyon oranı) parametreleri modellenerek, güven aralıkları tanımlanmıştır (29). Söz konusu modelde τ (*Tau; mutasyonel zaman birimi*) parametresi kullanılarak popülasyonun olası yaşı hesaplanabilmektedir. Bu parametre hesaplanırken kullanılan insan mitokondriyel nükleotid diverjans oranı güven aralığı milyon yılda %2 - %4 olarak belirlenmiştir. %4 diverjans oranına karşılık mutasyon oranı $u = 15 \times 10^{-4}$, %2 diverjans oranına karşılık mutasyon oranı $u = 7.5 \times 10^{-4}$, jenerasyon sayısı $1/2u$ ve jenerasyon yaşı da her jenerasyon başına 25 yıl olarak tanımlanmıştır. Sonuç olarak model de, zaman parametresi için güven aralığı 8333–16666 yıl olarak hesaplanmaktadır. Elde edilen zaman parametresi ile tau değerinin çarpımı popülasyon yaşını yaklaşık olarak göstermektedir (29).

2.4.5. Tarafsızlık (*Neutrality*) Testleri

Çalışmamızda Arlequin yazılımı kullanılarak, yayınlanmış verilere Tajima'nın D ve Fu'nun F_S tarafsızlık testleri uygulanmıştır. Tajima D testi, popülasyon içindeki ayrık site sayısını (S) ve haplotipler arasında ortalama çift farklılıklarını (π) referans alan iki teta (θ) değerini karşılaştırmaktadır (30-32). Diğer bir deyişle, Tajima D testi popülasyonların standart nötral modelden sapmasını test etmektedir. İstatistiksel olarak anlamlı ($p < 0.05$) kabul edilen Tajima D değeri nötral modelde doğal seçim, nüfussal dalgalanmalar ve diğer popülasyon dengesi bozucu nedenler ile beklenen S ve π parametrelerinin değişmesinden kaynaklanmaktadır. ($p > 0.05$) olduğu durum da ise, Tajima D değeri, popülasyonların nötral dengede olduğunu tahmin etmektedir. ($p > 0.05$) olduğunda popülasyonda nadir varyasyonların sayısı her zaman düşük çıkmaktadır (33). Tajima D değerinin pozitif olması; heterozigotluğun seçilimli bir avantaja sahip olduğunu, negatif olması ise belirli bir allelin diğer allel arasında seçilimli bir avantaja sahip olduğunu tanımlamaktadır. Ek olarak negatiflik, popülasyon nüfusunda hızlı bir artışı işaret etmektedir (34). Fu'nun F_S testi popülasyon nüfusunun artışına duyarlı bir belirteç parametresidir (35). Fu'nun F_S testinde, anlamlı ($p < 0.05$) p değerine sahip negatif Fu test sonucu, popülasyonların ortak nüfussal büyümeye sahip ve aynı gen havuzuna ait olduklarına işaret etmektedir (36).

3. MATERYAL VE METOT

Tez çalışmamızda model olarak kullanılan yayınlanmış haplotip verileri ilgili referans yayında ifade edildiği üzere [Hb D-Los Angeles [$\beta 121(GH4)Glu \rightarrow Gln$] mutasyonu taşıyan bireyler (40 adet) ile normal (59 adet) bireylere ait DNA örneklerinin beta globin gen bölgeleri, uygun primerler kullanılarak PCR yöntemi ile çoğaltılmış ve haplotip analizinin devam eden aşamalarında ise sırası ile enzim kesimi] uygulanarak elde edilmiş ve yayınlanmış verilerdir (7-9). Tez çalışmamızda bu verilerin istatistiksel değerlendirme aşamaları gerçekleştirilmiştir.

3.1. Arlequin; İstatistik Tabanlı Yazılım

Arlequin yazılımı, gensel ve gen kaynaklı sayısal verileri, test içeriğinde bulunan temel istatistik yöntemleri uygulayarak analizlerini gerçekleştirebilmektedir. Bu istatistik yöntemler ile DNA dizi, RFLP, DNA mikrodiziner, standart ve allel sıklığı gibi verilerin analizi gerçekleştirilmektedir (10,37). Yazılım literatür oranları göz önüne alındığında gensel verilerin anlamlandırılmasında araştırmacılar tarafından oldukça sık tercih edilmektedir (38). Arlequin yazılımı gensel çeşitlilik incelemeleri, F-istatistiği ve popülasyonlar arası gensel ilişkiler, bağlantı dengesizliği (*LD, Linkage Disequilibrium*) ve popülasyon içi çeşitlilik, çoklu genotiplere ait gametik faz değerlendirmeleri, nüfus dağılımları ve yaygınlık testleri, Hardy-Weinberg kesin testi, molekülse çeşitlilik analizi (*AMOVA*) gibi popülasyon gensel süreçlerini kapsayan konularda analiz çeşitliliği sunmaktadır (37, Tablo 3.1 ve 3.2).

Tablo 3.1. Popülasyon içi istatistiksel analiz yöntemleri; Arlequin yazılımı (7).

Popülasyon içi yöntem	Kısa Tanım
Standart endeksler	Polimorfik site sayısı ve gen çeşitliliği gibi bazı çeşitlilik ölçümleri.
Moleküler çeşitlilik	Nükleotid çeşitliliği ve popülasyon parametresi olan Teta (θ) değerinin farklı formlarını hesaplar.
Uyumsuzluk (<i>Mismatch</i>) dağılımı	Haplotipler arasındaki ikili (pairwise) farklılıkların dağılımından, popülasyonun nüfussal ve mekansal genişleme parametrelerini belirler.
Haplotip frekansı tahmini	Maximum Likelihood (ML) yöntemi ile popülasyon içinde bulunan haplotip sıklıklarını belirler.
Gametik faz tahmini	Pseudo-Bayesian yaklaşımını kullanarak, çoklu-lokus genotiplerinin en olası gametik fazını belirler (ELB algoritması).
Bağlantı dengesizliği (LD)	Farklı lokuslardaki allellerin rastgele-olmayan bağlantılarını test eder.
Hardy-Weinberg dengesi	Diploid bireylerde allellerin rastgele-olmayan bağlantılarını test eder.
Tajima' nın tarafsızlık testi (sonsuz site modeli)	Sonsuz site modeli altında DNA dizileri ve RFLP haplotiplerinin rastgele örneklem seçici tarafsızlık testidir.
Fu' nun <i>FS</i> tarafsızlık testi (sonsuz site modeli)	Sonsuz site modeli altında DNA dizileri ve RFLP haplotiplerinin rastgele örneklem seçici tarafsızlık testidir.
Ewens-Watterson tarafsızlık testi (sonsuz allel modeli)	Sonsuz site modeli altında Ewens örnekleme teorisi tabanlı seçici tarafsızlık testidir.
Chakraborty' nin birleştirme testi (sonsuz allel modeli)	Seçici tarafsızlık ve popülasyon homojenliği testidir. Bu test homojenliğin şüpheli olduğu durumlarda uygulanır.
Minimum yayılım ağı (<i>Minimum Spanning Network; MSN</i>)	Haplotipler arasında Minimum yayılım ağacı (<i>Minimum Spanning Tree; MST</i>) ve MSN hesaplar. Bu test AMOVA ile birlikte farklı popülasyonlardaki haplotiplerin bağlantılarını belirlemek için de kullanılabilir.

Tablo 3.2. Popülasyonlar arası istatistiksel analiz yöntemleri; Arlequin yazılımı (7).

Popülasyonlar arası yöntem	Kısa Tanım
Popülasyonlar arasında ortak olan haplotiplerin belirleme	Haplotip içeriklerine göre örneklerin popülasyonlar arası karşılaştırmasını yapar.
AMOVA	Popülasyonların gensel yapıları üzerinde molekülse çeşitlilik analizi yapar.
İkili gensel uzaklıklar (<i>Pairwise genetic distances</i>)	Kısa süre içerisinde gerçekleşmiş gensel çeşitlenmeler için F_{ST} tabanlı gensel mesafe analizi testidir.
Popülasyon farklılaşma kesin testi (<i>Exact test of population differentiation</i>)	Panmixia hipotezi altında popülasyonlar içindeki haplotiplerin rastgele olmayan dağılımlarını test eder.
Genotip belirleme testi	Olası allel sıklıkları ile popülasyonlardaki bireylerin genotiplerini belirler.
F-istatistiği ile seçilime uğrayan lokusları belirleme	F_{ST} dağılımı incelenmesi ile seçilime uğrayan lokusları ve "hierarchical island model" altında heterozigotluğu belirler.

3.2. Power Marker; İstatistik Tabanlı Yazılım

Power Marker, ilk resmi sürümü Ocak 2004'de piyasaya sürülen, gensel veri tabanlı bir yazılım olarak literatürde yer almaktadır. Popülasyonların gensel analizlerinde SSR / SNP verilerinin kullanımını sağlamaktadır. Elde edilen gensel veriler excel dosyası veya diğer veri dosyası formatları kullanılarak kolaylıkla analiz edilebilmektedir. Veriler ayrıca NEXUS ve Arlequin formatlarına da aktarılabilir. Bağlantı dengesizliği görselleştirmesi için bir '2D görüntüleyici' içerir. Kullanıcı, Power Marker yazılımı içindeki grafikleri düzenleyebilmekte veya yayınlamak için bunları yazılım dışına çıkarabilmektedir. Program doğruluk ve etkililik açısından kapsamlı bir şekilde test edilmiştir. Power Marker yazılımı kolay ulaşılır olmakla birlikte ücretsiz olarak paylaşılmaktadır. Ancak PHYLIP, TreeView ve Microsoft.net çerçeve sistemi (*tümüyle kullanıma açık*) ve Excel 2000 (*ücretli*) sürümü ile desteklenmektedir. Ek olarak yazılım, Windows 98 ve üzeri (*Macintosh veya diğer sistemler hariç*) sürümlerde çalışabilmektedir (11).

Son yıllarda, gensel işaretleyicilerin popülasyonların gensel analizleri için temel yöntem haline gelmesi ile güvenli veri analizi amaçlı bilgisayar yazılımına duyulan ihtiyaç artmaktadır. Power Marker yazılımı bu ihtiyacı, kullanıcı dostu ve tekrarlanabilir

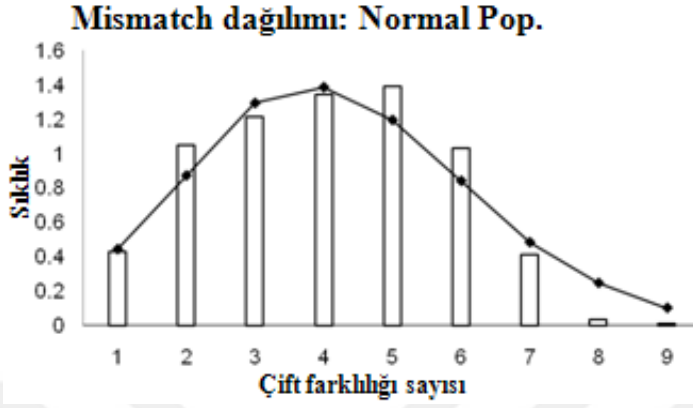
çözümler önererek karşılamaya çalışmaktadır. Yazılım, genetik işaretleyici veriler için sonuç odaklı, bütüncü bir analiz ortamı (*IAE-integrated analysis environment*) sunmaktadır. IAE, veri yönetimi, analiz ve görselleştirmeyi kullanıcı dostu bir grafik kullanıcı arabiriminde bütüncüştürmektedir. Analiz süresini kısaltarak, kullanıcıların süreç boyunca veri bütünlüğünü sürdürmesini sağlamaktadır. Yazılım ile haploid veya diploid olmak üzere çeşitli işaretleyici verilerin analizleri gerçekleştirilebilmektedir. Bilinen veya bilinmeyen diploid veriler için gametik faz hesaplamalarına olanak sağlamaktadır. Yazılım aynı zamanda mikrosatelit verileri, tek nükleotid polimorfizm (*SNP*) verileri ve RFLP verilerini de analiz edebilmektedir. Power Marker veri seti, frekans, uzaklık, ağaç, tablo ve metin olmak üzere altı farklı veri girişi seçeneği sunmaktadır. Her biri metin dosyalarından içe aktarılabilir ve metin dosyalarına aktarılabilir özellik göstermektedir. Dataset seçeneği, birkaç farklı genetik bilginin depolanmasını sağlayarak çoklu analizlerin gerçekleşmesini sağlamaktadır. Haplotip çeşitliliği frekans bilgisi, popülasyonların genetik farklılıklarının belirlenmesi için kullanılan analizlerin referansı olarak kullanılmaktadır (39).

Power Marker, χ^2 testleri, olasılık oranı testleri ve kesin testler de dahil olmak üzere Hardy-Weinberg dengesini ve bağlantı dengesini test etmek için tüm yaygın yöntemleri uygulamaktadır. Linkage Disequilibrium (*LD*) analizi D' ve r^2 parametreleri hesaplanarak gerçekleştirilmektedir. LD sonuçları Microsoft Excel kullanılarak program dışına aktarılabilir özellik göstermektedir. Oluşturulan LD matrisi 2D görüntüleyicisi tarafından dahili olarak görüntülenebilmekte ve daha ileri düzenleme için bir grafik dosyası olarak dışa aktarılabilir (39). Power Marker yazılımı her bir işaretleyici(marker) lokusu için allel numaralarını içeren, polimorfizm bilgi içeriği(*PIC*) ve mikro dizin veriler için birkaç adımlı olgu istatistiklerini hesaplamaktadır (40). Bireyler arasında analiz tekrar sayısını artırarak belirlenir. Power Marker yazılımındaki haplotip çeşitliliğinin hesaplanması bu belirteçlerin ikili özelliğinden yararlanarak polimorfizm verileri kullanılarak gerçekleştirilmektedir. Power Marker, bireylerin ilgili lokus çiftleri arasındaki ilişkiyi test etmek için 3 yöntem sunmaktadır. Bunlar allele ait kontrol testi, genotipe ait kontrol testi ve çok yönlü eğilim testi olarak ifade edilmektedir. Power Marker aynı zamanda nicelik özelliklerini test eden bir F-testi sunmaktadır. Bu test haplotip eğilim regresyonu ile beraber hem niceliğe hem de ikili lokuslara uygulanabilir özelliğindedir (36).

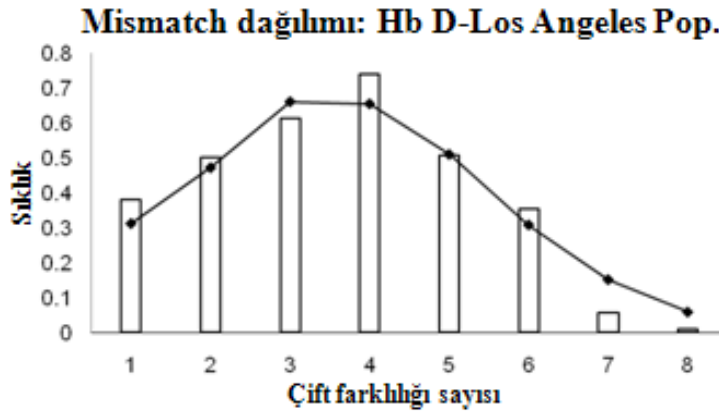
Tablo 3.3. Power Marker yazılımı ile gerçekleştirilebilen istatistiksel analizler (39).

Kategori	Kısa Tanım
Özet istatistikler	Allel sayısı, gen çeşitliliği, soydaşım katsayısı gibi toplam istatistikler; Allelik, genotipik ve haplotipik frekansının tahmini; Hardy-Weinberg dengesizliği ve Bağlantı dengesizliği
Tasarım	Çekirdek kümesi veya haplotip etiketleme işaretçilerinin seçimi
Popülasyon yapısı	Popülasyon farklılaştırma testi; Klasik F-İstatistiklerinin yanı sıra popülasyona özgü F-İstatistiği; Popülasyon yapısı tahmini
Filogenetik analiz	Frekans, uzaklık ve ağaç analizi
Birliktelik analizi	Farklı tasarımlar için ortaklık testi
Araçlar	SNP simülasyon ve tanımlama gibi yardımcı program araçları, Mantel Olasılık tabloları için test ve tam p-değerleri

4. BULGULAR



Şekil 4.1. Normal popülasyonun ani genişleme modelinde gözlenen çift farklılığı (çubuklar) ve beklenen uyumsuzluk dağılımları (düz çizgi). Arlequin 3.5 yazılımı kullanılarak elde edilmiştir.



Şekil 4.2. Hb D-Los Angeles popülasyonunun ani genişleme modelinde gözlenen çift farklılığı (çubuklar) ve beklenen uyumsuzluk dağılımları (düz çizgi). Arlequin 3.5 yazılımı kullanılarak elde edilmiştir.

Tablo 4.1. Hb D-Los Angeles popülasyonu için İkili Bağlantı Dengesizliği (*Pairwise Linkage Disequilibrium*); Arlequin sonuçları.

	Marker1	Marker2	Çok- Allelik D	Çok- Allelik D'	Exact p değeri	Ki-kare değeri	Ki-kare p-değeri	Tekrar sayısı
1	lokus 1	lokus 2	0.0086	0.0349	0.64347±0.00468	0.0976	0.7548	10100
2	lokus 1	lokus 3	0.0172	0.1209	0.49139±0.00474	0.4378	0.5082	10100
3	lokus 2	lokus 3	0.1828	1.0000	0.00000±0.00000	49.5238	0.0000	10100
4	lokus 1	lokus 4	0.0102	0.2889	0.36584±0.00427	0.5723	0.4494	10100
5	lokus 2	lokus 4	0.0352	1.0000	0.00000±0.00000	6.8571	0.0088	10100
6	lokus 3	lokus 4	0.0172	0.4074	0.16188±0.00339	1.8386	0.1751	10100
7	lokus 1	lokus 5	0.0266	0.1619	0.35248±0.00488	0.9786	0.3225	10100
8	lokus 2	lokus 5	0.1984	0.9407	0.00000±0.00000	54.6167	0.0000	10100
9	lokus 3	lokus 5	0.1781	0.8769	0.00000±0.00000	49.3675	0.0000	10100
10	lokus 4	lokus 5	0.0391	1.0000	0.00634±0.00088	8.8889	0.0029	10100
11	lokus 1	lokus 6	0.0961	0.7193	0.00040±0.00019	16.5759	0.0000	10100
12	lokus 2	lokus 6	0.0414	0.3985	0.11960±0.00284	3.0776	0.0794	10100
13	lokus 3	lokus 6	0.0647	0.8381	0.00396±0.00073	8.4264	0.0037	10100
14	lokus 4	lokus 6	0.0148	1.0000	0.08465±0.00295	1.6612	0.1974	10100
15	lokus 5	lokus 6	0.0766	0.8596	0.00099±0.00030	11.0486	0.0009	10100
16	lokus 1	lokus 7	0.0242	0.4921	0.13743±0.00304	2.3881	0.1223	10100
17	lokus 2	lokus 7	0.0383	1.0000	0.00208±0.00049	5.9665	0.0146	10100
18	lokus 3	lokus 7	0.0284	1.0000	0.03535±0.00181	3.6936	0.0546	10100
19	lokus 4	lokus 7	0.0055	1.0000	0.37693±0.00528	0.5114	0.4745	10100
20	lokus 5	lokus 7	0.0328	1.0000	0.01000±0.00114	4.6027	0.0319	10100
21	lokus 6	lokus 7	0.0042	0.0632	0.66891±0.00447	0.0985	0.7537	10100

Tablo 4.2. Hb D-Los Angeles popülasyonu için Linkage disequilibrium (*LD*) anlamlılık diyagramı; Arlequin sonuçları (*Hipotez sınırı = 0.0500, P < 0.05; (+), P > 0.05; (-)*).

	lokus 1	lokus 2	lokus 3	lokus 4	lokus 5	lokus 6	lokus 7
lokus 1	*	-	-	-	-	+	-
lokus 2	-	*	+	+	+	-	+
lokus 3	-	+	*	-	+	+	+
lokus 4	-	+	-	*	+	-	-
lokus 5	-	+	+	+	*	+	+
lokus 6	+	-	+	-	+	*	-
lokus 7	-	+	+	-	+	-	*

Tablo 4.3. Hb D-Los Angeles popülasyonu için Linkage disequilibrium (*LD*) anlamlılık diyagramı; Power Marker sonuçları.

	Marker1	Marker2	Karşılıklı Bilgi	Çok-Allelik D	Çok-Allelik D'	Ki-kare df	Ki-kare değeri	Ki-kare p-değeri
1	lokus 1	lokus 2	0.7056	0.2205	0.8959	1	64.2045	0.0000
2	lokus 2	lokus 3	0.5794	0.1828	1.0000	1	49.5238	0.0000
3	lokus 3	lokus 4	0.0060	0.0078	0.3846	1	0.3799	0.5377
4	lokus 4	lokus 5	0.1448	0.0391	1.0000	1	8.8889	0.0029
5	lokus 5	lokus 6	0.1703	0.0938	0.6318	1	16.5775	0.0000
6	lokus 6	lokus 7	0.0592	0.0208	1.0000	1	2.3894	0.1222

Tablo 4.4. Normal popülasyonu için İkili Bağlantı Dengesizliği (*Pairwise Linkage Disequilibrium*); Arlequin sonuçları.

	Marker1	Marker2	Çok-Allelik D	Çok-Allelik D'	Exact p değeri	Ki-kare değeri	Ki-kare p-değeri	Tekrar sayısı
1	lokus 1	lokus 2	0.0122	0.0545	0.582±0.004	0.2855	0.5931	10100
2	lokus 1	lokus 3	0.0006	0.0067	1.000±0.000	0.0012	0.9721	10100
3	lokus 2	lokus 3	0.0776	1.0000	0.000±0.000	21.9724	0.0000	10100
4	lokus 1	lokus 4	0.0022	0.0120	1.000±0.000	0.0102	0.9196	10100
5	lokus 2	lokus 4	0.1637	1.0000	0.000±0.000	57.9828	0.0000	10100
6	lokus 3	lokus 4	0.0017	0.0167	0.785±0.004	0.0124	0.9113	10100
7	lokus 1	lokus 5	0.0142	0.0680	0.557±0.005	0.3882	0.5332	10100
8	lokus 2	lokus 5	0.2159	0.8946	0.000±0.000	88.2472	0.0000	10100
9	lokus 3	lokus 5	0.0717	0.8943	0.000±0.000	18.8058	0.0000	10100
10	lokus 4	lokus 5	0.1692	1.0000	0.000±0.000	62.0554	0.0000	10100
11	lokus 1	lokus 6	0.0801	0.5124	0.000±0.000	15.2662	0.0001	10100
12	lokus 2	lokus 6	0.0358	0.2602	0.069±0.002	2.9980	0.0794	10100
13	lokus 3	lokus 6	0.0257	0.6027	0.088±0.003	2.9953	0.0835	10100
14	lokus 4	lokus 6	0.0562	0.6236	0.005±0.001	8.4623	0.0036	10100
15	lokus 5	lokus 6	0.0734	0.5530	0.001±0.000	12.6553	0.0004	10100
16	lokus 1	lokus 7	0.0303	0.1827	0.141±0.003	2.1077	0.1466	10100
17	lokus 2	lokus 7	0.0695	0.4768	0.000±0.000	10.9372	0.0009	10100
18	lokus 3	lokus 7	0.0368	0.8127	0.015±0.001	5.9157	0.0150	10100
19	lokus 4	lokus 7	0.0447	0.4677	0.019±0.001	5.1694	0.0230	10100
20	lokus 5	lokus 7	0.0730	0.5184	0.000±0.000	12.0766	0.0005	10100
21	lokus 6	lokus 7	0.0611	0.3107	0.002±0.000	10.4873	0.0012	10100

Tablo 4.5. Normal popülasyonu için Linkage disequilibrium (*LD*) anlamlılık diyagramı; Arlequin sonuçları (Hipotez sınırı = 0.0500, $P < 0.05$; (+), $P > 0.05$; (-)).

	lokus 1	lokus 2	lokus 3	lokus 4	lokus 5	lokus 6	lokus 7
lokus 1	*	-	-	-	-	+	-
lokus 2	-	*	+	+	+	-	+
lokus 3	-	+	*	-	+	-	+
lokus 4	-	+	-	*	+	+	+
lokus 5	-	+	+	+	*	+	+
lokus 6	+	-	-	+	+	*	+
lokus 7	-	+	+	+	+	+	*

Tablo 4.6. Normal popülasyon için Linkage disequilibrium (*LD*) anlamlılık diyagramı; Power Marker sonuçları.

	Marker1	Marker2	Karşılıklı Bilgi	Çok- Allelik D	Çok- Allelik D'	Ki-kare df	Ki-kare değeri	Ki-kare p-değeri
1	lokus 1	lokus 2	0.6932	0.2195	0.9610	1.0000	91.9150	0.0000
2	lokus 2	lokus 3	0.2190	0.0776	1.0000	1.0000	21.9724	0.0000
3	lokus 3	lokus 4	0.1244	0.0491	1.0000	1.0000	10.0890	0.0015
4	lokus 4	lokus 5	0.5005	0.1692	1.0000	1.0000	62.0554	0.0000
5	lokus 5	lokus 6	0.0162	0.0323	0.2195	1.0000	2.4446	0.1179
6	lokus 6	lokus 7	0.0415	0.0437	0.5262	1.0000	5.3496	0.0207

Tablo 4.7. Hb D-Los Angeles popülasyonu haplotip çeşitliliği; Arlequin sonuçları.

	Haplotip	Frekans	s.d.
1	+ - - - - + +	0.346909	0.057740
2	- + + - + + +	0.298632	0.054972
3	+ - - - - +	0.125000	0.038259
4	+ - - - - + -	0.065592	0.031056
5	- + - - + + +	0.025000	0.017043

Maksimum-olasılık haplotip frekansları Arlequin 3,5 yazılımı ile oluşturulmuştur.

Listelenen 14 frekansın toplamı: 1.000000 / Numunedeki gen kopya sayısı: 80 / S.d. : Standart sapma

Tablo 4.8. Hb D-Los Angeles popülasyonu haplotip çeşitliliği; Power Marker sonuçları.

	Haplotip	Ölçüm	Ortalama	s.d.	2,5% l.b.	97,5% u.b.
1	+-----++	0.34694	0.40286	0.08993	0.23744	0.53315
2	-++-+++	0.29860	0.23097	0.09321	0.04159	0.36124
3	+-----+	0.12496	0.08814	0.04949	0.00000	0.17500
4	+-----+-	0.06559	0.05159	0.03189	0.00000	0.11027
5	-+---+++	0.02500	0.02762	0.02811	0.00000	0.10000
6	-+-+---+	0.02424	0.02416	0.01929	0.00000	0.06250
7	-+-+---	0.02191	0.02202	0.01669	0.00000	0.06070
8	-+-+---+	0.01635	0.01408	0.01887	0.00000	0.05785
9	-++-++	0.01387	0.01444	0.01367	0.00000	0.04093
10	-----+--	0.01311	0.01325	0.01345	0.00000	0.04877
11	++----++	0.01250	0.01313	0.01310	0.00000	0.03750
12	+++----+	0.01250	0.01225	0.01317	0.00000	0.03750
13	-+-----+	0.01250	0.01100	0.01209	0.00000	0.03750
14	-----++	0.01189	0.00881	0.01046	0.00000	0.03643
15	-++-+-+	0.00004	0.04229	0.05993	0.00000	0.18518
16	-++-+-+	0.00000	0.01439	0.02254	0.00000	0.06702
17	-----+++	0.00000	0.00000	0.00001	0.00000	0.00000
18	-++-+-+	0.00000	0.00037	0.00278	0.00000	0.00000
19	-++-+-+	0.00000	0.00112	0.00796	0.00000	0.00000
20	--+-----+	0.00000	0.00012	0.00125	0.00000	0.00000
21	+++--++	0.00000	0.00140	0.00510	0.00000	0.01273
22	-+-----++	0.00000	0.00150	0.00541	0.00000	0.02500
23	-----++	0.00000	0.00171	0.00628	0.00000	0.02392
24	-+-+---++	0.00000	0.00175	0.00641	0.00000	0.02500
25	-+-+---++	0.00000	0.00100	0.00423	0.00000	0.01250

Listelenen 25 frekansın toplamı: 1.000000 / Numunedeki gen kopya sayısı: 80 / S.d. : Standart sapma

Tablo 4.9. Normal popülasyon haplotip çeşitliliği; Arlequin sonuçları.

	Haplotip	Frekans	s.d.
1	+-----++	0.266653	0.044010
2	-++-+++	0.146688	0.036171
3	+-----+-	0.125874	0.034472
4	-++-+++	0.062640	0.024809
5	+-----	0.054387	0.027891

Maksimum-olasılık haplotip frekansları Arlequin 3,5 yazılımı ile oluşturulmuştur.

Listelenen 14 frekansın toplamı: 1.000000 / Numunedeki gen kopya sayısı: 118 / S.d. : Standart sapma

Tablo 4.10. Normal popülasyon haplotip çeşitliliği; Power Marker sonuçları.

	Haplotip	Ölçüm	Ortalama	s.d.	2,5% l.b.	97,5% u.b.
1	+ - - - - ++	0.24751	0.25235	0.07021	0.11564	0.36559
2	- + - + + + +	0.16464	0.17316	0.05267	0.06573	0.27095
3	+ - - - - + -	0.12952	0.11234	0.04357	0.04187	0.19538
4	+ - - - - - +	0.08545	0.07777	0.04553	0.00000	0.16872
5	- + + - + + +	0.07326	0.06367	0.03906	0.00000	0.14358
6	- + - + + - -	0.05660	0.04019	0.03231	0.00000	0.09897
7	- + - + + - +	0.03027	0.02988	0.02947	0.00000	0.08635
8	- + + - + - +	0.02983	0.03780	0.03160	0.00000	0.09926
9	- + - + + + -	0.02473	0.02864	0.02670	0.00000	0.08673
10	- + + + + - -	0.01799	0.01520	0.01390	0.00000	0.04689
11	- + + - - - +	0.01732	0.01470	0.01376	0.00000	0.05085
12	+ - - - + + -	0.01446	0.01505	0.01131	0.00000	0.03674
13	+ - - - + + +	0.01150	0.01081	0.01053	0.00000	0.03214
14	+ - - - - - -	0.01111	0.02546	0.03101	0.00000	0.10318
15	+ + - + + + -	0.01065	0.01133	0.01256	0.00000	0.04237
16	+ + - - + + +	0.00951	0.00698	0.00907	0.00000	0.02692
17	- - - - - + -	0.00892	0.00926	0.00888	0.00000	0.02755
18	+ + - + + - +	0.00865	0.01035	0.00896	0.00000	0.02755
19	- + + - - + -	0.00865	0.00407	0.00706	0.00000	0.01929
20	+ + - + + + +	0.00850	0.00919	0.00989	0.00000	0.02601
21	- + - - + + +	0.00847	0.00933	0.00974	0.00000	0.03026
22	+ + - - - - +	0.00847	0.00648	0.00859	0.00000	0.02542
23	- + - - - - -	0.00847	0.00564	0.00891	0.00000	0.02542
24	+ + + - + - -	0.00550	0.00564	0.00786	0.00000	0.02418
25	- + - - - + -	0.00000	0.00444	0.00895	0.00000	0.03385
26	- + + - - - -	0.00000	0.00263	0.00545	0.00000	0.01695
27	+ - + - - - -	0.00000	0.00127	0.00543	0.00000	0.00847
28	- + + - + + -	0.00000	0.00442	0.01061	0.00000	0.03310
29	- + + + + - +	0.00000	0.00094	0.00550	0.00000	0.00067
30	- + + + + + -	0.00000	0.00113	0.00538	0.00000	0.01628
31	+ + - + + - -	0.00000	0.00015	0.00096	0.00000	0.00051
32	- + + - - + +	0.00000	0.00276	0.00740	0.00000	0.02227
33	- + - + - - +	0.00000	0.00018	0.00178	0.00000	0.00000
34	- + - - + + -	0.00000	0.00008	0.00082	0.00000	0.00000
35	+ - - + + - +	0.00000	0.00018	0.00124	0.00000	0.00000
36	- + - + - + +	0.00000	0.00017	0.00169	0.00000	0.00000
37	- + - + - - -	0.00000	0.00009	0.00089	0.00000	0.00000
38	- + - - - - +	0.00000	0.00008	0.00085	0.00000	0.00000
39	+ + - - + - +	0.00000	0.00309	0.00643	0.00000	0.02542
40	+ + - - - + +	0.00000	0.00309	0.00643	0.00000	0.02542

Listelenen 40 frekansın toplamı: 1.000000 / Numunedeki gen kopya sayısı: 118/ S.d. : Standart sapma Maksimum-olasılık haplotip frekansları Power Marker 3, 25 yazılımı ile oluşturulmuştur.

Tablo 4.11. Popülasyonlar arasında hesaplanan (AMOVA) F-istatistiği; Arlequin sonuçları.

			Uzaklık Metodu: Çift farklılığı		
Lokus	Değişim kaynağı	d.f.	Kareler toplamı	Varyans bileşenleri	Yüzde
Toplam	Popülasyonlar arasında	1	7.411	0.06333 Va	4.27
Toplam	Popülasyonlar içindeki bireyler arasında	97	133.144	-0.04854Vb	-3.27
Toplam	Nüfus içinde	99	145.500	1.46970 Vc	99.0
Toplam	Genel Toplam	197	286.056	1.48449	100

Fiksasyon Endeksleri: Anlamlılık testleri (1023 permütasyon)

F_{IS} : -0.03415	p-değeri = 0.75464 ± 0.01456
F_{ST} : 0.04266	p-değeri = 0.00391 ± 0.00185
F_{IT} : 0.00996	p-değeri = 0.61975 ± 0.01361

Haplotip Frekanslarına Dayalı Örnek Farklılaştırmanın Kesin Testi

Popülasyonlar arasındaki farklılık küresel testi, anlamlı $P \leq 0.05$

Farklılaştırma: Tam P değeri = 0.00003 ± 0.000001 (30000 Markov tekrar sayısı)

D.f. : serbestlik derecesi / F_{IS} : Akralılık Alt Popülasyon / F_{ST} : Alt Popülasyon Toplam

F_{IT} : Akralılık Toplam

Arlequin 3.5 yazılımı kullanılarak gerçekleştirilen moleküler varyans analizi (AMOVA)

Tablo 4.12. Popülasyonlar arasında hesaplanan (AMOVA) F-istatistiği; Power Marker sonuçları.

Lokus	Değişim kaynağı	Kareler toplamı	Yüzde
Tüm	Popülasyonlar arasında	14.8345	0.0259
Tüm	Popülasyon içinde (Hb D-Los Angeles)	89.4750	0.1563
Tüm	Popülasyon içinde Normal)	176.0339	0.3076
Tüm	Bireyler İçinde (Hb D-Los Angeles)	113.0000	0.1974
Tüm	Bireyler İçinde (Normal)	179.0000	0.3127
Tüm	Toplam	572.3434	1.0000

Tablo 4.13. Popülasyonlar için Özel F-İstatistiği; Power Marker sonuçları.

Lokus	Beta (tüm)	Alpha (tüm)	f(tüm)	Beta (Hb D-Los Angeles)	Alpha (Hb D-Los Angeles)	f (Hb D-Los Angeles)	Beta (Normal)	Alpha (Normal)	f (Normal)	Beta (Hb D-Los Angeles)/Normal)
lokus 1	-0.0100	0.0597	0.0691	-0.0092	0.0388	0.0476	-0.0106	0.0739	0.0837	0.0011
lokus 2	-0.0036	0.0918	0.0950	0.0035	0.0509	0.0476	-0.0084	0.1195	0.1268	0.0031
lokus 3	0.0226	-0.0028	-0.0259	-0.2952	-0.4599	-0.1272	0.2380	0.3072	0.0907	-0.1048
lokus 4	0.2383	0.3314	0.1223	0.7093	0.6936	-0.0541	-0.0810	0.0859	0.1544	0.1560
lokus 5	0.0182	-0.1726	-0.1944	0.0531	-0.2009	-0.2683	-0.0055	-0.1534	-0.1472	0.0119
lokus 6	0.0078	-0.2355	-0.2453	0.0689	-0.2104	-0.3000	-0.0336	-0.2525	-0.2118	0.0207
lokus 7	0.1764	0.1045	-0.0872	0.5788	0.5437	-0.0833	-0.0964	-0.1932	-0.0883	0.1338
Tüm	0.0589	0.0236	-0.0376	0.1526	0.0648	-0.1037	-0.0046	-0.0044	0.0002	0.0315

Tablo 4.14. İki popülasyon için molekül çeşitliliğin özeti; Arlequin sonuçları.

Pop.	n	Hap. no	k (95%CI)	θ_s	h	π	Tajima's D		Fu's Fs		Mismatch dağılımı		
							D	P	F _s	P	τ (95%CI)	θ_0	θ_1
Normal	59	24	3.03 (2.19- 4.55)	1.30	0.93 ± 0.02	0.43 ± 1.76	3.00	0.99	-16.88	0.00	3.46 (4.71-2.07)	0.01	25.8 2
Hb-D Los Angeles	40	14	2.56 (1.87- 4.25)	1.41	0.87 ± 0.02	0.36 ± 1.54	1.97	0.95	-6.37	0.00	3.16 (5.08-1.58)	0.00	12.6 4

Her popülasyon için bireylerin sayısı (n), haplotip sayısı, bireyler arasındaki ortalama çift farklılığı (k), ayırım konumlarının sayısı (S), haplotip çeşitliliği ($h \pm$ standart sapma), nükleotid çeşitliliği ($\pi \pm$ standart sapma). Tajima'nın D ve Fu'nun FS, karşılık gelen P-değeri ve her bir popülasyon için uyumsuzluk dağılım parametresi tahminleri. D Tajima'nın D tahmin nüfus genişlemesi, Fs Fu'nun Fs tahmini nüfus genişlemesi. τ , θ_0 ve θ_1 için değerler, sırasıyla, mutasyon zamanının birimleri cinsinden ifade edilen, genişlemenin yaşı, genişlemeden önceki nüfus büyüklüğü ve genişlemeden sonraki nüfus büyüklüğüdür. Anlamlı; $P \leq 0.05$.

Tablo 4.15. Mismatch dağılımı test istatistiklerinin SSD ve rg değerleri; Arlequin sonuçları.

Uygunluk testleri				
Popülasyonlar	SSD	SSD-P değeri	rg	P değeri
Normal pop.	0.0035	0.29	0.0227	0.54
Hb-D pop.	0.0030	0.57	0.0213	0.80

P (SSD), τ , θ_0 ve θ_1 parametreleriyle tanımlanan popülasyonun demografik tarihi için gözlenen ve uyumsuzluk dağılımı arasında şans eseri daha az uyum sağlama ihtimalidir.

SSD: kare sapmaların toplamı / rg: Harpending'in düzensizliği

Tablo 4.16. Normal popülasyondaki tüm lokuslar için Hardy-Weinberg Eşitliği (HWE) testi; Arlequin sonuçları.

Lokus	Genotip sayısı	Göz. Het.	Bek. Het.	P-değeri	Standart Sapma	Tekrar Sayısı
1	59	0.44068	0.49718	0.43522	0.00053	1001000
2	59	0.44068	0.50413	0.43454	0.00053	1001000
3	59	0.23729	0.26076	0.60500	0.00049	1001000
4	59	0.37288	0.44039	0.24873	0.00043	1001000
5	59	0.57627	0.50297	0.30386	0.00046	1001000
6	59	0.49153	0.40635	0.19169	0.00040	1001000
7	59	0.45763	0.42083	0.54780	0.00048	1001000

Her nüfus içindeki her lokus için HWE için yapılan testler, Fisher'in kesin testine benzer bir HWE testi kullanıldı. P değerleri, Arlequin 3.5 yazılımı kullanılarak elde edildi.
Göz.Het. : Gözlemlenen heterozigotluk / Bek.Het. : Beklenen heterozigotluk

Tablo 4.17. Normal popülasyondaki tüm lokuslar için Hardy-Weinberg Eşitliği (HWE) testi; Power Marker sonuçları.

Lokus	ki-kare değeri	ki-kare d.f.	ki-kare p-değeri	LRT değeri	LRT d.f.	LRT p-değeri	exact p-değeri
1	0.3335	1	0.5636	0.3335	1	0.5636	0.4378
2	0.8270	1	0.3632	0.8289	1	0.3626	0.3052
3	0.3989	1	0.5277	0.3660	1	0.5452	0.6044
4	1.2586	1	0.2619	1.2330	1	0.2668	0.1303
5	1.4272	1	0.2322	1.4339	1	0.2311	0.2944
6	2.8547	1	0.0911	3.1886	1	0.0742	0.1059
7	0.5520	1	0.4575	0.5693	1	0.4505	0.3483

Her nüfus içindeki her lokus için HWE için yapılan testler, Fisher'in kesin testine benzer bir HWE testi kullanıldı. P değerleri, Power Marker 3,25 yazılımı kullanılarak elde edildi.
LRT: Olasılık oranı testi

Tablo 4.18. Hb-D Los Angeles pöüasyonundaki tüm lokuslar için Hardy-Weinberg Eşitliđi (HWE) testi; Arlequin sonuçları.

Lokus	Genotip sayısı	Göz. Het.	Bek. Het.	P-deđeri	Standart Sapma	Tekrar Sayısı
1	40	0.47500	0.49842	1.00000	0.00001	1001000
2	40	0.47500	0.49842	1.00000	0.00001	1001000
3	40	0.50000	0.44430	0.48953	0.00049	1001000
4	40	0.12500	0.11867	1.00000	0.00001	1001000
5	40	0.60000	0.47468	0.10553	0.00030	1001000
6	40	0.47500	0.36677	0.08208	0.00027	1001000
7	40	0.17500	0.16171	1.00000	0.00001	1001000

Her nüfus içindeki her lokus için HWE için yapılan testler, Fisher'in kesin testine benzer bir HWE testi kullanıldı. P deđerleri, Arlequin 3.5 yazılımı kullanılarak elde edildi.

Göz.Het .: Gözlemlenen heterozigotluk / Bek.Het .: Beklenen heterozigotluk

Tablo 4.19. Hb-D Los Angeles pöüasyonundaki tüm lokuslar için Hardy-Weinberg Eşitliđi (HWE) testi; Power marker sonuçları.

Lokus	ki-kare deđer	ki-kare d.f.	ki-kare p-deđer	LRT deđer	LRT d.f.	LRT p-deđer	exact p-deđer
1	0.0488	1	0.8252	0.0488	1	0.8252	1.0000
2	0.0488	1	0.8252	0.0488	1	0.8252	1.0000
3	0.7795	1	0.3773	0.8069	1	0.3690	0.4860
4	0.1778	1	0.6733	0.3336	1	0.5636	1.0000
5	3.1360	1	0.0766	3.2956	1	0.0695	0.0940
6	3.8807	1	0.0488	6.0176	1	0.0142	0.0750
7	0.3678	1	0.5442	0.6723	1	0.4123	1.0000

Her nüfus içindeki her lokus için HWE için yapılan testler, Fisher'in exact testine benzer bir HWE testi kullanıldı. P deđerleri, Power Marker 3,25 yazılımı kullanılarak elde edildi.

LRT: Olasılık oranı testi

Tablo 4.20. Her bir lokus için normal popülasyona ait allel frekans hesaplamaları; Power Marker sonuçları.

İşaretleyici	Allel	Sayı	Frekans	Varyans	s.d.	2,5% l.b.	97,5% u.b.
lokus 1	-	53	0.4492	0.00225437	0.0475	0.3729	0.5339
lokus 1	+	65	0.5508	0.00225437	0.0475	0.4661	0.6271
lokus 2	-	60	0.5085	0.00236879	0.0487	0.4068	0.6017
lokus 2	+	58	0.4915	0.00236879	0.0487	0.3898	0.5847
lokus 3	-	100	0.8475	0.00118561	0.0344	0.7712	0.9068
lokus 3	+	18	0.1525	0.00118561	0.0344	0.0932	0.2203
lokus 4	-	80	0.6780	0.00212047	0.0460	0.5847	0.7458
lokus 4	+	38	0.3220	0.00212047	0.0460	0.2458	0.4068
lokus 5	-	62	0.5254	0.00178451	0.0422	0.4322	0.5932
lokus 5	+	56	0.4746	0.00178451	0.0422	0.3983	0.5678
lokus 6	-	33	0.2797	0.00133168	0.0365	0.2203	0.3475
lokus 6	+	85	0.7203	0.00133168	0.0365	0.6356	0.7797
lokus 7	-	35	0.2966	0.00159705	0.0400	0.2119	0.3559
lokus 7	+	83	0.7034	0.00159705	0.0400	0.6441	0.7797

s.d. : Standart sapma / Power Marker 3,25 yazılımı kullanılarak veriler elde edilmiştir

Tablo 4.21. Her bir lokus için Hb D-Los Angeles popülasyonuna ait allel frekans hesaplamaları; Power Marker sonuçları.

İşaretleyici	Allel	Sayı	Frekans	Varyans	s.d.	2,5% l.b.	97,5% u.b.
lokus 1	-	35	0.4375	0.00318359	0.0564	0.3125	0.5250
lokus 1	+	45	0.5625	0.00318359	0.0564	0.4625	0.6750
lokus 2	-	45	0.5625	0.00318359	0.0564	0.4500	0.6625
lokus 2	+	35	0.4375	0.00318359	0.0564	0.3375	0.5500
lokus 3	-	54	0.6750	0.00235938	0.0486	0.6000	0.7625
lokus 3	+	26	0.3250	0.00235938	0.0486	0.2250	0.3875
lokus 4	-	75	0.9375	0.00068359	0.0261	0.8875	0.9875
lokus 4	+	5	0.0625	0.00068359	0.0261	0.0125	0.1125
lokus 5	-	50	0.6250	0.00210938	0.0459	0.5250	0.7000
lokus 5	+	30	0.3750	0.00210938	0.0459	0.2875	0.4625
lokus 6	-	19	0.2375	0.00155859	0.0395	0.1500	0.3125
lokus 6	+	61	0.7625	0.00155859	0.0395	0.6750	0.8375
lokus 7	-	7	0.0875	0.00090234	0.0300	0.0375	0.1625
lokus 7	+	73	0.9125	0.00090234	0.0300	0.8375	0.9625

Tablo 4.22. Her bir lokus için normal popülasyona ait allel çifti frekans hesaplamaları; Power Marker sonuçları.

İşaretleyici	Allel	Normal
lokus 1	--	0.2203
lokus 1	+ -	0.4576
lokus 1	++	0.3220
lokus 2	--	0.2881
lokus 2	+ -	0.4407
lokus 2	++	0.2712
lokus 3	--	0.7288
lokus 3	+ -	0.2373
lokus 3	++	0.0339
lokus 4	--	0.4915
lokus 4	+ -	0.3729
lokus 4	++	0.1356
lokus 5	--	0.2373
lokus 5	+ -	0.5763
lokus 5	++	0.1864
lokus 6	--	0.0339
lokus 6	+ -	0.4915
lokus 6	++	0.4746
lokus 7	--	0.0678
lokus 7	+ -	0.4576
lokus 7	++	0.4746

Tablo 4.23. Her bir lokus için Hb D-Los Angeles popülasyonuna ait allel çifti frekans hesaplamaları; Power Marker sonuçları.

İşaretleyici	Allel	Hb D-Los Angeles
lokus 1	--	0.2000
lokus 1	+ -	0.4750
lokus 1	++	0.3250
lokus 2	--	0.3250
lokus 2	+ -	0.4750
lokus 2	++	0.2000
lokus 3	--	0.4250
lokus 3	+ -	0.5000
lokus 3	++	0.0750
lokus 4	--	0.8750
lokus 4	+ -	0.1250
lokus 5	--	0.3250
lokus 5	+ -	0.6000
lokus 5	++	0.0750
lokus 6	+ -	0.4750
lokus 6	++	0.5250
lokus 7	+ -	0.1750
lokus 7	++	0.8250

Tablo 4.24. Popülasyonlar için Haplotip Eğilim Regresyonu; Power Marker sonuçları.

Pencere	Haplotip	Frekans	Ortalama
lokus 1- lokus 2	++	0.0410	0.7477
lokus 1- lokus 2	+ -	0.5145	0.5784
lokus 1- lokus 2	- +	0.4287	0.6118
lokus 1- lokus 2	--	0.0158	0.3436
lokus 2- lokus 3	++	0.2222	0.4091
lokus 2- lokus 3	+ -	0.2475	0.8163
lokus 2- lokus 3	--	0.5303	0.5714
lokus3- lokus 4	+ -	0.2222	0.4091
lokus3- lokus 4	- +	0.2172	0.8837
lokus3- lokus 4	--	0.5606	0.5586
lokus 4- lokus 5	++	0.2172	0.8837
lokus 4- lokus 5	- +	0.2172	0.4186
lokus 4- lokus 5	--	0.5657	0.5536
lokus 5- lokus 6	++	0.2582	0.6647
lokus 5- lokus 6	+ -	0.1762	0.6313
lokus 5- lokus 6	- +	0.4792	0.5377
lokus 5- lokus 6	--	0.0865	0.6413
lokus 6- lokus 7	++	0.5594	0.5066
lokus 6- lokus 7	+ -	0.1780	0.8196
lokus 6- lokus 7	- +	0.2285	0.5943
lokus 6- lokus 7	--	0.0341	0.9054

Tablo 4.25. Popülasyonlar için Farklılık Testi; Power Marker sonuçları.

İşaretleyici	P-değeri
lokus 1	0.7745
lokus 2	0.4071
lokus3	0.0027
lokus 4	0.0000
lokus 5	0.1817
lokus 6	0.5137
lokus 7	0.0001
Toplam	0.0000

5. TARTIŞMA

Gensel çeşitliliklerin tanımlanması, çeşitliliklere neden olan parametrelerin belirlenmesi ve işlerlerinin ortaya çıkarılmasına yönelik araştırmalar son 50 yılda birçok gensel araştırmanın temelini oluşturmuştur. Üzerinde çalışılan genom bölgesinde yer alan polimorfizm odaklarına dayalı yüzde allel sıklıklarının belirlenmesi ve değerlendirilmesi geçmişe dönük gensel araştırmalar için oldukça değerli sonuçlar üretmektedir. Örnek olarak, HW kanunu allel sıklığı parametresi kullanılarak, popülasyona özgü gen dağılımı çeşitliliğinin nesilden nesile kalıtılabileceğini hesaplamaktadır. İkinci olarak HW eşitliği, bir genotip sıklığının bilinmesi ile bağlantılı olan diğer genotip ve heterozigotluk sıklıklarını hesaplayabilmektedir. Popülasyonların gensel analizleri yapılırken, HW kanunu temelinde allel, homozigot ve heterozigot sıklıklarının hesaplanması, ilişki (*association*), bağlantı (*linkage*), ayırım (*segregation*) gibi tanımların yapılabilmesi, istatistik tabanlı yazılım modellerinin yardımı ile gerçekleştirilmektedir (20).

Bu kapsamda, popülasyonların haplotip analizi ile belirlenen allel sıklıklarının ve haplotip çeşitlerinin Arlequin ve Power Marker yazılımları kullanılarak karşılaştırmalı analizinin yapılması, sonuçlar açısından olası farklılıkların veya uyumlulukların tartışılması bu tez çalışmasının amacını oluşturmaktadır.

Tez çalışmamızda elde edilen bulgular incelendiğinde (Tablo 4.7 ve Tablo 4.8), Arlequin ve Power Marker yazılımları, her iki popülasyonda da mutasyon ile ilişkili haplotip sıklıklarını eşit frekans oranlarına sahip biçimde hesaplamıştır. Tablo 4.8’ de gözlemlendiği üzere Power Marker yazılımı oransal olarak daha düşük yüzdelere sahip çok sayıda (25 adet) haplotip çeşidini içeren sonuçlar sağlamaktadır. Arlequin yazılımında 14 adet haplotip çeşidi listelenmektedir. Bu durum bir üstünlük olarak değerlendirilebilir ve popülasyon içinde yer alan düşük sıklıktaki gensel çeşitliliğin belirlenmesini sağlamakta faydalı olacağını düşündürmektedir. Tablo 4.9 ve Tablo 4.10 göz önüne alındığında iki popülasyon arasında haplotip çeşitliliği sayısında ve sıralamasında fark olduğu saptanmıştır. Normal popülasyon Arlequin tablosunda (Tablo 4.9) 4. ve 5. sıralarda yer alan haplotip çeşitleri, Power Marker tablosunda (Tablo 4.10) daha düşük sıklıklarda yer almaktadır. Arlequin yazılımı kendi içinde olasılığı düşük ve odak çeşitliliği benzer olan haplotipleri birleştirdiği için söz konusu farklılıklar ortaya çıkmaktadır. Araştırmanın amacı popülasyonda yer alan düşük frekanslı haplotipleri

belirlemek olduğunda, Arlequin yazılımının gerçekleştirdiği bu hesaplamanın dezavantajlı olduğu düşünülebilmektedir.

Arlequin mismatch dağılımı testi sonuçlarına göre popülasyonlara ait bulgular Tablo 4.13 ve 4.14' de sunulmaktadır. Popülasyonun yaşını hesaplamak için kullanılan ve mutasyonel zaman birimini ifade eden Tau değeri, θ_0 ve θ_1 değerleri iki popülasyon için Arlequin tarafından hesaplanmıştır. Bulgulara göre, popülasyonlar için rg indekslerini düşük değerlerde ve SSD p-değerlerini > 0.05 olarak belirlenmiştir. Bu sonuçlar her iki popülasyonun nüfussal olarak ortak geçmişe sahip olduğunu göstermektedir (28). Ek olarak elde edilen verilerin unimodal dağılım göstermesi, nüfus artışlarının her iki popülasyon için de düzgün ve herhangi bir nüfus azalışı veya göç etkisinde kalmadan değiştiğini göstermektedir. Arlequin yazılımından elde edilen sonuçlar her iki popülasyon için SSD ve rg indeksi p değerini, $p > 0.05$ olarak göstermiş ve dağılımları unimodal olarak belirlenmiştir (Şekil 4.1, Şekil 4.2). Arlequin tarafından sunulan tüm bu nüfussal gelişme ve popülasyon yaşı hesaplamaları, Power Marker yazılımı tarafından hesaplanamamaktadır. Moleküler çeşitlilik (Tablo 4.13) ve Mismatch dağılım parametreleri (Tablo 4.14) dikkate alındığında Arlequin yazılımı popülasyonların tarihsel gelişim süreçlerinin belirlenmesinde önemli avantaj sağlayabilmektedir.

Her iki yazılımdan elde edilen haplotip analizi sonuçlarına göre, iki popülasyonda da odakların Hardy-Weinberg dengesinde ($p > 0.05$) olduğu saptanmıştır (bkz. Tablo 4.16, 4.17, 4.18 ve 4.19). Benzer olarak Hb D Los Angeles popülasyonu 6. odak p değerleri hem Arlequin hem de Power Marker sonuçlarında sınır değerlere yaklaşmaktadır (Tablo 4.18 ve 4.19). Power Marker sonuçlarında p değerindeki bu azalmanın daha çok sayıda haplotip çeşidi sunmasından kaynaklandığı düşünülmektedir. HWE ve allel çeşitlenmesi arasındaki ilişki incelenen toplumdaki kuşaklar arasındaki allel dağılımının belirlenmesi açısından önemlidir. Bir allelin doğal seçim, akraba evlilikleri veya popülasyon dışından göç gibi baskıların etkisinde kalması bir sonraki kuşakta görülme sıklığını artırır. Ancak popülasyon bu etkilerden bağımsız gelişir ise sonraki kuşaklarda allel dağılımları rastgele düzenlenir. Sonuç olarak, etki altında kalmadan gelişen popülasyonlar için Hardy-Weinberg dengesinde olduğu ifade edilmektedir. Ölçülen ile beklenen heterozigotluk değerleri arasında anlamlı bir farkın olmaması “sıfır hipotezi” (H_0) olarak ifade edilmektedir. Arlequin yazılımında sıfır hipotezi, bu varsayımla kurulmuştur. Power Marker yazılımının

Hardy-Weinberg denkliđi sonuçlarını birden fazla istatistiksel testle karşılaştırması arařtırıcılar ađısından bir avantaj sunmaktadır.

Çalıřmamızda her iki popülasyonun haplotip analizi sonuçlarına, Arlequin ve Power Marker yazılımları kullanılarak bađlantı dengesizliđi (*Linkage Disequilibrium, (LD)*) testleri uygulanmıřtır (bkz. Tablo 4.1, 4.2, 4.3, 4.4, 4.5 ve 4.6). LD analizleri, komřu genler ve üzerlerindeki polimorfik odakların sonraki nesiller boyunca bir arada kalıtılacađı prensibine dayalı olarak uygulanmaktadır. Olası bađlantılar nedeni ile sonraki nesilde yeni haplotipler veya rekombinant ürünler ortaya çıkmamaktadır. Bu olaya bađlantı (*linkage*) adı verilir. Bađlantı dengesizliđi (LD) ise aynı kromozom üzerinde ya da farklı iki kromozom üzerinde yer alan iki ya da daha fazla gen bölgesindeki allellerin rastlantısal olmayan ortak kalıtılma özellikleridir. Arlequin yazılımı için, LD sonuçlarının istatistiksel anlamlılık testi, Ki-Kare test istatistiđi ile gerçeleřtirilmektedir. Ki-Kare testinin sıfır hipotezi “iki lokus arasında LD vardır” şeklinde ifade edilmektedir. İstatistiksel olarak $p < 0.05$ řartı sađlandığında H_0 hipotezi kabul edilir. Çalıřmamızdaki Arlequin yazılımı kullanılarak elde edilen bađlantı dengesizliđi sonuçlarına bakıldıđında, allel çiftleri arasında istatistiksel olarak anlamlı ($p < 0.05$, +) veya anlamsız ($p > 0.05$) olmak üzere LD belirlenmiřtir (bkz. Tablo 4.2 ve 4.5). Ek olarak Arlequin yazılımı LD sonuçları her bir lokusu komřu lokus ve komřu olmayan lokuslar ile karşılaştırarak sunmaktadır. Ancak Power Marker yazılımında, LD sonuçları, sadece komřu lokuslar arası bađlantıyı test etmektedir. LD test istatistiđi ađısından deđerlendirildiđinde Arlequin yazılımının hem kullandıđı test parametreleri hem de LD sonuçlarını kolay anlaşılır tablo biçiminde sunmasından dolayı daha kullanıcı dostu olduđu görölmektedir.

Arlequin yazılımı uygulanarak elde edilen molekölsele çeřitlilik analizi (AMOVA) sonuçlarına göre iki popülasyon arasında % 4.27 ($F_{ST} = 0.0427$, $P < 0.05$) gensele farklılık ve popülasyonlar içinde % 99 gensele çeřitlilik oranı saptanmıřtır (Tablo 4.11). $F_{ST} \leq 0.05$ olarak belirlenmesi, karşılaştırılan popülasyonların gensele olarak çok benzer ve aralarında yüksek gen akıřı parametresinin olduđunu göstermektedir. Bu sonuç akrabalık dıřı çiftleřen (*outcrossing*) popülasyonların ortak özelliđi olarak, popülasyon içi yüksek deđerde gensele çeřitlilik ve popülasyonlar arası düşük deđerde gensele farklılıđı açıklamaktadır (42). Popülasyonlar arası yüksek gen akıřı (*gene flow*) parametresi nedeni ile gensele farklılık azalmaktadır. Fiksasyon indeksleri (*F-istatistiđi*) bulguları dikkate alındıđında, genel F_{IS} deđerı -0.03415 olarak hesaplanmıřtır. F_{IS} deđerinin istatistiksel olarak anlamlı olmaması ($p > 0.05$), popülasyonların Hardy–

Weinberg dengesinde olduğunu kanıtlamaktadır (43). F_{IS} değerine göre popülasyonların HW dengesinde olması popülasyonlarda rastgele çiftleşme (*random mating*) olgusunun hakim olduğu sonucunu göstermektedir. Ek olarak F_{IT} ($F_{IT} = 0.00996$, $p > 0.05$) değerinin sıfıra yakın olarak hesaplanması ve istatistiksel olarak anlamlı olmaması, popülasyonların Hardy-Weinberg dengesinde olduğunu yani gensel sürüklenme olayının olmadığını ifade etmektedir. AMOVA test istatistiği Power Marker yazılımı sonuçları göz önüne alındığında (Tablo 4.12 ve Tablo 4.13), iki popülasyon arasında % 2.59 oranında gensel farklılık belirlenmiştir. Gözlenen % 1.68 oranındaki düşük fark yine Power Marker yazılımının haplotip çeşitliliği açısından Arlequin yazılımına göre daha fazla sayıda haplotip türü sunmasından kaynaklanmaktadır. Farklı olarak Power Marker yazılımı, her bir lokus için F istatistiği parametrelerini hesaplar (Tablo 4.13), F_{IS} ve F_{IT} gibi istatistiksel parametreleri hesaplamamaktadır.

Son olarak Tablo 4.20 - Tablo 4.25 arasındaki istatistiksel parametreler dikkate alındığında elde edilen sonuçların hepsinin Power Marker yazılımı ile üretildiği görülmektedir. Her bir lokus için her iki popülasyona ait allel frekans hesaplamaları, allel çifti frekans hesaplamaları, haplotip eğilim regresyonu, ve farklılık testi gibi parametreler Power Marker yazılımına özgü testler olarak sunulmaktadır. Bu test sonuçları incelendiğinde haplotip çeşidi odaklarının oluşumunda ayrıntılı sonuçlar elde edilebilmektedir. Her lokus için elde edilen bu sonuçlar Arlequin yazılımına göre oldukça ayrıntılı olarak sunulmaktadır.

6. SONUÇ VE ÖNERİLER

Tüm bu değerlendirmeler ışığında, karşılaştırdığımız iki farklı biyoinformatik analiz yazılımlarından elde edilen bulgular ile yazılımların birbirlerine göre üstünlük ve eksiklikleri sunulmaktadır. Buna göre Power Marker yazılımı, veri girişi açısından sağladığı kullanıcı dostu destekle Arlequin yazılımına üstünlük sağlamaktadır. Bununla beraber Power Marker özellikle haplotip çeşitliliği ve buna bağlı istatistiksel parametrelerde bariz bir şekilde daha verimli sonuçlar üretmektedir. Ancak Power Marker yazılımı, analiz listesinde bulunan bazı testleri gerçekleştirebilmek için birkaç diğer yazılımın desteğine ihtiyaç duymaktadır. Diğer yandan Arlequin yazılımı ise popülasyonun nüfusal büyümesi, popülasyon yaşı tahmini ve AMOVA gibi popülasyonlar arası farklılık testlerinde daha anlaşılır sonuçlar sunmaktadır. Ek olarak, Arlequin tüm testlerini kendi içinde gerçekleştirebilen tek yazılım paketi ile ücretsiz olarak internet ortamında sunulmaktadır. Kullanım kolaylığı açısından değerlendirildiğinde literatürde yer alan araştırmalarda çok sık kullanılması sonucunda, test istatistikleri çeşitlerinin ve örneklerinin bulunması nedeni ile Arlequin yazılımının daha tercih edilir olduğu düşünülmektedir. Çalışmamızın tartışma bölümünde ifade edildiği üzere, her iki programın içeriğinde yer alan ortak istatistiksel testlerin sonuçları uyumluluk göstermektedir. Özellikle haplotip çeşitliliği ve sıklıklarının ifade edildiği sonuçların her iki yazılımda da uyumlu olması güvenilirlik açısından yeterli olduklarını göstermektedir. Her bir odak için her iki popülasyona ait allel frekans hesaplamaları, allel çifti frekans hesaplamaları, haplotip eğilim regresyonu, ve farklılık testi gibi parametreleri kullanılarak bağlantı ağaç gösterimi sunabilen Power Marker yazılımı, bağlantısal genetik veri analizi yapacak olan araştırmacılar için tercih edilebilir olmasını sağlayabilmektedir. Örnek sayısı yüksek olan popülasyon çalışmaları için Excel ve word formatlarında veri girişini gerçekleştirebilen Power Marker yazılımı, araştırmacılar tarafından tercih edilebilir avantajlar taşımaktadır. Karşılaştırılan yazılımların erişime açık olan kullanıcı kılavuzları değerlendirildiğinde, Arlequin kılavuz içeriğinin, test istatistiklerini ayrıntılı açıklaması, içinde yer alan referans kaynakların bilimsel zenginliği, örnek veri girişi tabloları sunması ve kılavuzun güncellenme sıklığının yüksek olması gibi nedenlerle daha kolay ve anlaşılır olduğu düşünülmektedir.

Sonu olarak, tez alıřmamızda sunulan bulguların biyoinformatik ve gensel veri kullanarak yapılan alıřmalarda arařtırcıların uygun yazılım tercihleri yapabilmeleri iin deęerli katkılar saęlayabileceęi dūřünölmektedir.



KAYNAKLAR

1. Giardine B, Borg J, Higgs DR, Peterson KR, Philipson S, Maglott D, et al. Systematic documentation and analysis of human genetic variation in hemoglobinopathies using the microattribution approach. *Nat Genet* 2011, 20; 43(4): 295–301.
2. Benson DA, Karsch-Mizrachi I, Lipman DJ, Ostell J, Rapp BA, Wheeler DL. GenBank, *Nucleic Acids Res Jan* 2002, 1; 30(1): 17-20.
3. Kuonen D. Challenges in Bioinformatics for Statistical Data Miners, *Bulletin of the Swiss Statistical Society*, 2003; 46: 10-17.
4. Kumaresan V, Bhatt P, Palanisamy R, Gnanam AJ, Pasupuleti M, Arockiaraj J, A Murrel Cysteine Protease, Cathepsin L. Bioinformatics Characterization, Gene Expression and Proteolytic Activity, *Biologia* 2014; 69(3): 395-406.
5. Oppenheimer, S. Out-of-Africa, the peopling of continents and islands: tracing uniparental gene trees across the map. *Philos Trans R Soc* 2012 B 367, 770–84.
6. Ozturk O, Arıkan S, Atalay A, Atalay EO. Analysis of the population genetic structure of Hb D-Los Angeles [β 121 (GH4) Glu→Gln GAA→CAA] in Denizli, Turkey; genetic diversity, historical demography and estimation of the mutation rates based on haplotype variation. *Am J Hum Bio* 2016; 28(4): 476-83.
7. Öztürk O. Denizli yöresinde gözlenen Hb D-Los Angeles mutasyonunun beta globin gen ailesi haplotip analizi. Sağlık Bilimleri Enstitüsü, Biyofizik Anabilim Dalı, Yüksek lisans tezi, Denizli: Pamukkale Üniversitesi, 2007.
8. Oztürk O, Atalay A, Koseler A, Ozkan A, Koyuncu H, Bayram J, Demirtepe S, Aksoy K, Atalay EO. Beta globin gene cluster haplotypes of abnormal hemoglobins observed in Turkey. *Turk J Haematol* 2007; 24(4): 146-54.
9. Atalay EO, Atalay A, Ustel E, Yıldız S, Oztürk O, Koseler A, Bahadır A. Genetic origin of Hb D-Los Angeles according to beta globin gene cluster haplotypes. *Hemoglobin* 2007, 31: 387–91.
10. Excoffier L, Laval G, and Schneider S. An Integrated Software Package for Population Genetics Data Analysis, Arlequin ver 3.5.1.3 user manual 2011. 2006, <http://cmpg.unibe.ch/software/arlequin3>
11. Liu K. PowerMarker: New Genetic Data Analysis Software, Version 3.0. 2003, Free program distributed by the author over Internet at <http://www.powermarker.net>

12. Chen LZ, Eastal S, Board GP, and Kirk RL. Evolution of β -globin haplotypes in human populations. *Mol Biol Evol* 1990, 7(5): 423–37.
13. Currat M, Trabuchet G, Rees D, Perin P, Harding RM, Celegg JB, Langaney A, and Excoffier L. Molecular analysis of the β -globin gene cluster in the Niokholo Mandenka population reveals a recent origin of the β^s Senegal mutation. *J Hum Genet* 2002, 70: 207–23.
14. Itano HA. The hemoglobins. *Annu Rev Biochem* 1956, 25: 331–48.
15. Welch SG, Bateman C. Hb D-Neath or β 121 (GH4) Glu→Ala: a new member of the Hb D family. *Hemoglobin* 1993, 17(3): 255-9.
16. Jahnes E, Muller DJ, Schulze TG, Windemuth C, Cichon S, Ohlraun S, Fangerau H, Held T, Maier W, Propping P, Nothen MM, Rietschel M. Association study between two variants in the DOPA Decarboxylase gene in bipolar and unipolar affective disorder. *Am J Med Genet* 2002, 114: 519–22.
17. Salanti G, Amountza G, Ntzani EE, Ioannidis JP. Hardy–Weinberg equilibrium in genetic association studies: an empirical evaluation of reporting, deviations, and power. *Eur J HumGenet* 2005, 13: 840–8.
18. Zou GY, Donner A. The merit of testing Hardy–Weinberg equilibrium in the analysis of unmatched case control data: a cautionary note. *Ann Hum Genet* 2006, 70: 923–33.
19. Chatterjee N, Carroll RJ. Semiparametric maximum likelihood estimation exploiting gene–environment independence in case–control studies. *Biometrika* 2005, 92: 399.
20. Klug WS, Cummings MR. Hardy-Weinberg Kanunu, Genetik Kavramlar(Çeviri Editörü: Öner CA), Ankara, *Palme Yayıncılık* 2003, 686–92.
21. Excoffier L, Smouse P, and Quattro J. Analysis of molecular variance inferred from metric distances among DNA haplotypes: Application to human mitochondrial DNA restriction data. *Genetics* 1992, 131: 479–91.
22. Akarsu N. Haplotip bağlantı (Linkage) analizi. *Türk Farmakoloji Derneği Eğitim Sempozyumu*, Yüzüncü Yıl Üniversitesi Tıp Fakültesi, Van, 2006 20–3.
23. Ardlie KG, Kruglyak L, Seielstad M. Patterns of Linkage disequilibrium in the human genome. *Nat Rev Genet* 2002, 3(4): 299–309.
24. Bansal, A., et al. . *Bioinformatics for Geneticists*. Edited by Michael R. Barnes and Ian C. Gray , Chichester, West Sussex, United Kingdom: John Wiley & Sons Ltd, ISBNs: 0-470-84393-4 (HB); 0-470-84394-2 (PB), 2003, Chapter 11, 218-45.

25. Slatkin M, Hudson RR. Pairwise comparisons of mitochondrial DNA sequences in stable and exponentially growing populations. *Genetics* 1991, 129(2): 555–62.
26. Rogers AR, Harpending H. Population growth makes waves in the distribution of pairwise genetic differences. *Mol Biol Evol* 1992, 9(3): 552–69.
27. Schneider S, Excoffier L. Estimation of past demographic parameters from the distribution of pairwise differences when the mutation rates vary among sites: application to human mitochondrial DNA. *Genetics* 1999, 152(3): 1079–89.
28. Xiao Yongshuang, Yan Zhang, Takashi Yanagimoto, Jun Li, Zhizhong Xiao, Tianxiang Gao, Shihong Xu, Daoyuan Ma. Population genetic structure of the point-head flounder, *Cleisthenes herzensteini*, in the Northwestern Pacific. *Genetica* 2011, 139(2): 187–98.
29. Rogers AR. Genetic evidence for a Pleistocene population explosion. *Evolution* 1995, 49(4): 608–615.
30. Tajima F. (1989a). Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* 1989, 123(3): 585–95.
31. Tajima F. The effect of change in population size on DNA polymorphism. *Genetics* 1989, 123(3): 597–601.
32. Tajima F. Measurement of DNA polymorphism. In: Mechanisms of Molecular Evolution. Introduction to Molecular Paleopopulation Biology, edited by Takahata, N. and Clark, A.G., Tokyo, Sunderland, MA: *Japan Scientific Societies Press, Sinauer Associates, Inc.*, 1993, 37-59.
33. Hasan AU , Suguri S, Fujimoto C, Itaki RL, Harada M, Kawabata M, Bugoro H, Albino B. Genetic diversity in two sibling species of the *Anopheles punctulatus* group of mosquitoes on Guadalcanal in the Solomon Islands. *BMC Evol Biol* 2008, 8: 318.
34. Stephens JC, Julie A. Schneider, Debra A. Tanguay, Julie Choi, Tara Acharya, Scott E et al. Haplotype Variation and Linkage Disequilibrium in 313 Human Genes. *Science Express*, 2001,10.1126/science.1059431.
35. Fu Y. Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics* 1997, 147(2): 915–25.
36. Li YL, Kong XY, Yu ZN, Kong J, Ma S, Chen LM .Genetic diversity and historical demography of Chinese shrimp *Fenneropenaeus chinensis* in Yellow Sea and Bohai Sea based on mitochondrial DNA analysis. *Afr J Biotechnol* 2009, 8(7): 1193–1202.

37. Excoffier L and Heckel G. Computer programs for population genetics data analysis: a survival guide. *Nat Rev Genet* 2006, 7(10): 745–58.
38. Weir BS, Anderson, AD, and Hepler AB. Genetic relatedness analysis: modern data and new challenges. *Nat Rev Genet* 2006, 7(10): 771–80.
39. Lui J, Power Marker ver 3.0 user manual 2005. <http://www.powermarker.net>
40. Excoffier L and Slatkin M. Maximum-Likelihood estimation of molecular haplotype frequencies in a diploid population. *Mol Biol Evol* 1995, 12(5): 921–7.
41. Zaykin DV, Westfall PH, Young SS, Karnoub MC, Wagner MJ and Ehm MG. Testing association of statistically inferred haplotypes with discrete and continuous traits in samples of unrelated individuals. *Hum Hered* 2002, 53(2): 79–91.
42. Hamrick JL and Loveless MD. Isozym variation in tropical trees; procedures a preliminary results. *Biotropica* 1986, 18, 201-7.
43. Silva EF, Oliveira CAM, Silva ACBL, Rodal MJN. Diversity and genetic Structure of natural fragmented population of *Tapirira Guianensis* Aubl. In northeastern Brazil. *Bioremediation biodiversity an bioavailability* 2, (special issue 1) 2008, 35–40.

EKLER

EK.1. ÖZGEÇMİŞ

1989 yılında Kars' ta doğdu. İlkokul, ortaokul ve lise eğitimini Kars' ta tamamladı. Anadolu Üniversitesi Fen Fakültesi Fizik Bölümü lisans programından 2012 yılında mezun oldu. 2013 yılında ÖYP programı kapsamında Adıyaman Üniversitesi Tıp Fakültesi Biyofizik Anabilim Dalına Araştırma Görevlisi olarak atandı. 2014 yılında İnönü Üniversitesi Sağlık Bilimleri Enstitüsü Biyofizik Anabilim Dalında YL Bilimsel Hazırlık Programı'na başladı ve halen yüksek lisans eğitimine devam etmektedir.



EK-2. ETİK KURUL ONAYINA GEREK OLMADIĐINA DAİR BELGE

ETİK KURUL ONAYINA GEREK OLMADIĐINA DAİR BELGE

13 Nisan 2013 tarih ve 28617 sayı ile TC Resmi Gazetede yayınlanan “ Klinik arařtırmalar hakkında yönetmelik’ in 1. Fıkrası (bu yönetmelik, biyoyararlanım ve biyoeřdeęerlilik alıřmaları dahil, ruhsat veya izin alınmamıř olsa dahi insanlar üzerinde yapılacak olan ila, tıbbi ve biyolojik rnler ile bitkisel rnlerin klinik arařtırmaları, klinik arařtırma yerlerinin ve bu arařtırmaları gerekleřtirecek gerek veya tzel kiřileri kapsar.) gereęince bu alıřmanın, bir klinik arařtırma olmaması ve yayınlanmıř gensel verilerin istatistiksel analizlerinin yapılması sebebiyle etik kurul onayı alınmamıřtır.

Yrd. Do. Dr. Onur ZTRK