



## Genome-wide Analysis of Protein Diversity in Insects

Mehmet Dayı<sup>1,a,\*</sup>

<sup>1</sup>Forestry Vocational School, Düzce Üniversitesi, Düzce, Türkiye

\*Corresponding author

ARTICLE INFO	ABSTRACT
<p><i>Research Article</i></p> <p>Received : 22-08-2023 Accepted : 16-09-2023</p> <p><b>Keywords:</b> Genomics Insect species Pprotein family and variation Evolution Bioinformatics</p>	<p>Insects are one of the most successful species that can adapt to many different habitats. This feature of insects shows their evolutionary solid skills. Approximately, more than 80% of species described so far belong to Insecta. Advancements in DNA sequence technology and low cost allowed researchers to sequence entire genomes of many insect species. The comparative genomics approach is one of the powerful tools to uncover molecular and evolutionary mechanisms underlying the rapid and successful adaptation of insects. Protein families and their copy numbers are one of the key factors to uncover the evolutionary need of species. Several studies on insect evolution have been performed using different insect taxa. However, these studies focused on gene family evolution and phylogenetic relations. In this study, genomes of twenty insect species were examined to identify protein families and their copy numbers, and their variations in insects. The results showed that insects share fundamental protein families (Receptor proteins, Pkinase, Trypsin) with similar copy numbers to perform essential life tasks. Additionally, several protein families were found to have different copy numbers in some species, which showed that the adaptation need of each species differs. This study also highlighted the variation of several proteins in insects.</p>

Türk Tarım – Gıda Bilim ve Teknoloji Dergisi, 11(9): 1715-1720, 2023

## Böceklerde Protein Çeşitliliğinin Genom Düzeyinde Analizi

MAKALE BİLGİSİ	ÖZ
<p><i>Araştırma Makalesi</i></p> <p>Geliş : 22-08-2023 Kabul : 16-09-2023</p> <p><b>Anahtar Kelimeler:</b> Genomiks Böcek türleri Protein ailesi ve varyasyon Evrım Biyoenformatik</p>	<p>Böcekler, birçok farklı habitata uyum sağlayabilen en başarılı türlerden biridir. Böceklerin bu özelliği onların evrimsel açıdan güçlü becerilerini göstermektedir. Şimdiye kadar tanımlanan türlerin yaklaşık %80'inden fazlası Insecta'ya aittir. DNA dizisi teknolojisindeki gelişmeler ve düşük maliyet, araştırmacıların birçok böcek türünün tüm genomlarını dizilemesine izin verdi. Karşılaştırmalı genomik yaklaşımı, böceklerin hızlı ve başarılı adaptasyonunun altında yatan moleküler ve evrimsel mekanizmaları ortaya çıkarmak için güçlü araçlardan biridir. Protein aileleri ve kopya sayıları, türlerin evrimsel ihtiyaçlarını ortaya çıkarmak için kilit faktörlerden biridir. Farklı böcek taksonları kullanılarak böcek evrimi üzerine çeşitli çalışmalar yapılmıştır. Bununla birlikte, bu çalışmalar gen ailesinin evrimi ve filogenetik ilişkilere odaklandı. Bu çalışmada, protein familyalarını ve bunların kopya sayılarını ve böceklerdeki varyasyonlarını belirlemek için yirmi böcek türünün genomları incelenmiştir. Sonuçlar, böceklerin temel protein ailelerini (Reseptör proteinleri, Pkinaz, Tripsin) temel yaşam görevlerini yerine getirmek için benzer kopya numaralarıyla paylaştığını gösterdi. Ek olarak, birkaç protein ailesinin bazı türlerde farklı kopya sayılarına sahip olduğu bulundu, bu da her türün adaptasyon ihtiyacının farklı olduğunu gösterdi. Bu çalışma aynı zamanda böceklerdeki birkaç proteinin varyasyonunu da vurguladı.</p>

[mehmetdayi@duzce.edu.tr](mailto:mehmetdayi@duzce.edu.tr)

<https://orcid.org/0000-0002-5367-918X>



This work is licensed under Creative Commons Attribution 4.0 International License

## Giriş

Böcekler, Dünya'daki en çeşitli organizma gruplarından biri olarak kabul edilmektedir (Zhang ve ark., 2007). Böcekler, en eski kara hayvanları arasındadırlar ve dünyadaki varlıkları en az 400 milyon yıl öncesine dayanmaktadır (Zhang ve ark., 2007). Bugüne kadar yaklaşık bir milyon böcek türü tanımlanmış, sınıflandırılmış ve adlandırılmış olmasına rağmen, gerçek sayılarının 2,5 ila 10 milyon aralığında olduğu düşünülmektedir (Zhang ve ark., 2007). Uzun evrimsel tarihleri boyunca çok çeşitli habitatlarda kolayca uyum sağlamışlardır (Gaunt ve Miles 2002. Bunun arkasındaki böceklerin evrimsel başarısı, ekolojik, çevresel ve biyolojik zorlukların üstesinden gelmelerini sağlayan muazzam fenotipik ve genetik çeşitliliğin sonucudur. Böceklerdeki genetik çeşitliliğin, uzak gruplara kıyasla bir takım içinde bile daha yüksek olduğu bildirilmiştir. Örneğin, bir böcek takımı olan Diptera'daki genetik çeşitlilik (Severson ve ark., 2004) , insan ve zebra balığı gibi iki uzak türün genetik çeşitliliğine kıyasla çok daha yüksek olduğu ortaya konmuştur (Barbazuk ve ark., 2000).

Tam genom dizilemedeki son gelişmeler, araştırmacıların, böcekler de dahil olmak üzere organizmaların tüm genomunu ortaya çıkartarak genetik, sağlık başta olmak üzere pek çok alanda geniş çaplı çalışmalara olanak sağlamıştır.

Protein ailelerindeki varyasyonlar, moleküler evrimin ve uyumun temelini oluşturarak organizmaların hayatta kalmak için zorlu çevresel nişlere adaptasyon ihtiyacını yansıtır (Lees ve ark., 2016). Bazı protein aileleri, çeşitli ortamlara uyum sağlamada kilit oyuncularlardır. Bu protein ailelerinden bazıları farklı kaynaklardan (konukçu) gıda moleküllerinin sindirilmesi için Tripsin, çevresel kaynaklardan (konukçu, aynı türün bireyleri vb.) salınan koku maddelerini tespit etmek için 7TM grubu (7tm\_1, 7tm\_6 ve 7tm\_7) (yedi transmembran) reseptörleri olarak sıralanabilir (Liu ve ark., 2021).

Böceklerde hayatta kalmak ve üremek için değişen ortamlara yanıt vermek ve uyum sağlamak çok önemlidir. Bunun için, böceklerde çoğunlukla transkripsiyon faktörleri, Zn (Laity ve ark., 2001) ve sinyal proteinleri (GPCR) (Bleuven ve Landry 2016) önemli rol oynar. Başarılı adaptasyon ve hayatta kalma için bir diğer kritik faktör, konakçılar veya böcek ilaçları gibi sentetik ürünler tarafından salınan çeşitli toksik kimyasal bileşenlerin üstesinden gelmektir. Bir protein ailesi olan Sitokrom p450, monooksijenazlar gibi detoksifikasyon enzimleri içerir ve bu toksik bileşenlere karşı böcek savunma mekanizmalarında önemli roller gerçekleştirir (Feyereisen 1999) . Proteinlerin ekspresyonunun düzenlenmesine ek olarak, bu proteinlerin kopya sayısı varyasyonları da adaptasyon için ve dolayısıyla türlerin ekolojik özelliklerini belirlemek için çok önemlidir (Gerstein ve Berman 2015).

Böcek türlerinde protein ailesi çeşitliliğini ve protein ailelerinin varyasyonunu karşılaştıran yalnızca birkaç çalışma vardır. Bununla birlikte, bu çalışmalar yalnızca birkaç protein ailesi üzerine odaklanmıştır. Bu nedenle, protein ailelerinin böcekler arasında nasıl farklılaştığı ve böceklerde protein ailelerinin kopya sayısı varyasyonunun nasıl olduğu hala büyük ölçüde bilinmemektedir. Bu çalışmada, Ulusal Biyoteknoloji Bilgi Merkezi (NCBI) veritabanında bulunan tüm genom dizilemesinden üretilen 7 böcek takımından 20 böcek türünün protein verilerini kullanarak böcekler arasındaki protein ailesi çeşitliliğini ortaya konulması amaçlandı.

## Materyal ve Yöntem

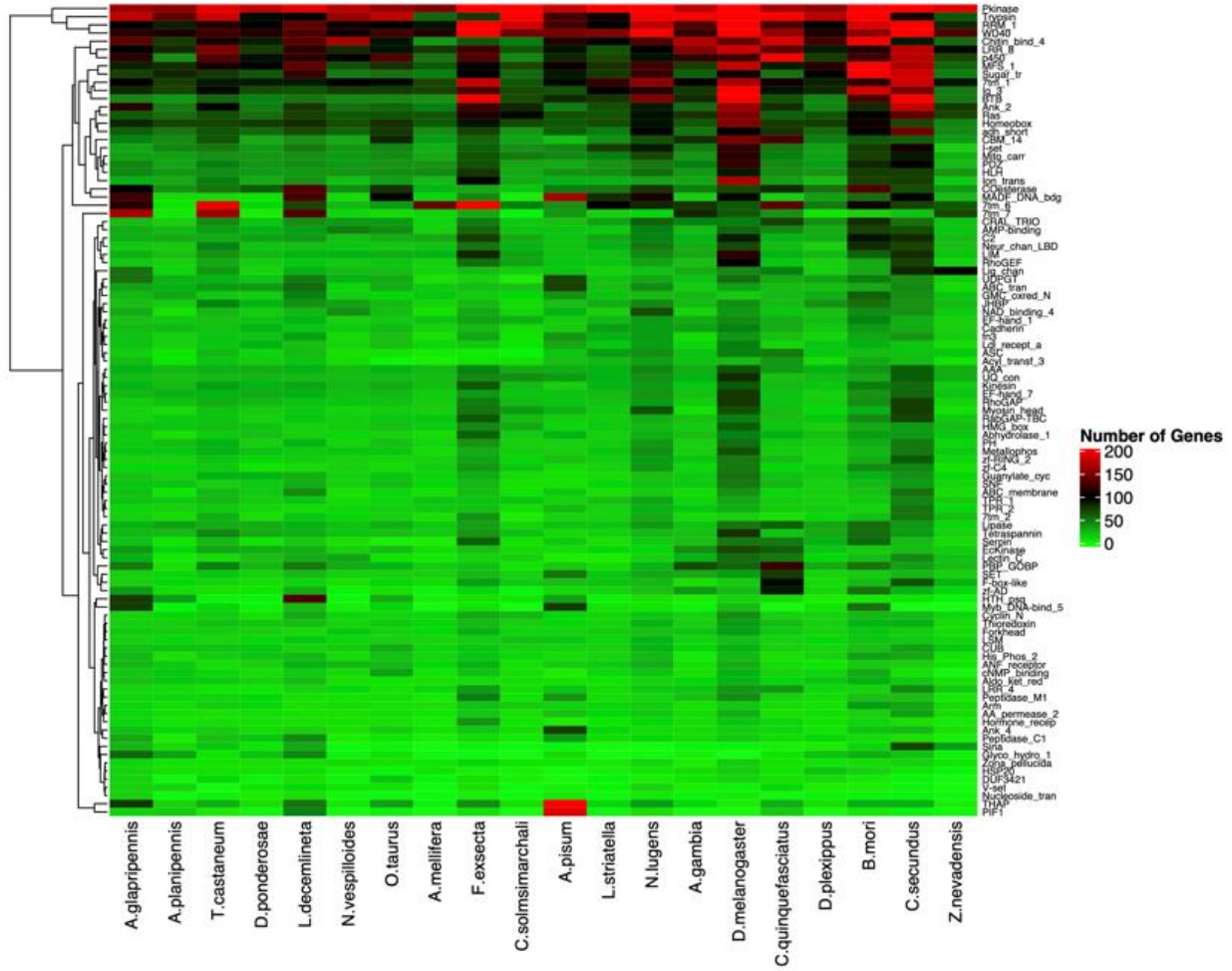
### Materyal

Bu çalışmada 7 takıma (Hemiptera, Homoptera, Isoptera, Hymenoptera, Diptera, Lepidoptera ve Coleoptera) ait 20 böcek türünün proteom verileri NCBI'den alınmıştır (Çizelge 1). Böcek türlerinin genomik özellikleri Çizelge 1'de verilmiştir.

Çizelge 1. Bu Çalışmada Kullanılan Böcek Türlerine Ait Bilgiler

Table 1. Features of insect species used in this study

Takım	Aile	Tür	Gen sayısı
Coleoptera	Cerambycidae	<i>Anoplophora glapripennis</i>	21859
	Buprestidae	<i>Agrilus planipennis</i>	15497
	Curculionidae	<i>Tribolium castaneum</i>	18534
	Curculionidae	<i>Dendroctonus ponderosae</i>	13457
	Chrysomelidae	<i>Leptinotarsa decemlineta</i>	24830
	Silphidae	<i>Nicrophorus vespilloides</i>	13516
	Scarabidae	<i>Onthophagus taurus</i>	17483
Hymenoptera	Apidae	<i>Apis mellifera</i>	15314
	Agaonidae	<i>Ceratosolen solmsi marchali</i>	12833
	Formicidae	<i>Formica exsecta</i>	22509
Hemiptera	Delphacidae	<i>Laodelphax striatellus</i>	17512
	Delphacidae	<i>Nilaparvata lugens</i>	24319
Homoptera	Aphididae	<i>Acyrtosiphon pisum</i>	36195
Diptera	Culicidae	<i>Anopheles gambia</i>	14102
	Drosophilidae	<i>Drosophila melanogaster</i>	30504
	Culicidae	<i>Culex quinquefasciatus</i>	18883
Lepidoptera	Nymphalidae	<i>Danaus plexippus</i>	15128
	Bombycidae	<i>Bombyx mori</i>	22510
Isoptera	Termopsidae	<i>Zootermopsis nevadensis</i>	14610
	Kalotermitidae	<i>Cryptotermes secundus</i>	26728



Şekil 1. Böceklerdeki Protein Ailelerinin Kümeleme Analiz Sonuçları  
Figure 1. Results of analysis of clustering protein families in insects

### Protein Ailelerinin Tanımlanması

Böceklerdeki protein ailelerinin tanımlanmasında hmmer aracı v3.1 (Finn ve ark., 2011) ile PFAM (Protein Family) protein veritabanı v32'ye karşı 0.01 eşik değeri (e-value) ile arandı. Arama sonucunda, en yüksek benzerlik oranına sahip olan protein aileleri seçilerek böceklerdeki protein ailesi varyasyonlarının analiz edilmesinde kullanıldı.

### Protein Ailelerinin Kümelenmesi ve Korelasyonu

R v3.6.1'de (Team 2013) complexHeatmap R paketi (Gu ve ark., 2016) kullanılarak böceklerdeki protein aileleri, clustering\_distance\_rows = "euclidean", clustering\_distance\_cols = "euclidean", clustering\_method = "pearson" seçeneklerle kümelendi. Böcek türlerinin protein ailelerinin çift yönlü korelasyonu (böcek türleri arasındaki ilişki ve protein aileleri arasındaki ilişki), R v3.6.1'deki (Team 2013) yerleşik cor fonksiyonu kullanılarak Pearson yöntemine göre hesaplanmıştır. Analiz sonuçları R v3.6.1'de (Team 2013) factoextra v1.0.7 R paketi kullanılarak görselleştirilmiştir.

### Protein Ailelerinin Temel Bileşen Analizleri

R v3.6.1'de (Team 2013) yerleşik olarak bulunan prcomp fonksiyonu kullanılarak Temel Bileşen Analizi yapıldı ve sonuçlar factoextra ve PCAtools R paketleri (Team 2013) kullanılarak görselleştirildi.

### Sonuçlar

#### Böceklerde Protein Ailesi Çeşitliliği

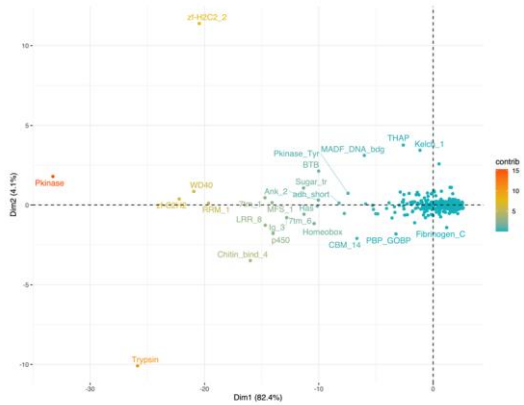
Protein ailelerinin sayısı böcekler arasında büyük farklılıklar göstermektedir (Şekil 1). Çeşitli metabolik/hücrel süreçlerde yer alan protein aileleri tüm böcek türlerinde tespit edildi. Bu protein aileleri, Pkinaz, Tripsin, zf-C2H2, zf-H2C2\_2, RRM\_1, WD40, LRR\_8, Ank\_2, BTB, Chitin\_bind\_4 ve Homeobox, 7tm\_1, 7tm\_6 ve 7tm\_7, MFS\_1, Sugar\_tr, p450, Ras ve Ig\_3 (Şekil 1) yer almaktadır. Her protein ailesindeki kopya sayısı karşılaştırmaları, çeşitli protein ailelerinde böcek takımına ve türe özgü protein ailesinin kopya sayısının artış gösterdiği tespit edildi. Örneğin, Tripsin'in Diptera takımında artış gösterdiği, MADF\_DNA\_bdg, THAP, Kelch\_1 ve PIF1 protein ailelerinin *Acyrtosiphon pisum*'da, BTB protein ailesi *Cryptotermes secundus*, *Drosophila melanogaster* ve *Formica exsecta* türlerinde diğer türlere kıyasla daha fazla kopya sayısına rastlanıldı (Şekil 2). Türe özgü protein ailesinin kopya sayısındaki artışın diğer örnekleri ise, reseptör bir protein ailesi olan 7tm\_6'nın *F. exsecta*'da ve Ig\_3 protein ailesinin *D. melanogaster*'da, 7tm\_7 protein ailesinin iki Coleoptera türünde (*Anoplophora glapripennis* ve *T. castaneum*) görülmesidir (Şekil 2). Genel olarak, bu sonuçlar, böcek türlerinin temel yaşam gereksinimlerini karşılayacak çekirdek protein ailelerine (her böcek türünde olan) sahip olduğu, bazı protein ailelerinin (sindirimden sorumlu,

konukçu/eş arama görevi gören vb.) protein ailelerinde protein ailesinin kopya sayısının artış gösterdiği tespit edilmiştir. Bu durum, böceklerin yaşam alanları, konukçu aralığı (tercih ettiği konukçu ağaç türü sayısı) yada eş seçimindeki tercihlerinin farklılıklarından kaynaklı olabileceği öne sürülebilir.

### Böceklerde Protein Ailesi Varyasyonu ve Korelasyonu

Protein ailelerindeki kopya sayılarındaki varyasyon ve korelasyona bakıldığında, en yüksek varyasyon PC1'de (%47,26) gözlemlendi ve bunu PC2 (%16,81), PC3 (%9,52), PC4 (%7,82) ve PC5 (%4,79) izledi (Şekil 2).

PBP\_GOBP, CBM\_14, Homeobox, 7tm\_6, COesteraz, Ras, Ig\_3, LRR\_8, p450 ve Chitin\_bind\_4 protein ailelerinin bir grup olarak pozitif korelasyon gösterdiği



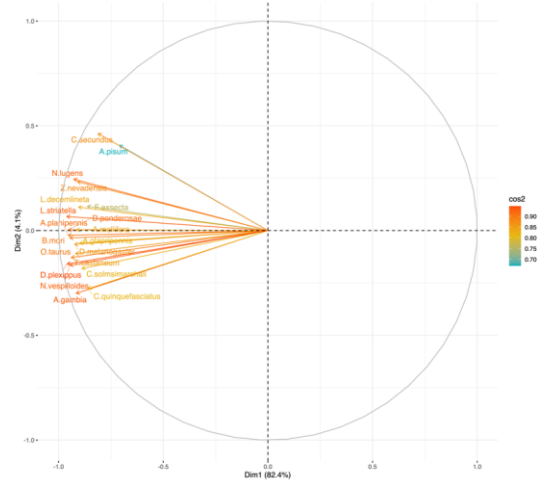
Şekil 2. Böceklerdeki Protein Ailelerinin Temel Bileşen Analizi Sonuçları.

Analizde Değişken Olarak Protein Aileleri Seçilmiştir.

Figure 2. Results of Principal Component Analysis of Protein families. Protein families were selected as variable in the analysis

bulundu, Kelc\_1, THAP, MADF\_DNA\_blog, BTB, Pkinase\_Tyr, I-set, adh\_shor, Ank\_2, Sugar\_tr, 7tm\_1, RRM\_1, WD40 ve zf\_C2H2'nin bir grupta birlikte pozitif korelasyona sahip olduğu bulundu (Şekil 2).

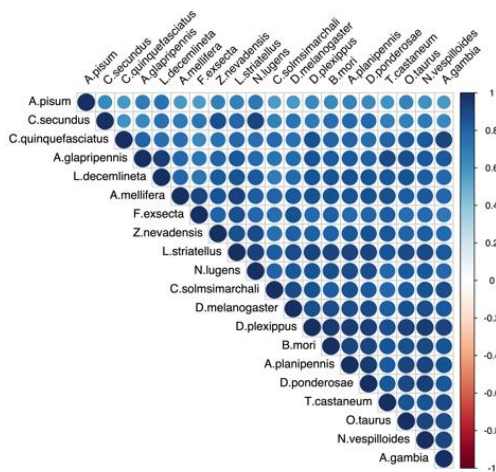
Böcekler arasındaki ikili protein ailesi karşılaştırmalarına bakıldığında, *C. secundus* dışında diğer böcekler arasında güçlü bir pozitif korelasyon tespit edildi (Şekil 3). Böcek türleri (değişkenler) arasındaki protein varyasyon karşılaştırmalarında ise, böcekleri birinci ve ikinci boyutta sırasıyla %82,4 ve %4,1'lik değişkenliklerle iki boyutta gruplandırmıştır (Şekil 3). Böceklerde protein familyalarındaki değişkenliğe en yüksek katkı *Anopheles gambia*'da gözlemlenmiştir (Şekil 3). Bu da protein ailelerinin kopya sayısındaki farklılık en fazla bu türde olduğunu göstermektedir. Değişkenliğe en düşük katkı *A. pisum*'da gözlemlenmiştir (Şekil 3).



Şekil 3. Böceklerdeki Protein Ailelerinin Temel Bileşen Analizi Sonuçları.

Analizde Değişken Olarak Böcek Türleri Seçilmiştir.

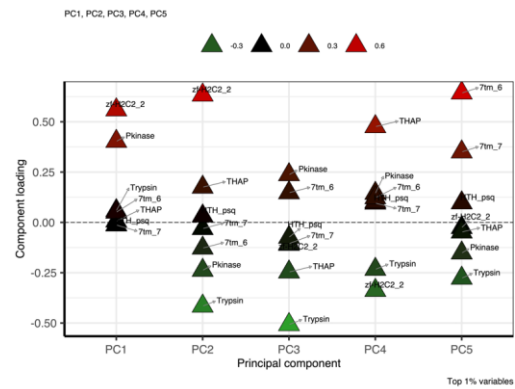
Figure 3. Results of Principal Component Analysis of Protein families. Insect species were selected as variable in the analysis



Şekil 4. Böceklerdeki Protein Ailelerinin Böcek Türlerinde Korelasyonu Sonuçları.

Bu Analizde Böcek Türleri Birbirleriyle İkili Kıyaslama Yapılmıştır.

Figure 4. Results of correlation of protein families in insects. In this analysis, pair-wise comparative of insect species were carried out



Şekil 5. Böceklerde En Çok Değişkenlik Gösteren Protein Ailelerinin Analiz Sonuçlarını Gösteren Yükleme Grafikleri.

Figure 5. Loading plots of analysis on the most variable protein families in insects

*C. secundus* ve *A. pisum*'un diğer türlere kıyasla daha düşük bir pozitif korelasyon göstermesi dışında, protein ailelerindeki genel kopya sayısı varyasyonlarının böceklerde benzer olduğu ve böcekler arasında yüksek bir pozitif korelasyon gösterdiği tespit edilmiştir (Şekil 4).

Protein ailelerinin böceklerde varyasyon üzerine yapılan analizde her bir ana bileşende (PC), *zf-H2C2\_2*, *Tripsin*, *Pkinaz*, *7tm\_6*, *THAP*, *H\_psq* ve *7tm\_7* protein ailelerinin, neredeyse tüm PC'lerde değişkenliğe en yüksek katkıyı yaptığı bulundu (Şekil 5).

## Tartışma ve sonuç

Bu çalışmada, çeşitli taksonomik gruplardan toplam 20 böcek genomu, böceklerde protein ailelerin araştırılması ve protein ailelerin çeşitliliğini anlamak için incelenmiştir.

Böcekler, temel metabolik ve hücrel süreçlerini sürdürmek için çekirdek protein ailelerine sahiptir. Bu protein ailelerinin en önemlilerinden birisi, proteinleri peptidlere parçalayarak proteoliz aktivitesi yoluyla gıda moleküllerinin sindirilmesine katılan ve birçok hayvanın sindirim sisteminde bulunan bir serin proteaz olan *Tripsin*'dir (Rawlings ve Barrett 1994). Bu çalışmada, *Tripsin*'in tüm böcek türlerinde kopya sayısında artış olduğu bulundu. Bu bulgu, beslenmenin tüm böcek türleri için hayatta kalma ve üreme için gerekli olduğunu ve yaşam döngüleri boyunca hücrel süreçleri sürdürmek için enerji sağladığını göstermektedir. *Pkinaz*, çeşitli hücrel süreç ve metabolizmada, hücre hareketinde, transkripsiyon hücre döngüsü düzenlemesinde (Hanks ve Quinn 1991), embriyonik gelişimde ve çevresel uyarılara yanıtta etkin rol oynamaktadır (Scheeff ve Bourne, 2005).

Ayrıca *zf-C2H2\_2* protein ailesi de kopya sayısı açısından böceklerde artış halinde olduğu tespit edilmiştir. Bu protein ailesinin DNA, RNA ve proteinleri bağlama yoluyla bağlanma aktivitesine dahil olduğu bilinmektedir (Englbrecht ve ark., 2004) ve böceklerde fizyolojik süreçlere ve dış çevresel uyarılara bir yanıt olarak farklı fonksiyonel genlerin transkripsiyonunu kontrol edebilir (Guo ve ark., 2004). Bu nedenle, *zf-C2H2\_2* ailesi, genleri ihtiyaçlarına göre düzenleyerek böceklerin çeşitli çevre koşullarına uyum sağlamalarına en büyük katkılardan biri olabilir.

Kitin böceklerin dış iskeletinin birincil ve önemli bir bileşenidir. Bu polisakarit, çeşitli kutikül ve matris protein çeşitlerinin kombinasyonunda karmaşık yapılar oluşturur (Zhu ve ark., 2016). Böceklerde büyüme, morfogenez ve gelişme, kitinin biyosentezi ve modifikasyonuna aracılık eder (Zhu ve ark., 2016). *Chitin\_bind\_4*, eklem bacaklılarda kitini bağlayarak kutikulanın yapısal bileşeninde yer alan ve böceklerin yapısal materyallerini oluşturarak onların yaşam döngüleri boyunca koruma, destek ve hareket için gerekli olan bir protein ailesidir (Rebers ve Willis 2001). Deri değiştirme ve metamorfoz, böceklerin gelişimindeki temel ve önemli süreçlerdir ve yaşam döngüleri boyunca devam eder (Merzendorfer ve Zimoch 2003). Bu nedenle, bu protein ailesinin kopya sayısının tüm böceklerde genişletilmesi, deri değiştirme, morfogenez ve büyümenin böceklerde temel evrimsel korunmuş süreçler olduğunu gösterdi.

Tüm böceklerde protein kopya sayısının artış gösterdiği sinyal protein aileleri olan *7tm\_1*, *7tm\_6* ve *7tm\_7*, böceklerin hayatta kalması ve üremesi için, örneğin besin kaynağını belirleme ve bir eş seçimi için gereklidir

(Bleuven ve Landry 2016). Bu protein aileleri, böceklerde habitat seçimi ve popülasyon dinamikleri için önemli olabilir ve bu nedenle çeşitli ortamlara uyum sağlamak için çok önemlidir. *P450* protein ailesi, büyüme ve gelişme, detoksifikasyon yoluyla böcek savunması ve pestisit direnci ve bitkiler tarafından salınan toksinlere tolerans gibi ksenobiyotiklere karşı koruma gibi çok çeşitli süreçlerde yer alan monooksijenaz enzimlerinden oluşur (Feyereisen, 1999).

*Pkinaz*, *7tm\_6*, *THAP*, *H\_psq* ve *7tm\_7* protein ailelerinde gen sayısında yüksek varyasyon bulunması türe özgü adaptasyon gereksinimlerinin birbirlerinden farklı olduğunu göstermektedir.

Genel olarak, tüm böcek türlerinde aynı protein ailelerinde (temel yaşamsal aktivitelerinde rol alan protein aileleri) kopya sayısının artış göstermesine rağmen, bazı protein ailelerinin kopya sayısının böcek türleri arasında değiştiği bulundu. Bu, tüm böceklerin üreme, büyüme ve gelişmeyi içeren uygun bir yaşam döngüsünü sürdürmek ve çevresel uyarılara yanıt yoluyla aynı düzenin ve ekolojinin diğer üyeleriyle iletişim kurmak için temel biyolojik ve fizyolojik süreçlerde yer alan protein ailelerine ihtiyaç duyduğunu gösterir. Bazı protein familyalarının türe özgü artış yada azalış olarak gözlemlendiği ve protein ailelerinin kopya sayılarında varyasyon gözlemlendiği görülmüştür. Bu durum, böceklerde yüksek genetik çeşitliliğin bulunduğunu ve türe özgü protein ailesinin evriminin böceklerin yaşam döngülerinde önemli bir role sahip olduğunu gösterebilir. Ayrıca, protein ailelerindeki kopya sayılarındaki varyasyon, böceklerin evrimini ve çeşitliliğini olduğu kadar morfolojik ve fizyolojik özelliklerini de yansıtır. Bu çalışmanın mevcut bulguları, evrimsel olarak korunmuş protein aileleri ve böceklerdeki protein ailelerindeki kopya sayılarının değişimi hakkında fikir vermektedir.

## Çıkar çatışması

Yazar herhangi bir çıkar çatışması beyan etmemektedir.

## Yazar Katkısı

MD: Araştırmanın tasarlanması, organizasyonu, analizler ve makalenin yazımında katkı sağladı.

## Finansman

Finansman mevcut değil.

## Etik Onay

Bu makale, insan katılımcıları veya hayvanları içeren herhangi bir deneysel çalışma içermemektedir.

## Kaynaklar

- Barbazuk WB, Korf I, Kadavi C, Heyen J, Tate S, Wun E, Bedell JA, McPherson JD, Johnson, SL., 2000. The syntenic relationship of the zebrafish and human genomes. *Genome Research*, 10:1351-1358. DOI 10.1101/gr.144700.
- Bleuven C, Landry CR. 2016. Molecular and cellular bases of adaptation to a changing environment in microorganisms. *Proceedings Biological Science*, 283. DOI 10.1098/rspb.2016.1458.

- Englbrecht CC, Schoof H, Böhm S. 2004. Conservation, diversification and expansion of C2H2 zinc finger proteins in the *Arabidopsis thaliana* genome. BMC genomics 5:1-17.
- Feyereisen R. 1999. Insect P450 enzymes. Annual Review in Entomology, 44:507-533. DOI 10.1146/annurev.ento.44.1.507.
- Finn RD, Clements J, Eddy SR. 2011. HMMER web server: interactive sequence similarity searching. Nucleic acids research, 39:W29-W37.
- Gaunt MW, Miles, MA. 2002. An insect molecular clock dates the origin of the insects and accords with palaeontological and biogeographic landmarks. Molecular Biology and Evolution, 19:748-761. DOI 10.1093/oxfordjournals.molbev.a004133.
- Gerstein AC, Berman J, 2015. Shift and adapt: the costs and benefits of karyotype variations. Current Opinion in Microbiol 26:130-136. DOI 10.1016/j.mib.2015.06.010.
- Gu Z, Eils R, Schlesner M. 2016. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. Bioinformatics, 32:2847-2849. DOI 10.1093/bioinformatics/btw313.
- Guo Z, Qin J, Zhou X, Zhang Y. 2018. Insect transcription factors: a landscape of their structures and biological functions in *Drosophila* and beyond. International Journal of Molecular Science, 19. DOI 10.3390/ijms19113691.
- Hanks SK, Quinn AM. 1991. Protein kinase catalytic domain sequence database: identification of conserved features of primary structure and classification of family members. Methods in enzymology, 200:38-62
- Laity JH, Lee BM, Wright PE. 2001. Zinc finger proteins: new insights into structural and functional diversity. Current opinion in structural biology, 11:39-46.
- Lees JG, Dawson NL, Sillitoe I, Orengo CA. 2016. Functional innovation from changes in protein domains and their combinations. Current opinion in structural biology, 38:44-52
- Liu N, Li T, Wang Y, Liu S. 2021. G-Protein Coupled Receptors (GPCRs) in Insects-A Potential Target for New Insecticide Development. Molecules, 26. DOI 10.3390/molecules26102993.
- Merzendorfer H, Zimoch L. 2003. Chitin metabolism in insects: structure, function and regulation of chitin synthases and chitinases. Journal of Experimental Biology, 206:4393-4412. DOI 10.1242/jeb.00709.
- Rawlings ND, Barrett AJ. 1994. Families of serine peptidases. Methods in Enzymology, 244:19-61. DOI 10.1016/0076-6879(94)44004-2.
- Rebers JE, Willis JH. 2001. A conserved domain in arthropod cuticular proteins binds chitin. Insect Biochemistry and Molecular Biology, 31:1083-1093. DOI 10.1016/S0965-1748(01)00056-x.
- Scheeff ED, Bourne PE. 2005. Structural evolution of the protein kinase-like superfamily. PLoS Computational Biology, 1:e49.
- Severson DW, DeBruyn B, Lovin DD, Brown SE, Knudson DL, Morlais, I. 2004. Comparative genome analysis of the yellow fever mosquito *Aedes aegypti* with *Drosophila melanogaster* and the malaria vector mosquito *Anopheles gambiae*. J Hered 95:103-113. DOI 10.1093/jhered/esh023.
- Team RC., 2013. R: A language and environment for statistical computing.
- Zhang G, Wang H, Shi J, Wang X, Zheng H, Wong GK, Clark T, Wang W, Wang J, Kang L. 2007 . Identification and characterization of insect-specific proteins by genome data analysis. BMC Genomics, 8:93. DOI 10.1186/1471-2164-8-93
- Zhu KY, Merzendorfer H, Zhang W, Zhang J, Muthukrishnan S. 2016. Biosynthesis, Turnover, and Functions of Chitin in Insects. Annual Review in Entomology, 61:177-196. DOI 10.1146/annurev-ento-010715-023933