

T.C.  
HACETTEPE ÜNİVERSİTESİ  
SAĞLIK BİLİMLERİ ENSTİTÜSÜ

FRONTONAZAL NADİR BİR YÜZ MALFORMASYONU İLE İLİŞKİLİ  
ADAY GENLERDE PROTEİN DÜZEYİNDE *IN SILICO* ANALİZ

İsmihan ALP

Biyoinformatik Programı  
YÜKSEK LİSANS TEZİ

ANKARA

2022



T.C.  
HACETTEPE ÜNİVERSİTESİ  
SAĞLIK BİLİMLERİ ENSTİTÜSÜ

FRONTONAZAL NADİR BİR YÜZ MALFORMASYONU İLE İLİŞKİLİ ADAY  
GENLERDE PROTEİN DÜZEYİNDE *IN SILICO* ANALİZ

İsmihan ALP

Biyoinformatik Programı  
YÜKSEK LİSANS TEZİ

TEZ DANIŞMANI  
Doç. Dr. Ceren SUCULARLI

Ankara

2022

**HACETTEPE ÜNİVERSİTESİ**  
**SAĞLIK BİLİMLERİ ENSTİTÜSÜ**

**FRONTONAZAL NADİR BİR YÜZ MALFORMASYONU İLE İLİŞKİLİ ADAY GENLERDE**  
**PROTEİN DÜZEYİNDE *IN SILICO* ANALİZ**

**Öğrenci: İsmihan ALP**

**Danışman: Doç. Dr. Ceren SUCULARLI**

Bu tez çalışması 28.07.2022 tarihinde jürimiz tarafından Biyoinformatik Programı'nda yüksek lisans tezi olarak kabul edilmiştir.

**Jüri Başkanı:** Doç. Dr. Yeşim AYDIN SON (İmza)  
(Orta Doğu Teknik Üniversitesi)

**Tez Danışmanı:** Doç. Dr. Ceren SUCULARLI (İmza)  
(Hacettepe Üniversitesi)

**Üye:** Prof. Dr. Ayşe Nurten AKARSU (İmza)  
(Hacettepe Üniversitesi)

Bu tez Hacettepe Üniversitesi Lisansüstü Eğitim-Öğretim ve Sınav Yönetmeliğinin ilgili maddeleri uyarınca yukarıdaki jüri tarafından uygun bulunmuştur.

11 Ağustos 2022

Prof. Dr. Müge YEMİŞÇİ ÖZKAN  
Enstitü Müdürü

## YAYIMLAMA VE FİKRİ MÜLKİYET HAKLARI BEYANI

Enstitü tarafından onaylanan lisansüstü tezimin/raporumun tamamını veya herhangi bir kısmını, basılı (kağıt) ve elektronik formatta arşivleme ve aşağıda verilen koşullarla kullanıma açma iznini Hacettepe Üniversitesine verdiğimi bildiririm. Bu izinle Üniversiteye verilen kullanım hakları dışındaki tüm fikri mülkiyet haklarım bende kalacak, tezimin tamamının ya da bir bölümünün gelecekteki çalışmalarda (makale, kitap, lisans ve patent vb.) kullanım hakları bana ait olacaktır.

Tezin kendi orijinal çalışmam olduğunu, başkalarının haklarını ihlal etmediğimi ve tezimin tek yetkili sahibi olduğumu beyan ve taahhüt ederim. Tezimde yer alan telif hakkı bulunan ve sahiplerinden yazılı izin alınarak kullanılması zorunlu metinlerin yazılı izin alınarak kullandığımı ve istenildiğinde suretlerini Üniversiteye teslim etmeyi taahhüt ederim.

Yükseköğretim Kurulu tarafından yayınlanan **“Lisansüstü Tezlerin Elektronik Ortamda Toplanması, Düzenlenmesi ve Erişime Açılmasına İlişkin Yönerge”** kapsamında tezim aşağıda belirtilen koşullar haricince YÖK Ulusal Tez Merkezi / H.Ü. Kütüphaneleri Açık Erişim Sisteminde erişime açılır.

- Enstitü / Fakülte yönetim kurulu kararı ile tezimin erişime açılması mezuniyet tarihimden itibaren 2 yıl ertelenmiştir. <sup>(1)</sup>
- Enstitü / Fakülte yönetim kurulunun gerekçeli kararı ile tezimin erişime açılması mezuniyet tarihimden itibaren ... ay ertelenmiştir. <sup>(2)</sup>
- Tezimle ilgili gizlilik kararı verilmiştir. <sup>(3)</sup>

İsmihan ALP

<sup>1</sup>“Lisansüstü Tezlerin Elektronik Ortamda Toplanması, Düzenlenmesi ve Erişime Açılmasına İlişkin Yönerge”

- (1) Madde 6. 1. Lisansüstü teze ilgili patent başvurusu yapılması veya patent alma sürecinin devam etmesi durumunda, tez **danışmanın** önerisi ve **enstitü anabilim dalının** uygun görüşü üzerine **enstitü** veya **fakülte yönetim kurulu** iki yıl süre ile tezin erişime açılmasının ertelenmesine karar verebilir.
- (2) Madde 6. 2. Yeni teknik, materyal ve metotların kullanıldığı, henüz makaleye dönüşmemiş veya patent gibi yöntemlerle korunmamış ve internette paylaşılması durumunda 3. şahıslara veya kurumlara haksız kazanç imkanı oluşturabilecek bilgi ve bulguları içeren tezler hakkında tez **danışmanın** önerisi ve **enstitü anabilim dalının** uygun görüşü üzerine **enstitü** veya **fakülte yönetim kurulunun** gerekçeli kararı ile altı ayı aşmamak üzere tezin erişime açılması engellenebilir.
- (3) Madde 7. 1. Ulusal çıkarları veya güvenliği ilgilendiren, emniyet, istihbarat, savunma ve güvenlik, sağlık vb. konulara ilişkin lisansüstü tezlerle ilgili gizlilik kararı, **tezin yapıldığı kurum** tarafından verilir \*. Kurum ve kuruluşlarla yapılan işbirliği protokolü çerçevesinde hazırlanan lisansüstü tezlere ilişkin gizlilik kararı ise, **ilgili kurum ve kuruluşun önerisi** ile **enstitü** veya **fakültenin** uygun görüşü üzerine **üniversite yönetim kurulu** tarafından verilir. Gizlilik kararı verilen tezler Yükseköğretim Kuruluna bildirilir.
- Madde 7.2. Gizlilik kararı verilen tezler gizlilik süresince enstitü veya fakülte tarafından gizlilik kuralları çerçevesinde muhafaza edilir, gizlilik kararının kaldırılması halinde Tez Otomasyon Sistemine yüklenir

\* Tez **danışmanın** önerisi ve **enstitü anabilim dalının** uygun görüşü üzerine **enstitü** veya **fakülte yönetim kurulu** tarafından karar verilir.

## ETİK BEYAN

Bu alıřmadaki bütn bilgi ve belgeleri akademik kurallar erevesinde elde ettiđimi, grsel, iřitsel ve yazılı tm bilgi ve sonuları bilimsel ahlak kurallarına uygun olarak sunduđumu, kullandıđım verilerde herhangi bir tahrifat yapmadıđımı, yararlandıđım kaynaklara bilimsel normlara uygun olarak atıfta bulunduđumu, tezimin kaynak gsterilen durumlar dıřında zgn olduđunu, Do. Dr. Ceren SUCULARLI danıřmanlıđında tarafımdan retildiđini ve Hacettepe niversitesi Sađlık Bilimleri Enstits Tez Yazım Ynergesine gre yazıldıđını beyan ederim.

İsmihan ALP

## TEŞEKKÜR

Yüksek lisans eğitimim ve özellikle tez süresince desteğini esirgemeyen, her zaman bana benden çok inanan, her ihtiyacım olduğunda sabırla beni yönlendiren, akademik gelişimimde çok değerli katkıları olan sevgili danışmanım Doç. Dr. Ceren SUCULARLI'YA,

Tezime yaptıkları çok değerli katkılarıyla değerli jüri üyelerim Prof. Dr. Ayşe Nurten AKARSU ve Doç. Dr. Yeşim AYDIN SON'a ve tezimi inceleyerek önemli yorumlarını paylaşan değerli hocam Dr. Öğr. Üyesi Arda ÇETİNKAYA'ya,

Yüksek lisans eğitimimin ilk anından itibaren her konuda bana destek olan ve yönlendiren Can KOŞUKCU'ya, yüksek lisans döneminde tanıştığım çok değerli arkadaşlarım Muntadher JİHAD ve Hazal Yağmur ÇOPUR'a,

Beni bu günlere üstün çabalarıyla getiren, kazandığım her başarıda benimle birlikte imzası olan, hem eğitimimi bir üst seviyeye çıkarmam hem de lisans hayatımdan beri kendi ayaklarım üzerinde durabilmem için her zaman beni destekleyen biricik annem Latife ALP'e; küçüklüğümde beni hep bilime yönlendiren, bilimsel merakımı artırmak için elinden geleni yapan, kendime inancımı yitirdiğimde bana benden çok inanarak motivasyon ve güç veren biricik babam Ali Haydar ALP'e; yüksek lisans eğitimim boyunca maddi manevi yanımda olan, tez sürecimdeki en büyük destekçim canım ablalarım Gülşah ALP ve Günseli ALP'e; lisans döneminden beri her konuda yanımda olan canım arkadaşlarım Didem ENİS ve Leyla KANMAZ'a; tez sürecinde desteklerini esirgemeyen çok değerli yöneticim Merve SAY'a, çalışma arkadaşlarım Merve AYDOĞAN, Derya YAVUZ ve Garen HARYANYAN'a; yüksek lisans eğitimim boyunca idari işleri sorunsuz halletmemi sağlayan Menekşe TARLA'ya,

Son olarak sahip olduğum tüm imkanlar için sonsuz minnet ve sevgi duyup saygıyla andığım Gazi Mustafa Kemal ATATÜRK'e,

En içten duygularıyla teşekkürlerimi sunarım.

## ÖZET

**Alp İ., Frontonazal Nadir Bir Yüz Malformasyonu ile İlişkili Aday Genlerde Protein Düzeyinde *In Silico* Analiz, Hacettepe Üniversitesi, Sağlık Bilimleri Enstitüsü, Biyoinformatik Programı Yüksek Lisans Tezi, ANKARA, 2022.** Frontonazal malformasyonlar orta yüz bölgesinde gelişen ve değişken fenotipik özellikler gösteren hastalıklardır. Nadir bir frontonazal malformasyon fenotipi açısından mikrodizin ve tüm ekzom dizileme yöntemleri ile, tez çalışması öncesinde, 5 etkilenmiş ve 2 etkilenmemiş birey bulunan bir aile incelenmiştir. Bu ailede linkaj analizi ile belirlenen hastalıktan sorumlu olabilecek aday gen bölgelerinde tüm ekzom verisi ile 3 aday yanlış anlamlı mutasyon (UBXN4; p.E260D, ANK; p.R1604K, TRAM2; p.E282K) tespit edilmiştir. Bu tez çalışmasında, elde edilen 3 aday varyantın protein üzerindeki olası yapısal ve fonksiyonel etkilerinin *in silico* yöntemlerle araştırılması amaçlanmıştır. Bu amaçla öncelikle dizi tabanlı PredictSNP, MetaSNP, I-Mutant 3.0, MUpro ve HOPE araçları kullanılmıştır. Daha sonra proteinlerin deneysel olarak oluşturulmuş bütüncül 3 boyutlu yapıları UniProt veri tabanında bulunmadığından, 3 aday genin kodladığı proteinlerin 3 boyutlu yapıları farklı yöntemler kullanılarak modellenmeye çalışılmıştır. ANK2 proteinine ait tahmini 3 boyutlu yapı oluşturulamamışken, UBXN4 ve TRAM2 proteinlerinin 3 boyutlu yapıları ise AlphaFold veri tabanından elde edilmiştir. UBXN4 ve TRAM2 proteinlerinde gözlemlenen mutasyonların protein yapısı üzerindeki etkilerini araştırmak için yapı tabanlı Missense3D, COACH, COFACTOR ve PremPS araçları kullanılmıştır. Sonuç olarak, bu araçlara göre, 3 aday varyantın protein üzerinde olası bir yapısal ve fonksiyonel etkiye sahip olmadığı tahmin edilmiştir.

**Anahtar sözcükler:** *in silico* analiz, yanlış anlamlı mutasyon, protein modellemesi



## ABSTRACT

**Alp İ., Protein Level In Silico Analysis of A Rare Frontonasal Facial Malformation Associated Candidate Genes, Hacettepe University, Graduate School of Health Sciences, Bioinformatics Program Master Thesis, ANKARA, 2022.** Frontonasal malformations are diseases that develop in the median facial line and appear in variable phenotypic features. A family of 5 affected and 2 unaffected individuals were studied by microarray and whole-exome sequencing for a rare frontonasal malformation phenotype prior to the thesis study. In this family, 3 candidate missense mutations (UBXN4; p.E260D, ANK2; p.R1604K, TRAM2; p.E282K) were detected by whole-exome data in candidate gene regions that may be responsible from the disease determined by linkage analysis. In this thesis, it was aimed to investigate the possible structural and functional effects of the 3 candidate mutations on the proteins by in silico methods. For this purpose, primarily sequence-based PredictSNP, MetaSNP, I-Mutant 3.0, Mupro and HOPE tools were used. Afterwards, as the complete experimentally produced 3D structures of proteins could not be obtained from UniProt database, the 3D structures of proteins, encoded by the candidate three genes, were intended to be modelled with different approaches. While the predicted 3D structure of the ANK2 protein could not be formed, the 3D structures of the UBXN4 and TRAM2 proteins were extracted from the AlphaFold database. Structure-based Missense3D, COACH, COFACTOR and PremPS tools were used to survey the effects of mutations that were observed on UBXN4 and TRAM2 proteins on protein structure. In conclusion, 3 candidate variants were predicted to have no possible structural and functional effect on the protein, according to these tools.

**Keywords:** in silico analysis, missense mutation, protein modelling

**İÇİNDEKİLER**

<b>ONAY SAYFASI</b>	<b>iii</b>
<b>YAYIMLAMA VE FİKRİ MÜLKİYET HAKLARI BEYANI</b>	<b>iv</b>
<b>ETİK BEYAN</b>	<b>v</b>
<b>TEŞEKKÜR</b>	<b>vi</b>
<b>ÖZET</b>	<b>vii</b>
<b>ABSTRACT</b>	<b>viii</b>
<b>İÇİNDEKİLER</b>	<b>ix</b>
<b>SİMGELER VE KISALTMALAR</b>	<b>xi</b>
<b>ŞEKİLLER</b>	<b>xiii</b>
<b>TABLolar</b>	<b>xv</b>
<b>1. GİRİŞ</b>	<b>1</b>
<b>2. GENEL BİLGİLER</b>	<b>3</b>
2.1. Frontonazal Malformasyonlar	3
2.2. Aday Genler	7
2.2.1. <i>UBXN4</i> (UBX Domain Protein 4)	7
2.2.2. <i>ANK2</i> (Ankirin 2)	9
2.2.3. <i>TRAM2</i> (Translokasyon ilişkili Membran Protein 2)	11
2.3. Yanlış Anlamalı Mutasyonlar	13
2.4. Varyantların Protein Üzerindeki Etkilerinin Belirlenmesi için Kullanılan Araçlar	14
2.4.1. Dizi Tabanlı Araçlar	14
2.4.2. Üç Boyutlu Yapı Kullanan Araçlar	17
<b>3. GEREÇ VE YÖNTEM</b>	<b>24</b>
3.1. İncelenecek Aday Genlere Ait Transkript ve Protein Dizilerinin Belirlenmesi	24
3.2. Varyantların Yapısal ve Fonksiyonel Etkilerinin Dizi Tabanlı Araçlarla Değerlendirilmesi	25
3.3. Varyantların Yapısal ve Fonksiyonel Etkilerinin 3 Boyutlu Yapı Kullanılarak Değerlendirilmesi	26
3.3.1. Proteinlerin Üç Boyutlu Yapılarının Tahmini	26
3.3.2. Proteinlerin 3 Boyutlu Yapısı Kullanılarak Varyantların Etkilerinin Tahmini	27
<b>4. BULGULAR</b>	<b>29</b>

4.1. UBXN4 Proteini	29
4.1.1. UBXN4 Üzerindeki Varyantın Yapısal ve Fonksiyonel Etkilerinin Dizi Tabanlı Araçlarla Değerlendirilmesi	29
4.1.2. UBXN4 Üzerindeki Varyantın Yapısal ve Fonksiyonel Etkilerinin 3 Boyutlu Yapı Kullanılarak Değerlendirilmesi	31
4.2. ANK2 Proteini	39
4.2.1. ANK2 Üzerindeki Varyantın Yapısal ve Fonksiyonel Etkilerinin Dizi Tabanlı Araçlarla Değerlendirilmesi	39
4.2.2. ANK2 Üzerindeki Varyantın Yapısal ve Fonksiyonel Etkilerinin 3 Boyutlu Yapı Kullanılarak Değerlendirilmesi	41
4.3. TRAM2 Proteini	41
4.3.1. TRAM2 Üzerindeki Varyantın Yapısal ve Fonksiyonel Etkilerinin Dizi Tabanlı Araçlarla Değerlendirilmesi	41
4.3.2. TRAM2 Üzerindeki Varyantın Yapısal ve Fonksiyonel Etkilerinin 3 Boyutlu Yapı Kullanılarak Değerlendirilmesi	44
<b>5. TARTIŞMA</b>	<b>51</b>
5.1. UBXN4 p.E260D	53
5.2. ANK2 p.R1604K	55
5.3. TRAM2 p.E282K	56
5.4. Varyantların Protein Üzerindeki ya da Fenotip ile ilişkili Etkilerinin Gösterilememesi ile İlgili Olabilecek Etkenler	57
<b>6. SONUÇ VE ÖNERİLER</b>	<b>59</b>
6.1. Sonuç	59
6.2. Öneriler	59
<b>7. KAYNAKLAR</b>	<b>60</b>
<b>8. EKLER</b>	<b>67</b>
EK-1: Tez Çalışması ile İlgili Etik Kurul İzni	67
EK-2: Şekil 2.1. için Alınan Telif Hakkı	68
EK-3: Tez Çalışması Orijinallik Raporu	70
<b>9. ÖZGEÇMİŞ</b>	<b>72</b>

## SİMGELER VE KISALTMALAR

<b>Aa</b>	Amino Asit
<b>Å</b>	Angstrom
<b>α</b>	Alfa
<b>Arg</b>	Arjinin
<b>Asp</b>	Aspartik asit
<b>ALX</b>	<i>Aristaless-like homeobox</i>
<b>ANK2</b>	Ankirin 2
<b>β</b>	Beta
<b>CASP</b>	<i>Critical Assessment of Methods of Protein Structure Prediction</i>
<b>DDG</b>	Gibbs serbest enerji deęişim deęeri
<b>ER</b>	Endoplazmik Retikulum
<b>ERAD</b>	Endoplazmik Retikulumla ilişkili Protein Yıkımı ( <i>ER-associated protein degradation</i> )
<b>FND</b>	Frontonazal Displazi
<b>Glu</b>	Glutamik asit
<b>Lys</b>	Lizin
<b>NMR</b>	Nükleer Manyetik Rezonans
<b>PAE</b>	Tahmin Edilmiş Hizalama Hatası ( <i>Predicted Aligned Error</i> )
<b>RMSD</b>	<i>Root-Mean-Square Deviation</i>
<b>STRING</b>	<i>Search Tool for the Retrieval of Interacting Genes/Proteins</i>
<b>SVM</b>	Destek Vektör Makinesi ( <i>Support Vector Machine</i> )
<b>OMIM</b>	<i>Online Mendelian Inheritance in Man</i>
<b>PPI</b>	Protein protein etkileşimi ( <i>Protein Protein Interaction</i> )

**pLDDT** *predicted Local Distance Difference Test*

**TRAM2** Translokasyon ilişkili Membran Protein 2

**UBXN4** UBX Domain Protein 4

## ŞEKİLLER

Şekil	Sayfa
2.1. Yüz gelişiminin aşamaları ve ALX gen ailesi ile ilişkilendirilen bozukluklar. Şeklin üst kısmında embriyo gelişiminde yüz ve damak gelişimi, alt kısmında ile ALX gen ailesi ile ilişkilendirilen malformasyonlar görülmektedir. ALX3 ile ilişkili gen mutasyonları daha hafif bir yüz malformasyonu oluştururken üst dudakta izlenen filtrum bombeleri tanı koydurucudur (14).	4
2.2. Sunulan tez çalışmasına aday gen tespiti açısından kaynak teşkil eden kraniyörini ailesine ait soy ağacı.	6
2.3. <i>UBXN4</i> geninin etkileşimde olduğu genler. (ER'da protein işlenmesi sırasında görev alan genler kırmızı ile işaretlenmiştir.)	8
2.4. <i>UBXN4</i> proteinine ait domain bilgisi.	9
2.5. <i>ANK2</i> geninin etkileşimde olduğu genler. Sitosolik $Ca^{+2}$ seviyelerinin yükselmesinde rol oynayan genler kırmızı ile gösterilmiştir.	10
2.6. <i>ANK2</i> proteinine ait domain ve ait domain bilgisi ve deneysel olarak belirlenmiş 3 boyutlu PDB yapılarının protein üzerindeki konumları.	11
2.7. <i>TRAM2</i> geninin etkileşimde olduğu genler. P-tipi ATPazlar ile iyon taşınmasında rol alan genler ise kırmızı ile gösterilmiştir.	12
2.8. <i>TRAM2</i> proteinine ait domain bilgisi.	13
4.1. <i>UBXN4</i> p.E260D amino asit değişiminin şematik görüntüsü. Orijinal amino asit solda, mutant amino asit sağda gösterilmiştir.	31
4.2. A) AlphaFold tarafından oluşturulan <i>UBXN4</i> proteinin 3 boyutlu yapısı B) AlphaFold tarafından oluşturulan <i>UBXN4</i> proteinin 3 boyutlu yapısı üzerinde p.E260 amino asidinin konumu pembe renkli olarak gösterilmiştir. C) p.E260 amino asidinin odaklanmış görüntüsü.	33
4.3. <i>UBXN4</i> PAE grafiği. p.E260 amino asidinin (x,y) eksenindeki konumu yaklaşık olarak kırmızı ok ile gösterilmiştir.	34
4.4. <i>UBXN4</i> PremPS sonucu. A) Yabanıl tip p.E260 rezidüsünün diğer amino asitlerin yan zincirleri ile oluşturduğu kimyasal bağların gösterimi. B) Mutant tip p.D260 rezidüsünün diğer amino asitlerin yan zincirleri ile oluşturduğu bağların gösterimi. C) Yabanıl tip p.E260'ın $\alpha$ -heliks yapısı içindeki gösterimi. D) Mutant tip p.D260'ın $\alpha$ -heliks yapısı içindeki gösterimi.	39
4.5. <i>TRAM2</i> p.E282K amino asit değişiminin şematik görüntüsü. Orijinal amino asit solda, mutant amino asit sağda gösterilmiştir.	44
4.6. A) AlphaFold tarafından oluşturulan <i>TRAM2</i> proteinin 3 boyutlu yapısı B) AlphaFold tarafından oluşturulan <i>TRAM2</i> proteinin 3 boyutlu yapısı	

- üzerinde p.E282 amino asidinin konumu pembe renkli olarak gösterilmiştir. C) p.E282 amino asidinin odaklanmış görüntüsü. 45
- 4.7.** TRAM2 PAE grafiği. E282 amino asidinin (x,y) eksenindeki konumu yaklaşık olarak kırmızı ok ile gösterilmiştir. 46
- 4.8.** TRAM2 PremPS sonucu. A) Yabanıl tip E282 rezidüsünün diğer amino asitlerin yan zincirleri ile oluşturduğu kimyasal bağların gösterimi. B) Mutant tip K282 rezidüsünün diğer amino asitlerin yan zincirleri ile oluşturduğu bağların gösterimi. C) Yabanıl tip E282'nin 3 boyutlu yapı içindeki gösterimi. D) Mutant tip K282'nin 3 boyutlu yapı içindeki gösterimi. 50

## TABLOLAR

<b>Tablo</b>	<b>Sayfa</b>
<b>2.1.</b> Elde edilen aday varyantlara ait bilgiler.	7
<b>2.2.</b> ANK2 proteinine ait UniProt veri tabanından elde edilmiş deneysel 3B yapıları	11
<b>2.3.</b> PredictSNP'e entegre edilmiş 8 tahmin aracı ve kullandıkları yöntemler.	15
<b>4.1.</b> UBXN4 p.E260D PredictSNP sonucu.	29
<b>4.2.</b> UBXN4 p.E260D Meta-SNP sonucu.	29
<b>4.3.</b> UBXN4 p.E260D I-Mutant 3.0 ve MUpuro sonucu.	30
<b>4.4.</b> UBXN4 Missense3D sonucu.	36
<b>4.5.</b> UBNX4 COACH sonucu.	37
<b>4.6.</b> UBXN4 COFACTOR sonucu.	37
<b>4.7.</b> UBXN4 p.E260D PremPS sonucu.	38
<b>4.8.</b> ANK2 p.R1604K PredictSNP sonucu.	40
<b>4.9.</b> ANK2 p.R1604K Meta-SNP sonucu.	40
<b>4.10.</b> ANK2 p.R1604K MUpuro ve I-Mutant 3.0 sonucu.	41
<b>4.11.</b> TRAM2 p.E282K PredictSNP sonucu.	42
<b>4.12.</b> TRAM2 p.E282K Meta-SNP sonucu.	42
<b>4.13.</b> TRAM2 p.E282K MUpuro ve I-Mutant 3.0 sonucu.	43
<b>4.14.</b> TRAM2 Missense3D sonucu.	47
<b>4.15.</b> TRAM2 COACH sonucu.	48
<b>4.16.</b> TRAM2 COFACTOR sonucu.	48
<b>4.17.</b> TRAM2 E282K PremPS sonucu.	49
<b>5.1.</b> 3 Aday genin kodladığı proteinlere ve varyantlara ait in silico araçlardan elde edilen bütün bulgular.	52



## 1. GİRİŞ

Frontonazal malformasyonlar; burun kökünün genişliği, oküler hipertelorizm, orta yüz hattını (burun, üst dudak ya da damak) etkileyen yarıklar, burun ucunun oluşmaması, kafatası yarığı gibi birçok klinik karakteristik özelliğin birleşimi olarak tanımlanır (1). Frontonazal displazi terimi ise 1970 yılında Sedano ve arkadaşları tarafından tanımlanmıştır (2). Otozomal resesif kalıtım kalıbı gösteren bu sendrom ALX protein ailesi ile ilişkilendirilmiştir. ALX proteinlerini kodlayan 3 genin hafiften şiddetliye seyreden fenotiplere neden olduğu bildirilmiştir. ALX3 proteinindeki fonksiyon kaybının, insanlarda "frontorini" olarak adlandırılan bir fenotipe neden olduğu gösterilmiştir (3). 1991 yılında Ali Nihat Mındıkoğlu'nun olgu raporunda ilk defa sunduğu kraniyosinostoz, basık ve geniş burun, antevort burun delikleri, üst dudakta simetrik kist benzeri oluşumlar daha önce bildirilen ve ALX3 gen mutasyonu ile ortaya çıkan frontorini malformasyonu ile bazı yönlerden benzerlik göstermektedir. Ancak bu sendrom otozomal dominant olarak kalıtılmaktadır (4).

DNA dizisinde, protein kodlayan bölgelerde (ekzonik) meydana gelen tek nükleotid değişimleri protein dizisinde farklı bir amino asidin (aa) kodlanmasına neden olabilir. Bu tür değişimler yanlış anlamlı mutasyon olarak adlandırılır. Yanlış anlamlı mutasyonlar protein üzerinde fonksiyon kaybına ya da protein-protein etkileşimlerinin bozulmasına neden olabilirler. Birçok çalışma, yanlış anlamlı mutasyonların, protein üzerindeki fonksiyon kaybı ya da protein-protein etkileşimlerinin bozulması gibi etkilerle kalıtsal ya da kompleks pek çok farklı hastalığa neden olduğunu deneysel ya da *in silico* yöntemler kullanarak göstermiştir (5-10). Bu nedenle yanlış anlamlı mutasyonların, protein yapısı ve fonksiyonu (stabilite, ligand bağlanma bölgeleri, katlanma, katalitik aktiviteler gibi) üzerindeki etkisini araştırmak önemlidir. Ancak deneysel yöntemlerle mutasyonların protein üzerindeki etkilerini araştırmak zaman alan ve maliyetli çalışmalardır.

Çalışmamızda kullanılan, ALX3 mutasyonlarında görülen fenotipe benzer bir fenotip gösteren ve daha önce olgu raporunda sunulan ailede, bu tez çalışması

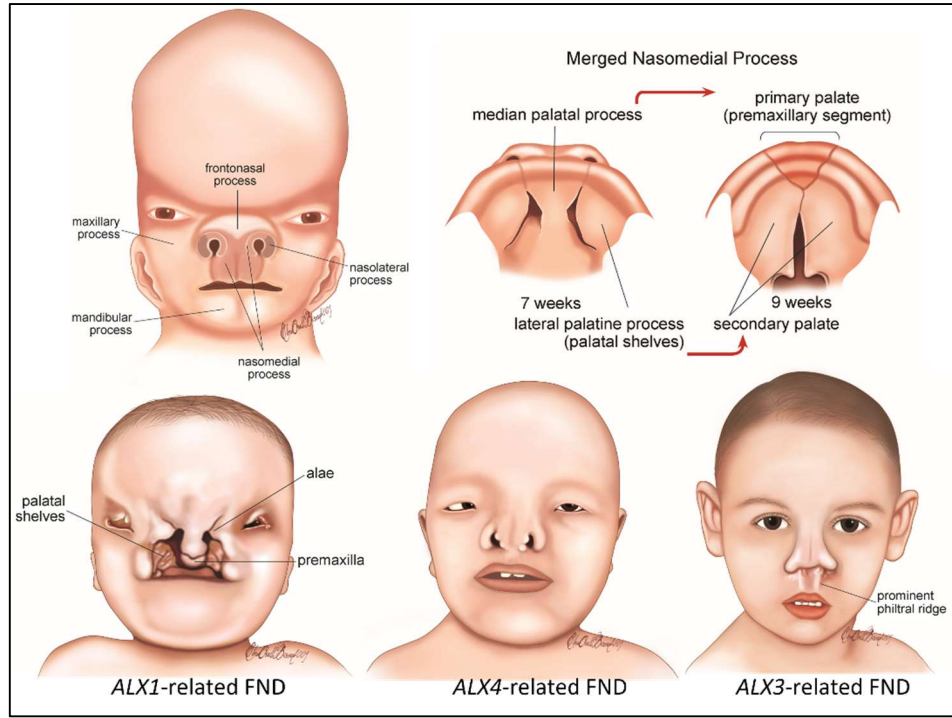
öncesinde yapılan, haplotip, bağlantı analizi ve tüm ekzom dizileme (*Whole Exome Sequencing*, WES) sonucunda fenotipi açıklayacak ve daha önceden bir genetik hastalıkla ilişkilendirilmiş genlerde herhangi bir değişikliğe rastlanmamıştır. Fakat yapılan analizlerde fenotip ile ilişkili olabilecek 3 aday gen saptanmıştır. Bu tez çalışmasında nadir yüz malformasyonu ile ilişkili olduğu ön görülen 3 aday gen (*UBXN4*, *ANK2*, *TRAM2*) üzerindeki 3 yanlış anlamlı mutasyonun (sırasıyla; p.E260D, p.R1604K, p.E282K) protein üzerindeki etkileri *in silico* yöntemlerle araştırılmıştır. Bu amaçla varyantların protein üzerindeki fonksiyonel ve yapısal etkileri dizi tabanlı araçlar kullanılarak tahmin edilmiş daha sonrasında bu proteinlerin 3 boyutlu yapıları modellenmeye çalışılarak yapı tabanlı araçlarla varyantların etkileri incelenmiştir.

## 2. GENEL BİLGİLER

### 2.1. Frontonazal Malformasyonlar

İnsanlarda yüz gelişimi kompleks bir süreç olup embriyoda 4. ve 5. haftalarda yüz tomurcuklarının oluşumu ile başlamaktadır (11). Burun kökü frontonazal tomurcuk tarafından oluşturulurken burnun dış kanatları bir çift nazolateral tomurcuktan gelişmektedir (Şekil 2.1.). Burun ucu, iki burun deliği arasındaki septum, üst dudakın bulunduğu ve filtrum olarak adlandırılan bölge ise nazomedial tomurcuk tarafından oluşturulmaktadır. Nazomedial tomurcuk aynı zamanda ön-üst damak yapısına da katılmaktadır. Buna ek olarak yanakların bulunduğu bölge bir çift maksiller tomurcuk ve çene bölgesi de yine bir çift mandibular tomurcuktan meydana gelmektedir. Bir kez tomurcuklar oluştuktan sonra epitel-mezenkim etkileşmesi sonucu yüz tomurcuklarının şekillenmesi birbirleri ile birleşecek şekilde devam ederek yüz gelişimi tamamlanmaktadır. Programlı ilerleyen bu sürece katkı yapan gen bozuklukları farklı derecelerde yüz yarıkları ve malformasyonlar ile sonlanmaktadır.

Orta hat yüz yarıkları'nı (medyan kleft) tanımlamada farklı disiplinlerce farklı isimlendirme sistemleri kullanılmakla birlikte genetik açıdan kabul gören Sedano ve ark. tarafından önerilen "frontonazal displazi (FND)" tanımlamasıdır (1, 2). FND, pariyetal kemikleşmede yetersizlik ve orta hat kemik defektleri (anterior cranium bifidum occultum), V- şeklinde frontal saç çizgisi, burun tabanının genişliği, oküler hipertelorizm, orta yüz hattı yarıkları, burun ucu formasyonunda yetersizlik, burun kanadında tek ya da çift taraflı çentiklenme/yarık burun ve üst dudak anomalileri ile karakterli bir malformasyondur (12, 13). Bu bulgulardan iki ya da daha fazlasının olması bir hastaya FND tanısı konması için yeterli olmaktadır.



**Şekil 2.1.** Yüz gelişiminin aşamaları ve ALX gen ailesi ile ilişkilendirilen bozukluklar. Şeklin üst kısmında embriyo gelişiminde yüz ve damak gelişimi, alt kısmında ise ALX gen ailesi ile ilişkilendirilen malformasyonlar görülmektedir. ALX3 ile ilişkili gen mutasyonları daha hafif bir yüz malformasyonu oluştururken üst dudakta izlenen filtrum bombeleri tanı koydurucudur (14).

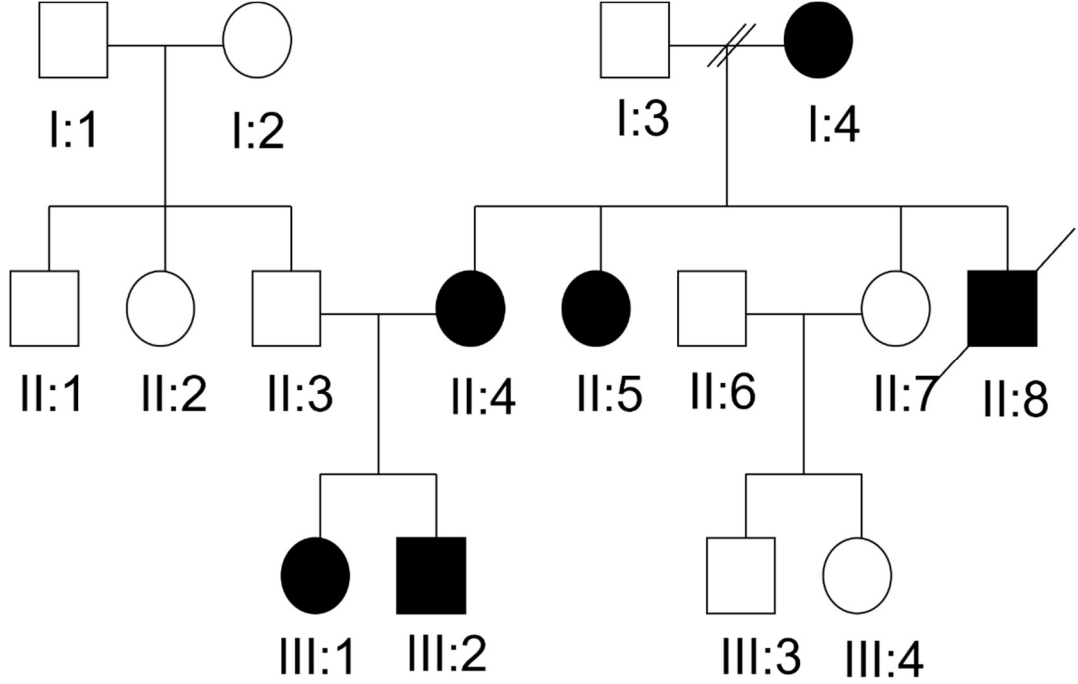
FND ile ilişkilendirilen ilk gen X kromozomunda lokalize olan *EFNB1* (Ephrin B1) genidir ve bu gendeki bozukluklar X'e bağlı kalıtılan “*Kraniyofrontonazal Sendrom*” (CFNS; OMIM numarası: 304110) ile ilişkilendirilmiştir. CFNS, üzerinde en çok çalışma yapılan hastalık olmakla birlikte burun gelişimi hakkında asıl bilgileri elde etmemiz ALX gen ailesi ile ilişkili genlerde gözlenen mutasyonlar sayesinde olmuştur (3, 14, 15). Aristaless-like homeobox (ALX) gen ailesi her üç üyesi (*ALX1*, *ALX3* ve *ALX4*) de kraniyofasiyal gelişimde rol alan bir transkripsiyon faktör ailesine mensuptur (16). Her üç gendeki mutasyonların da farklı bulgularla ortaya çıkan FND fenotipine neden olduğu gösterilmiştir ve keşfedildikleri sıra ile FND1 (OMIM numarası: 136760), FND2 (OMIM numarası: 613451) ve FND3 (OMIM numarası: 613456) olarak adlandırılmışlardır. Bu bozukluklardan ilk keşfedilene ve en hafif bulgularla seyredenine *ALX3* genindeki homozigot mutasyonlarla ortaya çıkan fenotiptir (FND1) (3, 14).

*ALX3* ilişkili FND (frontorini), hipertelorizm, geniş burun kökü, burun ucunda yarıklanma (bifid burun ucu), uzun filtrum bölgesi, üst dudak ve alveolde orta hatta çentiklenme ve üst dudak kıvrımlarının belirgin oluşu (filtrum bombeleri) ile karakterlidir (3). Üst dudakta görünen iki taraflı yumuşak doku kıvrımlarının belirgin oluşu hastalığa tanı koydurucudur. Yarık dudak, corpus callosum'da lipom ve işitme kayıpları da hastalığa eşlik edebilir. Kranial bölge anomalisi ya da kraniyosinostoz bulgulara eşlik etmez. Otozomal resesif kalıtım kalıbı göstermektedir ve heterozigot taşıyıcılarda herhangi bir bulgu bildirilmemiştir.

Kraniyorini (OMIM numarası: 123050) ise ilk kez 1991 yılında Mındıkoğlu ve ark. tarafından yayınlanmış henüz geni bulunmamış otozomal dominant bir frontonazal displazidir (4). Burun kökü açısının kaybı, kraniyosinostoz, geniş burun, dışa dönük burun kanatları ve burun kanatlarının üst dudakla birleşim yerlerinde simetrik kistik yumuşak doku oluşumları ile karakterlidir (4). Üst dudaktaki kistik yumuşak doku katlantıları nedeni ile frontorini malformasyonuna benzer olduğu düşünülse de; 1) yüz malformasyonuna eşlik eden kraniyosinostoz birlikteliği bulunması, 2) yüz dismorfik bulgularında farklılık ve 3) otozomal dominant kalıtım kalıbı göstermesi nedeniyle frontorini malformasyonundan ayrılmaktadır (3).

Sunulan tez çalışmasına, aday gen tespiti açısından kaynak teşkil eden kraniyorini ailesi ise Mındıkoğlu tarafından bildirilen orijinal vakanın bulunduğu orijinal ailedir (Şekil 2.2.). Tez öncesinde bir indeks vaka üzerinden Sanger dizileme ile *ALX3* geni dışlanmıştır. Ardından tez öncesindeki çalışmalarda elde edilen genom boyu tek nükleotid polimorfizm paneli ile genotipleme, parametrik linkage analizi ve WES verileri incelenmiştir. Bu inceleme sonucunda, linkaj analizi ile uyumlu bölgelerde bulunan, nadir protein kodlayan veya bir kırılma bölgesinde bulunan, *SELP*, *UBXN4*, *NEB*, *ANK2*, *PKHD1* ve *TRAM2* genlerindeki bazı varyantların hastalık ile ilişkili olabileceği düşünülmüştür. Bu genlerden *NEB* ve *PKHD1* genleri ile daha önceden ilişkilendirilmiş genetik hastalıkların tamamen farklı klinik özellikleri olması, *SELP* genindeki varyantın ise normal popülasyon veri tabanlarında (gnomAD (17)) bazı bireylerde homozigot olarak dahi görülmesi sebebi ile elenmiştir. Aday olarak

belirlenen *UBXN4*, *ANK2* ve *TRAM2* genlerinin kodladığı proteinlerin 3 boyutlu yapılarında aday varyant etkilerinin *in silico* incelemesi bu tezin kapsamını oluşturmaktadır.



**Şekil 2.2.** Sunulan tez çalışmasına aday gen tespiti açısından kaynak teşkil eden kraniyorini ailesine ait soy ağacı.

Bu aday genlerin kodladığı proteinlerin üzerinde tespit edilen 3 aday varyanta ilişkin bilgiler Tablo 2.1.'de gösterilmiştir.

**Tablo 2.1.** Elde edilen aday varyantlara ait bilgiler.

Gen/Varyant	dbSNP ID	MAF*	Transkript Numarası	Translasyon Numarası	Amino Asit Sayısı
<i>UBXN4</i> : p.Glu260Asp (E260D)	rs748759954	<0.001	NM_014607.4	NP_055422.1	508
<i>ANK2</i> : p.Arg1604Lys (R1604K)	rs776426910	<0.001	NM_001148.6	NP_001139.3	3957
<i>TRAM2</i> : p.Glu282Lys (E282K)	rs373852555	<0.001	NM_012288.4	NP_036420.1	370

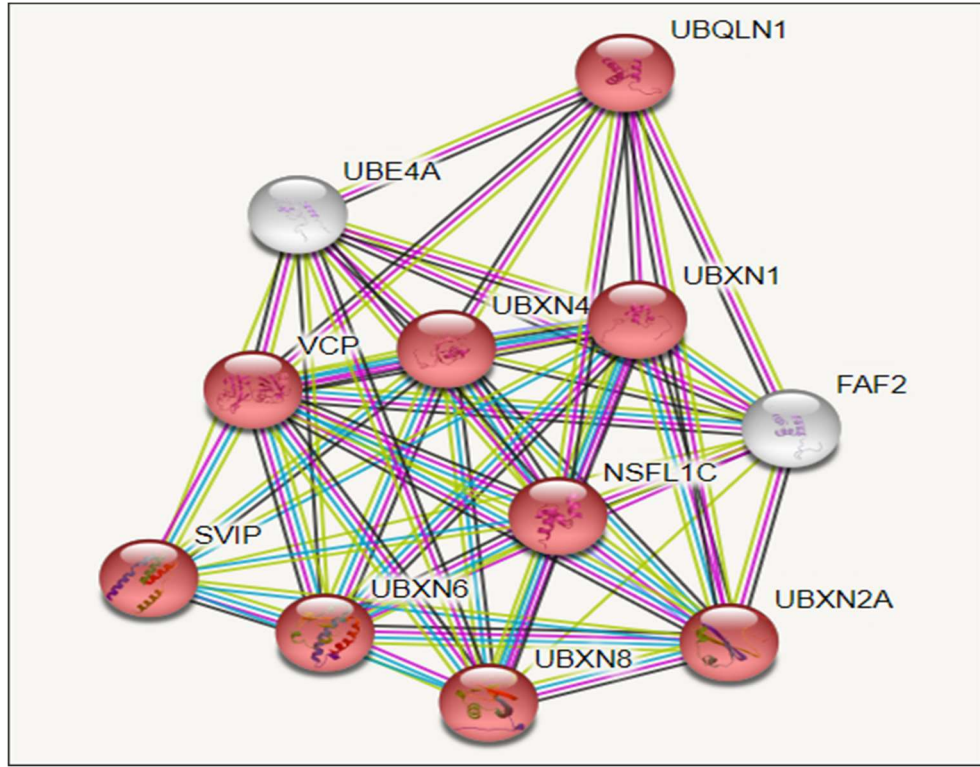
\*MAF: GnomAD veri tabanından (18) elde edilmiş Minör Alel Frekansı

## 2.2. Aday Genler

### 2.2.1. *UBXN4* (UBX Domain Protein 4)

*UBXN4*, endoplazmik retikulumun (ER) vasolin içeren proteini bağlayan bir integral proteindir ve endoplazmik retikulumla ilişkili protein yıkımı metabolizmasında (*ER-associated protein degradation*, ERAD) yer alır (19, 20). İnsan UBX domaini içeren *UBXN4* proteini, memeliler arasında yüksek korunmuşluğa sahiptir. *UBXN4*'ün, ERAD sırasında görev alan VCP (p97/vasolin içeren protein)'yi UBX domaini ile bağladığı gösterilmiştir. (19).

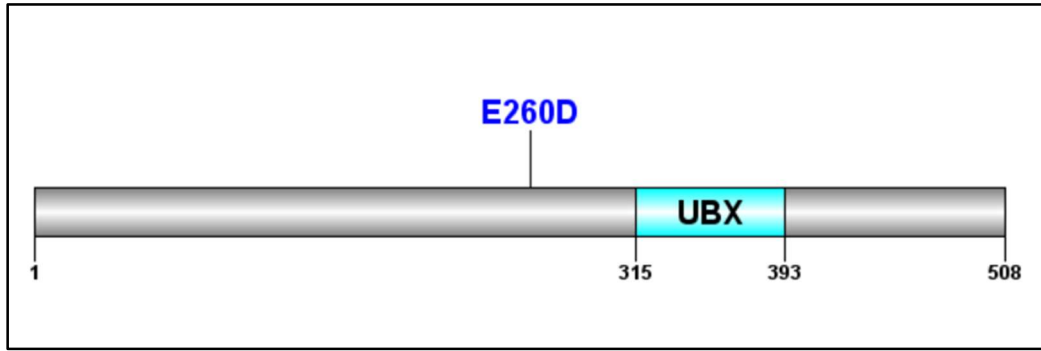
Genlerin etkileşimde oldukları diğer genleri ve buldukları yollarla ilgili bilgi sağlayan STRING 11.5 (*Search Tool for the Retrieval of Interacting Genes/Proteins*) (21) veri tabanında *UBXN4* geni için elde edilen etkileşim ağı Şekil 2.3.'te gösterilmiştir. ER'de protein işleme sırasında *UBXN4* ile etkileşimde olan genler ise kırmızı ile gösterilmiştir.



**Şekil 2.3.** *UBQN4* geninin etkileşimde olduğu genler. (ER’da protein işlenmesi sırasında görev alan genler kırmızı ile işaretlenmiştir.)

İnsan genleri ve gen-fenotip ilişkisinin yer aldığı bir veri tabanı olan OMIM (Online Mendelian Inheritance in Man) veri tabanında (22) ve literatürde *UBQN4* geni henüz herhangi bir hastalıkla ilişkilendirilmemiştir (OMIM numarası: 611216). Sitogenetik olarak 2q21.3 kromozomal bölgede yer alan *UBQN4* geni 508 aa uzunluğundadır. *UBQN4* proteinin Evrensel Protein Veri Tabanı olan UniProt’taki (23) domain bilgisini içeren ([https://www.uniprot.org/uniprotkb/Q92575/entry#family\\_and\\_domains](https://www.uniprot.org/uniprotkb/Q92575/entry#family_and_domains)) (24) ve aday varyantın belirtildiği şematik görüntü Şekil 2.4.’te gösterilmiştir.





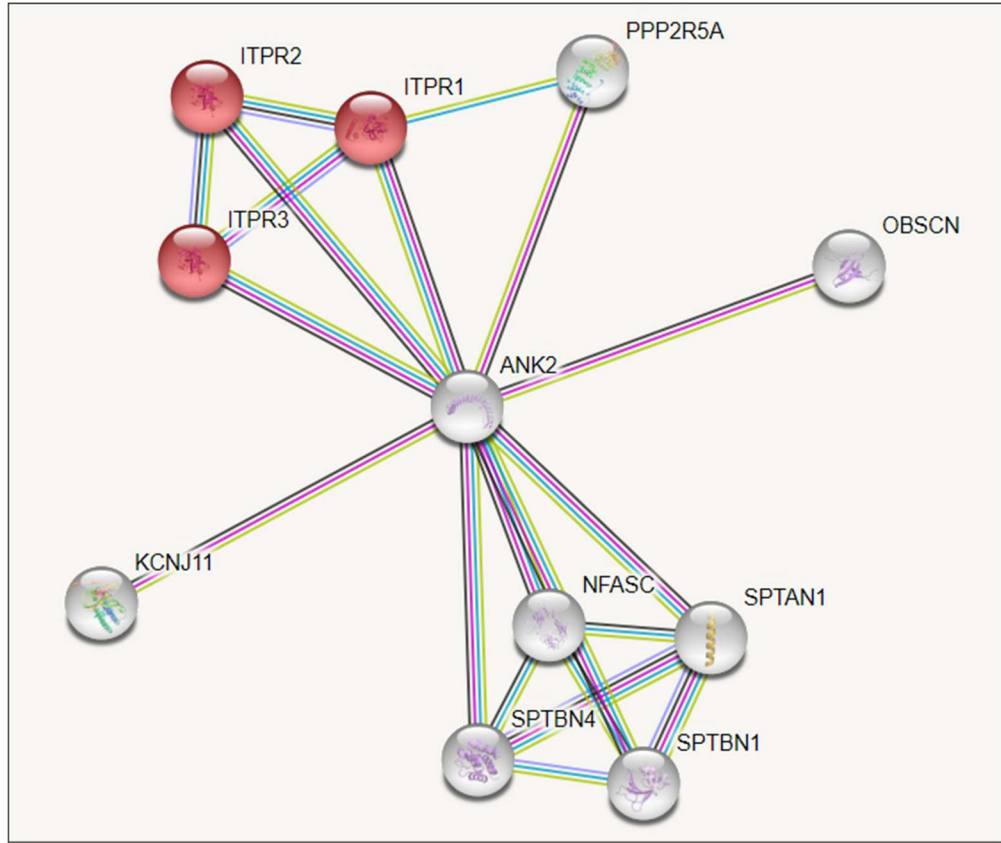
**Şekil 2.4.** UBXN4 proteinine ait domain bilgisi.

İnsan UBXN4 proteinindeki UBX domainine ait Nükleer Manyetik Rezonans (NMR) yöntemiyle çözülmüş 3 boyutlu yapısı PDB veri tabanında 2KXJ (<https://www.rcsb.org/structure/2KXJ>) (25) girdisinde bildirilmiştir. 2KXJ yapısı, 508 amino asitten 317 ile 397 amino asit arasındaki bölüme ve özellikle proteinin UBX domainini kapsamaktadır.

### 2.2.2. ANK2 (Ankirin 2)

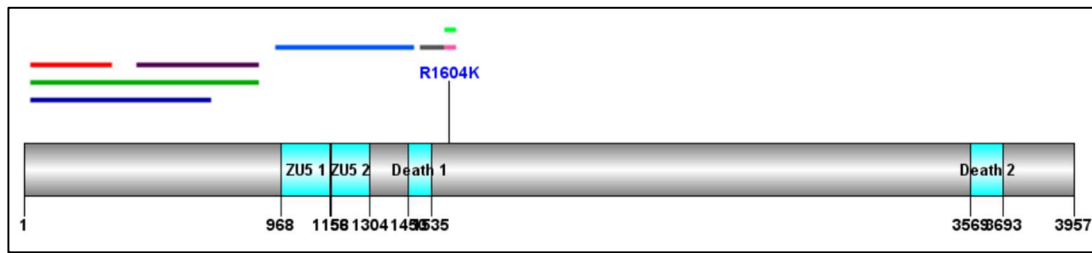
*ANK2* geni, ankirin protein ailesinin bir üyesi olan, spektrin-aktin hücre iskeletine bağlanan integral membran proteinlerini kodlar. Ankirinler hücre mobilitesi, aktivasyonu, çoğalması, bağlantısı ve membran domainlerinin onarımı için kilit rol oynarlar. Çoğu ankirin tipik olarak, çoklu ankirin tekrarları içeren bir amino-terminal, yüksek oranda korunmuş bir spektrin bağlanma alanına sahip merkezi bölge ve karboksi-terminal düzenleyici domain olmak üzere üç yapısal domainden oluşur (26).

*ANK2* geninin etkileşimde olduğu genler ve sitosolik  $Ca^{+2}$  seviyelerinin yükselmesinde rol alan genlerle etkileşimi STRING 11.5 veri tabanından elde edilmiş ve Şekil 2.5.'te gösterilmiştir.



**Şekil 2.5.** ANK2 geninin etkileşimde olduğu genler. Sitosolik  $Ca^{+2}$  seviyelerinin yükselmesinde rol oynayan genler kırmızı ile gösterilmiştir.

Bu gen tarafından kodlanan protein, kardiyomiyositlerde Na/Ca deęiřtirici'nin hedeflenmesi ve stabilitesi için gereklidir. Bu gendeki heterozigot mutasyonların, otozomal dominant kalıtılan uzun QT sendromu 4 ve kardiyak aritmi sendromuna neden olduęu bildirilmiřtir (OMIM numarası: 600919) (27, 28). Sitogenetik olarak 4q25-q26 kromozomal bölgede yer alan ANK2 geni 46 ekzondan oluřmaktadır (29). Ayrıca, farklı izoformları kodlayan birden fazla transkripti bulunmaktadır (<https://www.ncbi.nlm.nih.gov/gene/287>) (30). ANK2 proteininin UniProt veri tabanındaki domain bilgisini ([https://www.uniprot.org/uniprotkb/Q01484/entry#family\\_and\\_domains](https://www.uniprot.org/uniprotkb/Q01484/entry#family_and_domains)) (31) içeren, aday varyantın ve deneysel olarak belirlenmiř 3 boyutlu PDB yapılarının (Tablo 2.2.) belirtildięi řematik görüntü Şekil 2.6.'da gösterilmiřtir.



**Şekil 2.6.** ANK2 proteinine ait domain ve ait domain bilgisi ve deneysel olarak belirlenmiş 3 boyutlu PDB yapılarının protein üzerindeki konumları.

**Tablo 2.2.** ANK2 proteinine ait UniProt veri tabanından elde edilmiş deneysel 3B yapıları

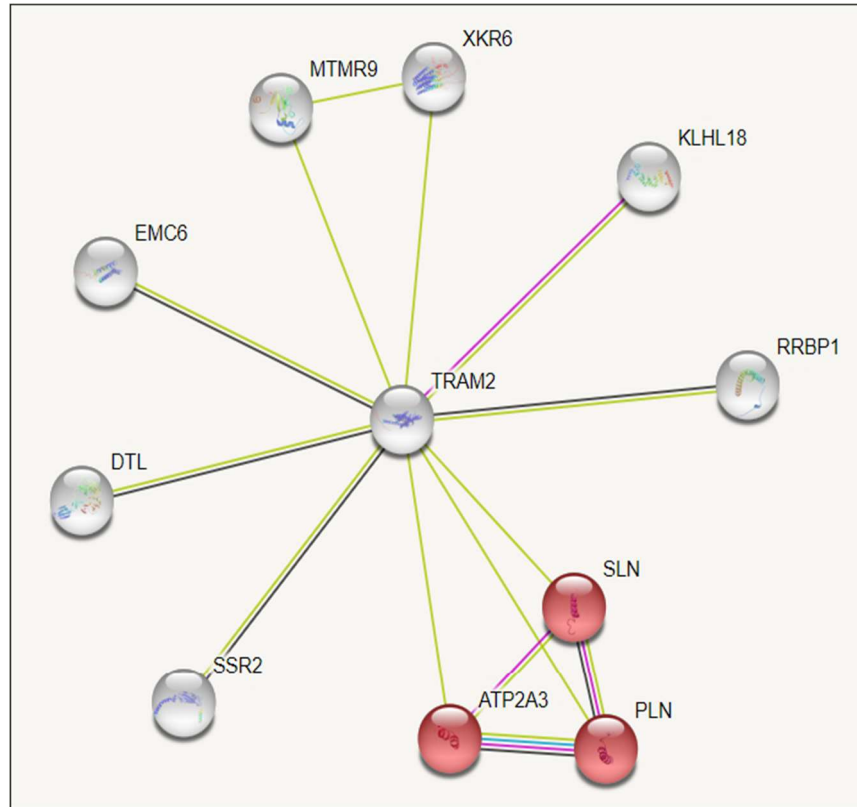
PDB Numarası	Yöntem	Zincir	Pozisyon	Bağlantı
4D8O	X-ray	-	-	<a href="https://www.rcsb.org/structure/4D8O">https://www.rcsb.org/structure/4D8O</a>
4RLV [— ]	X-ray	A	28-873	<a href="https://www.rcsb.org/structure/4RLV">https://www.rcsb.org/structure/4RLV</a>
4RLY [— ]	X-ray	A	28-318	<a href="https://www.rcsb.org/structure/4RLY">https://www.rcsb.org/structure/4RLY</a>
5Y4D [— ]	X-ray	A	28-693	<a href="https://www.rcsb.org/structure/5Y4D">https://www.rcsb.org/structure/5Y4D</a>
5Y4E	X-ray	-	-	<a href="https://www.rcsb.org/structure/5Y4E">https://www.rcsb.org/structure/5Y4E</a>
5Y4F [— ]	X-ray	A/B	430-873	<a href="https://www.rcsb.org/structure/5Y4F">https://www.rcsb.org/structure/5Y4F</a>
5YIR [— ]	X-ray	C/G/H	1588-1614	<a href="https://www.rcsb.org/structure/5YIR">https://www.rcsb.org/structure/5YIR</a>
5YIS [— ]	X-ray	C/D	1588-1614	<a href="https://www.rcsb.org/structure/5YIS">https://www.rcsb.org/structure/5YIS</a>
6KZJ [— ]	X-ray	A	1499-1570	<a href="https://www.rcsb.org/structure/6KZJ">https://www.rcsb.org/structure/6KZJ</a>
6M3Q [— ]	X-ray	E	951-1458	<a href="https://www.rcsb.org/structure/6M3Q">https://www.rcsb.org/structure/6M3Q</a>

### 2.2.3. TRAM2 (Translokasyon ilişkili Membran Protein 2)

TRAM2, ER zarında yeni oluşan salgı ve zar proteinlerinin translasyon sonrası işlenmesini kontrol eden kapılı bir makromoleküler kanal olan translokunun bir bileşenidir (32). Karaciğer yıldızlı hücreleri ve fibroblastlarda yapılan kollajen tip I ekspresyonundaki TRAM2 proteinin rolünün araştırıldığı çalışmada, aktive olmuş karaciğer yıldızlı hücreleri, fibrotik karaciğerde kollajen üreten hücrelerdir ve aktif

olan bu hücrelerde TRAM2 proteininin ekspresyonunun arttığı gösterilmiştir (32). TRAM2, translokon bileşeni olan TRAM proteini ile %53 amino asit benzerliği gösterdiği ve yalnızca %15 benzerlik gösteren C-terminal kısmının ER'nin SERCA2b'nin  $Ca^{+2}$  pompası ile etkileşime girdiği bildirilmiştir. TRAM2'nin C-terminal bölgesinin silinmesi karaciğer yıldız hücrelerinin aktivasyonu sırasında tip I kollajen sentezini inhibe edebileceği gösterilmiştir (32).

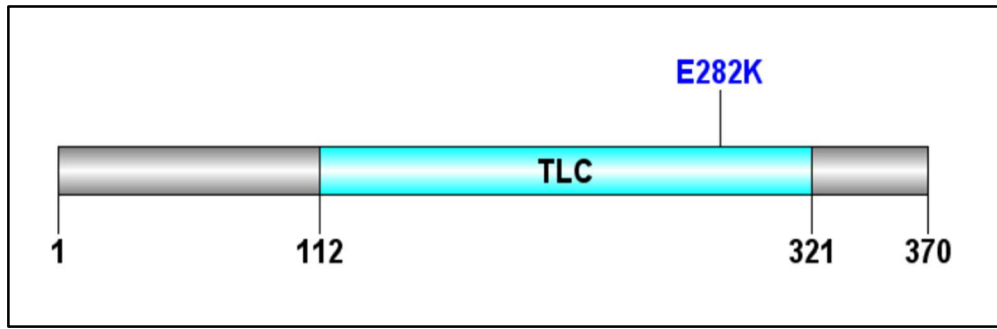
TRAM2 proteini polipeptid membran proteini olan *TM4SF20*'nin ilk transmembran sarmal N-terminalinin ER lümenine uygun şekilde yerleştirilmesi için gereklidir ve düzenlenmiş alternatif translokasyon (RAT) için bir seramid sensörü olarak işlev gördüğü öne sürülmüştür (33).



**Şekil 2.7.** *TRAM2* geninin etkileşimde olduğu genler. P-tipi ATPazlar ile iyon taşınmasında rol alan genler ise kırmızı ile gösterilmiştir.

STRING veri tabanından elde edilen *TRAM2* geninin etkileşimde olduğu genler Şekil 2.7.'de gösterilmiştir. P-tipi ATPazlar ile iyon taşınmasında rol alan genler ise kırmızı ile gösterilmiştir.

OMIM (OMIM numarası: 608485) (34)'de ve literatürde *TRAM2* geni henüz herhangi bir hastalıkla ilişkilendirilmemiştir. Sitogenetik olarak 6p12.2 kromozomal bölgede yer alan *TRAM2* geni 370 aa uzunluğunda bir proteini kodlamaktadır. Onuchic ve arkadaşları tarafından *TRAM2* geninin 11 ekzon içerdiği belirlenmiştir (35). UniProt'a (23) göre *TRAM2*'nin TLC domaini 112 ile 321 amino asitleri arasında bulunur ([https://www.uniprot.org/uniprot/Q15035#family\\_and\\_domains](https://www.uniprot.org/uniprot/Q15035#family_and_domains)) (36). SMART veri tabanında TLC domaini en az 5 transmembran  $\alpha$ -heliks içeren bir domain olarak bildirilmiştir ([http://smart.embl-heidelberg.de/smart/show\\_motifs.pl?ID=Q15035&](http://smart.embl-heidelberg.de/smart/show_motifs.pl?ID=Q15035&)) (37). *TRAM2* proteininin domain bilgisini içeren ve aday varyantın belirtildiği şematik görüntü Şekil 2.8.'de gösterilmiştir.



**Şekil 2.8.** *TRAM2* proteinine ait domain bilgisi.

İnsan *TRAM2* proteinine ait deneysel olarak belirlenmiş herhangi bir 3 boyutlu yapı bulunmamaktadır.

### 2.3. Yanlış Anlamli Mutasyonlar

DNA dizisindeki tek bir baz çiftinde meydana gelen nükleotit değişimlerinden, amino asit kodlayan üçlü kodonda farklı bir amino asit dönüşümüne neden olan mutasyonlara yanlış anlamli mutasyonlar denir. Bu mutasyonlardan bazıları hastalık etkeni olabilmektedir. Bu nedenle hastalıkların moleküler mekanizmalarının anlaşılabilmesi için yanlış anlamli mutasyonların etkilerinin araştırılması önemlidir (38). Bu etkilerin tahmini için genellikle dizi tabanlı ve yapı tabanlı *in silico* araçlar kullanılmaktadır.

Bu çalışmalara örnek olarak Sudarshan ve arkadaşlarının yaptığı bir çalışma verilebilir (39). Bu çalışmada otozomal dominant kalıtım kalıbı gösteren, tuberoskleroz kompleksine neden olan *TSC2* genlerindeki yanlış anlamlı mutasyonların hastalık yapıcılığı, farklı biyoinformatik araçlar ve bilgisayar tabanlı protein modelleme yöntemleri kullanılarak araştırılmıştır. Bunun için seçilen mutasyonlar proteinin fonksiyonel domainde olma durumuna göre değerlendirilmiş ve araştırılan mutasyonların protein yapısında, heliks yapıları arasındaki hidrofobik etkileşimleri bozduğu ve etkilenen domainin polaritesini değiştirdiği bulunmuştur. Bu durumun protein yapısını dengesizleştirdiği ve protein etkileşimi etkilediği belirtilmiştir. Proteinin domaininde yer alan mutasyonların protein yapısında; amino asidin yükünü, hidrofobik etkileşimlerini değiştirerek katlanmasını ve fonksiyonunu bozabileceği gösterilmiştir (39). 20 yanlış anlamlı mutasyon ile yapılan farklı bir çalışmada kullanılan araçlar arasında çelişkili sonuçlar elde edilmiş ve mutasyonların etkilerinin tahmini için birçok aracı birleştirerek kullanmanın önemi vurgulanmıştır (40).

## **2.4. Varyantların Protein Üzerindeki Etkilerinin Belirlenmesi için Kullanılan Araçlar**

### **2.4.1. Dizi Tabanlı Araçlar**

Yanlış anlamlı mutasyonların yapısal ve fonksiyonel etkilerini tahmin etmek için literatürde çeşitli makine öğrenmesi yöntemlerini kullanan birden fazla tahmin algoritmaları bildirilmiştir. Bu yöntemler girdi olarak protein dizi bilgisini, değişen amino asidi ve değişimin meydana geldiği kodonu kullanır.

#### **PredictSNP1.0**

Bir sınıflandırıcı web sunucusu olan PredictSNP1.0 (41), varyantların protein üzerindeki etkilerini tahmin etmek için yaygın olarak kullanılan bir araçtır. PredictSNP, farklı makine öğrenmesi algoritmalarını kullanan ve mutasyonun zararlı ya da nötral olup olmadığını tahmin eden 8 farklı tahmin aracını (MAPP, nsSNPAnalyzer,

PANTHER, PhD-SNP, Polyphen-1 (PHH-1), Polyphen-2 (PHH-2), SIFT ve SNAP) birleştiren bir konsensüs sınıflandırıcıdır (Tablo 2.1.).

Bu 8 farklı tahmin aracı, bir kıyaslama veri seti ve iki test setinden oluşan veri setleri kullanılarak değerlendirilmiştir. Değerlendirme sonucunda en iyi performans gösteren 6'sının ve bu araçlardan elde edilen sonuçların birleştirilmesiyle PredictSNP'in konsensüs sınıflandırma algoritması oluşturulmuştur. Buna göre her aracın gözlenen doğruluk değerleri %0-100 arasında değişen karşılaştırılabilir tek bir skora dönüştürülmüştür. Diğer araçların sonuçlarıyla karşılaştırmanın kolay yapılabilmesi için PredictSNP konsensüs skoru da tahminini yüzdelik olarak verir (41).

**Tablo 2.3.** PredictSNP'e entegre edilmiş 8 tahmin aracı ve kullandıkları yöntemler.

Tahmin Araçları	Kullandığı Yöntem	Referanslar
MAPP	Fizikokimyasal özellikler ve Hizalama skoru	(42)
PANTHER	Gizli Markov Modeli ve hizalama skoru	(43)
nsSNPAnalyzer	Rastgele Orman algoritması	(44)
PhD-SNP	Destek Vektör Makineleri (SVM)	(45)
Polyphen-1	Empririk kurallar	(46)
Polyphen-2	Naive Bayes sınıflandırıcısı	(47)
SIFT	Hizalama skorunu	(48)
SNAP	Sinir Ağları	(49)

### Meta-SNP

Meta-SNP 4 farklı tahmin aracını (PANTHER, PhD-SNP, SIFT and SNAP) birleştirerek, hastalıkla ilişkili yanlış anlamlı varyantların tespiti için rastgele orman metodunu kullanır (50). Bu 4 farklı tahmin araçlarının çıktılarını, her biri 4 elementten oluşan 2 grubu, 8 elementli bir öznitelik vektörüne çevirerek girdi olarak kullanır. Meta-SNP, belirli bir yanlış anlamlı varyantın hastalıkla ilgili olma olasılığını verir; puan > 0.5 verilen varyantın hastalığa neden olabileceğini, <0.5 ise varyantın etkisinin nötral olabileceğini gösterir (50). Meta-SNP, her iki tahmin için de 0 ile 10 arasında bir değer alan Güvenirlik İndeksi (*Reliability index*) hesaplar. Meta-SNP tarafından veri tabanlarından oluşturulan veri setine uygulanan tahminlere göre  $G \geq 5$  yaklaşık olarak %85 doğru olduğu bildirilmiştir (50).

### I-Mutant3.0 (Suite)

I-Mutant3.0 (Suite) (45, 51, 52), amino asit dizisinden ya da proteinin 3 boyutlu yapısından başlayarak, yanlış anlamlı mutasyonların protein stabilitesi üzerindeki değişikliklerini tahmin eder. Bunun için SVM kullanır ve çıktı olarak tahmini, Gibbs serbest enerji değişim değerini (DDG;  $\Delta\Delta G$ ) ve Artış veya Azalma olarak tahminin işaretini verir. DDG yanlış anlamlı varyantların protein stabilitesi üzerine olan etkisini ölçmek için kullanılır. Bu değer, mutasyona uğramış ve yabancı tip proteinlerinin karşılık gelen katlanması açılmış serbest enerji değerlerinin farkı alınarak kcal/mol cinsinden hesaplanır (Eşitlik 2.1) ([https://folding.biofold.org/i-mutant//pages/I-Mutant2.0\\_Details.html](https://folding.biofold.org/i-mutant//pages/I-Mutant2.0_Details.html)) (53). DDG>0, artmış protein stabilitesi ve DDG <0, azalmış protein stabilitesi anlamına gelir.

$$\Delta\Delta G_{\text{Mutant/Wild type}} = \Delta G_{\text{Mutant}} - \Delta G_{\text{Wild type}} \quad (2.1)$$

### MUpro

MUpro (54), amino asit değişikliklerinin protein stabilitesi üzerindeki etkisini tahmin etmek için SVM ve sinir ağını kullanır. Tahminin güvenilirliği olarak -1 ile 1 arasında bir skor hesaplar. Güven skoru <0, mutasyonun protein stabilitesini azalttığını gösterir ve buradaki değer ne kadar küçükse o kadar güvenilir bir tahmindir; güven skoru >0, mutasyonun protein stabilitesini arttırdığı ve buradaki değer ne kadar büyükse o kadar güvenilir bir tahmin olduğu anlamına gelir.

### HOPE

HOPE (<https://www3.cmbi.umcn.nl/hope/>) (55), doğal proteinin yapısal özelliklerini ve varyant modellerini incelemek için kullanılan otomatik bir mutant analiz sunucusudur. HOPE, mutasyona uğramış proteinlerin 3 boyutlu olarak yapısal görselleştirmesini sağlar ve sonuçları UniProt ve DAS tahmin sunucularını kullanarak verir. HOPE sunucusu, çıktıyı mutant ve yabancı tip amino asitler arasındaki yapısal varyasyon şeklinde tahmin eder.



## 2.4.2. Üç Boyutlu Yapı Kullanan Araçlar

### Proteinlerin Üç Boyutlu Modellerinin Tahmini için Kullanan Araçlar

#### Protein Modelleme Yöntemleri

DNA tarafından kodlanan proteinler 20 farklı amino asitten oluşur. Her bir amino asit birbirine bir kovalent bağ olan peptid bağlarıyla bağlanır. Bu yapı temel olarak polipeptit ana zinciri olarak adlandırılır. Bu 20 farklı amino asidi benzersiz kılan polipeptit ana zincir yapısına bağlanan yan zincirlerdir. Bir proteinin birincil yapısı amino asit dizilimi ile oluşur, amino asitler ise amino ucu (N-terminal), karboksil ucu (C-terminal) ve yan zincir olarak 3 bileşenden oluşur. Amino asitler bu yan zincirlerin özelliklerine göre polar, non-polar, pozitif ve negatif yüklü gibi gruplara ayrılır (56).

Bir proteinin şekli, kendisini oluşturan amino asitlere göre belirlenir. Polipeptit ana zincirindeki amino asitlerin yan zincirlerindeki fonksiyonel grupların diğer amino asitlerin yan zincirleriyle elektrostatik çekimler, hidrojen bağları ya da Van der Waals bağları gibi kovalent olmayan bağlarla etkileşime girerler. Bu etkileşimler proteinin katlanmasına yardım eden etkileşimlerdir. Bu katlanmaları yöneten önemli bir faktör; yan zincirlerin polar ya da non polar olma durumuna bağlı olarak proteinin içine doğru yönelme ya da dışında bir yerde düzenlenme eğiliminde olmalarıdır. Örneğin non-polar özellikli olanlar hücre içindeki su moleküllerinden kaçınmak için molekülün içine doğru kümelenme eğiliminde olur. Buna karşın polar amino asitler su molekülleriyle ya da polar moleküllerle etkileşime girebilecekleri molekülün dışında bir yerde düzenlenme eğiliminde olur. Bu etkileşimlerin sonucu olarak; çoğu protein, zincirindeki amino asitlerin sırasına göre belirlenen üç boyutlu özel bir yapıya sahiptir (56).

Proteinlerin ikincil yapısında temel olarak iki katlanma şekli gözlenmektedir:  $\alpha$ -heliks ve  $\beta$ -tabaka. Bu iki katlanma modeli, amino asitlerin yan zincirlerini içermeyen polipeptit omurgasındaki N-H ve C=O grupları arasındaki hidrojen bağından kaynaklandığı için özellikle yaygındır. Bu bağlar sonucu oluşan yapıya

proteinin ikincil yapısı denir. Proteinin üçüncül yapısı ise; amino asitlerin yan zincirleri arasında oluşan ve en düşük enerji seviyesine ulaşarak kararlı bir yapı oluşturma eğiliminde olan atomların yaptığı hidrojen bağı, van der Waals etkileşimleri, iyonik bağlar ve disülfid köprüsü gibi etkileşimler sonucu oluşur. Birden fazla polipeptit zincirinden oluşan üçüncül yapıların bir araya gelmesiyle de kompleks dördüncül yapı oluşur (56). Amino asitler, yan zincirlerinin özelliklerine göre proteinin ikincil yapısında  $\alpha$ -heliks ya da  $\beta$ -tabakalar içinde bulunmayı tercih edebilirler. Örneğin, hidrofobik bir amino asit olan izlösin,  $\alpha$ -heliks yerine  $\beta$ -tabakalarının içinde bulunma eğilimindedir (57).

UniProt'un 2021 yılında yayınladığı makaleye göre (23); yaklaşık 220 milyon protein dizi bilgisi arşivlenmiştir. Deneysel olarak çözülmüş 3 boyutlu protein yapılarının bulunduğu veri tabanı olan Protein Veri Bankası'nda ise (PDB) (<https://www.rcsb.org/>) 55.000'den fazla farklı protein için yalnızca 180.000'den biraz fazla 3 boyutlu yapı bulunmaktadır (58, 59). Proteinlerin 3 boyutlu yapıları ile protein dizi bilgisi arasındaki bu üssel farklılık araştırmacıları, deneysel yöntemlerden (X-ray kristallografisi, NMR gibi.) ziyade direkt dizi bilgisi kullanarak 3 boyutlu yapıları tahmin etmeye yöneltmiştir. Proteinlerin 3 boyutlu yapılarının modellenmesi için kullanılan başlıca 3 yöntem aşağıda özetlenmiştir:

### **Homoloji Modelleme**

Bir protein, dizi benzerliği olan (>%50) başka bir proteinle benzer katlanmalara ve benzer 3 boyutlu yapıya sahip olabilir. Dizi benzerliği azaldıkça yapı benzerliği de azalmaktadır (60). Karşılaştırmalı modelleme yöntemi olarak da bilinen homoloji modelleme yöntemi bu temel bilgiye dayanır. Hedef dizi kullanılarak bilinen yapılar (deneysel olarak tanımlanmış) üzerinden proteinin 3 boyutlu yapısını tahmin etmeyi amaçlar. Bu yöntem temel olarak 4 adımdan oluşur (61):

i) Hedef dizi ile bilinen en az bir şablon arasındaki benzerliğin tanımlanarak proteinin katlanmasının atanması: Bu adımın amacı; hedef diziyile benzerlik gösteren bütün yapıları tanımlamak ve en uygun olanını şablon olarak kullanmaktır. Protein

dizisiyle benzer katlanmalar gösteren yapıları bulmak için hedef diziler protein yapı veri tabanlarında aratılır. Bunun için kullanılan yöntemlerden PSI-BLAST (62) çoklu dizi karşılaştırması yöntemini kullanır. Bu yöntemle hedef dizi ve şablon yapı arasındaki benzerliğin düşük olduğu durumlarda bile daha hassas sonuçlar alınır. Hedef diziye homolog dizilerinin bulunmasına ek olarak hedef diziye homoloji gösteren ve modellerken şablon olarak kullanılacak protein yapılarını PDB’de aratır. Hedef dizi ile şablon dizi arasındaki benzerlik ne kadar yüksek olursa o kadar iyi bir şablon elde edilir (61).

ii) Hedef dizinin ve şablon (lar)ın hizalanması: %40’tan fazla benzerliğe sahip yakın protein dizileri için hizalama neredeyse doğru olarak yapılır. Bu oran azaldıkça hizalamada hatalar ve yüksek miktarda boşluklar meydana gelir (61). Ayrıca hedef dizinin en az %40’nın şablon dizi ile eşleşmesi gerekmektedir (63).

iii) Seçilen şablon (lar)la hizalamaya dayalı bir model oluşturulması: Hizalama oluşturulduktan sonra modelin oluşturulması için birçok farklı yöntem kullanılmaktadır. Dünya çapında en yaygın kullanılan ve otomatikleştirilmiş bir ara yüz sunan homoloji modelleme sunucusu SWISS-MODEL (64), model oluşturulması adımıyla hizalama sonucunda tanımlanan korunmuş atomik koordinatlarının aktarılmasıyla modeli oluşturur. Hizalamada eklenmiş ya da silinmiş rezidüleri karşılık gelen koordinatlar döngü modelleme ile oluşturulur ve korunmamış amino asit yan zincirlerinin oluşturulmasıyla tam atomlu protein elde edilir.

iv) Model hatalarının tahmin edilmesi: Oluşturulan modelin değerlendirilmesi için skorlama fonksiyonları kullanılır (65). Yerel kalite tahminleri, bulunan tüm şablon yapılarından topluluk bilgilerini temsil eden ikili mesafe kısıtlamaları ile geliştirilir.

### **Katlanma Tanıma (*Threading/Fold recognition*)**

Katlanma tanıma yöntemi, homoloji modelleme yöntemi gibi şablon tabanlı bir protein modelleme yöntemidir. Bu yöntem; hedef ve şablon protein arasında evrimsel bir ilişki olmadığında bile sorguya benzer kıvrımları tanımak amacıyla, farklı

evrimsel kökenlerden gelen proteinler benzer yapıya sahip olabileceğinden, hedef dizisini doğrudan diğer çözülmüş proteinlerin 3 boyutlu yapıları ile eşleştirmek için tasarlanmıştır (66, 67).

### ***Ab initio***

*Ab initio* modelleme yöntemi ise hedef diziye yönelik herhangi bir şablon kullanmadan en baştan proteinin 3 boyutlu modelini oluşturmak için kullanılan yöntemdir. Bu yöntem Afinsen (68) tarafından ortaya konan bir termodinamik hipotezine dayanır: bir proteinin doğal yapısı, belirli koşullar altında serbest enerjisinin global minimumuna karşılık gelir. Bu nedenle tahmini bir proteinin yapısı, uygun bir potansiyel enerji fonksiyonunun global minimumunun araştırılmasıyla elde edilir (69).

### **Protein Model Tahmin Araçları**

Protein modelleme yöntemlerini kullanarak çoğunlukla bir ara yüz aracılığıyla kullanıcıya sunulan çeşitli protein model tahmin araçları ve algoritmaları bulunmaktadır. Bu algoritmaların gelişmesine yardımcı olmak için düzenli olarak 2 yılda bir düzenlenen Protein Yapısı Tahmin Yöntemlerinin Kritik Değerlendirmesi (*Critical Assessment of Methods of Protein Structure Prediction, CASP*)'nde, yapısı bilinmeyen proteinlerin dizileri verilerek değerlendirmeye katılan algoritmalarından yapıların tahmin edilmesi istenir. Bu tez çalışmasında ise CASP değerlendirmeleri sonucunda katlanma tanıma (*threading*) ve *ab initio* yöntemlerini kullanan ve birçok CASP'ta en iyi sonucu alan I-TASSER (70) ve Phyre2 (71) sunucuları kullanılmıştır.

2020 yılında yayınlanan son CASP'ta ([https://www.predictioncenter.org/casp14/doc/CASP14\\_Abstracts.pdf](https://www.predictioncenter.org/casp14/doc/CASP14_Abstracts.pdf)) (72) hedef proteinleri yeni bir yapay zeka algoritması kullanarak kendinden bir önceki en iyi yönteme göre yüksek doğrulukta tahmin den AlphaFold (73), protein yapılarının tahmini problemine bir çözüm olarak kabul edilmiştir.

Yukarda belirtilen protein tahmin araçları ile ilgili bilgiler aşağıda sunulmuştur:

## I-TASSER

I-TASSER (*Iterative Threading ASSEmbly Refinement*, <https://seq2fun.dcmf.med.umich.edu/I-TASSER/>) (70) sunucusu, dizi bilgisine dayanarak otomatikleştirilmiş yapı ve fonksiyon tahmini yapan entegre bir platformdur. Amino asit dizisinden başlayarak yinelemeli yapısal birleştirme simülasyonları ve çoklu katlanma tanıma (*fold recognition/threading*) hizalamalarından 3 boyutlu yapıyı oluşturur.

## Phyre2

Proteinlerin 3 boyutlu yapılarını tahmin etmek için kullanılan başka bir araç, bir çevrim içi platform olan Phyre2'dur (<http://www.sbg.bio.ic.ac.uk/~phyre2/html/page.cgi?id=index>) (71). Phyre2, hizalamanın doğruluğunu arttırmak için, HHsearch aracılığıyla gizli Markov modellerinin hizalanmasını kullanır. Bilinen yapılarla tespit edilebilir bir şablon yapısı olmayan proteinleri modellemek için ise Poing adlı yeni bir *ab initio* katlanma simülasyonunu da içerir.

## AlphaFold

DeepMind tarafından geliştirilen AlphaFold, proteinin 3 boyutlu yapısını tahmin etmek için bir yapay zeka algoritması olan derin sinir ağını kullanır. Bu algorithmada kısaca çoklu dizi hizalamalarından ve homolog yapı şablonlarından türetilen özellikler birleştirilir. DeepMind, Avrupa Biyoinformatik Enstitüsü'nün (EMBL-EBI) ev sahipliğinde insan proteomuna ait bütün 3 boyutlu model setini <https://alphafold.ebi.ac.uk/> veri tabanında kullanılabilir hale getirmiştir (73, 74). AlphaFold tarafından oluşturulan modellerin güvenilirliği, tahmin edilen yerel mesafe farkı testi (*predicted Local Distance Difference Test*, pLDDT) skoru ile değerlendirilir. Bu skor 0 ile 100 arasında bir değer alır ve yüksek değerler güvenilir bir model olduğunu gösterir (74).

## Proteinlerin 3 Boyutlu Yapısı Kullanılarak Varyantların Etkilerinin Tahmini

### Missense3D

Missense3D (75), yanlış anlamlı mutasyonların yapısal olarak zararlı etkisini deneysel ve tahmin edilmiş yapılar kullanarak değerlendirir. 606 deneysel yapı üzerinde yaptıkları çalışmada; 1965 hastalık ilişkili yanlış anlamlı mutasyonlardan %40'ının mutant yapıda yapısal olarak zararlı değişikliğe yol açtığı gösterilmiştir. 2134 nötr varyantın ise yalnızca %11'i yapısal zarar verdiği tahmin edilmiştir. Phyre2 kullanılarak tahmin edilen 1052 model için de benzer sonuçlar elde edilmiştir. Böylelikle yanlış anlamlı mutasyonların etkisinin yapı tabanlı analizleri homoloji modellerine uygulanabilirliği gösterilmiştir. Missense3D hem kullanıcının kendi oluşturduğu modeli hem de AlphaFold tarafından oluşturulan 3 boyutlu protein modellerini girdi olarak kullanır (75). Değişikliğe uğramış amino asidin protein yapısı üzerindeki etkisini tahmin etmek için SCWRL4 (76) yöntemiyle mutant yapı oluşturulur. Hedef rezidüden 5 Å içindeki rezidüer yan zincirden çıkarılıp, değişen amino asiti içeren yan zincir SCWRL4 kullanılarak tekrar paketlenir (75).

Mutant ve yabanıl tip yapılar daha önceki çalışmalarda belirlenmiş, protein yapısının prensiplerine dayanan 16 yapısal özellik kullanılarak değerlendirir (75). Bu özellikler; disülfid bağlarının kırılması, gömülü prolin amino asidinin ortaya çıkması, çarpışma, gömülü hidrofilik rezidünün ortaya çıkması, gömülü rezidünün yüksüz rezidüden yüklü olarak değişmesi, ikincil yapının değişmesi, gömülü rezidünün yükünün değişmesi, izin verilmeyen phi/psi açıları, gömülü rezidünün yüklü konumdan yüksüz konuma değişmesi, gömülü rezidünün glisine dönüşmesi, gömülü hidrojen bağlarının kırılması, gömülü tuz köprüsünün kırılması, kavite hacminin değişmesi, gömülü ya da açıkta kalan rezidünün göreceli yüzey erişilebilirliklerindeki (RSA, Relative Solvent Accessibility) farklara göre değerlendirilmesi, Cis pozisyonunda prolin amino asidinin oluşması ve bir kıvrım bölgesinde glisin amino asidiyle değişmesi olarak sıralanabilir.

## **COFACTOR**

COFACTOR (77, 78), protein moleküllerinin biyolojik fonksiyon açıklamaları için yapı, dizi ve protein-protein etkileşim (*Protein-protein Interactions*, PPI) tabanlı bir yöntemdir. COFACTOR, 3 boyutlu yapısal modelden başlayarak, fonksiyonel bölgeleri ve homolojileri belirlemek için yerel ve global yapı eşleşmeleri ile BioLiP (79) protein fonksiyon veri tabanı üzerinden ilgilenen proteinlerle ilgili bilgileri işler.

## **COACH**

COACH, protein ligand bağlanma bölgelerini tahmin etmek için bir meta sunucu yaklaşımıdır (80). Hedef proteinlerin verilen yapısından başlayarak, COACH, bağlanmaya özgü alt yapı ve dizi profili karşılaştırmaları ile BioLiP (79) protein fonksiyon veri tabanından ligand bağlama şablonlarını tanıyan TM-SITE ve S-SITE adlı iki karşılaştırmalı yöntemi kullanarak tamamlayıcı ligand bağlanma yeri tahminleri üretir.

## **PremPS**

PremPS, Gibbs serbest enerjisinin açılmasındaki nicel değişiklikleri hesaplayarak tek nükleotit mutasyonların protein stabilitesi üzerindeki etkilerini değerlendirir. Tahminler protein yapısına dayanmaktadır (81). PremPS modeli, 131 proteinden 5296 mutasyon için Gibbs serbest enerji değişikliklerinin ( $\Delta\Delta G$ ) açılmasının deneysel verileri üzerinde eğitim veren rastgele orman (RF) regresyon skorlama fonksiyonunu kullanır.

### 3. GEREÇ VE YÖNTEM

Bu tez çalışması için etik izne ihtiyaç duyulmadığına dair Hacettepe Üniversitesi Girişimsel Olmayan Klinik Araştırmalar Etik Kurulu'nun karar belgesi Ek 1'de sunulmuştur.

#### 3.1. İncelenecek Aday Genlere Ait Transkript ve Protein Dizilerinin Belirlenmesi

Daha önceki çalışmalarda incelemeye alınmış, otozomal dominant kalıtım kalıbı gösteren ve frontozal nadir bir yüz malformasyonu olan bir ailede toplam 5 etkilenmiş ve 2 etkilenmemiş bireylere 250k Nspl SNP array'e ait haritalama ve WES yapılmıştır. Etkilenmiş 2 bireye ait WES verisine öncelikli olarak nadir varyant filtreleme stratejisi uygulanmıştır. Daha sonra SNP array'den dominant kalıtım modeline göre haplotip ve linkaj analiziyle hastalıkla ilişkili olabileceği düşünülen aday bölgeler elde edilmiştir. Bu bölgeler dikkate alınarak WES verisinde 3 aday gen ve bu genler üzerinde gözlenmiş 3 aday varyant bulunmuştur. Tespit edilen varyantlar arasında proteinlerin birden fazla amino asidini etkileyen başka bir varyant bulunmamaktadır.

Bu genler ve genlerde gözlenen varyantlara ilişkin bilgiler NCBI veri tabanından (<https://www.ncbi.nlm.nih.gov/gene>) elde edilmiştir (Bkz. Tablo 2.1.). Yalnızca ANK2 geninin 73 adet transkripti olduğu için, p.R1604. pozisyonu kapsayan ve ilgili kodonda Arjinin amino asidini (Arg) barındıran protein dizileri tek tek taranmış ve bu pozisyondaki Arg yalnızca NP\_001139.3 numaralı translasyonda saptandığı için çalışmalara NP\_001139.3 amino asit dizisi ile devam edilmiştir.

Her aday gene ait protein dizisi ilgili translasyon numarası dikkate alınarak (Bkz. Tablo 2.1.) NCBI veri tabanından FASTA formatında indirilip analizlere başlanmıştır.



### 3.2. Varyantların Yapısal ve Fonksiyonel Etkilerinin Dizi Tabanlı Araçlarla Değerlendirilmesi

Bu tez çalışmasında varyantların protein düzeyinde yapısal ve fonksiyonel etkilerini dizi bilgisi kullanarak *in silico* olarak değerlendiren 5 farklı tahmin aracı kullanıldı. Bunlardan MUpro (54) ve I-Mutant3.0 (Suite) (45, 51, 52) yanlış anlamlı varyantların protein stabilitesi üzerindeki etkisinin araştırılması için kullanıldı.

#### PredictSNP1.0

Hastalık ilişkili mutasyonların tahmini için bir konsensüs sınıflandırıcı olan PredictSNP (<https://loschmidt.chemi.muni.cz/predictsnp/>) (41) tahmin aracına, aday genlere ait protein dizisi FASTA formatında yüklendi ve her bir protein dizisindeki yabancı tip amino asit ve bu bölgede meydana gelen yanlış anlamlı değişim ile oluşan yeni amino asit bilgisi girildi. Sınıflandırıcı içinde bulunan bütün tahmin araçları seçildi.

#### Meta-SNP

Hastalıkla ilişkili yanlış anlamlı varyantların tespiti için rastgele orman metodunu kullanan Meta-SNP (<https://snps.biofold.org/meta-snp/index.html>) (50) tahmin aracına her bir aday gene ait protein dizisi, yabancı tip amino asit ve bu bölgede meydana gelen yanlış anlamlı değişim ile oluşan yeni amino asit bilgisi girilerek analiz başlatıldı.

#### I-Mutant 3.0 (Suite)

I-Mutant Suite (<http://gpcr2.biocomp.unibo.it/cgi/predictors/I-Mutant3.0/I-Mutant3.0.cgi>) (45, 51, 52) protein dizisinden ya da varsa 3 boyutlu yapıdan başlayarak tek nükleotit değişikliğinin protein stabilitesi üzerine etkisini, SVM kullanarak tahmin eder. Bu web sunucusuna, varsayılan parametreler kullanılarak (derece:25, pH:7), her bir aday gene ait protein dizileri ve yabancı tip amino asit ve bu bölgede meydana gelen yanlış anlamlı değişim ile oluşan yeni amino asit bilgisi girildi. I-Mutant'ın bu versiyonunda 3 boyutlu yapı bilindiğinde doğruluk oranı %68, dizi

bilindiğinde doğruluk oranı %64 tahmin edildiği belirtilmiştir ancak bu iki doğruluk oranı arasında önemli ölçüde bir fark gözlenmediği ve her aday proteinin 3 boyutlu yapısının bulunmama durumu göz önüne alındığından yalnızca dizi bilgisi kullanılarak analize devam edildi.

### **MUpro**

MUpro (<http://mupro.proteomics.ics.uci.edu/>) (54), amino asit değişikliklerinin protein stabilitesi üzerindeki etkisini tahmin etmek için SVM ve sinir ağını kullanan başka bir web sunucusudur. Girdi olarak protein dizisi ya da varsa 3 boyutlu yapı kullanan ve bu iki girdinin doğruluk oranları arasında (sırasıyla %84,2 ve %84,5) önemli bir fark bulunmayan program için analizler amino asit dizisi kullanılarak gerçekleştirildi. MUpro'ya yabancı tip amino asit ve bu bölgede meydana gelen yanlış anlamlı değişim ile oluşan yeni amino asit bilgisi girildi.

### **HOPE**

Yalnızca protein dizi bilgisini girdi olarak kabul eden HOPE sunucusunda (55) protein dizileri ve üzerindeki bölgedeki yabancı tip amino asit ve değişim sonucunda oluşan amino asitler seçilerek UBXN4 ve TRAM2 protein dizileri ile analiz başlatıldı. Ancak ANK2 proteinine ait amino asit dizisi çok uzun geldiği için bu programa yüklenemedi.

## **3.3. Varyantların Yapısal ve Fonksiyonel Etkilerinin 3 Boyutlu Yapı Kullanılarak Değerlendirilmesi**

### **3.3.1. Proteinlerin Üç Boyutlu Yapılarının Tahmini**

Aday genlere ait proteinlerin 3 boyutlu yapılarının tahmini için I-TASSER (70), PHYRE2 (71) ve AlphaFold (73, 74) programları kullanıldı.

## I-TASSER

Bir web sunucu olan I-TASSER'a 500 ila 1500 aa uzunluğunda olan protein dizileri yüklenebildiğinden aday proteinlerden yalnızca UBXN4 ve TRAM2'ye ait diziler yüklendi (Bkz. Tablo 3.1.). Analiz başlatılmadan önce UBXN4 proteini için daha önce PDB'de tanımlanmış bir yapı olan 2KXJ'yi (<https://www.rcsb.org/structure/2KXJ>) (25) hedef-şablon hizalama bilgisi olmadan, şablon olarak kullanılması için özellikle belirtildi. TRAM2 için PDB'de tanımlanmış bir yapı olmadığından yalnızca dizi bilgisi girilerek analiz edildi. Her iki protein için de varsayılan parametreler kullanılarak analiz başlatıldı.

## Phyre2

Phyre2 platformuna her bir aday gene ait amino asit dizisi yüklendi. Modelleme yöntemi olarak çoklu şablonları ve *ab initio* teknikleri kullanılarak bütün proteini modelleyen "Intensive" seçildi.

## AlphaFold

Bu tez çalışmasında kullanılan aday genler, organizma olarak *Homo Sapiens* dikkate alınarak, söz konusu veri tabanında aratıldı. 2700 aa'ten uzun proteinler bu veri tabanında dışlandığı için ANK2 proteinine ait bir veri bulunamadı.

### 3.3.2. Proteinlerin 3 Boyutlu Yapısı Kullanılarak Varyantların Etkilerinin Tahmini

#### Missense3D

Yanlış anlamlı mutasyonların proteinin 3 boyutlu yapısı üzerinde verdiği hasarı, deneysel ve tahmin edilmiş yapılar kullanarak değerlendiren Missense3D (<http://missense3d.bc.ic.ac.uk/~missense3d/>) (75) web ara yüzüne AlphaFold veri tabanından indirilen UBXN4 ve TRAM2'ye ait .pdb uzantılı ve içinde zincir bilgisi bulunan dosyalar, yabancı ve mutant tip rezidüleri girilerek yüklendi.

## COACH

Protein ligand bağlanma bölgesi tahmini için bir meta sunucu yaklaşımı olan COACH (<https://seq2fun.dcmdb.med.umich.edu/COACH/>) (79, 80) kullanıldı. COACH, tamamlayıcı bir ligand bağlanma tahmini için TM-SITE ve S-SITE adlı iki karşılaştırmalı yöntemi kullanır. Bu yöntemler; dizi profil karşılaştırmaları ve bağlanma spesifik alt yapıları ile BioLip protein fonksiyon veri tabanından (79) ligand bağlanma şablonlarını tanırlar. Hedef proteine ait yapı bilgisi kabul eden sunucuya yine AlphaFold veri tabanından indirilen UBXN4 ve TRAM2'ye ait .pdb uzantılı dosyalar yüklendi.

## COFACTOR

Proteinlerin 3 boyutlu yapılarını BioLip protein fonksiyon veri tabanında (79) lokal ve global yapı eşleştirmesi yoluyla fonksiyonel bölgeleri ve homolojileri belirlemek için işleyen başka bir yöntemdir. Bir web sunucu olan COFACTOR (<https://seq2fun.dcmdb.med.umich.edu/COFACTOR/>) (77-79) aracına AlphaFold veri tabanından indirilen UBXN4 ve TRAM2'ye ait .pdb uzantılı dosyalar herhangi bir parametre girilmeden yüklendi.

## PremPS

Protein yapısına dayanarak yanlış anlamlı mutasyonların protein stabilitesi üzerine etkisini tahmin eden PremPS (<https://lilab.jysw.suda.edu.cn/research/PremPS/>) (81) aracına girdi olarak yalnızca .pdb uzantılı dosya kabul eden bu araca ilgili dosyalar yüklendi. Yüklenen dosyadaki zincir bilgisi girildikten sonra varyasyon bilgisi girilerek analize devam edildi.

## 4. BULGULAR

### 4.1. UBXN4 Proteini

#### 4.1.1. UBXN4 Üzerindeki Varyantın Yapısal ve Fonksiyonel Etkilerinin Dizi Tabanlı Araçlarla Değerlendirilmesi

##### PredictSNP

Aday genlerden *UBXN4*'e ait protein dizisi üzerinde gözlenen p.E260D varyantının hastalık ile ilişkili olup olmadığına dair PredictSNP analizi sonucunda; entegre programlardan bazıları varyantı “Zararlı” olarak sınıflandırmasına rağmen konsensüs değerlendirmede %60 beklenen doğruluk oranında varyantın etkisi “Nötr” olarak saptanmıştır. Detaylı PredictSNP sonucu Tablo 4.1.'de gösterilmiştir.

**Tablo 4.1.** UBXN4 p.E260D PredictSNP sonucu.

Mutasyon	PredictSNP	MAPP	PhD-SNP	PolyPhen-1	PolyPhen-2	SIFT	SNAP	nsSNPAnalyzer	PANTHER
p.E260D	%60	%63	%72	%59	%60	%71	%61	-	%67

■ Zararlı
 ■ Nötr

##### MetaSNP

MetaSNP sonucuna göre bütün entegre programlarla benzer şekilde varyantın hastalık yapıcı etkisi güvenilirlik endeksi (Gi) 3 ile “Nötr” olarak saptanmıştır (Tablo 4.2.).

**Tablo 4.2.** UBXN4 p.E260D Meta-SNP sonucu.

Mutasyon	PANTHER	PhD-SNP	SIFT	SNAP	Meta-SNP	Gi
p.E260D	0.445	0.287	0.230	0.270	0.328	3

■ Zararlı
 ■ Nötr

PANTHER, PhD-SNP, SNAP ve Meta-SNP: 0 ile 1 arasında, eğer 0.5'ten büyükse mutasyon “Zararlı” olarak sınıflandırılır. SIFT: Eğer 0.05'ten büyükse mutasyon “Nötr” olarak sınıflandırılır. Gi: Güvenirlik endeksi 0 ile 10 arasında bir değer alır.

### I-Mutant3.0 (Suite)

Mutasyonun protein stabilitesi üzerindeki etkisini tahmin etmek için kullanılan I-Mutant sonucuna göre, 6 güvenilirlik endeksi ve DDG değeri -0.56 ile varyantın protein stabilitesini azalttığı tahmin edilmiştir (Tablo 4.3.).

### MUpro

Protein stabilite tahmini için kullanılan diğer bir sunucu olan MUpro, DDG değeri -1.02, SVM güven skoru -1 ve sinir ağı güven skoru -0.999 olarak varyantın protein stabilitesini azalttığını tahmin etmiştir (Tablo 4.3.).

**Tablo 4.3.** UBXN4 p.E260D I-Mutant 3.0 ve MUpro sonucu.

Mutasyon	dbSNP ID	MUpro			I-Mutant 3.0		
		Destek Vektör Makineleri (SVM)	SVM Güven Skoru	Sinir Ağı Güven Skoru	Stabilite	Serbest Değişim Enerji Değeri (DDG)	
p.Glu260Asp (E260D)	rs748759954	Azalma	-1	Azalma	-0.998	Azalma	-0.56

### HOPE

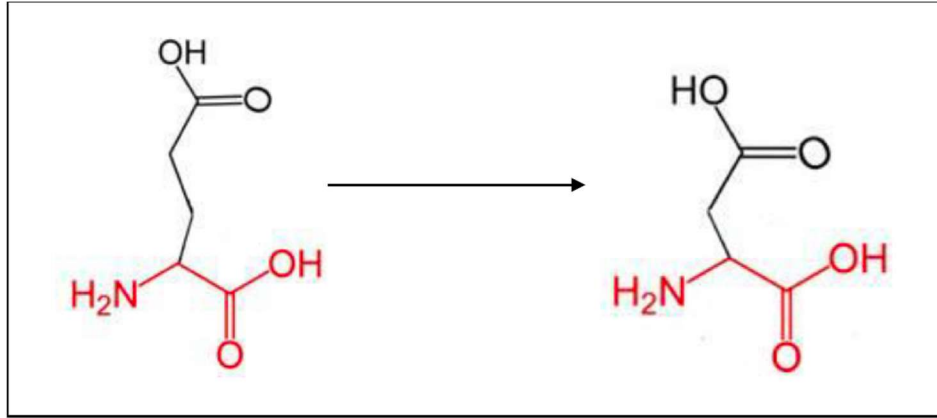
HOPE web sunucusu UBXN4 proteini için herhangi bir 3 boyutlu yapı ya da modellenmiş şablon bulamadığından yalnızca UniProt veri tabanı ve Reprof yazılımından gelen tahmin sonuçlarını vermiştir.

Glutamik asit (orijinal amino asit) (solda) ile Aspartik asidin (mutant amino asit) (sağda) şematik görüntüsü Şekil 4.1.'de gösterilmiştir. Amino asitlerdeki temel zincirler kırmızı, yan zincirler siyah ile belirtilmiştir. Buna göre;

- i) İki amino asit için de temel zincirler aynıdır ancak yan zincirler kendilerine özgüdür.
- ii) Varyant rezidü yabancı tip'ten daha küçüktür.

iii) Yabanıl tip rezidünün  $\alpha$ -heliks yapının içinde olduğu tahmin edilmiş (Reprof yazılımına göre) ve mutant rezidünün (Aspartik asit) ise ikincil yapı olarak  $\alpha$ -heliks bir yapıya yönelmediği tahmin edilmiştir.

iv) Mutasyonun muhtemelen zarar verici bir etkisi olmadığı tahmin edilmiştir.



**Şekil 4.1.** UBXN4 p.E260D amino asit değişiminin şematik görüntüsü. Orijinal amino asit solda, mutant amino asit sağda gösterilmiştir.

#### 4.1.2. UBXN4 Üzerindeki Varyantın Yapısal ve Fonksiyonel Etkilerinin 3 Boyutlu Yapı Kullanılarak Değerlendirilmesi

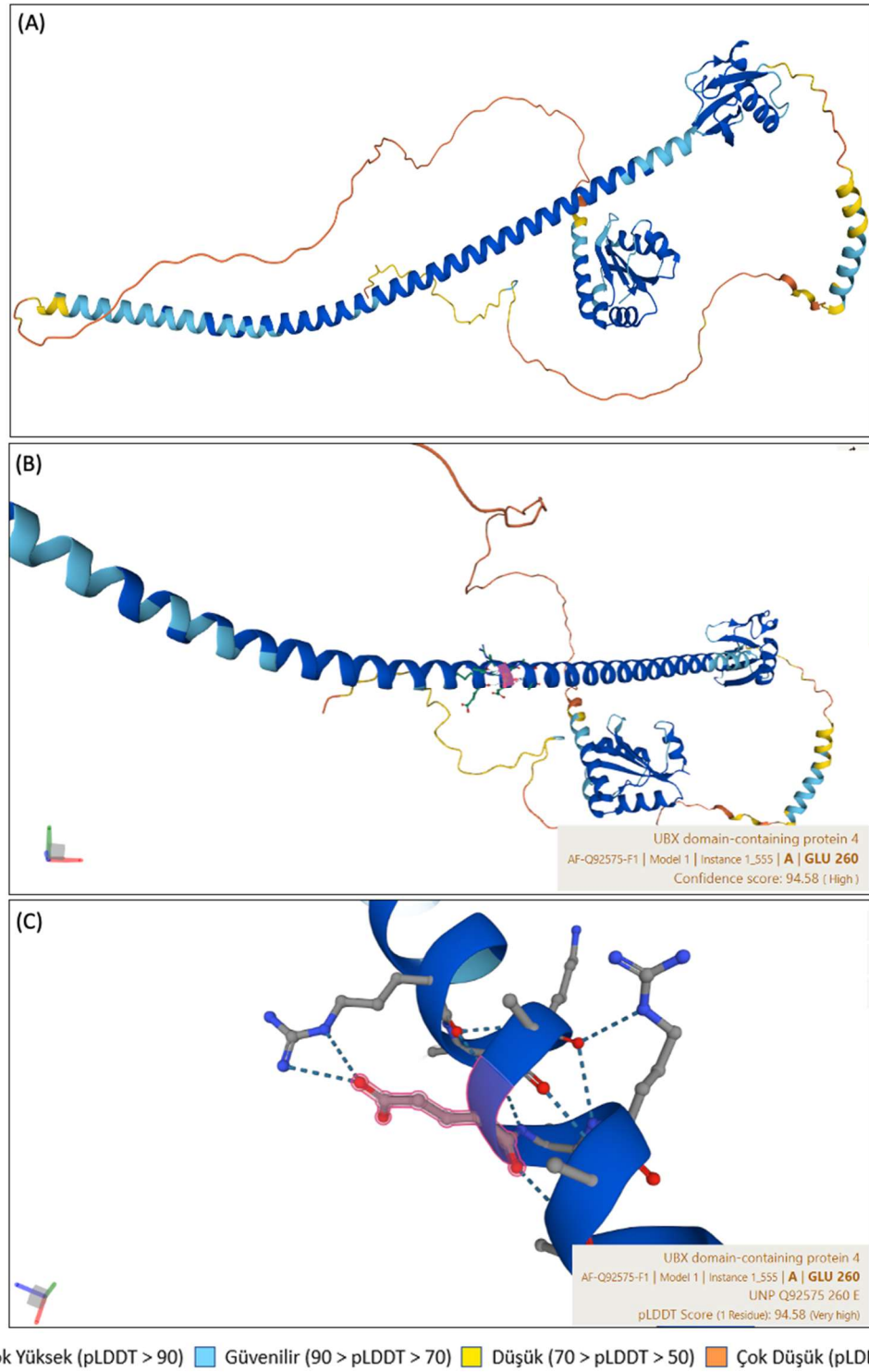
##### UBXN4'ün Üç Boyutlu Yapısının Tahmini

UBXN4 proteinin 3 boyutlu yapısı farklı modelleme yöntemleri uygulayarak protein yapısını tahmin eden programlar kullanılarak tahmin edilmiştir. UBXN4 proteinin 3 boyutlu yapısını tahmin etmek için kullanılan Phyre2 analizi sonucunda; 508 amino asitlik bu proteinin yalnızca %52'si %90'nın üzerinde bir güven yüzdesiyle modellendiği için sonraki araştırmalarda kullanılmamıştır. I-TASSER sonucunda ise, tahmin edilen 5 model arasında en iyisi olan, birinci modelin C skoru -3.40, TM skoru  $0.34 \pm 0.11$  olarak tespit edilmiştir. C skoru -5 ile 2 aralığında bir değer alır. Yüksek bir değere sahip C skoru, yüksek güvenilirliğe sahip bir modeli belirttiğinden, birinci model ileri araştırmalarda kullanılmamıştır.

## **AlphaFold**

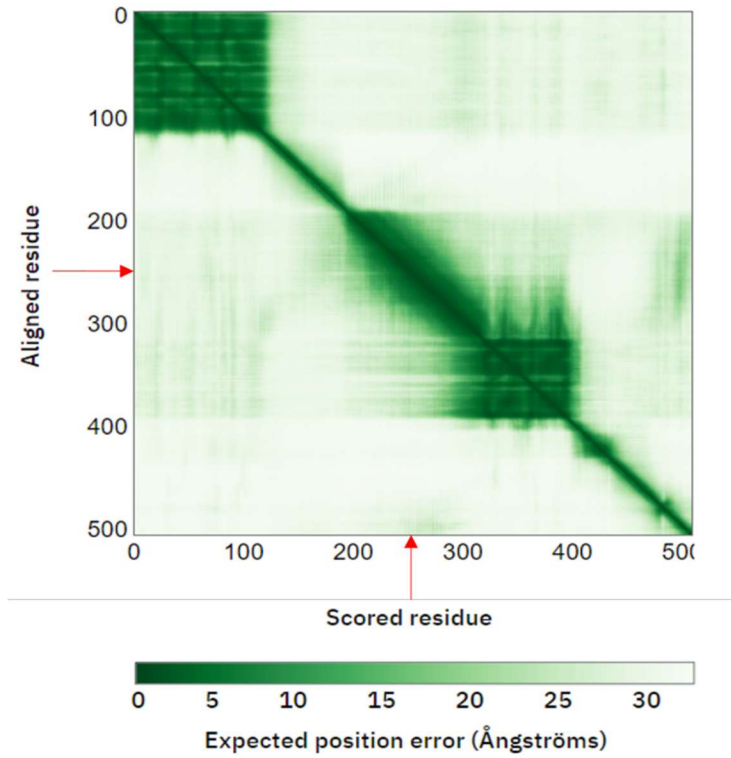
UBXN4 proteini için AlphaFold protein yapı veri tabanından elde edilen 3 boyutlu yapısı pLDDT skorlarıyla birlikte Şekil 4.2.'de gösterilmiştir.





**Şekil 4.2.** A) AlphaFold tarafından oluşturulan UBXN4 proteinin 3 boyutlu yapısı B) AlphaFold tarafından oluşturulan UBXN4 proteinin 3 boyutlu yapısı üzerinde p.E260 amino asidinin konumu pembe renkli olarak gösterilmiştir. C) p.E260 amino asidinin odaklanmış görüntüsü.

Buna ek olarak, 3 boyutlu yapıdan bağımsız, proteinin domainlerinin görel konumlarını nasıl tahmin ettiğine dair bir fikir sunan tahmin edilmiş hizalama hatası (Predicted Aligned Error/PAE) grafiği Şekil 4.3.'te gösterilmiştir. Grafikte; (x, y)'deki renk, tahmin ve gerçek yapı y rezidüsü üzerinde hizalandığında, x rezidüsünün konumunda beklenen mesafe hatasına karşılık gelir. Düşük hata oranı (koyu yeşil renk), yapı içindeki rezidülerin göreceleri konumlarının ne kadar doğru olduğunu gösterir. p.E260 konumu PAE grafiğinde kırmızı ok ile yaklaşık olarak işaretlenmiştir. AlphaFold veri tabanında interaktif bir şekilde kesin konuma denk gelen bölge işaretlenebilmektedir.



**Şekil 4.3.** UBXN4 PAE grafiği. p.E260 amino asidinin (x,y) eksenindeki konumu yaklaşık olarak kırmızı ok ile gösterilmiştir.

İleri araştırmalar için UBXN4'ün AlphaFold tarafından tahmin edilen üç boyutlu modeli kullanılmıştır.

## **Proteinlerin 3 Boyutlu Yapısı Kullanılarak Varyantların Etkilerinin Tahmini**

### **Missense3D**

Missense3D sonucunda göre varyant, proteinin 3 boyutlu yapısı üzerinde herhangi bir yapısal hasara sebep olmadığı saptanmıştır. Missense3D'den elde edilen detaylı bilgi Tablo 4.4.'te gösterilmiştir.

**Tablo 4.4.** UBXN4 Missense3D sonucu.

Özellik	Açıklama
Disülfid bağlarının kırılması	Yabanıl tip rezidü sistein olmadığı için disülfid bağı oluşturamaz.
Gömülü prolin amino asidinin ortaya çıkması	Bu değişim sonucunda prolin oluşmaz.
Çarpışma	Bu değişim çarpışmayı tetiklemez. Yabanıl ve mutant tipin lokal çarpışma skoru 0.00'dır.
Gömülü hidrofilik rezidünün ortaya çıkması	Bu değişim, gömülü bir hidrofobik rezidüyü bir hidrofilik rezidü ile değiştirmez. Yabanıl tip rezidü GLU, %54,6 RSA* ile hidrofilik ve mutant kalıntı ASP, %53,9 RSA ile hidrofilik olarak açıkta kalır.
Gömülü rezidünün yüksüz rezidüden yüklü olarak değişmesi	Bu değişim, gömülü yüksüz kalıntı uyarısını oluşturmaz. Yabanıl tip kalıntı GLU, RSA %54.6 ile yüklenerek, mutant kalıntı ASP, RSA ile %53,9 yüklenerek açıkta kalır.
İkincil yapının değişmesi	Bu değişim ikincil yapıyı değiştirmez.
Gömülü rezidünün yükünün değişmesi	Bu değişim, gömülü rezidünün yükünü değiştirmez. Yabanıl tip kalıntı GLU, RSA %54.6 ile negatif yüklü ve mutant kalıntı ASP, RSA %53.9 ile negatif yüklü olarak açıkta kalır.
İzin verilmeyen phi/psi açıları	Bu değişim, izin verilmeyen phi/psi açılarını tetiklemez. Phi/psi açıları, yabanıl ve mutant tip rezidü için tercih edilen bölgededir.
Gömülü rezidünün yüklü konumdan yüksüz konuma değişmesi	Bu değişim, gömülü yüklü bir rezidüyü yüklenmemiş bir rezidü ile değiştirmez. Yabanıl tip kalıntı GLU, RSA %54.6 ile yüklenmiş olarak, mutant kalıntı ASP, RSA ile %53,9 yüklenerek açıkta kalır.
Gömülü rezidünün glisine dönüşmesi	Bu değişim gömülü Gly rezidüsünün yerini almaz.
Gömülü hidrojen bağlarının kırılması	Bu değişim H bağlarının tamamen bozulmasına neden olmaz ve yabanıl tip rezidü gömülü değildir (RSA %54.6).
Gömülü tuz köprüsünün kırılması	Glu260 rezidüsü ile Arg256 arasında tuz köprüsü bulunmaktadır. Ancak mutant rezidü olan Asp260 da Arg256 ile tuz köprüsü oluşturmaktadır.
Kavite hacminin değişmesi	Mutant yapıdaki kavite hacmi hesaplanamamıştır.
Gömülü ya da açıkta kalan rezidünün RSA*'larındaki farklara göre değerlendirilmesi	Yabanıl tip rezidü Glu %54.6 RSA ile, mutant rezidü Asp %53.9 RSA ile açıkta kalır.
Cis pozisyonunda prolin amino asidinin oluşması	Yabanıl tip rezidü cis pozisyonundaki prolin değildir.
Bir kıvrım bölgesinde glisin amino asidinin bulunması	Yabanıl tip rezidü Gly değildir.

\*RSA: Göreceli Çözücü Erişilebilirliği (*Relative Solvent Accessibility*)

## COACH

Protein ligand bağlanma bölgelerinin tahmini için kullanılan COACH meta sunucu sonuçlarına göre UBXN4 proteini üzerindeki bazı rezidüler bağlanma bölgesi olarak tahmin edilmesine rağmen p.E260D varyantının olduğu rezidü için herhangi bir bağlanma bölgesi saptanamamıştır (Tablo 4.5.).

**Tablo 4.5.** UBXN4 COACH sonucu.

Sıralama	C-Skoru <sup>(a)</sup>	PDB	Ligand Adı <sup>(b)</sup>	Konsensüs Bağlanma Rezidüleri
1	0.05	2c3oA	SF4	323, 334, 335, 336, 386, 387
2	0.05	2d4cA	CA	257, 261
3	0.05	2vrzA	ZN	253, 256
4	0.03	5ccgC	CA	242, 246
5	0.02	4zfzA	ZN	218, 221

(a) **C-Skoru** tahminin güven skorudur ve [0-1] arasında değişir. Yüksek skor daha güvenilir bir tahmini göstermektedir.

(b) **Ligand Adı** muhtemelen bağlanan ligandın adını göstermektedir.

## COFACTOR

COFACTOR sonucuna göre; varyantın gözlemlendiği kodona denk gelen herhangi bir bağlanma bölgesi saptanamamıştır (Tablo 4.6.).

**Tablo 4.6.** UBXN4 COFACTOR sonucu.

C-Skoru <sup>(a)</sup>	PDB	TM-Skoru <sup>(b)</sup>	RMSD <sup>(c)</sup>	BS-Skoru <sup>(d)</sup>	Benzerlik <sup>(e)</sup>	Ligand Adı	Konsensüs Bağlanma Rezidüleri
0.01	3eifA	0.266	7.19	0.73	0.041	NA	12, 13, 14, 15
0.01	2uzaB	0.213	7.22	0.65	0.036	SF4	323, 334, 336, 386, 387
0.01	1b0pA	0.212	7.31	0.50	0.041	SF4	366, 367, 384, 386, 388, 399, 391
0.01	1u2fA	0.237	7.69	0.47	0.021	III	21, 22, 93, 94, 95, 102, 106, 109

(a) **C-Skoru**, tahmin edilen bağlanma bölgesinin güven puanıdır. C-skorlarının değerleri [0-1] arasında değişir; burada daha yüksek bir puan, daha güvenilir bir ligand bağlama yeri tahminini gösterir.

(b) **TM-Skoru**, sorgu ve şablon protein arasındaki küresel yapısal benzerliğin bir ölçüsüdür.

- (c) **RMSD**, TM hizalaması ile yapısal olarak hizalanmış rezidüer arasındaki RMSD.
- (d) **BS-Skoru**, şablon bağlama bölgesi ile sorgu yapısındaki tahmin edilen bağlama bölgesi arasındaki yerel benzerliğin (dizi ve yapı) bir ölçüsüdür. >1 bir BS-skoru tahmin edilmiş ve şablon bağlama bölgesi arasında önemli ölçüde yerel eşleşmenin olduğunu gösterir.
- (e) **Benzerlik**, yapısal olarak hizalanmış bölgedeki yüzde dizi benzerliğidir.

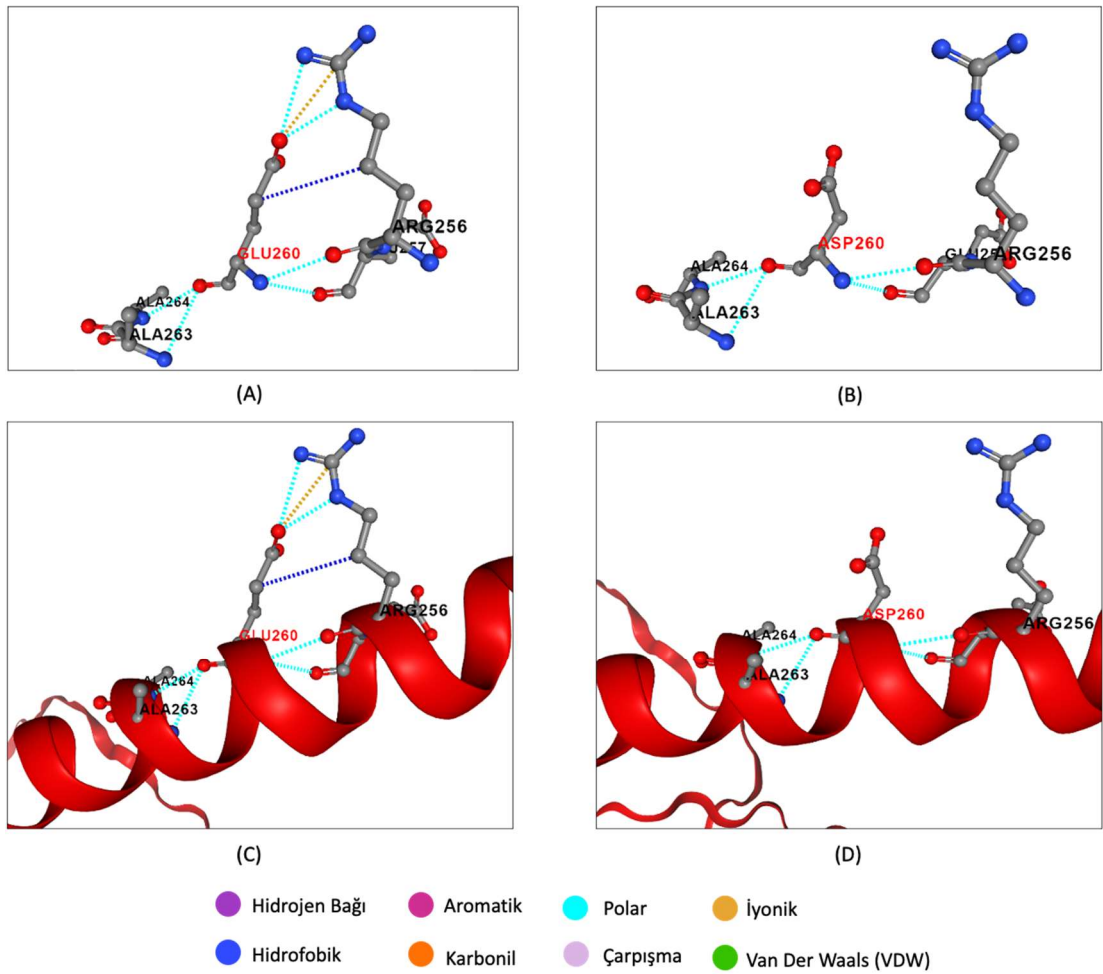
### PremPS

Mutasyonun protein stabilizasyonu üzerine etkisini tahmin etmek için kullanılan PremPS sonucuna göre p.E260D varyantı UBXN4 proteinin stabilizasyonunu bozucu etkisi olduğunu 0.38 DDG değeri ile tahmin edilmiştir. Ayrıca bu araç mutasyonun bulunduğu bölge hakkında da bilgi vermektedir. Eğer protein içindeki bir rezidünün, proteindeki ve uzatılmış tripeptitteki çözücü erişilebilir yüzey alanı oranı 0,2'den az ise, rezidü protein çekirdeğinde gömülü olarak tanımlanır, aksi takdirde proteinin yüzeyinde bulunur. Buna göre p.E260D varyantının protein yüzeyinde olduğu tahmin edilmiştir (Tablo 4.7.).

**Tablo 4.7.** UBXN4 p.E260D PremPS sonucu.

Mutasyona Uğramış Zincir	Mutasyon	$\Delta\Delta G$	Lokasyon
A	p.E260D	0.38	Yüzey

Bunlara ek olarak mutant ve yabancıl tip rezidüerlerin protein üzerindeki diğer amino asitler ile etkileşimleri ve kimyasal bağları Şekil 4.4.'te gösterilmiştir.



**Şekil 4.4.** UBXM4 PremPS sonucu. A) Yabanıl tip p.E260 rezidüsünün diđer amino asitlerin yan zincirleri ile oluşturduđu kimyasal bađların gösterimi. B) Mutant tip p.D260 rezidüsünün diđer amino asitlerin yan zincirleri ile oluşturduđu bađların gösterimi. C) Yabanıl tip p.E260'ın  $\alpha$ -heliks yapısı içindeki gösterimi. D) Mutant tip p.D260'ın  $\alpha$ -heliks yapısı içindeki gösterimi.

## 4.2. ANK2 Proteini

### 4.2.1. ANK2 Üzerindeki Varyantın Yapısal ve Fonksiyonel Etkilerinin Dizi Tabanlı Araçlarla Deđerlendirilmesi

#### PredictSNP

ANK2 proteininde gözlenen p.R1604K varyantının hastalık ilişkili olup olmadığına dair PredictSNP analizi sonucunda; varyant MAPP ve nsSNPAnalyzer dışındaki bütün entegre programlar tarafından "Nötr" olarak sınıflandırmış ve

PredictSNP konsensüs doğruluk oranı yüksek bir değer olan %83 olarak bulunmuştur (Tablo 4.8.).

**Tablo 4.8.** ANK2 p.R1604K PredictSNP sonucu.

Mutasyon	PredictSNP	MAPP	PhD-SNP	PolyPhen-1	PolyPhen-2	SIFT	SNAP	nsSNPAnalyzer	PANTHER
p.R1604K	%83	-	%89	%67	%87	%90	%77	-	%70

■ Zararlı
 ■ Nötr

### MetaSNP

MetaSNP sonucuna göre; p.R1604K varyantı için entegre programlardan PANTHER ve SIFT araçlarından sonuç alınamamıştır. Bununla birlikte diğer araçlar RI 9 ile varyantın hastalık yapıcı etkisini “Nötr” olarak tahmin etmişlerdir (Tablo 4.9.).

**Tablo 4.9.** ANK2 p.R1604K Meta-SNP sonucu.

Mutasyon	PANTHER	PhD-SNP	SIFT	SNAP	Meta-SNP	Gİ
p.R1604K	-	0.098	-	0.295	0.061	9

■ Zararlı
 ■ Nötr

PANTHER, PhD-SNP, SNAP ve Meta-SNP: 0 ile 1 arasında, eğer 0.5'ten büyükse mutasyon “Zararlı” olarak sınıflandırılır. SIFT: Eğer 0.05'ten büyükse mutasyon “Nötr” olarak sınıflandırılır. Gİ: Güvenirlik endeksi 0 ile 10 arasında bir değer alır.

### I-Mutant3.0 (Suite)

I-Mutant sonucuna göre p.R1604K varyantının protein stabilitesini, 5 güvenirlik endeksi ve DDG değeri -0.41 ile, azalttığı tahmin edilmiştir (Tablo 4.10.).

### MUpro

MUpro sonucuna göre ise, DDG değeri -1.08, SVM güven skoru -0.93 ve sinir ağı güven skoru -0.89 olarak varyantın protein stabilitesini azalttığını tahmin etmiştir (Tablo 4.10.).



**Tablo 4.10.** ANK2 p.R1604K MUpuro ve I-Mutant 3.0 sonucu.

Mutasyon	dbSNP ID	MUpuro				I-Mutant 3.0	
		Destek Vektör Makineleri (SVM)	SVM Güven Skoru	Sinir Ağı	Sinir Ağı Güven Skoru	Stabilite	Serbest Değişim Enerji Değeri (DDG)
p.Arg1604Lys (R1604K)	rs776426910	Azalma	-0.93	Azalma	-0.89	Azalma	-0.41

#### 4.2.2. ANK2 Üzerindeki Varyantın Yapısal ve Fonksiyonel Etkilerinin 3 Boyutlu Yapı Kullanılarak Değerlendirilmesi

ANK2 proteinin 3 boyutlu yapısı farklı modelleme yöntemleri uygulayarak protein yapısını tahmin eden programlar ile oluşturulmaya çalışılmıştır. NCBI PSI-BLAST (62) ve HHPred (82) ile yapılan blast (varsayılan parametreler kullanılarak) ile, ANK2 proteinine ait yeterli sayıda homolog yapı bulunamadığından homoloji modelleme yöntemi uygulanamamıştır. Bunun yanında ANK2 transkripti diğer yöntemler kullanarak protein yapısını tahmin eden programlar için çok uzun olduğundan protein yapısı tahmin edilememiştir.

#### 4.3. TRAM2 Proteini

##### 4.3.1. TRAM2 Üzerindeki Varyantın Yapısal ve Fonksiyonel Etkilerinin Dizi Tabanlı Araçlarla Değerlendirilmesi

###### PredictSNP

TRAM2 proteininde gözlenen p.E282K varyantının hastalık ilişkili olup olmadığına dair PredictSNP analizi sonucunda; entegre programlardan biri olan SNAP dışında kalan tüm araçlar varyantı "Nötr" olarak sınıflandırmıştır. PredictSNP sonucu varyantı konsensüs doğruluk oranı %74 ile "Nötr" olarak tahmin etmiştir (Tablo 4.11.).

**Tablo 4.11.** TRAM2 p.E282K PredictSNP sonucu.

Mutasyon	PredictSNP	MAPP	PhD-SNP	PolyPhen -1	PolyPhen -2	SIFT	SNAP	nsSNPAnalyzer	PANTHER
p.E282K	%74	%64	%66	%67	%61	%76	%56	-	%56

■ Zararlı
 ■ Nötr

### Meta-SNP

p.E282K varyantının hastalık yapıcı etkisi, Meta-SNP programına entegre programlardan SNAP dışında RI 4 ile “Nötr” olarak saptanmıştır (Tablo 4.12.).

**Tablo 4.12.** TRAM2 p.E282K Meta-SNP sonucu.

Mutasyon	PANTHER	PhD-SNP	SIFT	SNAP	Meta-SNP	Gİ
p.E282K	0.312	0.299	0.690	0.550	0.298	4

■ Zararlı
 ■ Nötr

PANTHER, PhD-SNP, SNAP ve Meta-SNP: 0 ile 1 arasında, eğer 0.5'ten büyükse mutasyon “Zararlı” olarak sınıflandırılır. SIFT: Eğer 0.05'ten büyükse mutasyon “Nötr” olarak sınıflandırılır. Gİ: Güvenirlik endeksi 0 ile 10 arasında bir değer alır.

### I-Mutant3.0 (Suite)

I-Mutant sonucuna göre p.E282K varyantının protein stabilitesini, 9 güvenirlik endeksi ile, DDG değeri -0.81 olarak azalttığı tahmin edilmiştir (Tablo 4.13.).

### MUpro

MUpro sonucuna göre ise, DDG değeri -1.03, SVM güven skoru -1 ve sinir ağı güven skoru -0.999 olarak varyantın protein stabilitesini azalttığını tahmin etmiştir (Tablo 4.13.).

**Tablo 4.13.** TRAM2 p.E282K MUpro ve I-Mutant 3.0 sonucu.

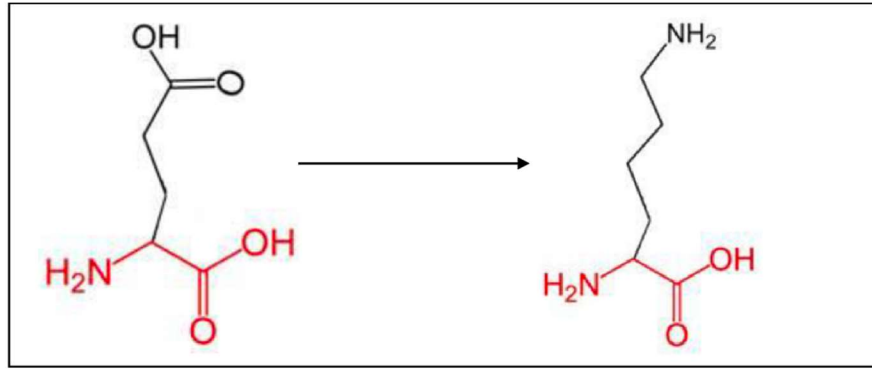
Mutasyon	dbSNP ID	MUpro				I-Mutant 3.0	
		Destek Vektör Makineleri (SVM)	SVM Güven Skoru	Sinir Ağı	Sinir Ağı Güven Skoru	Stabilite	Serbest Değişim Enerji Değeri (DDG)
p.Glu282Lys (E282K)	rs373852555	Azalma	-1	Azalma	-0.999	Azalma	-0.81

### HOPE

HOPE web sunucusu TRAM2 proteini için herhangi bir 3 boyutlu yapı ya da modellenmiş şablon bulamadığından yalnızca UniProt veri tabanı ve Reprof yazılımından gelen tahmin sonuçlarını vermiştir.

Glutamik asit (orijinal amino asit) (solda) ile Lizin'in (mutant amino asit) (sağda) şematik görüntüsü Şekil 4.5.'te gösterilmiştir. Amino asitlerdeki temel zincirler kırmızı, yan zincirler siyah ile belirtilmiştir. Buna göre;

- i) İki amino asit için de temel zincirler aynıdır ancak yan zincirler kendilerine özgüdür.
- ii) Mutasyona uğramış rezidü yabancı tip'ten daha büyüktür.
- iii) Yabancı tip rezidü negatif yüklü, mutant rezidü ise pozitif yüklüdür.
- iv) Mutasyon (p.E282K), proteinin TLC domaininde yer almaktadır.
- v) Yabancı tip rezidü korunmuş bölgede bulunmamaktadır. Mutant rezidü, mutasyonun muhtemelen proteine zarar verdiğini gösterebilecek diğer homolog dizilerde bu konumda gözlemlenen kalıntı türleri arasında değildir.



**Şekil 4.5.** TRAM2 p.E282K amino asit deęişiminin şematik görüntüsü. Orijinal amino asit solda, mutant amino asit sağda gösterilmiştir.

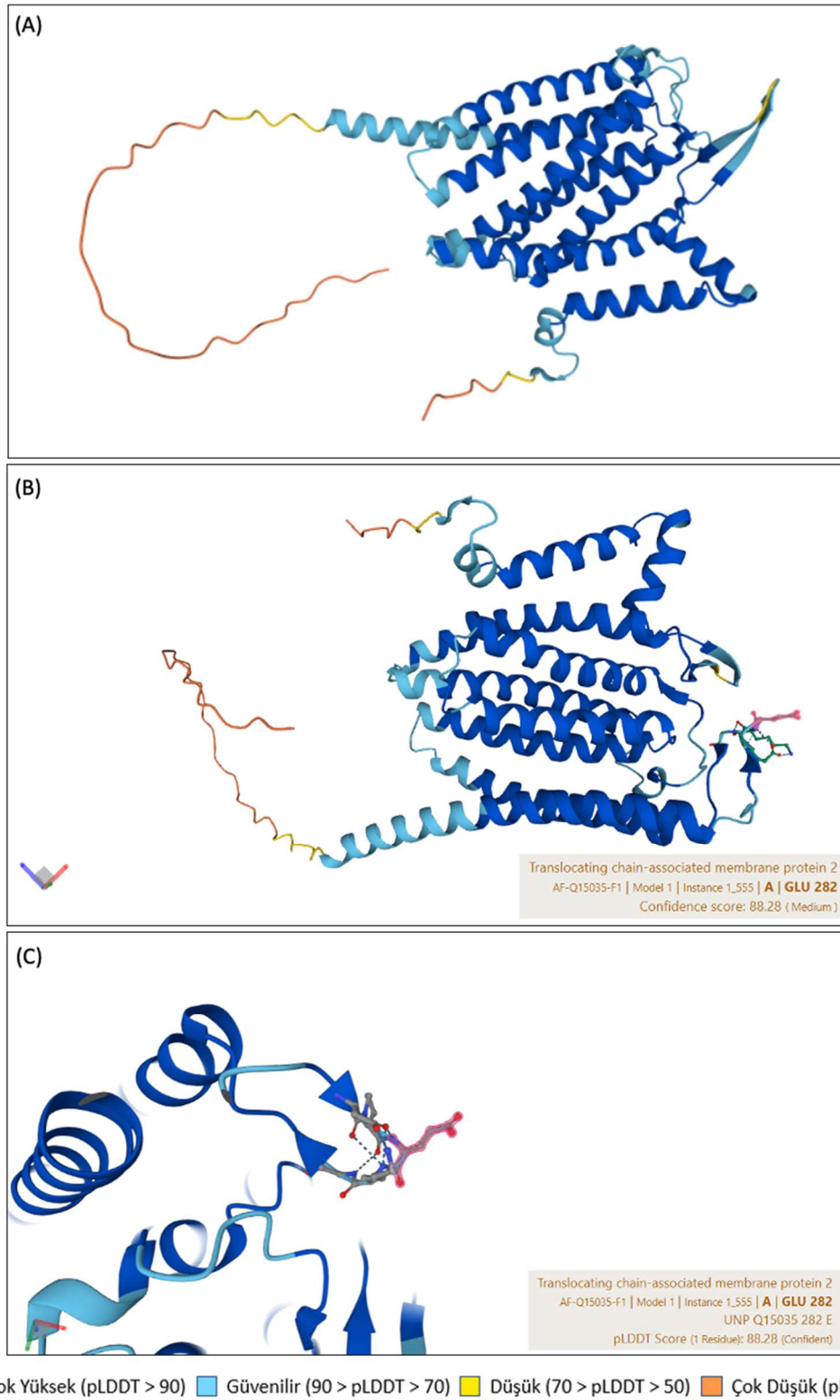
### 4.3.2. TRAM2 Üzerindeki Varyantın Yapısal ve Fonksiyonel Etkilerinin 3 Boyutlu Yapı Kullanılarak Deęerlendirilmesi

#### TRAM2'nin Üç Boyutlu Yapısının Tahmini

TRAM2 proteinin 3 boyutlu yapısını tahmin etmek için kullanılan Phyre2 analizi sonucunda; proteinin hiçbir rezidüsü %90'nın üzerinde bir güven yüzdesiyle modellenmemiştir. Bu nedenle sonraki araştırmalarda kullanılmamıştır. I-TASSER sonucunda ise, tahmin edilen 5 model arasında en iyisi olan, birinci modelin C skoru -1.45, TM skoru  $0.54 \pm 0.15$  olarak tespit edilmiştir. C skoru -5 ile 2 arasında deęişir ve daha yüksek bir deęere sahip C skoru, yüksek güvenirlilięe sahip bir modeli belirttiğinden, birinci model ileri araştırmalarda kullanılmamıştır.

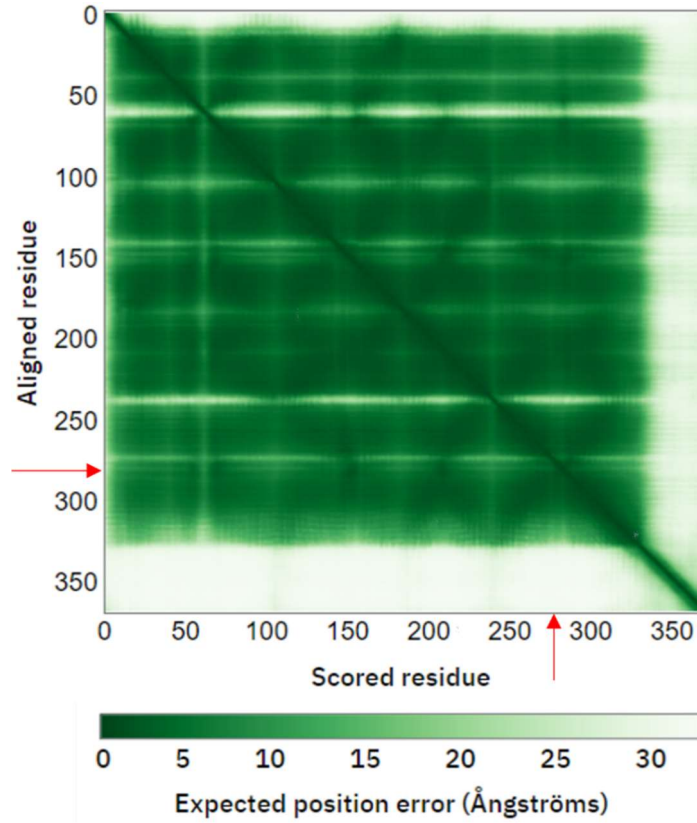
#### AlphaFold

TRAM2 proteini için AlphaFold protein yapı veri tabanından elde edilen 3 boyutlu yapısı pLDDT skorlarıyla birlikte Şekil 4.6.'da gösterilmiştir.



**Şekil 4.6.** A) AlphaFold tarafından oluşturulan TRAM2 proteinin 3 boyutlu yapısı B) AlphaFold tarafından oluşturulan TRAM2 proteinin 3 boyutlu yapısı üzerinde p.E282 amino asidinin konumu pembe renkli olarak gösterilmiştir. C) p.E282 amino asidinin odaklanmış görüntüsü.

Buna ek olarak, 3 boyutlu yapıdan bağımsız, proteinin domainlerinin görel konumlarını nasıl tahmin ettiğine dair bir fikir sunan tahmin edilmiş hizalama hatası (Predicted Aligned Error/PAE) grafiği Şekil 4.7.'de gösterilmiştir. Grafikte; (x, y)'deki renk, tahmin ve gerçek yapı y kalıntısı üzerinde hizalandığında, x kalıntısının konumunda beklenen mesafe hatasına karşılık gelir. Düşük hata oranı (koyu yeşil renk), yapı içindeki rezidülerin göreceleri konumlarının ne kadar doğru olduğunu gösterir. p.E282 konumu PAE grafiğinde kırmızı ok ile yaklaşık olarak işaretlenmiştir. AlphaFold veri tabanında interaktif bir şekilde kesin konuma denk gelen bölge işaretlenebilmektedir. İleri araştırmalar için TRAM2 proteinin AlphaFold tarafından tahmin edilen üç boyutlu yapısı kullanılmıştır.



**Şekil 4.7.** TRAM2 PAE grafiği. E282 amino asidinin (x,y) eksenindeki konumu yaklaşık olarak kırmızı ok ile gösterilmiştir.

### Proteinlerin 3 Boyutlu Yapısı Kullanılarak Varyantların Etkilerinin Tahmini

#### Missense3D

Missense3D sonucunda göre mutasyon, proteinin 3 boyutlu yapısı üzerinde herhangi bir yapısal hasara sebep olmadığı saptanmıştır. Missense3D'den elde edilen detaylı bilgi Tablo 4.14.'te gösterilmiştir.

**Tablo 4.14.** TRAM2 Missense3D sonucu.

Özellik	Açıklama
Disülfid bağlarının kırılması	Yabani tip rezidü sistein olmadığı için disülfid bağı oluşturamaz.
Gömülü prolin amino asidinin ortaya çıkması	Bu değişim sonucunda prolin oluşmaz.
Çarpışma	Bu değişim çarpışmayı tetiklemez. Yabani ve mutant tipin lokal çarpışma skoru sırasıyla 14.94 ve 14.82'dir.
Gömülü hidrofilik rezidünün ortaya çıkması	Bu değişim, gömülü bir hidrofobik rezidüyü bir hidrofilik rezidü ile değiştirmez. Yabani tip rezidü GLU, %98.4 RSA* ile hidrofilik ve mutant kalıntı LYS, %100.0 RSA ile hidrofilik olarak açıkta kalır.
Gömülü rezidünün yüksüz rezidüden yüklü olarak değişmesi	Bu değişim, gömülü yüksüz kalıntı uyarısını oluşturmaz. Yabani tip kalıntı GLU, RSA %98.4 ile yüklenerek, mutant kalıntı LYS, RSA ile %100.0 yüklenerek açıkta kalır.
İkincil yapının değişmesi	Bu değişim ikincil yapıyı değiştirmez.
Gömülü rezidünün yükünün değişmesi	Bu değişim, gömülü rezidünün yükünü değiştirmez. Yabani tip kalıntı GLU, RSA %98.4 ile negatif yüklü ve mutant kalıntı LYS, RSA %100.0 ile negatif yüklü olarak açıkta kalır.
İzin verilmeyen phi/psi açıları	Bu değişim, izin verilmeyen phi/psi açı uyarısını tetiklemez. Phi/psi açıları, yabani ve mutant tip rezidü için tercih edilen bölgededir.
Gömülü rezidünün yüklü konumdan yüksüz konuma değişmesi	Bu değişim, gömülü yüklü bir rezidüyü yüklenmemiş bir rezidü ile değiştirmez. Yabani tip kalıntı GLU, RSA %98.4 ile yüklenmiş olarak, mutant kalıntı LYS, RSA ile %100.0 yüklenerek açıkta kalır.
Gömülü rezidünün glisine dönüşmesi	Bu değişim gömülü Gly rezidüsünün yerini almaz.
Gömülü hidrojen bağlarının kırılması	Bu değişim H bağlarının tamamen bozulmasına neden olmaz ve yabani tip rezidü gömülü değildir (RSA %98.4).
Gömülü tuz köprüsünün kırılması	Yabani tip rezidü herhangi bir tuz köprüsü oluşturmaz.
Kavite hacminin değişmesi	Mutant yapıdaki kavite hacmi hesaplanamamıştır.
Gömülü ya da açıkta kalan rezidünün RSA*'larındaki farklılara göre değerlendirilmesi	Yabani tip rezidü Glu %98.4 RSA ile, mutant rezidü Lys %100.0 RSA ile açıkta kalır.
Cis pozisyonunda prolin amino asidinin oluşması	Yabani tip rezidü cis pozisyonundaki prolin değildir.
Bir kıvrım bölgesinde glisin amino asidinin bulunması	Yabani tip rezidü Gly değildir.

\*RSA: Göreceli Çözücü Erişilebilirliği (*Relative Solvent Accessibility*)

## COACH

COACH sonuçlarına göre TRAM2 proteini üzerindeki bazı rezidüler bağlanma bölgesi olarak tahmin edilmesine rağmen mutasyonun (p.E282K) olduğu rezidü için herhangi bir bağlanma bölgesi tahmin edilmemiştir (Tablo 4.15.).

**Tablo 4.15.** TRAM2 COACH sonucu.

Sıralama	C-Skoru <sup>(a)</sup>	PDB	Ligand Adı <sup>(b)</sup>	Konsensüs Bağlanma Rezidüleri
1	0.09	2y04B	2CV	259, 260, 299, 303, 306
2	0.05	2bu8A	TF4	221, 252, 253, 256, 303, 306, 307, 310
3	0.04	4m70A	PEPTIDE	224, 227, 251, 255
4	0.04	4d2bA	78N	227, 235, 248
5	0.04	4a01A	DMU	133, 136, 139

(a) **C-Skoru** tahminin güven skorudur ve [0-1] arasında değişir. Yüksek skor daha güvenilir bir tahmini göstermektedir.

(b) **Ligand Adı** muhtemelen bağlanan ligandın adını göstermektedir.

## COFACTOR

COFACTOR sonucuna göre TRAM2 proteininde tahmin edilen bazı bağlanma rezidüleri Tablo 4.16.'da gösterilmiştir. Mutasyonun gözlemlendiği 282. kodon ile ilgili herhangi bir sonuç elde edilememiştir.

**Tablo 4.16.** TRAM2 COFACTOR sonucu.

C-Skoru <sup>(a)</sup>	PDB	TM-Skoru <sup>(b)</sup>	RMSD <sup>(c)</sup>	BS-Skoru <sup>(d)</sup>	Benzerlik <sup>(e)</sup>	Ligand Adı	Konsensüs Bağlanma Rezidüleri
0.01	1i1eA	0.422	6.59	0.52	0.033	DM2	147, 148, 211, 213, 214
0.01	1f31A	0.419	6.67	0.52	0.018	UUU	151, 207, 208, 211, 213, 214
0.01	1hkbA	0.399	5.41	0.63	0.052	CA	132, 133, 135, 139
0.01	3fg4C	0.383	5.95	0.92	0.037	ACD	164, 168, 202, 203, 206
0.01	1qhaB	0.393	5.45	0.74	0.060	MG	220, 224, 253, 256



- (a) **C-Skoru**, tahmin edilen bağlanma bölgesinin güven puanıdır. C-skorlarının değerleri [0-1] arasında değişir; burada daha yüksek bir puan, daha güvenilir bir ligand bağlama yeri tahminini gösterir.
- (b) **TM-Skoru**, sorgu ve şablon protein arasındaki küresel yapısal benzerliğin bir ölçüsüdür.
- (c) **RMSD**, TM hizalaması ile yapısal olarak hizalanmış rezidüler arasındaki RMSD.
- (d) **BS-Skoru**, şablon bağlama bölgesi ile sorgu yapısındaki tahmin edilen bağlama bölgesi arasındaki yerel benzerliğin (dizi ve yapı) bir ölçüsüdür. >1 bir BS-skoru tahmin edilmiş ve şablon bağlama bölgesi arasında önemli ölçüde yerel eşleşmenin olduğunu gösterir.
- (e) **Benzerlik**, yapısal olarak hizalanmış bölgedeki yüzde dizi benzerliğidir.

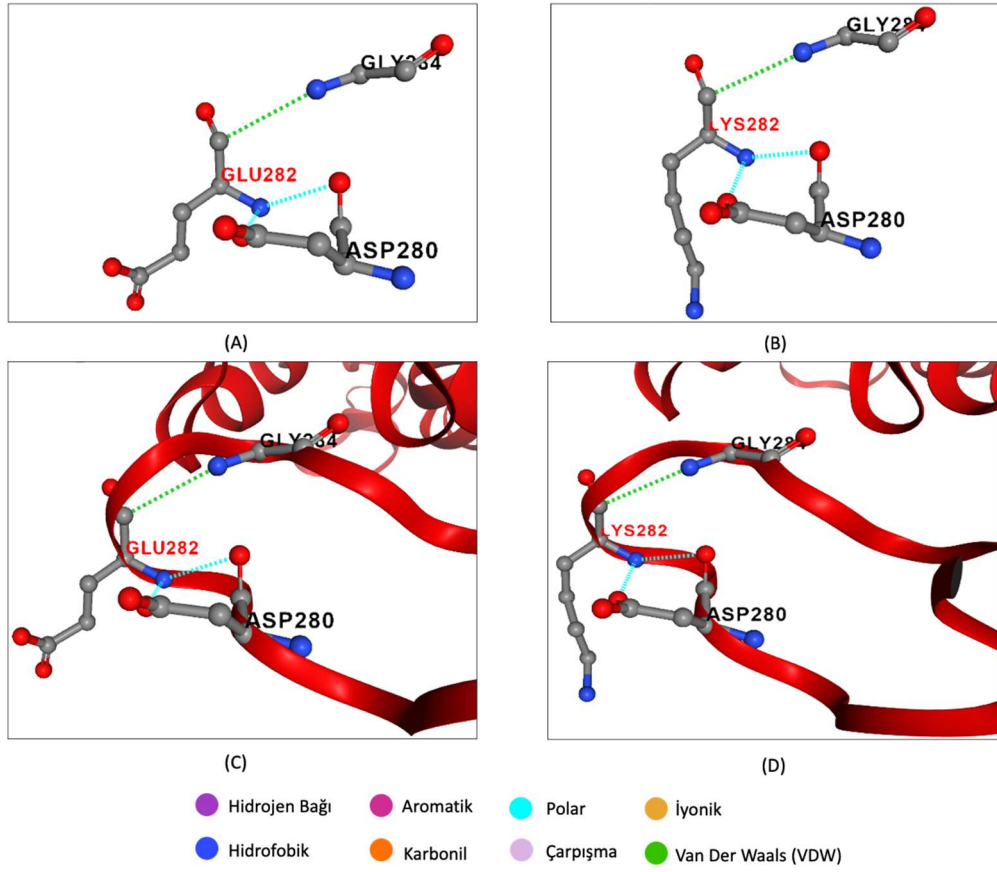
### PREMPS

PremPS sonucuna p.E282K varyantının TRAM2 proteinin stabilizasyonunu bozucu etkisi olduğu 0.09 DDG değeri ile göstermiştir. Ek olarak söz konusu mutasyonun protein yüzeyinde olduğu tahmin edilmiştir (Tablo 4.17.).

**Tablo 4.17.** TRAM2 E282K PremPS sonucu.

Mutasyona Uğramış Zincir	Mutasyon	$\Delta\Delta G$	Lokasyon
A	E282K	0.09	Yüzey

Mutant ve yabanıl tip rezidülerin protein üzerindeki diğer amino asitler ile etkileşimleri ve kimyasal bağları Şekil 4.8.'de gösterilmiştir.



**Şekil 4.8.** TRAM2 PremPS sonucu. A) Yabanıl tip E282 rezidüsünün diđer amino asitlerin yan zincirleri ile oluşturduđu kimyasal bađların gösterimi. B) Mutant tip K282 rezidüsünün diđer amino asitlerin yan zincirleri ile oluşturduđu bađların gösterimi. C) Yabanıl tip E282'nin 3 boyutlu yapı içindeki gösterimi. D) Mutant tip K282'nin 3 boyutlu yapı içindeki gösterimi.

## 5. TARTIŞMA

Frontonazal malformasyonlar orta yüz hattında gelişen ve deęişken fenotipik özellikler gösteren sendromlardır. Frontonazal nadir bir yüz malformasyonu fenotipi gösteren bir ailedeki etkilenmiş 2 bireye ait, tez çalışması öncesinde, yapılan analizler sonucunda 3 aday gen üzerindeki 3 aday yanlış anlamlı mutasyon tespit edilmiştir. Yanlış anlamlı mutasyonların proteinler üzerinde yapısal ve fonksiyonel etkileri dizi tabanlı ve yapı tabanlı araçlar kullanılarak *in silico* olarak araştırılmıştır. Her 3 proteine ait bütüncül 3 boyutlu yapıları PDB’de aratılmış ancak bu üç gene ait proteinlerin bütün bölgelerini kapsayan protein yapıları veri tabanında bulunamamıştır. Bu nedenle yapı tabanlı araçları kullanılmadan önce proteinlerin 3 boyutlu yapıları oluşturulması amaçlanmıştır. 3 aday varyantın dizi ve yapı tabanlı *in silico* değerlendirme sonuçları Tablo 5.1.’de özetlenmiş ve aşağıda her bir varyant için tartışılmıştır.

**Tablo 5.1.** 3 Aday genin kodladığı proteinlere ve varyantlara ait in silico araçlardan elde edilen bütün bulgular.

<b>Kullanılan Araçlar</b>	<b>UBXN4 p.E260D</b>	<b>ANK2 p.R1604K</b>	<b>TRAM2 p.E282K</b>
<b>Dizi Tabanlı Araçlar</b>			
PredictSNP	Nötr	Nötr	Nötr
Meta-SNP	Nötr	Nötr	Nötr
HOPE	Mutant, yabancı tipe göre daha küçük rezidüye sahip, yükleri değişmedi, zarar verici etkisi yok	Sonuç alınamadı.	Mutant, yabancı tipe göre daha büyük rezidüye sahip, yükleri değişti.
I-Mutant 3.0	Stabiliteyi azaltıcı etki	Stabiliteyi azaltıcı etki	Stabiliteyi azaltıcı etki
MUpro	Stabiliteyi azaltıcı etki	Stabiliteyi azaltıcı etki	Stabiliteyi azaltıcı etki
<b>Yapı Tabanlı Araçlar</b>			
Missense3D	Varyantın proteinin 3B yapısında zarar verici etkisi olmadığı tahmin edildi.	Uygulanamadı.	Varyantın proteinin 3B yapısında zarar verici etkisi olmadığı tahmin edildi.
PremPS	i) Destabilize etki ii) Rezidü yüzeyde yer alıyor. iii) Değişim sonucu hidrofobik ve karbonil etkileşimleri kayboluyor.	Uygulanamadı.	i) Destabilize etki ii) Rezidü yüzeyde yer alıyor. iii) Değişim sonucu kovalent olmayan etkileşimlerde bir farklılık bulunmuyor.
COACH	Varyantın bulunduğu rezidü için ligand bağlanma bölgesi tahmin edilmedi.	Uygulanamadı.	Varyantın bulunduğu rezidü için ligand bağlanma bölgesi tahmin edilmedi.
COFACTOR	Varyantın bulunduğu rezidü için ligand bağlanma bölgesi tahmin edilmedi.	Uygulanamadı.	Varyantın bulunduğu rezidü için ligand bağlanma bölgesi tahmin edilmedi.

### 5.1. UBXN4 p.E260D

Liang ve ark.'larının yaptığı çalışmada (19) UBXN4'ün ER ve nükleer zarın ayrılmaz bir membran proteini olduğu ve UBXD2'nin hem N- hem de C-terminalleri sitoplazmaya veya nükleoplazmaya bakacak şekilde yönlendirildiği bulunmuştur. UBXN4'ün ERAD sırasında görev alan VCP gibi proteinlerle aynı yolakta yer aldığı Şekil 2.3.'te görülmektedir.

UBXN4 proteini üzerinde aday olarak tespit edilmiş p.E260 kodundaki aspartik asit dönüşümünün yapısal ve fonksiyonel etkisini araştırmak için kullanılan dizi tabanlı araçlardan PredictSNP analizi sonucunda konsensüs değerlendirilmede "Nötr" tahmin edilmiştir. Ancak nötr sonucu %60 doğruluk oranı ile verilmiştir. Bununla birlikte entegre araçlardan bazıları yine yaklaşık %60 beklenen doğruluk oranı ile "Zararlı" olarak tahmin etmiştir.

Protein stabilite tahmin araçları sonucuna göre ise varyantın proteinin stabilitesini azalttığını tahmin etmesine rağmen kullanılan diğer araçların tahminleri ile değerlendirildiğinde mutasyonun zararlı bir etkisinin olmadığı tahmin edilmiştir.

Varyantın UBXN4 üzerindeki yapısal ve fonksiyonel etkisinin proteinin üç boyutlu yapısını girdi olarak kabul eden araçlar kullanılarak tahmin edilebilmesi için UBXN4 proteini, PDB veri tabanında aranmış, fakat proteine ait proteinin bütün bölgelerini kapsayan bir 3B yapı bulunamadığından proteinin modellenmesi yoluna gidilmiştir.

UBXN4 proteinin 3 boyutlu yapısını tahmin etmek için öncelikle yaygın olarak kullanılan homoloji modelleme yöntemleri denenmiştir. Ancak bu yöntem için kullanılacak hem dizi benzerliği yüksek (>%50) hem de proteinin %40'ını kapsayacak bir yapı bulunamamıştır. UBXN4 için şablon olarak kullanılacak ve yüksek dizi benzerliği gösteren PDB yapısı yalnızca yaklaşık 100 amino asitlik bölgeyi kapsamaktadır (2KXJ, amino asit aralığı: 317-397, (25)). Bu nedenle hedef dizi ile şablon yapı arasındaki dizi benzerliği düşük olduğundan (<%40) katlanma tanıma ve

ab initio yöntemleri denenmiştir. I-Tasser ve Phyre2 araçları bu iki yöntemi birlikte kullanan programlardır. Bu programlar kullanılarak oluşturulan protein modellerinin uzaysal ortamdaki 3 boyutlu yapıları birbirinden farklı çıkmıştır. Özellikle Phyre2 aracı tahmin için birden fazla kez tekrarlanmış, farklı 3 boyutlu yapılar elde edilmiş ve UBXN4'ün %48'lik bir alanı düşük güvenilirlikte modellenmiştir. Buna ek olarak I-Tasser'ın oluşturduğu en iyi modelin güven skoru, alt sınıra yakın bir değer olan -3.40 olarak bulunmuştur. I-Tasser'ın oluşturduğu modelde, UBXN4'ün PDB'de bulunan UBX domainine ait yapıyı (2KXJ) şablon olarak kullanılması gerektiği halde kullanılmadığı görülmüştür.

Kullanılan diğer araç AlphaFold ile elde edilen UBXN4 proteinin 3 boyutlu yapısı ve özellikle mutasyonun olduğu pozisyon genel olarak yüksek güvenilirlikte tahmin edilmiştir (Bkz. Şekil 4.2.(B)). Ayrıca PAE grafiğine göre rezidülerin göreceli konumlarının doğru bir şekilde hizalandığı görülmektedir (Bkz. Şekil 4.3.). AlphaFold tarafından oluşturulan yapıda yaklaşık 100 rezidülük bir bölgeyi kapsayan bir  $\alpha$ -heliks yapısı görünmektedir. (Bkz. Şekil 4.2.(A)). Ancak membran proteinlerinde lipit çift katmanını geçen bir  $\alpha$ -heliks yapısı yaklaşık 20-30 amino asit içerecek kadar uzundur (83). Bu nedenle Alphafold tarafından oluşturulan UBXN4'ün 3 boyutlu yapısı güvenilirlik açısından uygun bulursa da biyolojik olarak doğru bir katlanma olmayabilir.

Farklı protein modelleme yöntemleri denenmesi sonucunda en uygun model olan AlphaFold modeli kullanılarak varyantın proteinin üç boyutlu yapısı üzerindeki etkisini tahmin etmek için çeşitli araçlar kullanılmıştır.

Protein ligand bağlanma bölgelerini tahmin eden COACH ve COFACTOR analizleri sonucunda; 260. Rezidüye denk gelen bir ligand bulunamamıştır. UBXN4 proteini üzerinde farklı rezidülere denk gelen ligandlar tahmin edilmiştir ancak tahminin güvenilirliği düşük olarak gözlenmiştir (Bkz. Sırasıyla Tablo 4.5. ve 4.6.).

Varyantın 3 boyutlu yapı üzerindeki yakınında bulunan diğer amino asitlerin yan zincirleri ile yaptığı kovalent olmayan etkileşimleri de tahmin eden bir araç olan PremPS sonucunda ise Glutamik asitten aspartik asite dönüşen 260. rezidüde

karbonil, polar ve hidrofobik etkileşimlerinin kaybolduğu gözlenmiştir (Bkz. Şekil 4.4.). 260. rezidü UBXN4 3 boyutlu yapısı içerisinde membran zarından geçen  $\alpha$ -heliks yapısında bulunmaktadır (Bkz. Şekil 4.2.). Bu kovalent olmayan etkileşimlerin kaybolmasıyla UBXN4'ün 3 boyutlu yapısındaki uzun  $\alpha$ -heliks yapıyı kırabileceği düşünülebilir. Ancak Nilsson ve ark.'larının yaptığı bir çalışmaya göre transmembran sarmalına yerleştirilmiş Prolin rezidülerinin ancak yalnızca uca yakın yerleştirildiğinde ve yalnızca heliks yeterince uzun olduğunda bir transmembran sarmalını kırabileceğini bildirmişlerdir (84).

Sonuç olarak yapı tabanlı analizlerle birlikte değerlendirildiğinde UBXN4 üzerinde bulunan aday varyantın (p.E260D) protein üzerinde, yapısını değiştirecek ya da hastalığa neden olabilecek bir varyant olmayabileceği tahmin edilmiştir.

## 5.2. ANK2 p.R1604K

3957 amino asit uzunluğunda olan ANK2 proteini üzerindeki p.R1604K varyantı için yapılan dizi tabanlı analizler sonucunda; varyantın, proteinin stabilitesini azalttığı tahmin edilmiş ve PredictSNP ve MetaSNP meta sunucuları hastalık yapıcı etkisini yüksek güvenirlilikte "Nötr" olarak tahmin etmiştir. Bir diğer dizi tabanlı program olan HOPE, kendi bilgisayar kapasitesi nedeniyle ANK2 proteininin amino asit dizisini analiz edememiştir.

Varyantın ANK2 üzerindeki yapısal ve fonksiyonel etkisinin tahmin etmek için dizi bilgisini kullanan araçlardan sonra üç boyutlu protein yapısını girdi olarak alan araçların kullanılabilmesi için ANK2 proteini modellenmesi denenmiştir.

ANK2 proteininin 3 boyutlu yapısını tahmin etmek için öncelikle homoloji modelleme yöntemi değerlendirilmiştir. Ancak şablon olarak kullanılacak PDB yapıları, proteinin küçük bir kısmını kapsadığı, yaklaşık 2000 amino asitlik bir kısmı için herhangi bir şablon yapı bulunamadığından (Bkz. Şekil 2.6.) ANK2 proteininin 3 boyutlu yapısı karşılaştırmalı modelleme yöntemi ile oluşturulamamıştır. I- Tasser ve Phyre2 araçları ise en uzun 1500 amino asitlik proteinleri modellediğinden bu

programlar da kullanılamamıştır. AlphaFold veri tabanı ise en fazla 2700 amino asit uzunluğundaki proteinleri içermektedir. Bu nedenle ANK2 için 3 boyutlu yapı elde edilememiş, dolayısıyla yapı tabanlı araçlar kullanılamamıştır.

Bununla birlikte, OMIM’de yalnızca kalple ilgili fenotipik özelliklerle ilişkilendirilmiş *ANK2* geni, diğer aday varyantlara göre daha yüksek ihtimalle frontozal nadir bir yüz malformasyonuna neden olmayacağı düşünülmektedir.

### 5.3. TRAM2 p.E282K

TRAM2 proteini üzerinde tespit edilen p.E282K varyantı, dizi tabanlı meta sunucular tarafından “Nötr” olarak tahmin edilmiştir.

TRAM2’nin 3 boyutlu yapısını tahmin etmek için, homoloji modellemede şablon olarak kullanılacak bir yapı bulunamadığından, I-Tasser ve Phyre2 araçları kullanılmıştır. Phyre2 analizi sonucunda TRAM2’nin 3 boyutlu modeli düşük güvenilirlikte modellendiğinden ileri araştırmalar için seçilmemiştir. Benzer şekilde I-Tasser sonucunda da en iyi model olarak seçilen modelin güven skoru -1.45 olarak bulunmuştur. Sonuçlar değerlendirildiğinde TRAM2 için en uygun olabilecek model I-Tasser’ın oluşturduğu model olabilirdi ancak AlphaFold tarafından oluşturulan 3 boyutlu model çok yüksek güvenilirlikte tahmin edilmiştir (Bkz. Şekil 4.6.). Ayrıca PAE grafiği incelendiğinde rezidü başına hizalama hatasının çok düşük olduğu ve böylelikle rezidülerin göreceli konumlarının doğru olarak hizalandığı görülmektedir (Bkz. Şekil 4.7.).

TRAM2’nin 112 ile 321 amino asitler arasında TLC domaini, LAG1 (CERS1) ve CLN8 proteinlerinde TLC domainleriyle homoloji göstermektedir. Buna göre TRAM2’nin 3 boyutlu yapısında en az 5 transmembran  $\alpha$ -heliks yapısı görülmesi beklenebilir. AlphaFold veri tabanında LAG1 ve CLN8 proteinleri de TRAM2 ile benzer yapı göstermektedir (85-87) .

Varyantın proteinin 3 boyutlu yapısı üzerindeki etkisi araştırıldığında, protein ligand bağlanma bölgelerini tahmin eden COACH ve COFACTOR analizleri sonucunda;



282. Rezidüye denk gelen bir ligand bulunamamıştır. TRAM2 proteini üzerinde farklı rezidülere denk gelen ligandlar tahmin edilmiştir ancak tahminin güvenilirliğini gösteren skorlar değerlendirildiğinde bu tahminlerin güvenirligi düşük olarak gözlenmiştir (Bkz. Sırasıyla Tablo 4.15. ve 4.16.).

Yapı tabanlı analizlerle birlikte değerlendirildiğinde TRAM2 üzerinde bulunan aday varyantın (p.E282K) protein üzerinde, yapısını değiştirecek ya da hastalığa neden olabilecek bir varyant olmayabileceğini göstermektedir.

#### **5.4. Varyantların Protein Üzerindeki ya da Fenotip ile ilişkili Etkilerinin Gösterilememesi ile İlgili Olabilecek Etkenler**

Varyantların patojenik etkisini tahmin etmek için kullanılan farklı tahmin araçları bulunmaktadır ve bu farklı araçlar değerlendirildiğinde bu araçların tahminlerinin kalitesinde sistematik farklılıklar olup olmadığı sorulabilir. Bu amaçla varyant patojenite tahmin araçlarının karşılaştırmalı değerlendirilmesinin yapıldığı bir çalışmada, bu karşılaştırmayı bozan bir döngüsellik tespit edilmiş ve bu döngüsellik dikkate alınmadan değerlendirilen araçların performanslarının yapay olarak artmasına neden olduğu gösterilmiştir (88). Bu yapay artış, makine öğrenme yöntemlerinde kullanılan eğitim ve test veri setlerindeki varyantların örtüşmesi ve aşırı uyum göstermesidir. Bu da bir aracın belirli bir veri setine yüksek oranda uyarlandığı ancak yeni varyantlarda daha kötü performans göstereceği anlamına gelebileceği öne sürülmüştür (88).

Meta-SNP ve PredictSNP'in de içinde olduğu klinikle ilişkili yanlış anlamlı mutasyonlar üzerinde tahmin araçlarının performanslarının değerlendirildiği güncel bir çalışmada (89), yukarıda bahsedilen karşılaştırmayı önleyici döngüden kaçınarak hazırlanmış veri setlerinde, bu araçlar genel olarak orta düzeyde bir performans gösterdiği bildirilmiştir. Bütün bu değerlendirmelere göre yalnızca dizi tabanlı araçlara göre varyantların etkilerinin değerlendirmesi yeterli olmayabilir. Bu nedenle kullanılan dizi tabanlı araçlara ek olarak yapı tabanlı araçların kullanılması önemli olabilir.

Frontonazal nadir bir yüz malformasyonu ile ilişkilendirilen ve yapılan analizler sonucunda aday olarak belirlenen bu üç gen ve varyant dışında hastalığa neden olabilecek farklı genetik değişiklikler söz konusu olabilir. Örneğin kopya sayısı değişiklikleri (Copy Number Variation, CNV) kalıtsal hastalıklarla ilişkilendirildiği (90) için bu fenotipe neden olabilecek farklı genetik değişikliklerden birisi olabilir.

Etkilenmiş 2 birey için gerçekleştirilen WES analizi, insan genomunda yalnızca protein kodlayan bölgelere yönelik bir dizileme yöntemidir. Literatürde protein kodlamayan (intronik) bölgelerde gözlenen genomik değişiklikler kalıtsal hastalıklarla ilişkilendirildiğinden (91) bu bölgelere yönelik tüm genom dizileme çalışmasıyla fenotiple ilişkili olabilecek genomik değişiklikler araştırılabilir.

## 6. SONUÇ VE ÖNERİLER

### 6.1. Sonuç

Çalışmamızda kullanılan *in silico* araçlardan elde ettiğimiz tahmin sonuçları, frontozal nadir bir yüz malformasyonu olan etkilenmiş 2 bireyde tespit edilen ve klinik önemi bilinmeyen aday varyantların protein yapısı ve fonksiyonunda herhangi bir etkisinin olmama olasılıklarını göstermektedir.

### 6.2. Öneriler

1. Genetik hastalıkların alta yatan mekanizması yalnızca yanlış anlamlı mutasyonlar değildir. Buna göre bu ailenin genomik DNA'sında kodlanmayan bölgelerin ve kopya sayısı değişikliklerinin analiz edilebilir.

2. Aday varyantların üzerinde oldukları proteinlerin etkileşimde olabilecekleri diğer proteinler ve moleküller araştırılarak kenetlenme (*docking*) simülasyonları ile bu varyantların kenetlenme üzerindeki etkileri araştırılabilir.

3. Güncel literatür taranarak, kullanılacak yeni araçlarla tekrar *in silico* analiz gerçekleştirilebilir. Ayrıca bu tez çalışmasında kullanılan araçların güncel versiyonları yayınlandığında da farklı sonuçlar elde edilebileceğinden, versiyon güncellemeleriyle birlikte proteinlerin etkisi yeniden araştırılabilir.

## 7. KAYNAKLAR

1. Sedano HO, Gorlin RJ. Frontonasal Malformation as a Field Defect and in Syndromic Associations. *Oral Surg Oral Med O.* 1988;65(6):704-10.
2. Sedano HO, Cohen MM, Jirasek J, Gorlin RJ. Frontonasal Dysplasia. *J Pediatr.* 1970;76(6):906-&.
3. Twigg SRF, Versnel SL, Nurnberg G, Lees MM, Bhat M, Hammond P, et al. Frontorhiny, a Distinctive Presentation of Frontonasal Dysplasia Caused by Recessive Mutations in the ALX3 Homeobox Gene. *Am J Hum Genet.* 2009;84(5):698-705.
4. Mindikoglu AN, Erginel A, Cenani A. An Unknown Syndrome of Nose Deformity, Oxycephaly, Aplasia of the Nasolacrimal Ducts, and Symmetrical Cyst Formation on the Upper Lip in Siblings - Craniorhiny. *Plast Reconstr Surg.* 1991;88(4):699-702.
5. Gardner JF, Cushion TD, Niotakis G, Olson HE, Grant PE, Scott RH, et al. Clinical and Functional Characterization of the Recurrent TUBA1A p.(Arg2His) Mutation. *Brain Sci.* 2018;8(8).
6. Sucularli C, Arslantas M. Computational prediction and analysis of deleterious cancer associated missense mutations in DYNC1H1. *Mol Cell Probe.* 2017;34:21-9.
7. Pecimonova M, Kluckova D, Csicsay F, Reblova K, Krahulec J, Prochazkova D, et al. Structural and Functional Impact of Seven Missense Variants of Phenylalanine Hydroxylase. *Genes-Basel.* 2019;10(6).
8. Zhang BB, Ma H, Khan T, Ma A, Li T, Zhang H, et al. A DNAH17 missense variant causes flagella destabilization and asthenozoospermia. *J Exp Med.* 2020;217(2).
9. Zeitz C, Mejecase C, Stevenard M, Michiels C, Audo I, Marmor MF. A Novel Heterozygous Missense Mutation in GNAT1 Leads to Autosomal Dominant Riggs Type of Congenital Stationary Night Blindness. *Biomed Res Int.* 2018;2018.
10. Kannengiesser C, Borie R, Menard C, Reocreux M, Nitschke P, Gazal S, et al. Heterozygous RTEL1 mutations are associated with familial pulmonary fibrosis. *Eur Respir J.* 2015;46(2):474-85.
11. Carlson B. *Human Embryology and Developmental Biology* Saint Louis: Mosby-Year Book, Inc.; 1994.
12. Ullah A, Kalsoom UE, Umair M, John P, Ansar M, Basit S, et al. Exome sequencing revealed a novel splice site variant in the ALX1 gene underlying frontonasal dysplasia. *Clin Genet.* 2017;91(3):494-8.
13. Ullah A, Umair M, e-Kalsoom U, Shahzad S, Basit S, Ahmad W. Exome sequencing revealed a novel nonsense variant in ALX3 gene underlying frontorhiny. *J Hum Genet.* 2018;63(1):97-100.
14. Uz E, Alanay Y, Aktas D, Vargel I, Gucer S, Tuncbilek G, et al. Disruption of ALX1 Causes Extreme Microphthalmia and Severe Facial Clefting: Expanding the Spectrum of

- Autosomal-Recessive ALX-Related Frontonasal Dysplasia. *Am J Hum Genet.* 2010;86(5):789-96.
15. Kayserili H, Uz E, Niessen C, Vargel I, Alanay Y, Tuncbilek G, et al. ALX4 dysfunction disrupts craniofacial and epidermal development. *Hum Mol Genet.* 2009;18(22):4357-66.
  16. Meijlink F, Beverdam A, Brouwer A, Oosterveen TC, Ten Berge D. Vertebrate aristaless-related genes. *Int J Dev Biol.* 1999;43(7):651-63.
  17. Karczewski KJ, Francioli LC, Tiao G, Cummings BB, Alföldi J, Wang Q, et al. The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature.* 2020;581(7809):434-43.
  18. gnomAD [Internet]. 2022 [Erişim Tarihi: 07 Ağustos 2022]. [Available from: <https://gnomad.broadinstitute.org/>].
  19. Liang J, Yin C, Doong H, Fang S, Peterhoff C, Nixon RA, et al. Characterization of erasin (UBXD2): a new ER protein that promotes ER-associated protein degradation. *J Cell Sci.* 2006;119(19):4011-24.
  20. Lim PJ, Danner R, Liang J, Doong H, Harman C, Srinivasan D, et al. Ubiquilin and p97/VCP bind erasin, forming a complex involved in ERAD. *J Cell Biol.* 2009;187(2):201-17.
  21. Szklarczyk D, Gable AL, Nastou KC, Lyon D, Kirsch R, Pyysalo S, et al. The STRING database in 2021: customizable protein-protein networks, and functional characterization of user-uploaded gene/measurement sets. *Nucleic Acids Res.* 2021;49(D1):D605-D12.
  22. OMIM UBXN4 [Internet]. 2022 [Erişim Tarihi: 16 Haziran 2022]. [Available from: <https://www.omim.org/entry/611216?search=ubxn4&highlight=ubxn4>].
  23. Bateman A, Martin MJ, Orchard S, Magrane M, Agivetova R, Ahmad S, et al. UniProt: the universal protein knowledgebase in 2021. *Nucleic Acids Res.* 2021;49(D1):D480-D9.
  24. UniProt UBXN4 Domain [Internet]. 2020 [Erişim Tarihi: 09 Mart 2020]. [Available from: [https://www.uniprot.org/uniprotkb/Q92575/entry#family\\_and\\_domains](https://www.uniprot.org/uniprotkb/Q92575/entry#family_and_domains)].
  25. RCSB PDB UBXN4 2KXJ [Internet]. 2020 [Erişim Tarihi: 09 Mart 2020]. [Available from: <https://www.rcsb.org/structure/2KXJ>].
  26. Mohler PJ, Gramolini AO, Bennett V. Ankyrins. *J Cell Sci.* 2002;115(8):1565-6.
  27. Mohler PJ, Schott JJ, Gramolini AO, Dilly KW, Guatimosim S, duBell WH, et al. Ankyrin-B mutation causes type 4 long-QT cardiac arrhythmia and sudden cardiac death. *Nature.* 2003;421(6923):634-9.
  28. OMIM ANK2 [Internet]. 2022 [Erişim Tarihi: 16 Haziran 2022]. [Available from: <https://www.omim.org/entry/106410?search=ank2&highlight=ank2>].
  29. Mohler PJ, Le Scouarnec S, Denjoy I, Lowe JS, Guicheney P, Caron L, et al. Defining the cellular phenotype of "ankyrin-B syndrome" variants - Human ANK2 variants

associated with clinical phenotypes display a spectrum of activities in cardiomyocytes. *Circulation*. 2007;115(4):432-41.

30. NCBI ANK2 [Internet]. 2020 [Eriřim Tarihi: 11 řubat 2020]. [Available from: <https://www.ncbi.nlm.nih.gov/gene/287>].
31. UniProt ANK2 Domain [Internet]. 2020 [Eriřim Tarihi: 09 Mart 2020]. [Available from: [https://www.uniprot.org/uniprotkb/Q01484/entry#family\\_and\\_domains](https://www.uniprot.org/uniprotkb/Q01484/entry#family_and_domains)].
32. Stefanovic B, Stefanovic L, Schnabl B, Bataller R, Brenner DA. TRAM2 protein interacts with endoplasmic reticulum Ca<sup>2+</sup> pump Serca2b and is necessary for collagen type I synthesis. *Mol Cell Biol*. 2004;24(4):1758-68.
33. Chen QY, Denard B, Lee CE, Han SW, Ye JS, Ye J. Inverting the Topology of a Transmembrane Protein by Regulating the Translocation of the First Transmembrane Helix. *Mol Cell*. 2016;63(4):567-78.
34. OMIM TRAM2 [Internet]. 2022 [Eriřim Tarihi: 16 Haziran 2022]. [Available from: <https://www.omim.org/entry/608485?search=tram2&highlight=tram2>].
35. Onuchic LF, Mrug M, Lakings AL, Muecher G, Becker J, Zerres K, et al. Genomic organization of the KIAA0057 gene that encodes a TRAM-like protein and its exclusion as a polycystic kidney and hepatic disease 1 (PKHD1) candidate gene. *Mamm Genome*. 1999;10(12):1175-8.
36. UniProt TRAM2 Domain [Internet]. 2020 [Eriřim Tarihi: 09 Mart 2020]. [Available from: [https://www.uniprot.org/uniprotkb/Q15035/entry#family\\_and\\_domains](https://www.uniprot.org/uniprotkb/Q15035/entry#family_and_domains)].
37. SMART TRAM2 Domain [Internet]. 2020 [Eriřim Tarihi: 09 Mart 2020]. [Available from: [http://smart.embl-heidelberg.de/smart/show\\_motifs.pl?ID=Q15035&](http://smart.embl-heidelberg.de/smart/show_motifs.pl?ID=Q15035&)].
38. Stefl S, Nishi H, Petukh M, Panchenko AR, Alexov E. Molecular Mechanisms of Disease-Causing Missense Mutations. *J Mol Biol*. 2013;425(21):3919-36.
39. Sudarshan S, Kumar M, Kaur P, Kumar A, Sethuraman G, Sapra S, et al. Decoding of novel missense TSC2 gene variants using in-silico methods. *Bmc Med Genet*. 2019;20(1).
40. Gyulkhandanyan A, Rezaie AR, Roumenina L, Lagarde N, Fremeaux-Bacchi V, Miteva MA, et al. Analysis of protein missense alterations by combining sequence- and structure-based methods. *Mol Genet Genom Med*. 2020;8(4).
41. Bendl J, Stourac J, Salanda O, Pavelka A, Wieben ED, Zendulka J, et al. PredictSNP: Robust and Accurate Consensus Classifier for Prediction of Disease-Related Mutations. *Plos Comput Biol*. 2014;10(1).
42. Stone EA, Sidow A. Physicochemical constraint violation by missense substitutions mediates impairment of protein function and disease severity. *Genome Res*. 2005;15(7):978-86.
43. Thomas PD, Kejariwal A. Coding single-nucleotide polymorphisms associated with complex vs. Mendelian disease: Evolutionary evidence for differences in molecular effects. *P Natl Acad Sci USA*. 2004;101(43):15398-403.

44. Bao L, Zhou M, Cui Y. nsSNPAnalyzer: identifying disease-associated nonsynonymous single nucleotide polymorphisms. *Nucleic Acids Res.* 2005;33:W480-W2.
45. Capriotti E, Calabrese R, Casadio R. Predicting the insurgence of human genetic diseases associated to single point protein mutations with support vector machines and evolutionary information. *Bioinformatics.* 2006;22(22):2729-34.
46. Ramensky V, Bork P, Sunyaev S. Human non-synonymous SNPs: server and survey. *Nucleic Acids Res.* 2002;30(17):3894-900.
47. Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, et al. A method and server for predicting damaging missense mutations. *Nat Methods.* 2010;7(4):248-9.
48. Sim NL, Kumar P, Hu J, Henikoff S, Schneider G, Ng PC. SIFT web server: predicting effects of amino acid substitutions on proteins. *Nucleic Acids Res.* 2012;40(W1):W452-W7.
49. Bromberg Y, Yachdav G, Rost B. SNAP predicts effect of mutations on protein function. *Bioinformatics.* 2008;24(20):2397-8.
50. Capriotti E, Altman RB, Bromberg Y. Collective judgment predicts disease-associated single nucleotide variants. *Bmc Genomics.* 2013;14.
51. Capriotti E, Fariselli P, Casadio R. I-Mutant2.0: predicting stability changes upon mutation from the protein sequence or structure. *Nucleic Acids Res.* 2005;33:W306-W10.
52. Capriotti E, Fariselli P, Calabrese R, Casadio R. Predicting protein stability changes from sequences using support vector machines. *Bioinformatics.* 2005;21:54-8.
53. I-Mutant [Internet]. 2022 [Erişim Tarihi: 16 Haziran 2022]. [Available from: <https://folding.biofold.org/i-mutant/i-mutant2.0.html>].
54. Cheng JL, Randall A, Baldi P. Prediction of protein stability changes for single-site mutations using support vector machines. *Proteins-Structure Function and Bioinformatics.* 2006;62(4):1125-32.
55. Venselaar H, te Beek TAH, Kuipers RKP, Hekkelman ML, Vriend G. Protein structure analysis of mutations causing inheritable diseases. An e-Science approach with life scientist friendly interfaces. *Bmc Bioinformatics.* 2010;11.
56. Bruce Alberts AJ, Julian Lewis, David Morgan, Martin Raff, Keith Roberts, and Peter Walter. *Proteins.* Mol Biol Cell. Sixth ed. United States of America: Garland Science, Taylor & Francis Group, LLC; 2015. p. 144.
57. MATTHEW J. BETTS RBR. Amino Acid Properties and Consequences of Substitutions. In: Michael R. Barnes ICG, editor. *BIOINFORMATICS FOR GENETICISTS.* England: John Wiley & Sons Ltd; 2003. p. 305-16.
58. Armstrong DR, Berrisford JM, Conroy MJ, Gutmanas A, Anyango S, Choudhary P, et al. PDBe: improved findability of macromolecular structure data in the PDB. *Nucleic Acids Res.* 2020;48(D1):D335-D43.

59. Varadi M, Bccrisford J, Deshpande M, Nair SS, Gutmanas A, Armstrong D, et al. PDBe-KB: a community-driven resource for structural and functional annotations. *Nucleic Acids Res.* 2020;48(D1):D344-D53.
60. Chothia C, Lesk AM. The Relation between the Divergence of Sequence and Structure in Proteins. *Embo J.* 1986;5(4):823-6.
61. Marti-Renom MA, Stuart AC, Fiser A, Sanchez R, Melo F, Sali A. Comparative protein structure modeling of genes and genomes. *Annu Rev Bioph Biom.* 2000;29:291-325.
62. Altschul SF, Madden TL, Schaffer AA, Zhang JH, Zhang Z, Miller W, et al. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 1997;25(17):3389-402.
63. Peitsch MC. ProMod and Swiss-model: Internet-based tools for automated comparative protein modelling. *Biochem Soc T.* 1996;24(1):274-9.
64. Waterhouse A, Bertoni M, Bienert S, Studer G, Tauriello G, Gumienny R, et al. SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res.* 2018;46(W1):W296-W303.
65. Benkert P, Biasini M, Schwede T. Toward the estimation of the absolute quality of individual protein structure models. *Bioinformatics.* 2011;27(3):343-50.
66. Bowie JU, Luthy R, Eisenberg D. A Method to Identify Protein Sequences That Fold into a Known 3-Dimensional Structure. *Science.* 1991;253(5016):164-70.
67. Jones DT, Taylor WR, Thornton JM. A New Approach to Protein Fold Recognition. *Nature.* 1992;358(6381):86-9.
68. Anfinsen CB. Principles That Govern Folding of Protein Chains. *Science.* 1973;181(4096):223-30.
69. Liwo A, Lee J, Ripoll DR, Pillardy J, Scheraga HA. Protein structure prediction by global optimization of a potential energy function. *P Natl Acad Sci USA.* 1999;96(10):5482-5.
70. Roy A, Kucukural A, Zhang Y. I-TASSER: a unified platform for automated protein structure and function prediction. *Nat Protoc.* 2010;5(4):725-38.
71. Kelley LA, Mezulis S, Yates CM, Wass MN, Sternberg MJE. The Phyre2 web portal for protein modeling, prediction and analysis. *Nature Protocols.* 2015;10(6):845-58.
72. Pereira J, Simpkin AJ, Hartmann MD, Rigden DJ, Keegan RM, Lupas AN. High-accuracy protein structure prediction in CASP14. *Proteins-Structure Function and Bioinformatics.* 2021;89(12):1687-99.
73. Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, et al. Highly accurate protein structure prediction with AlphaFold. *Nature.* 2021;596(7873):583-9.
74. Varadi M, Anyango S, Deshpande M, Nair S, Natassia C, Yordanova G, et al. AlphaFold Protein Structure Database: massively expanding the structural coverage of protein-sequence space with high-accuracy models. *Nucleic Acids Res.* 2022;50(D1):D439-D44.



75. Ittisoponpisan S, Islam SA, Khanna T, Alhuzimi E, David A, Sternberg MJE. Can Predicted Protein 3D Structures Provide Reliable Insights into whether Missense Variants Are Disease Associated? *J Mol Biol.* 2019;431(11):2197-212.
76. Krivov GG, Shapovalov MV, Dunbrack RL. Improved prediction of protein side-chain conformations with SCWRL4. *Proteins-Structure Function and Bioinformatics.* 2009;77(4):778-95.
77. Roy A, Yang JY, Zhang Y. COFACTOR: an accurate comparative algorithm for structure-based protein function annotation. *Nucleic Acids Res.* 2012;40(W1):W471-W7.
78. Zhang CX, Freddolino PL, Zhang Y. COFACTOR: improved protein function prediction by combining structure, sequence and protein-protein interaction information. *Nucleic Acids Res.* 2017;45(W1):W291-W9.
79. Yang JY, Roy A, Zhang Y. BioLiP: a semi-manually curated database for biologically relevant ligand-protein interactions. *Nucleic Acids Res.* 2013;41(D1):D1096-D103.
80. Yang JY, Roy A, Zhang Y. Protein-ligand binding site recognition using complementary binding-specific substructure comparison and sequence profile alignment. *Bioinformatics.* 2013;29(20):2588-95.
81. Chen YT, Lu HY, Zhang N, Zhu ZF, Wang SQ, Li MH. PremPS: Predicting the impact of missense mutations on protein stability. *Plos Comput Biol.* 2020;16(12).
82. Zimmermann L, Stephens A, Nam SZ, Rau D, Kubler J, Lozajic M, et al. A Completely Reimplemented MPI Bioinformatics Toolkit with a New HHpred Server at its Core. *J Mol Biol.* 2018;430(15):2237-43.
83. Bruce Alberts AJ, Julian Lewis, David Morgan, Martin Raff, Keith Roberts, and Peter Walter. *Membrane Structure.* Mol Biol Cell. Sixth ed. United States of America: Garland Science, Taylor & Francis Group, LLC; 2015. p. 614.
84. Nilsson I, Saaf A, Whitley P, Gafvelin G, Waller C, von Heijne G. Proline-induced disruption of a transmembrane alpha-helix in its natural environment. *J Mol Biol.* 1998;284(4):1165-75.
85. AlphaFold CERS1 [Internet]. 2022 [Erişim Tarihi: 06 Temmuz 2022]. [Available from: <https://alphafold.ebi.ac.uk/entry/P27544>].
86. AlphaFold CLN8 [Internet]. 2022 [Erişim Tarihi: 06 Temmuz 2022]. [Available from: <https://alphafold.ebi.ac.uk/entry/Q9UBY8>].
87. Winter E, Ponting CR. TRAM, LAG1 and CLN8: members of a novel family of lipid-sensing domains? *Trends Biochem Sci.* 2002;27(8):381-3.
88. Grimm DG, Azencott CA, Aicheler F, Gieraths U, MacArthur DG, Samocha KE, et al. The Evaluation of Tools Used to Predict the Impact of Missense Variants Is Hindered by Two Types of Circularity. *Hum Mutat.* 2015;36(5):513-23.
89. Qorri E, Takacs B, Graf A, Enyedi MZ, Pinter L, Kiss E, et al. A Comprehensive Evaluation of the Performance of Prediction Algorithms on Clinically Relevant Missense Variants. *Int J Mol Sci.* 2022;23(14).

90. McCarroll SA, Altshuler DM. Copy-number variation and association studies of human disease. *Nat Genet.* 2007;39:S37-S42.
91. Vaz-Drago R, Custodio N, Carmo-Fonseca M. Deep intronic mutations and human disease. *Hum Genet.* 2017;136(9):1093-111.

## 8. EKLER

## EK-1: Tez Çalışması ile İlgili Etik Kurul İzni



T.C.  
HACETTEPE ÜNİVERSİTESİ  
Girişimsel Olmayan Klinik Araştırmalar Etik Kurulu

Sayı : 16969557-500

Konu :

17.03.2020

**Dr. Öğr. Üyesi Ceren SUCULARLI**  
Sağlık Bilimleri Enstitüsü  
Biyoinformatik Anabilim Dalı  
Öğretim Üyesi

Sayın Dr. Öğr. Üyesi SUCULARLI,

Kurulumuza değerlendirilmek üzere sunduğunuz GO 20/231 kayıt numaralı ve "*Frontonazal Nadir Bir Yüz Malformasyonu ile İlişkili Aday Genlerde Protein Düzeyinde In Silico Analiz*" başlıklı proje Kurulumuzun 17.03.2020 tarihli toplantısında değerlendirilmiş olup, aday genlere ait yanlış anlamlı mutasyonların protein düzeyindeki etkilerinin bilgisayar temelli yöntemlerle değerlendirileceği anlaşılmıştır. Gönüllü insanlar üzerinde gerçekleştirilecek nitelikte olmayan bu tip çalışmalar Etik Kurulların kapsamı dışında kalmaktadır.

Bu yazı ilgili protokolün etik açıdan incelendiğini belirtmek için Etik Kurul kararı yerine geçmek üzere hazırlanmıştır.

Prof. Dr. Ayşe Lale DOĞAN  
Başkan

EK \_\_\_\_\_ ;  
Toplantı Katılım Tutanağı.

## EK-2: Şekil 2.1. için Alınan Telif Hakkı

ELSEVIER LICENSE  
TERMS AND CONDITIONS

Aug 11, 2022

---

This Agreement between Hacettepe University -- Ismihan Alp ("You") and Elsevier ("Elsevier") consists of your license details and the terms and conditions provided by Elsevier and Copyright Clearance Center.

License Number	5365430342805
License date	Aug 10, 2022
Licensed Content Publisher	Elsevier
Licensed Content Publication	The American Journal of Human Genetics
Licensed Content Title	Disruption of ALX1 Causes Extreme Microphthalmia and Severe Facial Clefting: Expanding the Spectrum of Autosomal-Recessive ALX-Related Frontonasal Dysplasia
Licensed Content Author	Elif Uz, Yasemin Alanay, Dilek Aktas, Ibrahim Vargel, Safak Gucer, Gokhan Tuncbilek, Ferdinand von Eggeling, Engin Yilmaz, Ozgur Deren, Nicole Posorski, Hilal Ozdag, Thomas Liehr, Sevim Balci, Mehmet Alikasifoglu, Bernd Wollnik, Nurten A. Akarsu
Licensed Content Date	May 14, 2010
Licensed Content Volume	86
Licensed Content Issue	5
Licensed Content Pages	8
Start Page	789
End Page	796
Type of Use	reuse in a thesis/dissertation
Portion	figures/tables/illustrations
Number of figures/tables/illustrations	1
Format	both print and electronic

Are you the author of this Elsevier article? No  
Will you be translating? No  
Title Master of Science  
Institution name Hacettepe University  
Expected presentation date Aug 2022  
Portions Figure 3  
Hacettepe University  
Yukarı Bahçelievler Mah.

Requestor Location

ÇANKAYA, ANKARA 06490  
Turkey  
Attn: Hacettepe University

Publisher Tax ID

GB 494 6272 12

Total

0.00 USD

**EK-3: Tez Çalışması Orijinallik Raporu****Dijital Makbuz**

Bu makbuz ödevinizin Turnitin'e ulaştığını bildirmektedir. Gönderiminize dair bilgiler şöyledir:

Gönderinizin ilk sayfası aşağıda gönderilmektedir.

Gönderen: İsmihan Alp  
Ödev başlığı: İsmihan Alp Yüksek Lisans Tez  
Gönderi Başlığı: FRONTONAZAL NADİR BİR YÜZ MALFORMASYONU İLE İLİŞKİLİ...  
Dosya adı: KI\_LI\_ADAY\_GENLERDE\_PROTEI\_N\_DU\_ZEYI\_NDE\_IN\_SILICO\_A...  
Dosya boyutu: 4.97M  
Sayfa sayısı: 60  
Kelime sayısı: 10,445  
Karakter sayısı: 71,735  
Gönderim Tarihi: 11-Ağu-2022 12:27ÖS (UTC+0300)  
Gönderim Numarası: 1881319852



## Turnitin Orijinallik Raporu

İşleme konu: 11-Ağu-2022 12:32 +03

NUMARA: 1881319852

Kelime Sayısı: 10445

Gönderildi: 1

Benzerlik Endeksi

%3

Kaynağa göre Benzerlik

İnternet Sources: %3

Yayınlar: %2

Öğrenci Ödevleri: %0

FRONTONAZAL NADİR BİR YÜZ  
MALFORMASYONU İLE İLİŞKİLİ  
ADAY GENLERDE PROTEİN  
DÜZEYİNDE IN SILICO ANALİZ  
İsmihan Alp tarafından

< 1% match (03-Ağu-2022 tarihli internet)

[https://acikbilim.yok.gov.tr/bitstream/handle/20.500.12812/494418/yokAcikBilim\\_10254711.pdf?sequence=-1](https://acikbilim.yok.gov.tr/bitstream/handle/20.500.12812/494418/yokAcikBilim_10254711.pdf?sequence=-1)

< 1% match (10-Şub-2022 tarihli internet)

[https://acikbilim.yok.gov.tr/bitstream/handle/20.500.12812/40079/yokAcikBilim\\_10313412.pdf?isAllowed=y&sequence=-1](https://acikbilim.yok.gov.tr/bitstream/handle/20.500.12812/40079/yokAcikBilim_10313412.pdf?isAllowed=y&sequence=-1)

< 1% match (15-Haz-2022 tarihli internet)

[https://acikbilim.yok.gov.tr/bitstream/handle/20.500.12812/329089/yokAcikBilim\\_10292392.pdf?isAllowed=y&sequence=-1](https://acikbilim.yok.gov.tr/bitstream/handle/20.500.12812/329089/yokAcikBilim_10292392.pdf?isAllowed=y&sequence=-1)

< 1% match (13-Mar-2022 tarihli internet)

[https://acikbilim.yok.gov.tr/bitstream/handle/20.500.12812/146110/yokAcikBilim\\_10150323.pdf?sequence=-1](https://acikbilim.yok.gov.tr/bitstream/handle/20.500.12812/146110/yokAcikBilim_10150323.pdf?sequence=-1)

< 1% match (24-May-2022 tarihli internet)

[https://acikbilim.yok.gov.tr/bitstream/handle/20.500.12812/416457/yokAcikBilim\\_10157293.pdf?isAllowed=y&sequence=-1](https://acikbilim.yok.gov.tr/bitstream/handle/20.500.12812/416457/yokAcikBilim_10157293.pdf?isAllowed=y&sequence=-1)

< 1% match (yayınlar)

[Dóra Judit Kiss, Gáspár Pándy-Szekeres, György Miklós Keserű. "Computational Medicinal Chemistry to Target GPCRs", Elsevier BV, 2022](#)

< 1% match (yayınlar)

[Christy Priyadharshini J, Thirumal Kumar D, Sneha P, Siva R, Charles Emmanuel Jebaraj W, George Priya Doss C, Hatem Zayed. "An integrative bioinformatics pipeline to demonstrate the alteration of the interaction between the ALDH2\\*2 allele with NAD and Disulfiram", Journal of Cellular Biochemistry, 2019](#)

< 1% match (yayınlar)

[Himani Tanwar, D. Thirumal Kumar, C. George Priya Doss, Hatem Zayed. "Bioinformatics classification of mutations in patients with Mucopolysaccharidosis IIIA", Metabolic Brain Disease, 2019](#)

< 1% match (07-Oca-2022 tarihli internet)

<https://coek.info/pdf-molecular-contacts-between-antibiotics-and-the-30s-ribosomal-particle-.html>

< 1% match (yayınlar)

[Maisha Adiba, Tonmoy Das, Anik Paul, Ashish Das, Sajib Chakraborty, Md Ismail Hosen, A.H.M. Nurun Nabi. "In silico characterization of coding and non-coding SNPs of the androgen receptor gene", Informatics in Medicine Unlocked, 2021](#)

## 9. ÖZGEÇMİŞ



