**DNA ANALİZİ VE HASTALIK ÇIKARIMI**

Bilal Emre Taydaş

Kütahya Sağlık Bilimleri Üniversitesi, Mühendislik ve Doğa Bilimleri Fakültesi, Bilgisayar Mühendisliği Bölümü, Kütahya, Türkiye

[bemre.taydas@ogr.ksbu.edu.tr](mailto:bemre.taydas@ogr.ksbu.edu.tr)

**ÖZET**

Bu proje, araştırmacılara ve sağlık çalışanlarına, çeşitli sınıflandırma ve filtreleme seçenekleri sunarak genom dosyalarını etkili bir şekilde okuma ve analiz etme imkânı sağlamayı hedeflemektedir. Programın temel amacı, kullanıcının gerçekte üzerinde çalışmak istediği konuya odaklanmasını kolaylaştırarak zaman ve bilgi işlem gücünden tasarruf etmektir. Bu sayede, bireysel hastaların veritabanında sadece gerekli veriler tutulacak ve fazla bilgi depolaması yapmayacak; aynı zamanda, benzerlik oranlarını elde etmek için etkili karşılaştırmalar yapma olanağı sunacaktır.

**GİRİŞ**

Genom araştırmaları, uzun bir geçmişe sahip olup zaman içinde maliyetlerinin düşmesi ve yatırımların katlanarak artmasıyla dikkat çeken bir alan haline gelmiştir. Yeni ilaç üretimi ve tedavi yöntemlerinin keşfi gibi potansiyel avantajları nedeniyle büyük ilgi görmektedir. Bu sektör hızla genişlemekte ve artan yatırımlarla desteklenmektedir. Bu gelişmeler bizi etkiledi ve derinleştikçe, ekip olarak bu alanda bilgisayar ve yazılım teknolojilerinden faydalanmanın mümkün olduğunu fark ettik. Bu düşünceyle bir proje oluşturmaya karar verdik. Projemiz, bireylerin genetik analizi üzerinden hastalıklara yatkınlıklarını belirlemeyi hedefleyerek, benzer projelere alternatif bir yaklaşım sunmayı amaçlamaktadır.

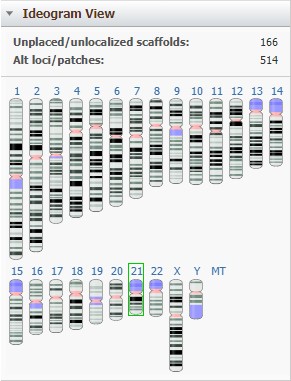
**YAPILAN ÇALIŞMALAR**

Projenin ana kaynak verisi olan NCBI hakkında detaylı araştırmalar yapılmıştır. NCBI'nın geniş veri tabanları ve genomik veriler üzerinde kapsamlı bir değerlendirme yapılmıştır. Bu süreçte, projemizle doğrudan ilgili ve potansiyel olarak değerli olan veri setlerini ve dosyaları incelenmiştir. Projemize entegre edilebilecek, geniş bir bilgi yelpazesi içeren bu kaynakları belirleyerek, gelecekteki geliştirmeler ve araştırmalar için sağlam bir temel oluşturmayı hedefledik. Bu süreç, projemizin bilgi tabanını güçlendirmek ve daha geniş bir perspektiften ele almak adına önemli bir adımdır.

***ARAŞTIRMALAR***

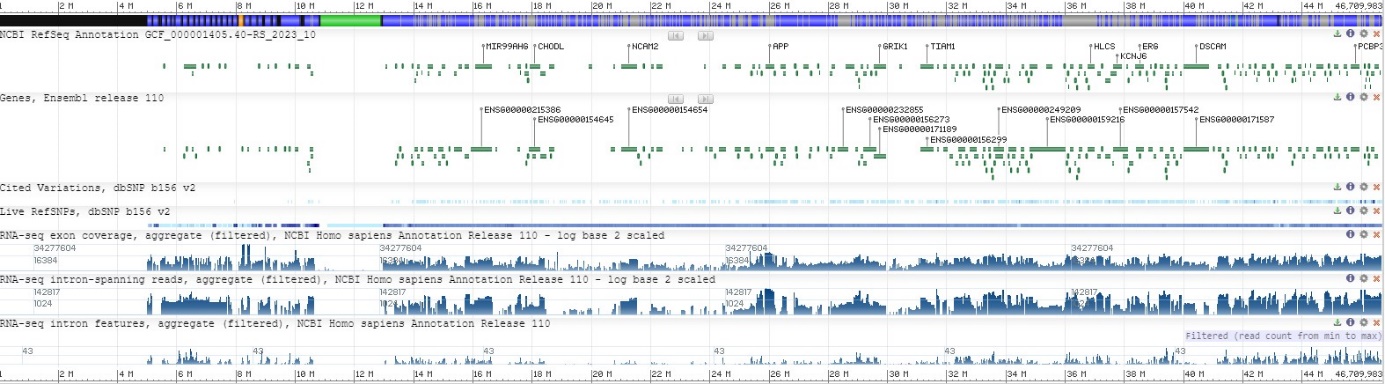
Resim1- Sayfaya Genel Bakış [2]

Buradan seçtiğimiz kromozomun NCBI verilerinde bulunan tüm verilerine ulaşmak mümkün. Bizim öncelik olarak belirlediğimiz 21. Kromozomun Genom Data Viewer’ı (GDV) var. Burada 21. kromozomun genetik bilgilerinin kapsamlı bir şekilde incelenmesine ve görselleştirilmesine imkân tanır [1].



Resim2- İdeogram Görünümü [2]

NCBI açıklamalı kromozom tabanlı bir düzenekteki kromozomları gösterir. Ekranda görüntülenen mevcut kromozomu yeşil bir vurgu çevreler. Bir kromozomun tıklanması Dizi Görüntüleyiciyi seçilen kromozoma güncelleyecektir ve Sequence Viewer içinde o kromozomun verileri çıkacaktır [1].

Resim3- Sequence Viewer [2]

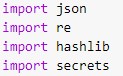
Bu kısım birçok NCBI sayfasında yerleşiktir. Sequence Viewer bizim istediğimiz birçok bilginin görselleştirilmiş ve grafiğe dökülmüş halidir. Buradan seçtiğimiz kromozomun NCBI verilerinde bulunan tüm verilerine ulaşmak mümkün [1].

NCBI’ın bizi ilgirendiren bir diğer bölümü ise Blast ve Premier Blasttır. BLAST, benzer dizilere sahip genlerin veya proteinlerin hızlı bir şekilde tespiti için idealdir, Primer-BLAST ise spesifik PCR amplifikasyonu için uygun primerleri tasarlamak için oldukça kullanışlıdır. Bu da bizim NCBI’dan aldığımız dosyaları çevirmek için kullanabiliriz [1].

***UYGULAMA KODU***

Kod kısmında kullanıcının Gui içine girebilmesi için kayıt olmasını, giriş yapmasını ve şifresini unutursa şifre yenileme işlemini gerçekleştiriyor.

Kullanılan bazı kütüphaneler:



**Re:** Gelişmiş string işleme için kurallı ifade araçları sağlar. Karmaşık eşleme ve manipülasyon için, kurallı ifadeler kısa ve optimize edilmiş çözümler sunar [7].

**Json:** Kullanıcı bilgilerini json dosyasında tutmamızı yarar [7].

**Hashlib:** Bu modül birçok farklı güvenli karma ve mesaj özeti algoritmasına ortak bir arayüz uygular [7].

**Secrets:** Şifreler, hesap kimlik doğrulaması, güvenlik belirteçleri ve ilgili sırlar gibi verileri yönetmeye uygun kriptografik olarak güçlü rastgele sayılar üretmek için kullanılır [7].

******Resim4- Kayıt ol fonksiyonu

Kayıt ol fonksiyonu (Resim4) kullanıcıdan aldığı mail ve şifreyi kontrol eder. Aynı mail ve iki kere girilen şifre birbirine uyuyorsa bir tuz oluşturulur. Bu tuz her kullanıcı için rastgele oluşturulan ve şifrenin güvenliğini artıran bir değerdir. Kullanıcının şifresi ve oluşturulan tuz birleştirilip SHA-256 hash fonksiyonu ile hashlenir.



Resim5- Giriş yap fonksiyonu

Giriş yap fonksiyonu (Resim5) mail ve şifreyi kontrol ederek kullanıcın kayıtlı olup olmadığını kontrol eder. Kullanıcı bilgileri kullanıcılar dosyasında varsa giriş yapma işlemi başarılı olur.

**SONUÇ**

Program, grafik arayüz, veri tabanı ve şifreleme işlemleri tamamlandığında kullanıma hazır hale gelecektir. Kullanıcılar, giriş ve kayıt ekranlarına erişim sağlayarak hastaları düzenlemek, eklemek veya kaldırmak için gerekli araçları içeren bir panelle karşılaşacaklardır. Bu panel, bireysel hasta profillerini içeren bir hasta listesini sunacak ve genom bilgisi yanı sıra kullanıcı notlarına da erişim imkânı sağlayacaktır. Kullanıcılar, genomun bir konumunu başka bir genomun farklı bir konumuyla karşılaştırabilecek ve benzerlik puanlarını hızlı bir şekilde alabileceklerdir. Ayrıca, bu karşılaştırmaların gen hizalamalarını da görebileceklerdir. Bu tasarımın temel hedefi, tespit işlemlerini daha hızlı, daha ekonomik ve daha az gereksinimle gerçekleştirmek üzerinedir.

**REFERANSLAR**

1. NCBI (2020) Genome Data Viewer (Tanımlar)(1)

<https://www.ncbi.nlm.nih.gov/genome/gdv/browser/help/#SEQVIEW>

1. NCBI Genome Data Viewer (Grafik) (2)

<https://www.ncbi.nlm.nih.gov/genome/gdv/browser/genome/?id=GCF_000001405.40>

1. NCBI BLAST

<https://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastn&PAGE_TYPE=BlastSearch&LINK_LOC=blasthome>

1. NCBI PREMER – BLAST

<https://www.ncbi.nlm.nih.gov/tools/primer-blast/index.cgi>

1. NCBI (2022) Grafiksel Görünüm Açıklamaları

<https://www.ncbi.nlm.nih.gov/tools/sviewer/legends/#anchor_1>

1. Darcan C., Türkyılmaz O. (2017). *Yeni Nesil Dizileme Teknolojisine Genel Bakış*. Bilecik Şeyh Edebali Üniversitesi Fen Bilimleri Dergisi, Cilt: 5, Sayı: 1, 2018
2. Python 3.12.1 belgeleri

https://docs.python.org/3/