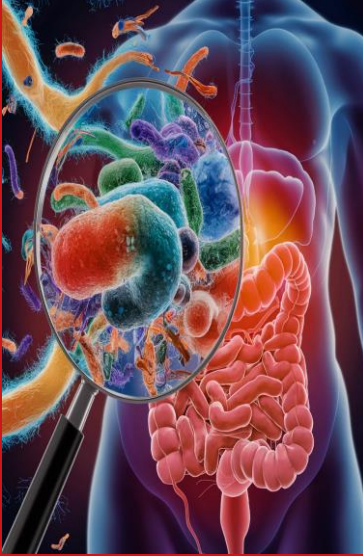


Ayşegül Koçak Ulutopcu



aysegul.kocak@agu.edu.tr

0009-0005-4677-7338



Assoc. Prof. Burcu Bakır Güngör

Burcu Bakır Güngör

burcu.gungor@agu.edu.tr

Development of a Machine Learning-Based System to Identify Disease Biomarkers From Human Gut

abstract

The human gut microbiota consists of a diverse ecosystem of organisms, encompasses billions of species. Recently developed next-generation sequencing methods have enabled researchers to examine the microbiota in greater detail, leading to new insights into its functions and dysfunctions.

This study aims to identify metagenomic biomarkers (Microorganism-Enzyme Pairs) for colorectal cancer (CRC). The tool that we used allows for the analysis of microorganisms and enzymes within the gut microbiota. It achieves this by initially clustering enzymes based on their correlations with species and subsequently utilizing these clustering results to evaluate the ability of groups to differentiate between patient and healthy cohorts.

By integrating species and enzymes, it is possible to identify pathogen microorganisms and enzyme clusters, that have the potential to distinguish cases (individuals with CRC) from controls (healthy individuals).

The identified enzyme clusters and associated species could potentially act as biomarkers for colorectal cancer (CRC), enabling early diagnosis and more effective treatment. This approach holds promise for further exploration of the gut microbiota and its importance in human health and illness.

keywords : Bioinformatics, Machine Learning, Colorectal Cancer Diagnosis

özet

İnsan bağırsak mikrobiyotası, milyarlarca türü kapsayan karmaşık bir organizma topluluğudur. Son zamanlarda geliştirilen yeni nesil dizileme yöntemleri, araştırmacıların mikrobiyotayı daha detaylı bir şekilde incelemesine olanak tanımış ve bu sayede işlevleri ve disfonksiyonları hakkında yeni bilgiler elde edilmiştir.

Bu çalışma, kolorektal kanser (CRC) için metagenomik biyobelirteçler (Mikroorganizma-Enzim Çiftleri) tanımlamayı amaçlamaktadır. Kullandığımız araç, bağırsak mikrobiyotası içindeki mikroorganizmaların ve enzimlerin analizine olanak tanır. Bunu, öncelikle enzimleri türlerle olan korelasyonlarına göre kümeleyerek ve daha sonra bu kümeleme sonuçlarını hasta ve sağlıklı kohortlar arasında ayırt etme yeteneği gruplarını değerlendirecek şekilde gerçekleştirir.

Türler ve enzimler entegre edilerek, vakaları (CRC'li bireyler) kontrollerden (sağlıklı bireyler) ayırma potansiyeline sahip patojen mikroorganizmaları ve enzim kümeleri belirlenebilir.

Tanımlanan enzim kümeleri ve ilişkili türler, CRC için potansiyel biyobelirteçler olarak hizmet edebilir, erken tanı ve daha etkili tedavi sağlayabilir. Bu yaklaşım, bağırsak mikrobiyotasının ve insan sağlığı ile hastalıkları üzerindeki rolünün daha fazla keşfedilmesi için umut vaat etmektedir.

anahtar kelime : *Biyoinformatik, Makine Öğrenimi, Kolorektal Kanser Teşhisi*