

Kapsamlı Çoklu-Omik Analizi ve Kişiselleştirilmiş Tıp Çalışmaları için Akciğer Kanseri Biyobanka Sisteminin Kurulması

Dilek Çeker^{1,2}, Volkan Baysungur^{3,4}, Serdar Evman⁴, İlker Kolbaş⁴, Abdurrahim Gördebil⁴, Yusuf Tambağ⁵, Ömer Kaçar¹, Ahmet Midi⁶, Hatice Aslanoğlu⁴, Nilüfer Kara⁴, Nilgün Algan⁴, Betül Karademir Yılmaz⁷, Ali Şahin⁷, Hivda Ülbeği Polat¹, Abidin Şehitoğulları⁸, İhsan Boyacı⁹, Mahmut Gümüş¹⁰, Uygur Halis Tazebay², Abdullah Karadağ^{1,11}

¹ TÜBİTAK Marmara Araştırma Merkezi, BİYB, MBAG, Moleküler Onkoloji Laboratuvarı, Kocaeli, Türkiye

² Gebze Teknik Üniversitesi, Moleküler Biyoloji ve Genetik Anabilim Dalı, Kocaeli, Türkiye

³ Sağlık Bilimleri Üniversitesi Tıp Fakültesi, Göğüs Cerrahisi Anabilim Dalı, İstanbul, Türkiye

⁴ T.C. Sağlık Bakanlığı, İstanbul Süreyyapaşa Göğüs Hastalıkları ve Göğüs Cerrahisi Eğitim ve Araştırma Hastanesi, İstanbul, Türkiye

⁵ TÜBİTAK BİLGEM, Ankara, Türkiye

⁶ Bahçeşehir Üniversitesi Tıp Fakültesi, Patoloji Anabilim Dalı, İstanbul, Türkiye

⁷ Marmara Üniversitesi Tıp Fakültesi, Biyokimya Anabilim Dalı, Genetik ve Metabolik Hastalıklar Araştırma ve İnceleme Merkezi, İstanbul, Türkiye

⁸ Sakarya Üniversitesi Tıp Fakültesi, Göğüs Cerrahisi Anabilim Dalı, Sakarya, Türkiye

⁹ İstanbul Medipol Üniversitesi Tıp Fakültesi, İç Hastalıkları Anabilim Dalı, İstanbul, Türkiye

¹⁰ İstanbul Medeniyet Üniversitesi Tıp Fakültesi, İç Hastalıkları Anabilim Dalı, İstanbul, Türkiye

¹¹ Gebze Teknik Üniversitesi, Biyoteknoloji Enstitüsü, Kocaeli, Türkiye

ÖZET

Günümüzde modern biyomedikal araştırmalar, hasta odaklı çalışmaları ön plana almaktadır. Bu nedenle hastalardan alınan örneklerin kapsamlı bir şekilde korunmasına ve araştırmaların yapılacağı merkezlerde toplanmalarına ihtiyaç vardır. Biz de LUNGBANK “Biyobelirteç temelli akciğer kanseri erken tanı ve tarama sistemi geliştirilmesi” projesi kapsamında bir biyobanka oluşturduk. Klasik biyobankalar daha çok hastaların tanısına yönelik örnek toplama yoluna gider ancak günümüzde omik teknolojilerinin gelişmesi ve ucuzlamasıyla birlikte hasta örneklerinden daha detaylı bilgiler elde edilmesi ve hastaların biyolojik haritalarının oluşturularak erken tanı ve tedavilerine yönelik sistemlerin geliştirilmesi ihtiyacı doğmuştur. Kısaca LUNGBANK olarak adlandırdığımız bu biyobanka sisteminin temelinde dinamik bir yaklaşımla doğru hastaların tespiti, kabulü, örneklerin toplanması, işlenmesi ve depolanması yatar. Etik kurul kararı ile ıslak imzalı gönüllü hasta onam formları eşliğinde çalışmaya dâhil edilen hastalardan alınan çeşitli örnekler, tarafımızca geliştirilen protokollere göre bu çalışmaya dâhil klinik ekip tarafından alındı. Alınan örnekler arasında akciğer kanseri tanısı almış hastaların normal akciğer dokusu, primer tümörleri, reaktif stroma (tümör mikroçevre dokusu), lokal lenf nodları, kan/serum/plazma, idrar, tükürük, balgam, bronkoalveolar sıvı ve dışkı sayılabilir. Bu örnekler doku bütünlüğü korunarak ve RNA muhafaza edilecek şekilde steril ve RNase-free ortamlarda soğuk zincir ile laboratuvara taşındı. Katı örnekler daha sonra homojenize edilip her biri en az beş ayrı tüpe alıktı ve her bir tüp tüm genom, transkriptom, proteom ve metabolom çalışmalarına ayrı ayrı alındı. Bu yaklaşım yeni nesil dizileme ve likit kromatografi ile tandem edilmiş kütle spektrometresi gibi yüksek hacimli teknolojilerin çoklu-omik analizlerinin aynı başlangıç materyaliyle yapılabilmesini mümkün kıldı. Bütün katı dokuların histolojik kesitleri H & E ile boyanarak patolojik tanıları doğrulandı. Hasta ve örnek bilgileri LUNGBASE adı verilen 96 terabayt kapasiteli bir veri tabanına kaydedildi ve erişim sadece ilgili personele açıldı. Veri tabanının yönetimi için de tarafımızca geliştirilen bir yazılım (LUNGSOFT) kullanıldı. Bu yaklaşımla LUNGBANK projesi başarılı bir şekilde tamamlanmış oldu ve elde edilen çok kıymetli verilerin yakın bir zamanda bilim camiasının erişimine açılması planlandı.

Anahtar kelimeler: Biyobanka, akciğer kanseri, kişiselleştirilmiş tıp, çoklu omik analizleri

Lung Cancer Biorepository: Advancing P-Medicine Through Comprehensive Multi-Omics Analysis

Dilek Çeker^{1,2}, Volkan Baysungur^{3,4}, Serdar Evman⁴, İlker Kolbaş⁴, Abdurrahim Gördebil⁴, Yusuf Tambağ⁵, Ömer Kaçar¹, Ahmet Midi⁶, Hatice Aslanoğlu⁴, Nilüfer Kara⁴, Nilgün Alğan⁴, Betül Karademir Yılmaz⁷, Ali Şahin⁷, Hivda Ülbeği Polat¹, Abidin Şehitoğulları⁸, İhsan Boyacı⁹, Mahmut Gümüşt¹⁰, Uygur Halis Tazebay², Abdullah Karadağ^{1,11}

¹ TÜBİTAK Marmara Research Center, VPLS, MBRG, Molecular Oncology Laboratory, Kocaeli, Türkiye

² Department of Molecular Biology and Genetics, Gebze Technical University, Kocaeli, Türkiye

³ Department of Thoracic Surgery, University of Health Sciences Medical School, İstanbul, Türkiye

⁴ Ministry of Health, İstanbul Süreyyapaşa Chest Diseases and Thoracic Surgery Training and Research Hospital, İstanbul, Türkiye

⁵ TÜBİTAK BILGEM, Ankara, Türkiye

⁶ Department of Pathology, Bahçeşehir University Faculty of Medicine, İstanbul, Türkiye

⁷ Department of Biochemistry, Genetic and Metabolic Diseases Research and Investigation Center, Marmara University Faculty of Medicine, İstanbul, Türkiye

⁸ Department of Thoracic Surgery, Sakarya University Faculty of Medicine, Sakarya, Türkiye

⁹ Department of Internal Medicine, Medipol İstanbul University Faculty of Medicine, İstanbul, Türkiye

¹⁰ Department of Internal Medicine, İstanbul Medeniyet University Faculty of Medicine, İstanbul, Türkiye

¹¹ Institute of Biotechnology, Gebze Technical University, Kocaeli, Türkiye

ABSTRACT

In response to the critical need for comprehensive specimen preservation in modern biomedical research, our initiative, LUNGBANK, emerged as a tailored biorepository system for completed lung cancer research studies. Traditional pathology units, primarily focused on diagnostics, often encountered limitations in preserving an adequate number of samples for emerging multi-omics analysis and P-medicine applications. Essential to effective specimen management was a dynamic and innovative approach that encompassed collection, processing, and storage. LUNGBANK exemplified this need through its systematic design. A dedicated clinical team meticulously collected diverse specimens from lung cancer patients, adhering to IRB-approved protocols and voluntary patient consent, spanning from normal lung tissue to various cancer-related samples. LUNGBANK's infrastructure extended beyond solid tissue, encompassing body fluids and stool specimens, all meticulously registered in the open-stack database, LUNGBASE. Managed through the specially designed software LUNGSOFT, LUNGBASE boasted a storage capacity of 96 TB. The collected samples, including solid tissues, underwent immediate snap-freezing upon transfer to the laboratory, followed by meticulous DNA/RNA isolation and quality controls. LUNGBANK established a novel paradigm for integrative multi-omics analysis, providing high-quality specimens for next-generation sequencing and mass spectrometry analyses, thus contributing to the advancement of lung cancer research and the translation of bench data to bedside applications in the pursuit of precision medicine. This meticulous approach ensured the highest quality of specimens and positioned LUNGBANK as a pivotal resource in unraveling the molecular intricacies of lung cancer initiation and progression. We aimed to foster the creation of advanced diagnostic and therapeutic systems to enhance patient outcomes through LUNGBANK.

Keywords: Biobank, Biorepository, Lung cancer, P-medicine, integrative multiomics

Akciğer Kanseri Terapötik Uygulamalarında Önemli Olan Bazı Proteinlerin Kannabidiol ile Olan Etkileşimlerinin *in Siliko* Yaklaşımlarla Belirlenmesi

Ali Aydın¹, Burçin Türkmenoğlu²

¹ Yozgat Bozok Üniversitesi Tıp Fakültesi, Temel Tıp Bilimleri Bölümü, Yozgat, Türkiye

² Erzincan Binali Yıldırım Üniversitesi Eczacılık Fakültesi, Temel Eczacılık Bilimleri Bölümü, Erzincan, Türkiye

ÖZET

Amaç: Akciğer kanseri, görülme sıklığı ve ölüm oranlarında en önde gelen kanserlerden biridir. Hastalığın tedavisini kolaylaştırma amacıyla fitokannabinoidlerin *in siliko* gen profillemeye ve protein-etkileşim analizleri yapılarak akciğer kanseri hastalığında kullanılabilirlik derecesinin ortaya çıkarılması amaçlanmıştır.

Gereç ve Yöntem: TCGA veri bankasından akciğer kanseri hastaları için alınan transkriptomik çıktılarının kannabinoid yolağıyla olan ilişkileri skorlanarak DGE analizine göre belirlenen survival genlerin STRING veri tabanı kullanılarak interaktom analizleri yapılmıştır. Kombinasyon puanı 0,9 ve üzeri olan en güçlü etkileşimler ADAMTS1, AGER, AKT1, ALK, CCNA2, CNR1, CNR2, FABP4, FKBP4 ve RET proteinleridir. Bu proteinlerin *in siliko* olarak yapılan biyoinformatik destekli protein-ligand docking çalışmalarında Schrödinger 2021-2 yazılımı kullanılmıştır. Akciğer kanseri hastalarında hedef olarak belirlenen proteinlerle kannabidiol etkileşimleri incelenmiştir.

Bulgular: İlaç adayı bileşiklerin tasarımında, hedef yapıdaki bileşiklerin bağlanma parametrelerinin değerleri *in siliko* yaklaşımlar için önemlidir. Bu sebeple farklı proteinler üzerinde etkili olabileceği öngörülen CBD fitokannabinoidinin bağlanma parametre değerleri hesaplanmıştır. En iyi bağlanma parametre değerlerine sahip olan bileşikler için 2D ve 3D etkileşim diyagramları sunulmuştur.

Sonuç: Akciğer kanseri hastalarında kannabinoid yolağı kemoterapötik uygulamalarda önemli olabilecek ve DGE survival analiziyle ortaya çıkarılan pek çok gen, pseudogen ve noncoding RNA gen ile etkileşim içinde olduğu anlaşılmaktadır. Bu anlamda kannabinoidlerin akciğer kanseri hastalarında kemoterapötik olarak ya da adjuvan tedavide kullanımları bu hastalığın genel sağ kalımı üzerinde olumlu etkiler yapabileceği açıkça görülmektedir. Henüz bakir olan pek çok kenevir türü, içerdiği çeşitli kannabinoidlerle açık kaynak kodlu bir ilaç ham madde platformu sunmaktadır.

Anahtar kelimeler: Akciğer kanseri, kannabidiol, moleküler docking, ilaç tasarımı

Determination of the Interactions of Certain Proteins, Crucial in Lung Cancer Therapeutic Applications, with Cannabidiol Through *in Silico* Approaches

Ali Aydın¹, Burçin Türkmenoğlu²

¹ Department of Basic Medical Sciences, Yozgat Bozok University Faculty of Medicine, Yozgat, Türkiye

² Department of Basic Pharmaceutical Sciences, Erzincan Binali Yıldırım University Faculty of Pharmacy, Erzincan, Türkiye

ABSTRACT

Objectives: Lung cancer ranks among the leading cancers in both incidence and mortality rates. To facilitate the treatment of the disease, it was aimed to assess the potential utility of phytocannabinoids in lung cancer through *in silico* gene profiling and protein interaction analysis.

Materials and Methods: The relationships of the transcriptomic outputs taken from the TCGA database for lung cancer patients and the cannabinoid pathway were scored. Additionally, survival genes identified through DGE analysis were further analyzed using the STRING database. The most robust interactions, scoring 0.9 and above involved the following proteins: ADAMTS1, AGER, AKT1, ALK, CCNA2, CNR1, CNR2, FABP4, FKBP4, and RET. Schrödinger 2021-2 software was employed for bioinformatics-supported, *in silico* protein-ligand docking studies of these proteins. Analysis was performed on the interactions between cannabidiol and proteins identified as targets in lung cancer patients.

Results: In the design of drug candidate compounds, the values of binding parameters in the target structure play a crucial role in in silico approaches. For this reason, the binding parameter values of the CBD phytocannabinoid, which is predicted to be effective on different proteins, were calculated. 2D and 3D interaction diagrams are presented for the compounds with the best binding parameter values.

Conclusion: It is understood that the cannabinoid pathway in lung cancer patients interacts with many genes, pseudogenes, and noncoding RNA genes that may be important in chemotherapeutic applications and revealed by DGE survival analysis. In this sense, it is clear that the use of cannabinoids as a chemotherapeutic or adjuvant treatment in lung cancer patients can positively impact the overall survival of the disease. The diverse cannabinoids found in several untapped hemp species offer an open-source platform for pharmaceutical raw materials.

Keywords: Lung cancer, cannabidiol, molecular docking, drug design

Makine Öğrenmesi Yöntemleri ile Antibiyotiğe Dirençli Bakterilerin DNA Dizilerinden Tanımlanması

Seda Şahin^{ID}, Asiye Bükre Çay^{ID}, Ramazan Gülmez^{ID}, Barış Yıldız^{ID}

Çankırı Karatekin Üniversitesi Mühendislik Fakültesi, Bilgisayar Mühendisliği Bölümü, Çankırı, Türkiye

ÖZET

Amaç: Bakterilerdeki antibiyotik direncinin halk sağlığı açısından önemli bir tehdit oluşturması, hızlı ve doğru tespit yöntemleri için yenilikçi yaklaşımlar gerektirmektedir. Bu çalışmada antibiyotiğe dirençli bakterileri, doğrudan DNA dizilerinden makine öğrenimi algoritmalarıyla tanımlamanın avantajını ele alıyoruz.

Gereç ve Yöntem: Çalışma, antibiyotik direnciyle ilişkili spesifik genetik belirteçleri tanımlamayı amaçlayan bakteriyel DNA örneklerini analiz etmek için biyoinformatik araçların gücünden yararlanıldı. *Neisseria gonorrhoeae* adı verilen bakteri türüne dayalı açık kaynaklı bir veri seti kullanıldı. Bu veri seti genom dizilimi ve antibiyotik direnci verilerini içermektedir. Bu çalışmada Elastic net, Support vector machine, XGBoost ve Random forest gibi makine öğrenimi yöntemleri kullanılmıştır.

Bulgular: Elastic net için duyarlılık (recall) değeri 0,9970, kesinlik (precision) değeri 0,9985, F1 değeri 0,9978 ve doğruluk (accuracy) değeri %99,56 olarak bulunmuştur. Support vector machine için duyarlılık (recall) değeri 1, kesinlik (precision) değeri 0,9985, F1 değeri 0,9992 ve doğruluk (accuracy) değeri %99,85 olarak bulunmuştur. XGBoost için duyarlılık (recall) değeri 0,9910, kesinlik (precision) değeri 0,9440, F1 değeri 0,9669 ve doğruluk (accuracy) değeri %94,24 olarak bulunmuştur. Random forest için duyarlılık (recall) değeri 0,9728, kesinlik (precision) değeri 0,9781, F1 değeri 0,9754 ve doğruluk (accuracy) değeri %96,12 olarak bulunmuştur.

Sonuç: Genomik veri analizi ve makine öğreniminin entegrasyonu yoluyla araştırmamız, yalnızca genetik düzeyde antibiyotik direnç mekanizmalarının anlaşılmasını geliştirmekle kalmayıp aynı zamanda klinik teşhis ve tedavi stratejileri için pratik bilgiler de sağlamaktadır.

Anahtar kelimeler: Antibiyotik direnci, genomik veri analizi, makine öğrenimi

Identification of Antibiotic-Resistant Bacteria From DNA Sequences Using Machine Learning Methods

Seda Şahin^{ID}, Asiye Bükre Çay^{ID}, Ramazan Gülmez^{ID}, Barış Yıldız^{ID}

Department of Computer Engineering, Çankırı Karatekin University Faculty of Engineering, Çankırı, Türkiye

ABSTRACT

Objectives: Antibiotic resistance in bacteria poses a significant threat to public health, necessitating innovative approaches for fast and accurate detection methods. In this study, we address the advantage of identifying antibiotic-resistant bacteria directly from DNA sequences using machine learning algorithms.

Materials and Methods: The study leverages bioinformatics tools to analyze bacterial DNA samples with the goal of identifying specific genetic markers associated with antibiotic resistance. An open-source dataset, focusing on the bacteria type *Neisseria gonorrhoeae*, was utilized, comprising genome sequences and antibiotic resistance data. The study employed various machine learning methods, including Elastic net, Support Vector Machine, XGBoost, and Random Forest.

Results: For Elastic net, the recall was found to be 0.9970, the precision was 0.9985, the F1 value was 0.9978 and the accuracy was 99.56%. For the Support Vector Machine, the recall was found to be 1, the precision was 0.9985, the F1 value was 0.9992 and the accuracy was 99.85%. For XGBoost, the recall was found to be 0.9910, the precision 0.9440, the F1 value was 0.9669, and the accuracy was 94.24%. For Random Forest, the recall was found to be 0.9728, the precision was 0.9781, the F1 value was 0.9754 and the accuracy was 96.12%.

Conclusion: Through the integration of genomic data analysis and machine learning, our research not only advances the understanding of antibiotic resistance mechanisms at the genetic level but also provides practical insights for clinical diagnosis and treatment strategies.

Keywords: Antibiotic resistance, genomic data analysis, machine learning

Yenidoğan ve Çocuk Yoğun Bakım Ünitesinde Hızla Kötüleşen Kritik Yenidoğan ve Süt Çocuklarında Hızlı Yeni Nesil Dizileme ile Genetik Tanı - İlk Türkiye Deneyimi

Bengisu Güner Yılmaz¹, **Özlem Akgün Doğan^{2,3}**, **Özkan Özdemir^{3,5}**, **Kaya Bilgüvar^{3,6}**, **Özden Hatırnaz Ng^{3,5}**, **Gülşah Şebnem Özköse⁴**, **Eylül Aydın⁴**, **Ayça Yiğit⁴**, **Aybike Bulut⁴**, **Atalay Demirel⁷**, **Selma Aktaş⁷**, **Baran Cengiz Arcagök⁷**, **Ebru Kazancı⁷**, **Sare Güntülü Şık⁸**, **İbrahim Bingöl⁸**, **Özge Umur⁸**, **Melike Ersoy Olbak⁹**, **Serdar Beken⁷**, **Uğur Işık¹⁰**, **Ayşe Toygar Korkmaz⁷**, **Agop Çıtak⁸**, **Adil Mardinoğlu¹¹**, **Uğur Özbek¹²**, **Yasemin Alanay^{2,3}**

¹ Acıbadem Mehmet Ali Aydınlar Üniversitesi Tıp Fakültesi, Çocuk Sağlığı ve Hastalıkları Anabilim Dalı, İstanbul, Türkiye

² Acıbadem Mehmet Ali Aydınlar Üniversitesi Tıp Fakültesi, Çocuk Sağlığı ve Hastalıkları Anabilim Dalı, Çocuk Genetik Bilim Dalı, İstanbul, Türkiye

³ Acıbadem Mehmet Ali Aydınlar Üniversitesi, Nadir Hastalıklar ve Yetim İlaçlar Uygulama ve Araştırma Merkezi (ACURARE), İstanbul, Türkiye

⁴ Acıbadem Mehmet Ali Aydınlar Üniversitesi, Sağlık Bilimleri Enstitüsü, Genom Çalışmaları Programı, İstanbul, Türkiye

⁵ Acıbadem Mehmet Ali Aydınlar Üniversitesi Tıp Fakültesi, Tıbbi Biyoloji Anabilim Dalı, İstanbul, Türkiye

⁶ Acıbadem Mehmet Ali Aydınlar Üniversitesi Tıp Fakültesi, Tıbbi Genetik Anabilim Dalı, İstanbul, Türkiye

⁷ Acıbadem Mehmet Ali Aydınlar Üniversitesi Tıp Fakültesi, Çocuk Sağlığı ve Hastalıkları Anabilim Dalı, Neonatoloji Bilim Dalı, İstanbul, Türkiye

⁸ Acıbadem Mehmet Ali Aydınlar Üniversitesi Tıp Fakültesi, Çocuk Sağlığı ve Hastalıkları Anabilim Dalı, Çocuk Yoğun Bakım Bilim Dalı, İstanbul, Türkiye

⁹ Bakırköy Dr. Sadi Konuk Eğitim ve Araştırma Hastanesi, Çocuk Sağlığı ve Hastalıkları Anabilim Dalı, Çocuk Beslenme ve Metabolizma Hastalıkları Bilim Dalı, İstanbul, Türkiye

¹⁰ Acıbadem Mehmet Ali Aydınlar Üniversitesi Tıp Fakültesi, Çocuk Sağlığı ve Hastalıkları Anabilim Dalı, Çocuk Nörolojisi Bilim Dalı, İstanbul, Türkiye

¹¹ SZA OMICS Genetik Tanı Merkezi, İstanbul, Türkiye

¹² IBG- İzmir Biyotıp ve Genom Merkezi, İzmir, Türkiye

ÖZET

Amaç: Nadir genetik durumlar, kritik çocuk hastaların yoğun bakım ünitesi (YBÜ) yatışı sırasında sıklıkla şüphelenilen ancak çoğunlukla tanımlanamayan; çocukluk çağı morbidite ve mortalitesinin önemli nedenleridir. Bu durumları aydınlatmak ve etkili hasta yönetimi sağlamak amacıyla, hızlı yeni nesil dizileme (hızlı-YND) testlerinin YBÜ'deki kritik hastalarda kullanımı dünya çapında son yıllarda oldukça yaygınlaşmış ve hatta bazı ulusal sağlık sistemlerine entegre edilmiştir. Bu ön çalışmada, Türkiye'de ilk kez YBÜ'de hızla kötüleşen, kritik yenidoğan ve süt çocuklarında hızlı-YND'nin kullanılabilirliğinin, tanısal gücünün, uygulanma süresinin ve klinik faydasının değerlendirilmesi amaçlanmıştır.

Gereç ve Yöntem: Çalışmaya, İstanbul Acıbadem hastanelerinin YBÜ'lerinde takip edilen bir yaş altı tanısı koyulamamış kritik hastalar prospektif olarak dâhil edilmiştir. Hastalar ve ebeveynlerinden alınan kan ile üçlü hızlı tüm genom dizileme (rWGS) yapılmıştır. Çalışma hastayı takip eden klinisyenler, konsültanlar, genetik ve biyoinformatik uzmanları tarafından oluşturulan multidisipliner ekip tarafından yürütülmüştür. Laboratuvarımızda geliştirilen akış hattı ile yapılan veri analizinde fenotipleme ve pedigrî bilgilerine göre varyant önceliklendirmesi yapılmıştır.

Bulgular: rWGS uygulanan 10 hastanın üçüne (%30) kesin tanı konuldu. Bu hastalarda IDH3A, SCO1 ve FBN1 genlerinde mutasyon tespit edildi. rWGS sonucu ile altı (%60) hastanın tanı belirsizliği ortadan kaldırıldı, beş (%50) hastanın klinik takip ve tedavi planında değişikliği gidildi. rWGS sürecinde örnek alımından sonuç raporuna kadar geçen ortalama süre 169 saattir (124-240 saat).

Sonuç: Hızlı yeni nesil dizileme testlerinin YBÜ'lerde kullanımı kritik hastaların tanı yolculuğunda ve klinik karar verme süreçlerinde umut vadeden güçlü bir araç olmaya başlamıştır. Bu pilot çalışmada, hızlı-YND'nin yararlılığı ve etkinliğinin yanı sıra multidisipliner yaklaşımla Türkiye'de uygulanabilirliği de gösterilmiştir ancak yaygınlaştırılması ve sağlık sistemine entegrasyonu için daha kapsamlı çalışmalara ihtiyaç duyulmaktadır.

Anahtar kelimeler: Hızlı yeni nesil dizileme, tüm genom dizileme, hızla kötüleşen kritik hasta, yenidoğan yoğun bakım ünitesi, çocuk yoğun bakım ünitesi

Genetic Diagnosis Through Rapid Next-Generation Sequencing for Acutely Deteriorating Critically Ill Newborns and Infants in Neonatal and Pediatric Intensive Care Units - The First Experience in Türkiye

Bengisu Güner Yılmaz¹, **Özlem Akgün Doğan^{2,3}**, **Özkan Özdemir³⁻⁵**, **Kaya Bilgüvar^{3,6}**, **Özden Hatırnaz Ng^{3,5}**, **Gülşah Şebnem Özköse⁴**, **Eylül Aydın⁴**, **Ayça Yiğit⁴**, **Aybike Bulut⁴**, **Atalay Demirel⁷**, **Selma Aktaş⁷**, **Baran Cengiz Arcagök⁷**, **Ebru Kazancı⁷**, **Sare Güntülü Şık⁸**, **İbrahim Bingöl⁸**, **Özge Umur⁸**, **Melike Ersoy Olbak⁹**, **Serdar Beken⁷**, **Uğur Işık¹⁰**, **Ayşe Toygar Korkmaz⁷**, **Agop Çıtak⁸**, **Adil Mardinoğlu¹¹**, **Uğur Özbek¹²**, **Yasemin Alanay^{2,3}**

¹ Department of Pediatrics, Acıbadem Mehmet Ali Aydınlar University Faculty of Medicine, İstanbul, Türkiye

² Division of Genetics, Department of Pediatrics, Acıbadem Mehmet Ali Aydınlar University Faculty of Medicine, İstanbul, Türkiye

³ Rare Diseases and Orphan Drugs Application and Research Center (ACURARE), Acıbadem Mehmet Ali Aydınlar University, İstanbul, Türkiye

⁴ Department of Genome Studies, Health Sciences Institute, Acıbadem Mehmet Ali Aydınlar University, İstanbul, Türkiye

⁵ Department of Medical Biology, Acıbadem Mehmet Ali Aydınlar University Faculty of Medicine, İstanbul, Türkiye

⁶ Department of Medical Genetics, Acıbadem Mehmet Ali Aydınlar University, Faculty of Medicine, İstanbul, Türkiye

⁷ Division of Neonatology, Department of Pediatrics, Acıbadem Mehmet Ali Aydınlar University Faculty of Medicine, İstanbul, Türkiye

⁸ Division of Intensive Care, Department of Pediatrics, Acıbadem Mehmet Ali Aydınlar University Faculty of Medicine, İstanbul, Türkiye

⁹ Division of Metabolism and Nutrition, Department of Pediatrics, Bakırköy Dr. Sadi Konuk Training and Research Hospital, İstanbul, Türkiye

¹⁰ Division of Neurology, Department of Pediatrics, Acıbadem Mehmet Ali Aydınlar University Faculty of Medicine, İstanbul, Türkiye

¹¹ Genetic Diagnosis Center, SZA OMICS, İstanbul, Türkiye

¹² IBG-İzmir Biomedicine and Genome Center, İzmir, Türkiye

ABSTRACT

Objectives: Rare genetic conditions, frequently suspected but often unidentified during intensive care admissions, are major contributors to pediatric mortality and morbidity. While rapid genomic testing in critical care has been adopted globally, this pilot study represents the inaugural effort to evaluate the feasibility and utility of rapid whole genome sequencing (rWGS) within the Turkish healthcare context.

Materials and Methods: In this pioneering effort, we conducted a prospective Trio rWGS on families with infants (under one year) with undiagnosed illnesses in intensive care units of İstanbul Acıbadem hospitals. A multidisciplinary team, consisting of primary care providers, specialists, and genetics and bioinformatics experts, was assembled. An in-house pipeline was tailored to focus on phenotype-driven variants in recognized disease-causing genes while considering pedigree information.

Results: Of the 10 infants subjected to Trio WGS, three (30%) received a definitive diagnosis. We identified mutations in the IDH3A, SCO1, and FBN1 genes. rWGS concluded the diagnostic journey for six (60%) patients and directly influenced the clinical care of five (50%). The average turnaround from sample collection to report was 169 hours (range= 124-240 hours).

Conclusion: Whole genome sequencing emerges as a promising diagnostic tool for critically ill children, crucial for both diagnosis and clinical decision-making. This inaugural study underscores the potential and feasibility of rapid genomics in Türkiye, achieved through a collaborative approach. However, its broader integration into the national healthcare system warrants further exploration.

Keywords: Rapid next-generation sequencing, whole genome sequencing, acutely deteriorating critically ill patient, pediatric intensive care unit, neonatal intensive care unit

Preeklampsi Hastalarında Plasental Patofizyolojinin Anlaşılmasına Yönelik Metabolomik ve Transkriptomik Verilerin Analizine Dayalı Multi-Omik Yaklaşım

Akın Mumcu^{1,2}, Erdiñ Sarıdoğın^{1,3}, Senem Arda Düz^{1,3}, Görkem Tuncay^{1,3}, Ali Erdoğan^{1,4}, Kadri Karer⁵, Taylan Onat^{1,3}, Abdullah Karaer^{1,3}, Berat Doğan^{1,4}

¹ İnönü Üniversitesi Üreme Bilimleri ve İleri Biyoinformatik Uygulama ve Araştırma Merkezi, Malatya, Türkiye

² İnönü Üniversitesi Bilimsel ve Teknolojik Araştırma Merkezi, NMR Laboratuvarı, Malatya, Türkiye

³ İnönü Üniversitesi Tıp Fakültesi, Kadın Hastalıkları ve Doğum Anabilim Dalı, Malatya, Türkiye

⁴ İnönü Üniversitesi Mühendislik Fakültesi, Biyomedikal Mühendisliği Bölümü, Malatya, Türkiye

⁵ Pamukkale Üniversitesi Tıp Fakültesi, Tıbbi Genetik Anabilim Dalı, Denizli, Türkiye

ÖZET

Amaç: Bu çalışmanın amacı, preeklampsi hastalarının plasentalarından alınan doku örneklerinin sağlıklı kontrollere kıyasla metabolomik ve transkriptomik düzeyde ne tür farklılıklar gösterdiğini tespit etmek ve elde edilen multi-omik veriyi sistem biyolojisi yaklaşımıyla analiz edip preeklampsi hastalığının patofizyolojisini daha iyi anlamaktır.

Gerçek ve Yöntem: Hasta ve kontrol gruplarından alınan doku örneklerinin 1H HRMAS NMR metabolomik ve RNA-seq transkriptomik analizleri gerçekleştirilmiştir. Elde edilen metabolomik ve transkriptomik veriler ayrı ayrı ve entegre olarak analiz edilmiştir.

Bulgular: Metabolomik analizler preeklampsi hastalarında laktat, glutamin, glutamat, myo-inositol, valin, lösin, alanin, izölösin, kreatin, kolin, gliserofosfokolin, fosfokolin, o-fosfoetanolamin gibi birçok metabolitin seviyesinde kontrol grubuna kıyasla artış olduğunu ortaya koymuştur. Bu değişiklikler Krebs çevrimi, enerji metabolizması, aminoasit metabolizması ve lipid metabolizması gibi birçok metabolik yolağı etkilemektedir. Transkriptomik analizler ise preeklampsili grup ile kontrol grubu arasında toplamda 47 genin (10 ekspresyonu artan, 37 ekspresyonu azalan) farksal olarak ($|\log_2FC > 1|$ ve $p\text{-adj} < 0,05$) ifade edildiğini ortaya koymuştur. Bu genler vazokonstriksiyon, anjiyogenez, hormonal denge, enflamasyon, oksidatif stres ve kolajen bütünlüğü gibi birçok biyolojik sürecin yürütülmesinde kritik roller üstlenen genlerdir. Multi-omik analiz ile bazı metabolitler ile diğer bazı genlerin ilişkili olduğu tespit edilmiştir. Tespit edilen genlerle oluşturulan bir protein-protein etkileşim ağının istatistiksel olarak anlamlı ($p\text{-değeri} < 1e-16$) etkileşimler içerdiği saptanmıştır.

Sonuç: Preeklampsi hastalarının plasenta doku örnekleri sağlıklı kontrollere kıyasla metabolomik ve transkriptomik düzeyde farklılıklar göstermektedir. Multi-omik analiz, standart analiz ile elde edilemeyen etkileşimleri açığa çıkararak hastalığın patofizyolojisini aydınlatmaya katkı sağlamaktadır.

Anahtar kelimeler: Preeklampsi, plasenta, metabolomik, transkriptomik, multi-omik

Multi-Omics Approach Based on the Analysis of Metabolomics and Transcriptomics Data for Understanding Placental Pathophysiology in Preeclampsia Patients

Akın Mumcu^{1,2}, Erdiñ Saridođan^{1,3}, Senem Arda Düz^{1,3}, Görkem Tuncay^{1,3}, Ali Erdođan^{1,4}, Kadri Karer⁵, Taylan Onat^{1,3}, Abdullah Karaer^{1,3}, Berat Dođan^{1,4}

¹ Reproductive Sciences & Advanced Bioinformatics Application & Research Center, İnönü University, Malatya, Türkiye

² Laboratory of NMR, İnönü University Scientific and Technological Research Center, Malatya, Türkiye

³ Division of Reproductive Endocrinology and Infertility, Department of Obstetrics and Gynecology, İnönü University Faculty of Medicine, Malatya, Türkiye

⁴ Department of Biomedical Engineering, İnönü University Faculty of Engineering, Malatya, Türkiye

⁵ Department of Medical Genetics, Pamukkale University Faculty of Medicine, Denizli, Türkiye

ABSTRACT

Objectives: This study aimed to determine the differences in metabolomics and transcriptomics levels in tissue samples obtained from the placentas of patients with preeclampsia compared to healthy controls and to analyze the obtained multi-omics data through a systems biology approach to better understand the pathophysiology of preeclampsia.

Materials and Methods: ¹H HRMAS NMR metabolomics and RNA-seq transcriptomics analyses were performed on tissue samples obtained from the two groups. Subsequently, a multi-omics analysis of the obtained data was performed to identify dysregulated metabolites and genes, and their association with each other.

Results: Metabolomics analyses revealed increased levels of several metabolites including lactate, glutamine, glutamate, myo-inositol, valine, leucine, alanine, isoleucine, creatine, choline, glycerophosphocholine, phosphocholine, and o-phosphoethanolamine in patients with preeclampsia compared to the healthy controls. These changes affect the Krebs cycle, energy metabolism, amino acid metabolism, and lipid metabolism. Transcriptomics analyses demonstrated a total of 47 genes (10 upregulated, 37 downregulated) that were differentially expressed ($|\log_2FC| > 1$ and $p\text{-adj} < 0.05$) between the two groups. These genes play critical roles in various biological processes such as vasoconstriction, angiogenesis, hormonal balance, inflammation, oxidative stress, and collagen integrity. Multi-omics analysis identified associations between certain metabolites and several other genes. A protein-protein interaction network created with the identified genes was found to have statistically significant interactions ($p\text{-value} < 1e-16$).

Conclusion: Placental tissue samples exhibit significant differences in metabolomics and transcriptomics levels between the two groups. Multi-omics analysis unveils complex interactions that cannot be obtained through standard analyses.

Keywords: Preeclampsia, placenta, metabolomics, transcriptomics, multi-omics

Akciğer Kanserine Karşı CAR-T Hücre Tedavisi için Tek Hücre RNA Dizileme Verileriyle Yeni Hedef Antijenlerin Belirlenmesi

Ebru Kocakaya¹, **Gizem Baydemir²**, **Yasin Kaymaz¹**

¹ Ege Üniversitesi Mühendislik Fakültesi, Biyomühendislik Anabilim Dalı, İzmir, Türkiye

² Bahçeşehir Üniversitesi Mühendislik ve Doğa Bilimleri Fakültesi, Moleküler Biyoloji ve Genetik Bölümü, İstanbul, Türkiye

ÖZET

Amaç: Akciğer kanserine yönelik henüz onaylanmış bir CAR-T hücre tedavisi bulunmadığından mevcut klinik araştırmalarda denenmekte olan antijenlerin yerine, kanser hücrelerine özgü, daha etkili yeni antijen adaylarının tek hücre RNA dizileme verilerinden veri madenciliğiyle keşfedilmesi hedeflenmiştir.

Gereç ve Yöntem: Çalışmamızda, 21 farklı çalışmadaki 298 hastadan 500'den fazla örneğin tek hücre RNA dizileme (scRNA-Seq) verileri toplanarak işlendi. Eşlenik sağlıklı dokulara ait transkriptom verileriyle birlikte üç milyondan fazla hücreyle veri tabanı oluşturuldu. Büyük potansiyeline rağmen scRNA-Seq'le elde edilen gen ekspresyon dağılımlarının limitli olması, rastgele kayıplar ve gürültü, diferansiyel ekspresyon gerçekleştiren genlerin tespitini kısıtlamaktadır. Bu nedenle geliştirdiğimiz iş akışıyla scRNA-Seq verileri işlenip benzer hücre gruplarının kümelenmesiyle, elde edildikleri dokuyu temsil eden kümülatif gen ekspresyonları (pseudo-bulk) hesaplandı. Tümöre has yüksek seviyelerde ifade edilen genlerin eşleniği sağlıklı dokulardan üretilen "epitel" hücrelerle diğer organların epitel hücreleri arasında karşılaştırma analizleri gerçekleştirildi. Aday antijenler seçilirken potansiyel CAR-T hücre tedavisinde muhtemel sitotoksiteyi engellemek için genlerin diğer tüm sağlıklı doku ve hücrelerdeki ifadesinin düşük olmasıyla hücre yüzeyinde fonksiyon göstermeleri dikkate alındı.

Bulgular: Akciğer kanseri alt türleri adenokarsinoma, skuamöz hücreli karsinom ve alt evreleri bazında gruplandırılan hastaların scRNA-Seq verileriyle belirlediğimiz kanser hücre transkriptom profilleri, normal epitel hücrelerle karşılaştırıldığında, hücre membran proteinlerini kodlayan genlerden SLC2A1, PSCA, UPK1B ve LY6D'nin gen ifadelerinin istatistik anlamlılık seviyesinde ($p < 0,05$) en az iki kat arttığı tespit edildi. Bu artışın diğer doku ve hücre türlerine nazaran daha yüksek seviyelerde olması ve CAR-T hücre tedavisindeki potansiyel kullanımı veri tabanlarındaki mevcut verilerle eşleştirildi.

Sonuç: Çalışmamız sonucunda, akciğer kanserine karşı CAR-T hücre tedavisi için aday olabilecek SLC2A1, PSCA, UPK1B ve LY6D genleri ve ifade ettikleri yüzey proteinleri bulunmuştur.

Anahtar kelimeler: Akciğer kanseri, CAR-T hücre tedavisi, tek hücre RNA dizileme, pseudobulk analizi

Identification of Potential Target Antigens for CAR-T Cell Therapy Against Lung Cancer Using scRNA-Seq and Pseudobulk Analysis

Ebru Kocakaya¹, Gizem Baydemir², Yasin Kaymaz¹

¹ Department of Bioengineering, Ege University Faculty of Engineering, İzmir, Türkiye

² Department of Molecular Biology and Genetics, Bahçeşehir University Faculty of Engineering and Natural Sciences, İstanbul, Türkiye

ABSTRACT

Objectives: Human single-cell RNA sequencing cell atlases were analyzed with the pseudobulk method to compare the transcriptome of lung cancer cells with non-cancerous cell types to identify potential target antigen candidates for CAR-T cell therapy against lung cancer.

Materials and Methods: Single-cell RNA sequencing (scRNA-Seq) data of 500 samples from 298 patients across 21 different studies were merged. Database with transcriptome data from matching healthy tissues had over three million cells. Despite the significant potential, and limited gene expression distributions, drop-outs from scRNA-Seq restrict the identification of differentially expressed genes. Our workflow clustered similar cells into groups to compute cumulative gene expressions (pseudo-bulk) to represent their origin tissue. Comparative analyses were performed between genes expressed at high levels in tumors and those expressed by epithelial cells in corresponding healthy tissues and other organs. Candidate antigens were identified considering low expression in healthy tissues, with their expression on the cell surface, to prevent cytotoxicity in CAR-T cell therapy.

Results: Transcriptomes of lung cancer patients were grouped into adenocarcinoma, and squamous cell carcinoma using scRNA-Seq data. Compared to normal epithelial cells, gene expressions of membrane proteins encoded by SLC2A1, PSCA, UPK1B, and LY6D were found to have increased by at least two-fold at a statistically significant level ($p < 0.05$) in cancer cells. This increase was higher compared to other tissue and cell types, and they were matched with metadata information from databases for potential candidates for CAR-T cell therapy.

Conclusion: We have identified SLC2A1, PSCA, UPK1B, and LY6D genes and their expressed proteins as candidates for CAR-T cell therapy against lung cancer.

Keywords: Lung cancer, pseudobulk analysis, scRNA-Seq, CAR-T cell therapy

Yeni Nesil Dizileme (YND) ile Çocukluk Çağı Kanserlerinde Genetik Yatkınlığın Belirlenmesi ve Klinik Yansıması

Gizem Önder^{1,2}, **Özkan Özdemir^{2,3}**, **Ufuk Amanvermez²⁻⁴**, **Cengiz Canpolat⁵**, **Fatih Erbey^{6,7}**, **Banu Oflaz Sözmen⁷**, **Fikret Asarcıklı⁷**, **Davut Albayrak⁸**, **Elif İnce⁹**, **Gül Nihal Özdemir¹⁰**, **Yunus Murat Akçabelen¹¹**, **Neşe Yaralı¹¹**, **Namık Yaşar Özbek¹¹**, **İkbal Okbozkaya¹¹**, **Turan Bayhan¹¹**, **Dilek Kaçar¹¹**, **Seda Aras¹²**, **Ömer Doğru¹²**, **Koray Yalçın¹³**, **Berk Ergün¹⁴**, **Eylül Aydın^{2,4}**, **Sezer Akyöney¹⁵**, **Fulya Taylan¹⁶**, **Ann Nordgren¹⁶**, **Müge Sayitoğlu¹⁷**, **Uğur Özbek¹⁸**, **Nihat Buğra Ağaoğlu¹⁹⁻²¹**, **Özden Hatırnaz Ng²²**

- ¹ Acıbadem Mehmet Ali Aydınlar Üniversitesi Sağlık Bilimleri Enstitüsü, Biyokimya ve Moleküler Biyoloji, İstanbul, Türkiye
² Acıbadem Mehmet Ali Aydınlar Üniversitesi Nadir Hastalıklar ve Yetim İlaçlar Uygulama ve Araştırma Merkezi (ACURARE), İstanbul, Türkiye
³ Acıbadem Mehmet Ali Aydınlar Üniversitesi Sağlık Bilimleri Enstitüsü, Türkiye
⁴ Acıbadem Mehmet Ali Aydınlar Üniversitesi Sağlık Bilimleri Enstitüsü, Genom Çalışmaları Yüksek Lisans programı, İstanbul, Türkiye
⁵ Acıbadem Altunizade Hastanesi, Çocuk Onkoloji Kliniği, İstanbul, Türkiye
⁶ Koç Üniversitesi Hastanesi, Çocuk Hematoloji ve Onkoloji Kliniği, İstanbul, Türkiye
⁷ Koç Üniversitesi Tıp Fakültesi, Çocuk Sağlığı Anabilim Dalı, İstanbul, Türkiye
⁸ Samsun Medical Park Hastanesi, Çocuk Hematolojisi Kliniği, Samsun, Türkiye
⁹ Ankara Üniversitesi Tıp Fakültesi, Çocuk Hematolojisi Anabilim Dalı, Çocuk Hematoloji ve Onkoloji Bilim Dalı, Ankara, Türkiye
¹⁰ Liv Hospital Ulus, Çocuk Hematolojisi Kliniği, İstanbul, Türkiye
¹¹ Ankara Şehir Hastanesi, Çocuk Hematoloji ve Onkoloji Kliniği, Ankara, Türkiye
¹² Marmara Üniversitesi Tıp Fakültesi Hastanesi, Çocuk Sağlığı ve Hastalıkları, İstanbul, Türkiye
¹³ Bahçeşehir Üniversitesi, Göztepe Medikal Park Hastanesi, Çocuk Hematolojisi Anabilim Dalı, İstanbul, Türkiye
¹⁴ Geniva Bilişim Sağlık Hizmetleri Anonim Şirketi, İstanbul, Türkiye
¹⁵ Acıbadem Mehmet Ali Aydınlar Üniversitesi, Sağlık Bilimleri Enstitüsü, Biyoinformatik ve Biyoistatistik Anabilim Dalı, İstanbul, Türkiye
¹⁶ Karolinska Enstitüsü, Moleküler Tıp ve Cerrahi Anabilim Dalı, Stockholm, İsveç
¹⁷ İstanbul Üniversitesi, Aziz Sanca Deneysel Tıp Araştırma Enstitüsü, İstanbul, Türkiye
¹⁸ İzmir Dokuz Eylül Üniversitesi, Uluslararası Biotıp ve Genom Enstitüsü (iBG), İzmir, Türkiye
¹⁹ Ümraniye Eğitim ve Araştırma Hastanesi, Kanser Genetiği Kliniği İstanbul, Türkiye
²⁰ IKF Institut für Klinische Krebsforschung GmbH, Frankfurt, Almanya
²¹ Dana Farber Kanser Enstitüsü, Boston, Amerika Birleşik Devletleri
²² Acıbadem Mehmet Ali Aydınlar Üniversitesi, Tıbbi Biyoloji Anabilim Dalı, İstanbul, Türkiye

ÖZET

Amaç: Çocukluk çağı kanserlerinin yaklaşık %15'inin kalıtsal patojenik varyant kaynaklı olduğu bildirilmiştir. Bu çalışmada, çocukluk çağı kanserlerine sahip vakalarda kanser yatkınlığına neden olabilecek genlerde eşey hattı (germline) değişimlerin yeni nesil dizileme tabanlı yöntemlerle belirlenmesi ve hastalığın tedavi yönetimi ile ailede taşıyıcı bireylerin belirlenerek koruyucu/önleyici yaklaşımlara yönlendirilmesi hedeflenmiştir.


















Gereç ve Yöntem: Türkiye'deki farklı merkezlerden birimize yönlendirilen 78 indeks [ortalama yaş= 12 (min-maks= 2/18, cinsiyet= 36K/42E)], 12 etkilenmiş akraba ve 151 etkilenmemiş aile üyesi çalışmaya dâhil edilmiştir. Vakaların 62'si lösemi/lenfoma tanıyken, diğer vakalar farklı solid tümör tanısı almıştır. Periferik kan/kemik iliği örnekleri ile uygun olan hastalardan alınan ağız içi swab/kıl kökü örneklerinden DNA elde edilerek, 24 indeks vakada ve 47 aile üyelerinde farklı tüm genom (WGS)/tüm ekzom (WES)/klinik ekzom (CES) dizileme teknikleri uygulanmıştır. WES/CES analizleri "GenNext" platformunda, WGS analizleri ise "Scout" platformunda gerçekleştirilmiş, aday varyant validasyonu ve segregasyon analizleri için Sanger dizileme yapılmıştır.

Bulgular: Dizilenen toplam n= 47 örnekten n= 23 adet aday varyant tespit edilmiştir. Bunlardan sekizi patojenik (MLH1, RET, BRCA1, TP53, JAK2, MSH6, WRN, BCHE), dokuzu olası patojenik (ERBB4, FLT3, ETV6, MAP2K2, ETTAA1, JAK3, MYH11, ATR, MVP), altısı ise klinik önemi belirsiz (BRCA2, ABCC6, RAD52, BCNP1, KSR1, TNFRSF9) varyantlardı ve sadece 13'ü ClinVar veri tabanında bildirilmiştir.

Sonuç: Kanserde genetik yatkınlık çalışmaları, hastanın tedavi stratejisi, nakil planlaması ve aile bireylerine koruyucu yaklaşımların sağlanması açısından önem taşımaktadır. Bu çalışmadan elde edilen sonuçlarla üç hastanın kemik iliği nakli için verici değişikliğine gidilmiş, bir hastanın yüksek toksisite nedeni belirlenerek tedavi yöntemi değiştirilmiş, patojenik/olası patojenik varyant taşıyan aile bireyleri olası riskleri açısından bilgilendirilmiş, koruyucu yaklaşım sağlanması açısından tüm ailelere genetik danışmanlık hizmeti verilmesi sağlanmıştır.

Anahtar kelimeler: Yeni nesil dizileme (YND), çocukluk çağı kanserleri, genetik yatkınlık, genetik danışmanlık

Next-Generation Sequencing (NGS) for Determining Genetic Predisposition in Childhood Cancers and Its Clinical Implications

Gizem Önder^{1,2}, Özkan Özdemir^{2,3}, Ufuk Amanvermez²⁻⁴, Cengiz Canpolat⁵, Fatih Erbey^{6,7}, Banu Oflaz Sözmen⁷, Fikret Asarcıklı⁷, Davut Albayrak⁸, Elif İnce⁹, Gül Nihal Özdemir¹⁰, Yunus Murat Akçabelen¹¹, Neşe Yaralı¹¹, Namık Yaşar Özbek¹¹, İkbâl Okbozkaya¹¹, Turan Bayhan¹¹, Dilek Kaçar¹¹, Seda Aras¹², Ömer Doğru¹², Koray Yalçın¹³, Berk Ergün¹⁴, Eylül Aydın^{2,4}, Sezer Akyöney¹⁵, Fulya Taylan¹⁶, Ann Nordgren¹⁶, Müge Sayitoğlu¹⁷, Uğur Özbek¹⁸, Nihat Buğra Ağaoğlu¹⁹⁻²¹, Özden Hatırnaz Ng²²

¹ Department of Biochemistry and Molecular Biology, Acıbadem Mehmet Ali Aydınlar University, Health Sciences Institute, İstanbul, Türkiye

² Rare Diseases and Orphan Drugs Application and Research Center (ACURARE), Acıbadem University, İstanbul, Türkiye

³ Institute of Healthy Science, Acıbadem Mehmet Ali Aydınlar University, İstanbul, Türkiye

⁴ Department of Genome Studies, Acıbadem Mehmet Ali Aydınlar University, Institute of Healthy Science, İstanbul, Türkiye

⁵ Department of Pediatric Oncology, Acıbadem Altunizade Hospital, İstanbul, Türkiye

⁶ Clinic of Pediatric Hematology and Oncology, Hospital of Koç University, İstanbul, Türkiye

⁷ Department of Pediatrics, Koç University Faculty of Medicine, İstanbul, Türkiye

⁸ Department of Pediatric Hematology, Samsun Medical Park Hospital, Samsun, Türkiye

⁹ Division of Pediatric Hematology, Department of Pediatric Hematology and Oncology, Ankara University Faculty of Medicine, Ankara, Türkiye

¹⁰ Department of Pediatric Hematology, Liv Hospital Ulus, İstanbul, Türkiye

¹¹ Department of Pediatric Hematology and Oncology, Ankara City Hospital, Ankara, Türkiye

¹² Department of Pediatric Diseases and Health, Marmara University Faculty of Medicine Hospital, İstanbul, Türkiye

¹³ Department of Pediatric Hematology, Bahçeşehir University Göztepe Medical Park Hospital, İstanbul, Türkiye

¹⁴ GENIVA Information Health Services Company, İstanbul, Türkiye

¹⁵ Department of Bioinformatics and Biostatistic, Acıbadem Mehmet Ali Aydınlar University, Institute of Healthy Science, İstanbul, Türkiye

¹⁶ Department of Molecular Medicine and Surgery, Center for Molecular Medicine Karolinska Institutet Stockholm, Sweden

¹⁷ Department of Genetics, Aziz Sancar Experimental Medicine, İstanbul University, İstanbul, Türkiye

¹⁸ International Biomedicine and Genome Institute (iBG), İzmir Dokuz Eylül University, İzmir, Türkiye

¹⁹ Clinic of Cancer Genetics, Ümraniye Training and Research Hospital, İstanbul, Türkiye

²⁰ IKF Institut für Klinische Krebsforschung GmbH, Frankfurt, Germany

²¹ Dana Farber Cancer Institute, Boston, USA

²² Department of Medical Biology, Acıbadem Mehmet Ali Aydınlar University, İstanbul, Türkiye

ABSTRACT

Objectives: It has been reported that approximately 15% of childhood cancers are associated with hereditary pathogenic variants. This study aimed to identify germline alterations in genes that could predispose to childhood cancers using NGS methods in cases with such cancers, and to determine carrier individuals in families for disease management, guiding them towards preventive approaches.

Materials and Methods: A total of seventy-eight cases (mean age= 12 years, range= 2-18 years, gender distribution= 36 males and 42 females), referred to our center from various locations in Türkiye, were included in the study. Additionally, the study involved 12 affected relatives and 151 unaffected family members. While 62 cases were diagnosed with leukemia/lymphoma, other cases had different solid tumor diagnoses. DNA was extracted from peripheral blood/bone marrow samples as well as oral swabs/hair root samples collected from appropriate patients. Different whole-genome (WGS)/whole-exome (WES)/clinical exome (CES) sequencing techniques were applied to 24 cases and 47 family members. WES/CES analyses were performed on the “GenNext” and WGS analyses were carried out on the “Scout” platform. Sanger sequencing was performed for candidate variant validation and segregation analyses.

Results: From the total of 47 sequenced samples, a total of 23 candidate variants were identified. Eight of these were pathogenic (MLH1, RET, BRCA1, TP53, JAK2, MSH6, WRN, BCHE), nine were likely pathogenic (ERBB4, FLT3, ETV6, MAP2K2, ETAA1, JAK3, MYH11, ATR, MVP), and six were of uncertain clinical significance (BRCA2, ABCC6, RAD52, BCNP1, KSR1, TNFRSF9) with only 13 reported in the ClinVar database.

Conclusion: Genetic susceptibility studies in cancer play a crucial role in shaping the patient’s treatment strategy, aiding transplant planning, and offering preventive approaches for family members. The outcomes of this study prompted donor changes for bone marrow transplants in three patients, modification of the treatment method for one patient due to high toxicity, and the communication of potential risks to family members carrying pathogenic/likely pathogenic variants.

Keywords: Next generation sequencing (NGS), childhood cancers, genetik predisposition, genetic counseling

Kanser Genom Atlası Transkriptomik Verilerinden Pan-Kanser Biyobelirteç Tayini

Hale Nur Korkmaz^{ID}, Müberra Gül Öztürk^{ID}, Muhammed Erkan Karabekmez^{ID}

İstanbul Medeniyet Üniversitesi Mühendislik ve Doğa Bilimleri Fakültesi, Biyomühendislik Anabilim Dalı, İstanbul, Türkiye

ÖZET

Amaç: Kanser somatik hastalıklar arasından sıkça rastlanan, ölümcül ve kompleks bir hastalıktır. Biyoinformatik alanındaki gelişmeler kanserin tespiti, tedavisi ve önlenmesi ile ilgili yeni bakış açıları sunmuştur. Bu çalışmada, Kanser Genom Atlası'ndan elde ettiğimiz transkriptomik verileri farklı araçlarla analiz edip, tekrar eden örüntülerin keşfedilmesiyle sağlanacak iç görüyü hedef olarak farklı kanserlerin teşhis ve tedavisi için yeni ve ortak biyobelirteçlerin tayini amaçlanmıştır.

Gereç ve Yöntem: Dünya Sağlık Örgütü verilerine göre insidans ve mortalite düzeyi yüksek olan kanser türlerine ait transkriptomik veriler, TCGA veri tabanından alındı ve Galaxy araçları yardımıyla analiz edildi. Yapılan istatistiksel analizlerin sonucunda, kanser türlerinin gen setlerindeki diferansiyel olarak ifade edilen genler belirlendi ve g:Profiler çevrim içi yazılım aracı ile fonksiyonel zenginleştirme analizleri yapıldı. TRRUST aracıyla transkripsiyon faktörleri ile düzenleyici ilişkileri tespit edildi, STRING veri tabanında tespit edilen proteinlerin etkileşim ağları haritalandı. Bu haritalar Cytoscape uygulamasına aktarıldıktan sonra CytoHubba eklentisi ile merkezi genler tespit edildi ve merkezi genlerin fonksiyonel zenginleşim analizleri ayrıca gerçekleştirildi.

Bulgular: T-test sonucuna göre veri sayısı yetersiz olan kanser tipleri elenmiştir. Mide, böbrek, tiroid ve karaciğer kanser türleri için yapılan analizler sonucunda kanserle ilgili olduğunu tespit ettiğimiz başlıca süreçler tiroid için otofaji, hücre gelişimi, DNA hasar tepkisi; böbrek için hücre dışı taşıma, hücre hareketliliği, protein metabolik süreci; karaciğer için makromolekül metabolik süreci, protein metabolik süreci, oksidatif fosforilasyon; mide için ribozom biyogenezi, hücresel metabolik süreçler, nükleik asit metabolik sürecidir.

Sonuç: Yapılan analizlerle, sıklıkla ifade edilen ve hastalık fenotipiyle ilişkisi olabilecek genler tespit edilmiş, biyobelirteç olarak hizmet edebilecek çok sayıda yeni gen ve yol tanımlamıştır, yeni çalışmalara ışık tutacak öngörülerde bulunulmuştur.

Anahtar kelimeler: Kanser Genom Atlası (TCGA), kanser, transkriptomik, veri analizi, biyoinformatik araçları

Pan-Cancer Biomarkers Determination from the Cancer Genome Atlas Transcriptomic Datasets

Hale Nur Korkmaz^{ID}, Müberra Gül Öztürk^{ID}, Muhammed Erkan Karabekmez^{ID}

Department of Bioengineering, İstanbul Medeniyet University Faculty of Engineering and Natural Sciences, İstanbul, Türkiye

ABSTRACT

Objectives: Cancer is a fatal and complex somatic disease. Bioinformatics advancements have opened up new possibilities for understanding cancer prevention, diagnosis, and treatment. This study aims to identify new and common biomarkers for the diagnosis and treatment of various cancers by analyzing the transcriptomic data obtained from the Cancer Genome Atlas using different tools, targeting the insight derived from the discovery of recurring patterns.

Materials and Methods: Transcriptomic data on cancer types characterized by elevated incidence and mortality levels, as per WHO statistics, were retrieved from the TCGA database. Subsequently, these data were analyzed using Galaxy tools. Following the statistical analyses conducted, differentially expressed genes in the gene sets of cancer types were identified. Functional enrichment analyses were then performed using the g:Profiler online software tool. The TRRUST tool was utilized to identify regulatory connections with transcription factors, and protein interaction networks were mapped in the STRING database. Following the transfer of these maps to the Cytoscape application, central genes were identified using the CytoHubba tool, and functional enrichment analyses of central genes were performed.

Results: Cancer types with insufficient data were excluded based on T-test results. Central genes associated with four types of cancer and their related biological processes were identified. Thyroid cancer showed associations with autophagy, cell development, and DNA damage response. Kidney cancer was linked to extracellular transport, cell mobility, and protein metabolic processes. Liver cancer demonstrated connections with the metabolism of macromolecules, proteins, and oxidative phosphorylation. Stomach cancer was found to be associated with ribosome biogenesis, cellular metabolic processes, and nucleic acid metabolism.

Conclusion: The analyses have identified genes frequently expressed and potentially associated with the disease phenotype. Numerous new genes and pathways have been defined, which could serve as biomarkers, providing insights that may illuminate future research endeavors.

Keywords: The cancer genome atlas, cancer, transcriptomic data analysis, bioinformatics tools

Tıpta Yapay Zekâ-Makine Öğrenmesi: Kalıtsal Kanselerde Genetik Test Yapmadan Kanser Riskinin Belirlenmesi ve Sağlık Harcamalarının Azaltılması

Hülya Yazıcı^{1,2}, Büşra Kurt Gültaşlar¹, Özge Şükrioğlu Erdoğan¹, Seda Kılıç Erciyas¹, Şeref Buğra Tunçer¹, Betül Çelik¹, Demet Akdeniz Ödemiş¹

¹ İstanbul Üniversitesi Onkoloji Enstitüsü, Kanser Genetiği Bilim Dalı, Temel Onkoloji Anabilim Dalı, İstanbul, Türkiye

² İstanbul Arel Üniversitesi Tıp Fakültesi, Tıbbi Biyoloji ve Genetik Anabilim Dalı, İstanbul, Türkiye

ÖZET

Amaç: Türkiye’de sağlık harcamaları her geçen yıl artmaktadır. Sağlık harcamalarında yapılan her tasarruf gelecekteki sağlık yatırımlarının geliştirilmesini sağlayacaktır. Tıpta yapay zekâ ve makine öğrenmesi (YZ-MÖ) uygulamaları yıllardır genetik dâhil olmak üzere farklı alanlarda ve amaçlarla kullanılmaktadır. Bu çalışmada, özellikle hastanelerde tedavi planlaması ve prevansiyon için yapılan bazı genetik test harcamalarının tarafımızca geliştirilen yapay zekâ-makine öğrenmesi algoritmasının kullanılarak azaltılıp azaltılamayacağı araştırılmıştır.

Gereç ve Yöntem: Ekibimiz tarafından BRCA1/BRCA2 negatifliğini tanımlamak üzere YZ-MÖ tekniği kullanılarak bir tahmin algoritması ile bu algoritmayı arka planda kullanan bir yazılım programı geliştirilmiştir. Bu yazılım programı ile İstanbul Üniversitesi Onkoloji Enstitüsü, Kanser Genetiği Polikliniğine test için başvuran 40 hastanın test öncesi tüm klinik, demografik ve genetik özelliklerine ait veriler yazılım programıyla sisteme girilerek algoritmanın negatif kişileri tahmin etmesi sağlanmış ve ardından tüm hastalar rutin genetik teste tabi tutulmuştur. Algoritmadan elde edilen sonuçlar ile rutin genetik test sonrasında çıkan sonuçlar karşılaştırılmak suretiyle değerlendirilmiştir.

Bulgular: Yapay zekâ ve makine öğrenmesi algoritması, yapılan değerlendirmede 40 hastanın 10’una genetik test yapılmasına gerek olmadığı tahmininde bulunurken 30 hastaya genetik test önermiştir. Algoritmanın genetik test önerdiği 30 kişinin 10’unda (%33) BRCA1/BRCA2 pozitifliği saptanırken, genetik test sonrasında algoritmasının genetik test önermediği 10 hastanın tamamında (%100) BRCA1/BRCA2 negatifliği saptanmıştır.

Sonuç: Geliştirilen YZ-MÖ algoritmasının BRCA1/BRCA2 negatif olguları belirleme oranının %100 olduğu görülmüştür. Yapay zekâ ve makine öğrenmesi algoritmasının BRCA1/2 gen testi yapılmaksızın BRCA1/BRCA2 negatiflik durumunun belirlenmesinde %100 doğrulukla kullanılabileceği ve rutin uygulamalarda genetik test harcamalarının %25 oranında azaltılabileceği gösterilmiştir. Söz konusu algoritmanın sağlık harcamalarını %25 azaltmasının yanı sıra hastaların sonuç bekleme stresinin olmaması, tedavi planlamalarının hızlı, ikinci cerrahi girişime gerek kalmadan doğru bir şekilde yapılması gibi ilave katkılarının da olduğu görülmüştür.

Anahtar kelimeler: Yapay zekâ-makine öğrenmesi, kalıtsal kanser, genetik test, sağlık harcamaları

Artificial Intelligence and-Machine Learning in Medicine: Determining Cancer Risk, Reducing Health Expenditures without Genetic Testing in Hereditary Cancers

Hülya Yazıcı^{1,2}, Büşra Kurt Gültaşlar¹, Özge Şükrüoğlu Erdoğan¹, Seda Kılıç Erciyas¹, Şeref Buğra Tunçer¹, Betül Çelik¹, Demet Akdeniz Ödemiş¹

¹ Division of Cancer Genetics, Department of Basic Oncology, İstanbul University Oncology Institute, İstanbul, Türkiye

² Department of Medical Biology and Genetics, İstanbul Arel University Faculty of Medicine, İstanbul, Türkiye

ABSTRACT

Objectives: Health expenditures in Türkiye have been increasing annually, and need to be reduced. Artificial intelligence and machine learning (AI-ML) applications in medicine have been employed for various purposes, including genetics, for years. This study aims to explore the potential of our AI-ML algorithm in reducing expenses associated with certain genetic tests performed for treatment planning and prevention in hospitals.

Materials and Methods: Our team developed a software program utilizing a prediction algorithm based on the AI-ML technique to identify BRCA1/BRCA2 negativity. This software was employed to input clinical, demographic, and genetic data of 40 patients seeking genetic testing at the İstanbul University Oncology Institute Cancer Genetics Clinic. The algorithm predicted negative results for certain individuals. Subsequently, all patients underwent routine genetic testing. The results obtained from the algorithm were compared with those obtained after routine genetic testing for further evaluation.

Results: The algorithm predicted that genetic testing was unnecessary for 10 out of 40 patients, while it recommended genetic testing for the remaining 30 patients. Among the 30 patients recommended for genetic testing by the algorithm, BRCA1/BRCA2 positivity was detected in 10 individuals (33%), whereas all 10 patients for whom the algorithm did not recommend genetic testing were confirmed to be BRCA1/BRCA2 negative (100%) after routine genetic testing.

Conclusion: The AI-ML algorithm demonstrated a 100% accuracy rate in identifying BRCA1/BRCA2-negative cases. The algorithm can be effectively employed in determining BRCA1/BRCA2 negativity without conducting a genetic test, leading to a 25% reduction in genetic testing expenditures in routine applications. Additionally, the algorithm contributes to alleviating patient stress related to waiting for results and facilitates prompt and accurate treatment planning without the need for a second surgical intervention.

Keywords: Artificial intelligence, machine learning, hereditary cancer, genetic testing, healthcare expenditures

Oral Kanser Hastalarında TP53 ve Notch Mutasyonlarının Cisplatin ve Paclitaxel Kullanımına Göre Karşılaştırılması

Kemal Selçuk Yücel¹, Emin Murat Canger¹, Rıdvan Akyol²

¹ Erciyes Üniversitesi Diş Hekimliği Fakültesi, Ağız Diş ve Çene Radyolojisi Anabilim Dalı, Kayseri, Türkiye

² Nuh Naci Yazgan Üniversitesi Diş Hekimliği Fakültesi, Ağız Diş ve Çene Radyolojisi Anabilim Dalı, Kayseri, Türkiye

ÖZET

Amaç: Oral skuamöz hücreli karsinomda (OSHK) gen ekspresyonu, kopya sayısı, metilasyon ve nokta mutasyonlarının kapsamlı genomik analizi şeklinde gerçekleştirilen entegre çalışmalarda başlıca ana sürücü yolu NOTCH ve TP53 olarak belirlenmiştir. Bu çalışmanın amacı OSHK'nin başlıca tedavisinde kullanılan cisplatin ve paclitaxelin TP53 ve NOTCH sürücü yolları üzerindeki mutasyonlarının karşılaştırılmasıdır.

Gereç ve Yöntem: Çalışmada veri toplama aracı olarak "cBioPortal for Cancer Genomic" aracı kullanılmıştır. Veriler TCGA (The Cancer Genome Atlas) veri portalından elde edilmiştir. Altı farklı çalışmadan toplam 1476 hasta 1478 örnekten elde edilen veriler indirilmiştir. Cisplatin ve paclitaxel kullanan radyoterapi görmeyen toplam 143 hasta çalışmaya dâhil edilmiştir. Yanlış okuma (missense), kısaltma (truncating), çerçeve kayması (inframe) ve kısaltma (splice) şeklinde meydana gelen mutasyonlar arasındaki farklılıkların karşılaştırılması için Fischer testi uygulanmıştır. Araştırmada içerik ve kapsam geçerliliği için cBioPortal tarafından otomatik olarak uygulanan Benjamini-Hochberg prosedürüne uyulmuştur.

Bulgular: Katılımcıların %68,5'i cisplatin, %32,5'i ise paclitaxel kullanmıştır. Cisplatin kullanan hastalarda TP53'te %70,97, NOTCH'ta %23,66 oranında mutasyon meydana geldiği bildirilmiştir. Paclitaxel kullanan hastalarda ise TP53'te %50, NOTCH'ta %10 mutasyon meydana geldiği bildirilmiştir. TP53 sürücü yolunda istatistiksel olarak anlamlı bir fark bulunurken ($p=0,0291$), NOTCH sürücü yolunda ise anlamlı bir farklılık bulunmamıştır ($p=0,0773$). Cisplatin kullanan hastalarda en sık mutasyon %56'lık oranla yanlış okuma şeklinde meydana gelirken, paclitaxel kullanan hastalarda da %58 oranla yanlış okuma şeklinde bulunmuştur.

Sonuç: Bu sonuçlar, cisplatin ve paclitaxel kullanımının OSHK geliştirme riski olan fonksiyonel yollardan NOTCH sürücü yoluna etkisinin olmadığını ancak TP53 sürücü yoluna etkisinin olduğunu göstermiştir. Dolayısıyla moleküler düzeyde yapılması planlanan çalışmaların sistem biyolojisi prensipleri çerçevesinde multi-omik çalışmalar ile planlanması önem taşımaktadır.

Anahtar kelimeler: Oral kanser, NOTCH, Cisplatin, Paclitaxel, TP53

Comparison of Mutations in TP53 and Notch Driver Pathways in Oral Cancer Patients Treated with Cisplatin and Paclitaxel

Kemal Selçuk Yücel¹, Emin Murat Canger¹, Rıdvan Akyol²

¹ Department of Oral and Maxillofacial Radiology, Erciyes University Faculty of Dentistry, Kayseri, Türkiye

² Department of Oral and Maxillofacial Radiology, Nuh Naci Yazgan University Faculty of Dentistry, Kayseri, Türkiye

ABSTRACT

Objectives: The major driver pathways in oral squamous cell carcinoma (OSCC) have been identified as NOTCH and TP53 through integrated studies, which encompass comprehensive genomic analyses, including gene expression, copy number, methylation, and point mutations. This study aimed to compare the mutations in the TP53 and NOTCH driver pathways induced by cisplatin and paclitaxel, which are the main treatments for oral squamous cell carcinoma (OSCC).

Materials and Methods: In the study, the “cBioPortal for Cancer Genomic” tool was used for data collection. Data were obtained from the TCGA (The Cancer Genome Atlas) data portal. Data obtained from 1476 patients and 1478 samples from six different studies were downloaded. A total of 143 patients who did not receive radiotherapy and were treated with cisplatin and paclitaxel were included in the study. The Fischer test was applied to compare the differences between mutations occurring as missense, truncating, inframe, and splice. For content and scope validity, the Benjamini-Hochberg procedure automatically applied by cBioPortal was followed.

Results: Cisplatin was used by 68.5% and paclitaxel by 32.5% of the participants. It was reported that 70.97% mutation in TP53 and 23.66% mutation in NOTCH occurred in patients using cisplatin. In patients using paclitaxel, it was reported that 50% mutation occurred in TP53 and 10% mutation occurred in NOTCH. While a statistically significant difference was found in the TP53 driver pathway ($p=0.0291$), no significant difference was found in the NOTCH driver pathway ($p=0.0773$). The most common mutation occurred in the form of misreading, with a rate of 56% in patients using cisplatin and 58% in patients using paclitaxel.

Conclusion: These results indicate that the use of cisplatin and paclitaxel had no effect on the NOTCH driver pathway but did impact the TP53 driver pathway, which is a functional pathway at risk of developing OSCC. Therefore, it is crucial to conduct planned studies at the molecular level, incorporating multi-omics approaches within the framework of systems biology principles.

Keywords: Oral cancer, NOTCH, cisplatin, paclitaxel, TP53

PEPD Geni Sık Rastlanan Haplotip Etkileşimlerinin Koroner Arter Ektazi Riski Üzerine Etkileri

Kübra Çiğdem Pekkoç Uyanık^{1,2}, **Ezgi Irmak Aslan¹**, **Onur Kılıçarslan³**, **Özgür Selim Ser³**,
Serhan Özyıldırım³, **Fatih Yanar^{1,4}**, **Ahmet Yıldız³**, **Oğuz Öztürk¹**, **Hülya Yılmaz Aydoğan¹**

¹ İstanbul Üniversitesi Aziz Sancar Deneysel Tıp Araştırma Enstitüsü, Moleküler Tıp Bölümü, İstanbul, Türkiye

² Haliç Üniversitesi Tıp Fakültesi, Tıbbi Biyoloji Anabilim Dalı, İstanbul, Türkiye

³ İstanbul Üniversitesi-Cerrahpaşa Kardiyoloji Enstitüsü, Kardiyoloji Bölümü, İstanbul, Türkiye

⁴ Boğaziçi Üniversitesi Fen Fakültesi, Moleküler Biyoloji ve Genetik Bölümü, İstanbul, Türkiye

ÖZET

Amaç: Prolidaz enzimi, kolajenin yapım ve yıkım döngüsünde aktif bir rol oynar. Kolajen, koroner arterin yapısındaki ana proteinlerden biridir. Prolidaz enzimi kolajen katabolizmasında rol oynadığından, koroner arter ektazisi (KAE) etiopatogenezinde önemli bir rol oynayabilir. PEPD geni prolidaz enzimini eksprese etmektedir. Bu çalışma, KAE'de PEPD gen varyasyonlarının haplotiplerini araştıran ilk çalışmadır.

Gereç ve Yöntem: PEPD gen dizi analizi, MiSeq cihazı kullanılarak yeni nesil dizileme tekniği ile 76 KAE hastasında ve 76 sağlıklı kontrolde gerçekleştirilmiştir. Daha sonra PEPD geni rs17570, rs17569, rs75449530, rs1061338 ve rs16968070 haplotiplerinin koroner arter ektazisi ile ilişkisi bir, iki, üç, dört ve beşli alel setleri olarak Haploview programıyla analiz edilmiştir.

Bulgular: Haploview programı ile yapılan analizler sonucunda, rs1061338 G aleli ve rs17570 G alelinin birlikte KAE riskini arttırdığını göstermiştir ($p=0,047$). Ayrıca, rs1061338 G aleli ve rs17570 G aleli diğer alellerle (rs17569 G ve rs75449530 T alelleri) birleştiğinde KAE riskini arttırmaktadır (sırasıyla $p=0,047$, $p=0,047$). Ancak rs1061338 G aleli, rs17569 G aleli ve rs75449530 T aleli birlikte KAE riskini nispeten azaltmaktadır ($p=0,059$). Tek alelde yalnızca rs1061338 A>G polimorfizmi, G alelinin istatistiksel olarak anlamlılığa yakın olduğu görülmüştür ($p=0,054$).

Sonuç: Bulgularımız, prolidaz enzimi ile ilişkili PEPD geni haplotiplerinin etkileşimlerinin koroner arter ektazi riskini etkileyebileceğini göstermektedir.

Anahtar kelimeler: Koroner arter ektazi, prolidaz, PEPD geni, haplotip

*Bu çalışma İstanbul Üniversitesi Bilimsel Araştırma Projeleri Birimi tarafından TDK-2018-31311 numaralı proje olarak desteklenmiştir.

Interactive Effects of Common Haplotypes of PEPD Gene on Coronary Artery Ectasia Risk

Kübra Çiğdem Pekkoç Uyanık^{1,2}, **Ezgi Irmak Aslan¹**, **Onur Kılıçarslan³**, **Özgür Selim Ser³**,
Serhan Özyıldırım³, **Fatih Yanar^{1,4}**, **Ahmet Yıldız³**, **Oğuz Öztürk¹**, **Hülya Yılmaz Aydoğan¹**

¹ Department of Molecular Medicine, İstanbul University Aziz Sancar Institute of Experimental Medicine, İstanbul, Türkiye

² Department of Medical Biology, Haliç University Faculty of Medicine, İstanbul, Türkiye

³ Department of Cardiology, İstanbul University-Cerrahpaşa Institute of Cardiology, İstanbul, Türkiye

⁴ Department of Molecular Biology and Genetics, Boğaziçi University Faculty of Medicine, İstanbul, Türkiye

ABSTRACT

Objectives: The prolidase enzyme plays an active role in the turnover of collagen, one of the main proteins in the structure of the coronary artery. Given the role of the prolidase enzyme in collagen catabolism, it may play a significant role in the etiopathogenesis of coronary artery ectasia (CAE). Prolidase is an enzyme encoded by the PEPD gene in humans. This study represents the first investigation into the haplotypes of the PEPD gene in CAE.

Materials and Methods: PEPD gene sequence analysis was performed in 76 CAE patients and 76 healthy controls using the next-generation sequencing technique with MiSeq. Common variations of the PEPD gene, including rs17570, rs17569, rs75449530, rs1061338, and rs16968070 haplotypes, were analyzed for their association with coroner artery ectasia in sets of one, two, three, four, and five alleles using Haploview software.

Results: The analyses conducted using the Haploview software revealed that the combination of the rs1061338 G allele and rs17570 G allele increases the risk of CAE ($p=0.047$). Additionally, the rs1061338 G allele and rs17570 G allele when combined with other alleles (rs17569 G and rs75449530 T alleles) increase the CAE risk ($p=0.047$, $p=0.047$, respectively). However, the combination of the rs1061338 G allele, rs17569 G allele, and rs75449530 T allele was found to relatively reduce the risk of CAE ($p=0.059$). In terms of single alleles, only in the rs1061338 A>G polymorphism, the G allele was observed to be associated with the risk for CAE, which was nearly statistically significant ($p=0.054$).

Conclusion: Our findings indicate that interactive effects of PEPD gene haplotypes associated with prolidase enzyme may affect coronary artery ectasia risk.

Keywords: Coronary Artery Ectasia, prolidase, PEPD gene, haplotype

*This study was supported by the İstanbul University Scientific Research Projects Unit under project number TDK-2018-31311.

Genomik Veriler Kullanılarak Herediter ve Solid Tümör Kanserlerinde Saptanan Patojenik Varyantların ve Alel Frekanslarının Hesaplanmasıyla Veri Tabanı Oluşturulması: Bir Toplum Çalışması

Kübra Damla Erol¹, **İlkem Özce Özçelik¹**, **Polat Olgun^{2,3}**, **Ömer Diker^{2,3}**, **Mahmut Çerkez Ergören¹**

¹ Yakın Doğu Üniversitesi Tıp Fakültesi, Tıbbi Genetik Anabilim Dalı, Lefkoşa, KKTC

² Yakın Doğu Üniversitesi Tıp Fakültesi, İç hastalıklar Anabilim Dalı, Lefkoşa, KKTC

³ Dr. Burhan Nalbantoğlu Devlet Hastanesi, Onkoloji Kliniği, Lefkoşa, KKTC

ÖZET

Amaç: Bir ada ülkesi olan KKTC’de ilk kez herediter kanserler ve solid tümörlerde görülen varyasyonların alel frekanslarının analizlerini gerçekleştirerek, topluma özgü varyantların hesaplanması.

Gereç ve Yöntem: Bu çalışma ile Yakın Doğu Üniversitesi Hastanesi Genetik Laboratuvarı genomik arşivleri kullanılarak, son beş yılda geçmişten günümüze moleküler kanser tanısı almış hastalarda retrospektif inceleme ile Kıbrıs Türk popülasyonundaki genetik yapının ve moleküler patolojinin mevcut durumunu saptama amaçlandı.

Bulgular: Yapılan veri analizi sonucunda, herediter kanserlerde incelenen genlerin %22’sinde mutasyon saptanmıştır, bunlardan en sık saptanan BRCA1 genindeki c.1444_1447del mutasyonudur. Solid kanserlerde 134 mutasyon saptanmıştır, saptanan mutasyonların sadece %14,7’sinin tedavilere endike olduğu bilinmektedir.

Sonuç: Sonuç olarak, çalışma raporu KKTC Sağlık Bakanlığı ile paylaşılmış ve çalışmanın ileride yapılacak kanser koruma programlarına öncü olabileceğini umuyoruz.

Anahtar kelimeler: Herediter kanser, solid tümör, BRCA, Kıbrıs Türk toplumu

Generating a Database by Calculating the Pathogenic Variants and Allele Frequencies Detected in Hereditary and Solid Tumor Cancers Using Genomic Data: A National Study

Kübra Damla Erol¹, **İlkem Özce Özçelik¹**, **Polat Olgun^{2,3}**, **Ömer Diker^{2,3}**, **Mahmut Çerkez Ergören¹**

¹ Department of Medical Genetics, Near East University Faculty of Medicine, Nicosia, TRNC

² Department of Internal Medicine, Near East University Faculty of Medicine, Nicosia, TRNC

³ Clinic of Oncology, Dr. Burhan Nalbantoğlu State Hospital, Nicosia, TRNC

ABSTRACT

Objectives: The aim of this study was to calculate population-specific variants by analyzing the allele frequencies of the variations seen in hereditary cancers and solid tumors for the first time in the TRNC, an island country.

Materials and Methods: This study aimed to determine the current status of the genetic structure and molecular pathology in the Turkish Cypriot population through a retrospective examination of patients diagnosed with molecular cancer in the last five years, using the genomic archives of the Near East University Hospital Genetics Laboratory.

Results: As a result of the data analysis, mutations were identified in 22% of the genes examined in hereditary cancers, with the most frequently detected mutation being the c.1444_1447del mutation in the BRCA1 gene. 134 mutations have been identified in solid cancers, with only 14.7% known to have indications for treatment.

Conclusion: The comprehensive study report has been shared with the TRNC Ministry of Health, and we aspire for it to serve as a pioneer in shaping future cancer prevention programs.

Keywords: Hereditary cancer, solid tumor, BRCA, Turkish Cypriot Community

Karşılaştırmalı, Dizileme Verisi Analiz Platformu

Mehmet Arif Ergün^{ID}, Ömer Çinal^{ID}, Berkant Bakışlı^{ID}, Mehmet Baysan^{ID}

İstanbul Teknik Üniversitesi Bilgisayar ve Bilişim Fakültesi, Bilgisayar Mühendisliği Anabilim Dalı, İstanbul, Türkiye

ÖZET

Amaç: Son yıllarda, dizileme teknolojilerindeki gelişmelerle birlikte genomik özelliklerin detaylı analizleri hız kazanmıştır. Günümüzde yeni nesil dizileme analiz programlarının çoğu, ya erişimi ve doğrulamayı çok zorlaştıran, tescilli kodlara sahip, pahalı kurumsal yazılımlardır ya da çoğunlukla komut satırı işlemlerine dayanan, çok sınırlı kullanıcı arayüzleri ve dokümantasyonu olan açık erişimli programlardır. Ayrıca, popüler haritalama ve varyant keşfi algoritmaları arasında yüksek düzeyde uyumsuzluk olduğunu gösteren birçok çalışma vardır ve bu da tek bir algoritmaya güvenmeyi olanaksız kılmaktadır. Bu nedenle, dizileme yaygınlaştıkça ve veri üretimi arttıkça, donanım hızlandırması ve alternatif algoritmalar sunan kullanıcı dostu açık kaynaklı yazılım araçlarına duyulan ihtiyaç artmaktadır.

Gereç ve Yöntem: Bu çalışma, karşılaştırmalı analiz platformu COSAP'ı tanıtmaktadır. Bu platform, tek nükleotid polimorfizmi, kopya sayısı değişiklikleri, yapısal değişimler, mikrosatelit dengesizliği ve gen füzyonları gibi çeşitli genetik varyasyonların saptanması ve anotasyonları için yaygın olarak kullanılan algoritmaları bir arada sunmaktadır. COSAP, işlevsel bir ağ arayüzü ve hem bireysel hem de kurumsal ölçekte kuruluma izin veren bir arka uç (backend) sunucusu ile sunulmaktadır. Ayrıca donanım hızlandırması için Clara Parabricks platformunu desteklemektedir. Kolay kuruluma olanak sağlamak için Docker görüntüleri de ek olarak sunulmaktadır.

Sonuç: COSAP, yaygın olarak kullanılan algoritmaları kullanıcı dostu bir şekilde sunarak DNA dizileme analizlerini basitleştirir ve hızlandırır. Modüler bir platformda sunulan alternatif algoritmalar, yeniden üretilebilirlik açısından oldukça önemli olan, farklı iş akışlarının sonuçları üzerindeki etkilerini analiz etmeyi kolaylaştırmaktadır.

Anahtar kelimeler: YND analizi, varyant keşfi, DNA-Seq

Comparative Sequencing Analysis Platform

Mehmet Arif Ergün^{ID}, Ömer Çinal^{ID}, Berkant Bakışlı^{ID}, Mehmet Baysan^{ID}

Department of Computer Engineering, İstanbul Technical University Faculty of Computer and Informatics Engineering, İstanbul, Türkiye

ABSTRACT

Objectives: With the advancements in sequencing technologies over the last decade, detailed profiling of genomic traits has accelerated. Currently, many of the related software programs are either expensive enterprise software with proprietary code, making access and validation challenging, or open-access ones that primarily rely on command-line operations with very limited user interfaces and documentation. Additionally, many studies show a high level of disagreement between popular mapping and variant calling algorithms making it unreliable to rely on a single algorithm. Therefore, user-friendly open-source software tools that offer alternative algorithms with hardware acceleration are needed as the sequencing data continues to grow.

Materials and Methods: This work introduces the Comparative Sequencing Analysis Platform (COSAP), an open-source platform that provides common sequencing algorithms for discoveries and annotations of single nucleotide and structural variants, copy number variations, microsatellite instability, and fusions. COSAP offers a functional web interface and a backend server, allowing for independent deployment at both individual and institutional scales. It also supports the hardware-accelerated Clara Parabricks framework out of the box. The platform is also packaged as Docker images to allow effortless deployment.

Results: COSAP simplifies and accelerates DNA sequencing analyses by providing commonly used algorithms in a user-friendly manner. Standardized implementations of popular algorithms on a modular platform make comparisons much easier to assess the impact of alternative pipelines, which is crucial in ensuring the reproducibility of sequencing analyses.

Keywords: NGS analysis, variant calling, DNA-Seq

Gaucher Geninin (GBA1) Parkinson Hastalığına Etkisi: Asya Minör'ün En Büyük Kohortu

Merve Koç Yekedüz¹, **Talha Abali²**, **Rezzak Yılmaz³**, **Görkem Kayış²**, **Elif Yüstra Unutmaz⁴**,
Ahmet Veli Karacan², **Sema Nur Kibrit²**, **Serdar Ceylaner⁵**, **Muhittin Cenk Akbostancı³**,
Fatma Tuba Eminoğlu¹

¹ Ankara Üniversitesi Tıp Fakültesi, Çocuk Sağlığı ve Hastalıkları Anabilim Dalı, Çocuk Beslenme ve Metabolizma Bilim Dalı, Ankara, Türkiye

² Ankara Üniversitesi Tıp Fakültesi, Ankara, Türkiye

³ Ankara Üniversitesi Tıp Fakültesi Nöroloji Anabilim Dalı, Ankara, Türkiye

⁴ Ankara Üniversitesi Dil ve Tarih-Coğrafya Fakültesi, Psikoloji Bölümü, Ankara, Türkiye

⁵ İntergen Genetik ve Nadir Hastalıklar Araştırma ve Uygulama Merkezi, Ankara, Türkiye

ÖZET

Amaç: Bu araştırma, GBA1 geninin Parkinson hastalığına etkisini araştırmayı amaçlamaktadır.

Gereç ve Yöntem: Katılımcılarda GBA gen analizi, beta-glukoserebrosidaz ve Lyso-Gb1 düzeyleri analiz edildi. Varyantların patojenitesi açısından gruplar arasında klinik veriler karşılaştırıldı.

Bulgular: Araştırmaya 513 Parkinson hastası (PH) ve 203 sağlıklı kontrol (SK) olmak üzere 716 kişi dâhil edildi. 67 PH'de (%13,0) ve 13 SK'de (%6,4) GBA varyantı saptandı. Diğer lizozomal hastalıklarla ilgili genler dışlandı. Hastalık süresi bir yıldan uzun olan 379 PH, hastalıkla ilgili parametreler açısından analiz edildi. 379 PH'den 30'u patojenik GBA varyantı olan Parkinson hastaları (PH-pat-GBA) grubundaydı, 25'i benign GBA varyantı olan Parkinson hastaları (PH-ben-GBA) ve 324'ü GBA varyantı negatif olan Parkinson hastaları (PH-GBA-neg) idi. PH-pat-GBA ve PH-GBA-neg grupların karşılaştırılması, PH-pat-GBA'da başlangıç yaşının daha düşük olduğunu gösterdi [ortalama (SS), 52,7 (10,5)'ye karşı 57,3 (11,9); p= 0,042]. Ayrıca levodopa eş değer günlük dozu (mg/gün), PH-pat-GBA'da PH-GBA-neg'e göre daha yüksekti [973,8 (506,8), 728,3 (434,5); p= 0,004]. PH-pat-GBA'da [n= 10 (%34,5)], PH-GBA-neg'e [n= 46 (%14,4)] göre halüsinasyon ve psikoz varlığı daha yüksekti (p= 0,005). Ortostatik hipotansiyon PH-pat-GBA'da [n= 16 (%53,3)] PH-GBA-neg gruba [n=102 (%31,9; p= 0,017)] göre daha yaygındı. Diğer değerlendirmeler önemli bir farklılık göstermedi. Enzim ve substrat seviyeleri açısından grupların karşılaştırılması, PH-pat-GBA grubundaki daha düşük enzim ve daha yüksek Lyso-Gb1 seviyelerine rağmen anlamlı bir fark göstermedi.

Sonuç: PH'deki GBA çeşitlerine ilişkin Asya Minör'den en geniş kohortu rapor ediyoruz. Patojenik GBA varyantları daha erken başlangıç yaşına sahiptir ve bazı klinik değerlendirmelerde daha kötü performans gösterirler.

Anahtar kelimeler: Beta-glukoserebrosidaz, gaucher hastalığı, GBA geni, Parkinson hastalığı

The Effect of the Gaucher Gene (GBA1) on Parkinson's Disease: The Largest Cohort of Asia Minor

Merve Koç Yekedüz¹, Talha Abali², Rezzak Yılmaz³, Görkem Kayış², Elif Yüstra Unutmaz⁴, Ahmet Veli Karacan², Sema Nur Kibrit², Serdar Ceylaner⁵, Muhittin Cenk Akbostancı³, Fatma Tuba Eminoğlu¹

¹ Department Pediatric Diseases, Division of Pediatric Nutrition and Metabolism, Ankara University Faculty of Medicine, Ankara, Türkiye

² Ankara University Faculty of Medicine, Ankara, Türkiye

³ Department of Neurology, Ankara University Faculty of Medicine, Ankara, Türkiye

⁴ Department of Psychology, Ankara University Faculty of Language and History-Geography, Ankara, Türkiye

⁵ InterGen Genetics and Rare Diseases Diagnosis Research & Application Center, Ankara, Türkiye

ABSTRACT

Objectives: The objective of this study was to investigate the effect of the GBA1 gene on Parkinson's disease.

Materials and Methods: Analysis of the GBA gene, beta-glucocerebrosidase, and Lyso-Gb1 levels was performed. Clinical assessments were then compared between the groups based on the pathogenicity of the variants.

Results: 716 participants [513 patients with Parkinson's disease (PwP) and 203 healthy control (HC)] were included in the study. Sixty-seven (13.0%) had a GBA variant in PwP and 13 (6.4%) in HC. Genes related to other lysosomal disorders were excluded. 379 PwP with a disease duration of over one year were analyzed for disease-related parameters. Of the 379 PwP, 30 were in the patients with Parkinson's disease-pathogenic GBA variant (PwP-pat-GBA) group, 25 in patients with Parkinson's disease-benign GBA variant (PwP-ben-GBA), and 324 in patients with Parkinson's disease-no GBA variant (PwP-no-GBA). Comparison of the PwP-pat-GBA and PwP-no-GBA groups showed that the age of onset was lower in the PwP-pat-GBA [mean (SD), 52.7 (10.5) vs. 57.3 (11.9); $p=0.042$]. Also, the levodopa equivalent daily dose (mg/day) was higher in PwP-pat-GBA than in PwP-no-GBA [973.8 (506.8), 728.3 (434.5); $p=0.004$]. In addition, the presence of hallucinations and psychosis was higher in PwP-pat-GBA [$n=10$ (34.5%)] than in PwP-no-GBA [$n=46$ (14.4%); $p=0.005$]. Orthostatic hypotension was also more prevalent in PwP-pat-GBA [$n=16$ (53.3%)] than in PwP-no-GBA [$n=102$ (31.9%); $p=0.017$]. Other assessments showed no significant differences. A comparison of the groups for enzyme and substrate levels showed no significant difference despite the lower enzyme and higher Lyso-Gb1 levels in the PwP-pat-GBA group.

Conclusion: We report the largest data on GBA variants in PwP from Asia Minor. Pathogenic GBA variants have an earlier age of onset and perform worse in some clinical assessments.

Keywords: Beta-glucocerebrosidase, gaucher disease, GBA gene, Parkinson's disease

Ankilozan Spondilit Metabolizmasının Aydınlatılması: Tek Hücreli RNA Dizileme Verileri ile Genom Ölçekli Metabolik Modellerin Oluşturması

Merve Yarıcı^{ID}, Muhammed Erkan Karabekmez^{ID}

İstanbul Medeniyet Üniversitesi Mühendislik ve Doğa Bilimleri Fakültesi, Biyomühendislik Anabilim Dalı, İstanbul, Türkiye

ÖZET

Amaç: Ankilozan spondilit (AS) başta omurgada olmak üzere eklemleri etkileyen otoimmün hastalıklardan biridir. Hastalık patogenezi henüz aydınlatılmamıştır ve hastalığın erken teşhisinde kullanılan etkili bir biyobelirteç yoktur. Tedavi geliştirme süreçlerinde hastalıkların hücre içi işleyiş mekanizmasının belirlenmesi ve moleküler düzeyde incelenmesi çok önemlidir. Tek hücreli RNA dizileme tekniği (scRNA-seq) hastalıkların immünolojik mekanizmalarının araştırılmasında kullanılan avantajlı bir yöntemdir. Ayrıca, hücresel metabolizmayı anlamak için genom ölçekli metabolik modellerin scRNA-seq verilerine uygulanması gen-reaksiyon bağlantılarını belirleyerek metabolik akışın tahmin edilmesi ile hastalıkların biyolojik mekanizmalarının anlaşılması için önemli ipuçları sağlayabilmektedir. Bu çalışmada Ankilozan Spondilit scRNA-seq veri kümeleri analiz edilerek belirlenen her hücre grubu için metabolik ağ haritası oluşturulması ve hastalığın biyolojik mekanizmasının daha anlaşılır bir şekilde aydınlatılması amaçlanmıştır.

Gereç ve Yöntem: AS hastalığına özgü periferik kan mononükleer hücrelerinin scRNA-seq verileri analiz edilmiş ve GIMME algoritması kullanılarak belirlenen her hücre tipi için ayrı ayrı genom ölçekli metabolik modeller oluşturulmuştur. Oluşturulan modellerdeki reaksiyonlar için akı örnekleme yapılarak değişen reaksiyon akıları analiz edilmiştir. Ayrıca, her hücre tipi için raportör metabolit analizleri gerçekleştirilerek hastalık durumunda hücrelerdeki metabolizma-gen bağlantıları incelenmiştir.

Bulgular: Ankilozan spondilitli hastaların CD14 Monocytes, CD4 Memory, CD4 Naive ve CD8 T hücrelerinde pürin metabolizması, doymamış yağ asitlerinin biyosentezi ve glikoliz üzerindeki akışların sağlıklı bireylere göre diferansiyel olarak arttığı görülmüştür. Genel olarak her hücre tipinde aminoasitler, su, proton ve oksijen gibi metabolitler raportör metabolit olarak tespit edilmiştir.

Sonuç: Sonuç olarak, AS hastalığının biyolojik mekanizması araştırılmış ve tedavi geliştirme süreçleri için alternatif hedefler belirlenmiştir. Ayrıca bu çalışmayla otoimmün hastalıklar ve kanser gibi karmaşık biyolojik süreçler için yeni biyobelirteçler bulmaya yönelik umut verici bir yaklaşım önerilmiştir.

Anahtar kelimeler: Ankilozan spondilit, scRNA-seq, genom ölçekli metabolik modelleme, GIMME, ağ biyolojisi

Unveiling Ankylosing Spondylitis Metabolism: Constructing Genome-Scale Metabolic Models Using Single-Cell RNA Sequencing Data

Merve Yarıcı^{ID}, Muhammed Erkan Karabekmez^{ID}

Department of Bioengineering, İstanbul Medeniyet University Faculty of Engineering and Natural Sciences, İstanbul, Türkiye

ABSTRACT

Objectives: Ankylosing spondylitis (AS) is an autoimmune disease that primarily affects the spine. The pathogenesis of the disease has not yet been elucidated, and no effective biomarker has been proposed to diagnose the disease. In the process of developing treatments, it is crucial to ascertain the intracellular mechanisms of diseases and thoroughly examine them at the molecular level. The single-cell RNA sequencing technique (scRNA-seq) is an advantageous method for investigating the immunological mechanisms of diseases. Furthermore, integrating genome-scale metabolic models with scRNA-seq data to understand cellular metabolism has the potential to predict metabolic flux. This involves identifying gene-reaction connections and offering essential insights into comprehending the biological mechanisms of diseases. This study aimed to analyze AS scRNA-seq datasets to construct a metabolic network map for each determined cell type and to elucidate the biological mechanism of the disease more clearly.

Materials and Methods: The scRNA-seq data from AS-specific peripheral blood mononuclear cells were analyzed, and genome-scale metabolic models were constructed separately for each cell type determined using the GIMME algorithm. Changing reaction fluxes were analyzed by flux sampling for the reactions in the constructed models. Additionally, by performing reporter metabolite analyses for each cell type, metabolism-gene connections in cells were examined in the disease context.

Results: It has been observed that the fluxes on purine metabolism, biosynthesis of unsaturated fatty acids, and glycolysis are differentially increased in CD14 Monocytes, CD4 Memory, CD4 Naive, and CD8 T cells of patients with AS compared to healthy individuals. Also, metabolites such as amino acids, water, proton, and oxygen were detected as reporter metabolites in each cell type.

Conclusion: As a result, the biological mechanism of AS was investigated, and alternative targets were identified for treatment development processes. Additionally, this study proposed a promising approach to finding new biomarkers for complex biological processes such as autoimmune diseases and cancer.

Keywords: Ankylosing spondylitis, scRNA-seq, genome-scale metabolic modeling, GIMME, network biology

SARS-CoV-2 Varyantlarının Tüm Genom Dizileme Yöntemi ile Filogenetik Analizi

Yasemin Coşgun¹, Süleyman Yalçın¹, Özlem Ünalı¹, Merve Yıldız¹, Meral Tiryaki¹, Ekrem Sağtaş¹, Sedat Kaygusuz²

¹ T.C. Sağlık Bakanlığı Halk Sağlığı Genel Müdürlüğü, Mikrobiyoloji Referans Laboratuvarları ve Biyolojik Ürünler Daire Başkanlığı, Ankara, Türkiye

² T.C. Sağlık Bakanlığı Halk Sağlığı Genel Müdürlüğü, Ankara, Türkiye

ÖZET

Amaç: COVID-19'un etiyolojik ajanı olan SARS-CoV-2, ilk olarak Aralık 2019'da Wuhan'da tanımlanmış ve küresel olarak yayılmıştır. Dünya genelinde COVID-19'un pandemik aşamadan epidemik aşamaya geçeceği düşünülmektedir. Ancak virüsün yeni varyantlarının ortaya çıkması pandeminin devam etme olasılığını artırabilir. Bu çalışmada, ülkemizdeki SARS-CoV-2 varyantlarını karakterize etmek amaçlanmıştır.

Gereç ve Yöntem: Ülkemizin 12 farklı şehirden 2023 Ocak ve Kasım ayları arasında Halk Sağlığı Genel Müdürlüğü Ulusal Viroloji Referans Laboratuvarına gönderilen 532 COVID-19 numunesi çalışmaya dâhil edilmiştir. Ulusal Moleküler Mikrobiyoloji Referans Laboratuvarında Illumina Miseq ve Illumina Nextseq platformları kullanılarak tüm genom dizileme yapılmıştır. Nexclade ve Pangolin sınıflamalarına göre varyantlar ve soylar belirlenmiştir. Varyantların MegaX programında filogenetik analizleri yapılmıştır.

Bulgular: Çalışılan numunelerde tüm varyantlar Omicron varyantı ve alt tipleri olarak belirlenmiştir. En sık görülen varyant Omicron'un XBB alt türleri olarak bulunmuştur. Bunu EG alt türleri takip etmiştir. Eris varyantı (EG5.1) ülkemizde ilk defa Eylül ayında saptanmıştır.

Sonuç: Pandeminin kontrolü, alınacak tedbirler, aşı ve tedavi planlamalarının yapılması gibi konularda SARS-CoV-2 varyantlarının izlenmesi gerekmektedir ve hangi varyantın baskın hâle geldiği takip edilmelidir.

Anahtar kelimeler: Pandemi, COVID-19, dizi

Phylogenetic Analysis of SARS-CoV-2 Variants Using Whole Genome Sequencing Method

Yasemin Coşgun¹, Süleyman Yalçın¹, Özlem Ünalı¹, Merve Yıldız¹, Meral Tiryaki¹, Ekrem Sağtaş¹, Sedat Kaygusuz²

¹ Microbiology Reference Laboratories and Biological Products Department, General Directorate of Public Health, Republic of Türkiye Ministry of Health, Ankara, Türkiye

² General Directorate of Public Health, Republic of Türkiye Ministry of Health, Ankara, Türkiye

ABSTRACT

Objectives: SARS-CoV-2, the etiological agent of COVID-19, was first identified in Wuhan in December 2019 and has since spread globally. It is believed that COVID-19 may transition from the pandemic phase to the epidemic phase worldwide. However, the emergence of new variants of the virus may lead to the potential continuation of the pandemic. This study aimed to characterize SARS-CoV-2 variants in our country.

Materials and Methods: 532 COVID-19 samples sent to the National Virology Reference Laboratory of the General Directorate of Public Health between January and November 2023 from 12 different cities of our country were included in the study. Whole genome sequencing was performed using Illumina Miseq and Illumina Nextseq platforms at the National Molecular Microbiology Reference Laboratory. Variants and lineages were identified according to the Nexclade and Pangolin classifications. Phylogenetic analyses of the variants were performed using the MegaX software.

Results: All variants in the studied samples were identified as Omicron variants and subtypes. The most prevalent variant identified was the XBB subtype of Omicron, followed by EG subtypes. Notably, the Eris variant (EG5.1) was first detected in our country in September.

Conclusion: Monitoring SARS-CoV-2 variants is crucial for pandemic control, determining necessary measures, planning vaccination and treatment strategies, and keeping track of the dominant variant.

Keywords: Pandemic, COVID-19, sequence

Enzim Komisyon Numaralarına Göre Gruplandırma Puanlama Modeli Geliştirerek Metagenomik Veriden Kolorektal Kansere İlişkili Enzim Gruplarının Keşfi

Nur Şebnem Ersöz¹, Malik Yousef^{2,3}, Burcu Bakır-Güngör^{4,5}

¹ Abdullah Gül Üniversitesi Fen Bilimleri Enstitüsü, Biyomühendislik Bölümü, Kayseri, Türkiye

² Zefat Akademik Koleji, İnfomasyon Sistemleri Bölümü, Zefat, İsrail

³ Zefat Akademik Koleji, Galilee Dijital Sağlık Araştırma Merkezi (GDH), Zefat, İsrail

⁴ Abdullah Gül Üniversitesi Mühendislik Fakültesi, Bilgisayar Mühendisliği Bölümü, Kayseri, Türkiye

⁵ Abdullah Gül Üniversitesi Yaşam ve Doğa Bilimleri Fakültesi, Biyomühendislik Bölümü, Kayseri, Türkiye

ÖZET

Amaç: Kolorektal kanser (KRK), dünya çapında kanserden ölümlerin önde gelen nedenidir. Son zamanlarda metagenomların fonksiyonel profillemesi, enzim kodlayan genlerin nisbi bolluk değerlerinin hesaplanmasına olanak sağlamaktadır. Bu çalışmanın temel amacı, biyolojik alan bilgisine dayalı öznelik seçimini kullanarak KRK ile ilişkili mikrobiyal enzimleri belirlemek için bir yöntem geliştirmektir.

Gereç ve Yöntem: Bu çalışmada kullanılan veri seti sekiz farklı ülkeden 1262 örneği (600 KRK ve 662 kontrol) içermektedir. Metagenomik verinin fonksiyonel profillemesini yapmak ve topluluk düzeyinde enzim komisyon (EK) bolluk değerlerinin tahmin edilmesi için HUMAnN3 kullanılmıştır. Bu çalışmada Gruplandırma Puanlama Modellemesi (G-S-M) yaklaşımına dayalı bir sınıflandırma modeli öneriyoruz. G-S-M modeline biyolojik alan bilgisi sağlamak için enzimleri EK numaralarına göre gruplandırdık. Deneylerimizde 10 kez Monte Carlo Çapraz Doğrulama kullanarak farklı yöntemlerle karşılaştırmalı değerlendirme yaptık.

Bulgular: Deneylerimizde KRK ile ilişkili yüksek skorlu enzim gruplarının glikozidazları (EK: 3.2.1), hidro-lyazları (EK: 4.2.1) ve CoA-transferazları (EK: 2.8.3) içerdiğini gözlemledik. Oligo-1,6-glukozidaz enzimi (EK: 3.2.1.10), KRK ile ilişkili en önemli enzim olarak bulundu. Ayrıca, G-S-M modelinin çeşitli sınıflandırıcılarla test edilen geleneksel öznelik seçim yöntemlerinden daha iyi performans gösterdiğini gösterdik.

Sonuç: Burada, KRK ile ilişkili enzim gruplarını bulan G-S-M tabanlı bir model geliştirerek KRK metagenomunun fonksiyonel analizini gerçekleştiriyoruz. Bu çalışma, metagenomik enzim aktivitelerinin ve mikrobiyal fonksiyonların KRK süreçlerine nasıl dahil olabileceğini aydınlatmaya yönelik gelecekteki araştırma çabalarına ilham verecektir.

Anahtar kelimeler: Kolorektal kanserin metagenomik analizi, makine öğrenimi, öznelik gruplaması, metagenomların fonksiyonel profilinin çıkarılması, topluluk düzeyinde enzim komisyon (EK) bollukları

Discovery of Colorectal Cancer-Associated Enzyme Groups from Metagenomic Data by Developing a Grouping Scoring Model Based on Enzyme Commission Numbers

Nur Şebnem Ersöz¹, Malik Yousef^{2,3}, Burcu Bakır-Güngör^{4,5}

¹ Department of Bioengineering, Abdullah Gül University Graduate School of Engineering and Science, Kayseri, Türkiye

² Department of Information Systems, Zefat Academic College, Zefat, Israel

³ Galilee Digital Health Research Center (GDH), Zefat Academic College, Zefat, Israel

⁴ Department of Computer Engineering, Abdullah Gül University Faculty of Engineering, Kayseri, Türkiye

⁵ Department of Bioengineering, Abdullah Gül University Faculty of Life and Natural Sciences, Kayseri, Türkiye

ABSTRACT

Objectives: Colorectal cancer (CRC) is the leading cause of cancer-related deaths worldwide. Recently, functional profiling of metagenomes has enabled the calculation of relative abundance values for enzyme-encoding genes. The primary objective of this study is to develop a method for identifying CRC-associated microbial enzymes through feature selection based on biological domain knowledge.

Materials and Methods: The dataset used in this study included 1262 samples (600 CRC and 662 controls) from eight different countries. HUMAnN3 was used to perform functional profiling of the metagenomic data and for the estimation of community-level enzyme commission (EC) abundance values. In this study, we propose a classification model based on the Grouping Scoring Modeling (G-S-M) approach. We grouped the enzymes according to their EC numbers to provide biological domain knowledge to the G-S-M model. In our experiments, we performed a comparative evaluation with different methods using 10 times Monte Carlo Cross-Validation.

Results: In our experiments, we observed that high-scoring CRC-associated enzyme groups include glycosidases (EC: 3.2.1), hydro-lyases (EC: 4.2.1), and CoA-transferases (EC: 2.8.3). Oligo-1,6-glucosidase enzyme (EC 3.2.1.10) was found as the most significant CRC-associated enzyme. Also, we have shown that the G-S-M model outperforms traditional feature selection methods with various classifiers.

Conclusion: Here, we performed a functional analysis of CRC metagenome by developing a G-S-M-based model that finds CRC-associated enzyme groups. This study aims to inspire future research efforts and shed light on the involvement of metagenomic enzyme activities and microbial functions in colorectal cancer (CRC) processes.

Keywords: Metagenomic analysis of colorectal cancer, machine learning, feature grouping, functional profiling of metagenomes, community-level enzyme commission (EC) abundances

İdiyopatik Obstrüktif Olmayan Azospermi için Bilinen ve Yeni Aday Genlerde Kopya Sayısı Varyantlarının Tüm Genom Dizileme ile Araştırılması

Oğuzhan Kalyon¹, Bilal Kamil Sulaiman Alobaidi¹, Lois Batty¹, Jonathan Coxhead², Kevin McEleny³, Godfried van der Heijden⁴, Liliana Ramos⁴, Manon Oud⁵, Miguel Xavier¹, Aneta Mikulasova¹, Giles Holt¹, Joris Andre Veltman¹

¹ Newcastle Üniversitesi, Biyobilimler Enstitüsü Tıbbi Bilimler Fakültesi, Newcastle upon Tyne, Birleşik Krallık

² Newcastle Üniversitesi, Tıbbi Bilimler Fakültesi, Genomik Çekirdek Tesisi, Newcastle upon Tyne, Birleşik Krallık

³ Newcastle Doğum Merkezi, The Newcastle upon Tyne Hastanesi, NHS Foundation Trust, Newcastle upon Tyne, Birleşik Krallık

⁴ Radboudumc Üniversite Hastanesi, Kadın Hastalıkları ve Jinekoloji Anabilim Dalı, Nijmegen, Hollanda

⁵ Radboudumc Üniversite Hastanesi, Donders Beyin, Biliş ve Davranış Enstitüsü, İnsan Genetiği Bölümü, Nijmegen, Hollanda

ÖZET

Amaç: Kromozomal anormallikler erkek infertilitesinin yaygın bir nedenidir ancak kopya sayısı varyantlarının (KSV) rolü hakkında çok az şey bilinmektedir. Bu çalışmada, kantitatif sperm defekti olan 200 hasta-ebeveyn üçlüsünde tüm genom sekanslama (TGS) gerçekleştirerek bilinen ve yeni erkek infertilite aday genlerini etkileyen KSV belirlendi.

Gereç ve Yöntem: İdiyopatik azospermisi olan 200 hasta ve ebeveynlere TGS yapılmıştır. CNVRobot, dysgu-SV ve AnnotSV kullanılarak KSV'ler tespit edilmiştir. Nadir *de novo* ve maternal olarak kalıtılan KSV'ler daha ileri çalışmalar için önceliklendirilmiştir. Replikasyon kohortu olarak kullanılmak üzere azospermisi olan 235 ek hastada tüm ekzom sekanslama (TES) gerçekleştirilmiştir.

Bulgular: Proband başına ortalama 1083 olmak üzere 200 üçlüde yaklaşık 212.000 KSV tespit edilmiştir. Çoğunlukla kodlamayan DNA'yı etkileyen, 2 kb ile 256 kb arasında değişen toplam 12 nadir *de novo* KSV tanımlanmıştır. Bu *de novo* delesyonlardan biri, akraba olmayan azospermik bir hastada bir duplikasyon KSV'sinden de etkilenen CTNNA3 genini etkilemektedir. CTNNA3'ün bozulması daha önce azospermik bir hastada tanımlanmıştır (PMID: 35017386). Buna ek olarak, 1863 nadir maternal kalıtlı KSV tanımlanmıştır. STRING analizi, bu KSV'lerden etkilenen ve birçoğunun spermatogenezde önemli bir rol oynadığı bilinen etkileşimli genlerin sayısında (p-değeri <4,99e-10) önemli bir zenginleşme olduğunu ortaya koymuştur.

Sonuç: Bu TGS çalışmasında, spermatogenezde yer alan genleri etkileyen ve potansiyel olarak erkek infertilitesini açıklayan birçok KSV keşfettik. Bunların etkisini daha fazla belirlemek ve erkek infertilitesi genetiğini daha iyi anlamak için büyük ölçekli replikasyon çalışmaları devam etmektedir.

Anahtar kelimeler: Genomik, yapısal varyasyonlar, erkek infertilitesi

Genome Sequencing Reveals Copy Number Variants in Known and Novel Candidate Genes for Idiopathic non-Obstructive Azoospermia

Oğuzhan Kalyon¹, Bilal Kamil Sulaiman Alobaidi¹, Lois Batty¹, Jonathan Coxhead², Kevin McEleny³,
Godfried van der Heijden⁴, Lilitiana Ramos⁴, Manon Oud⁵, Miguel Xavier¹, Aneta Mikulasova¹,
Giles Holt¹, Joris Andre Veltman¹

¹ Newcastle Üniversitesi, Biyobilimler Enstitüsü Tıbbi Bilimler Fakültesi, Newcastle upon Tyne, Birleşik Krallık

² Newcastle Üniversitesi, Tıbbi Bilimler Fakültesi, Genomik Çekirdek Tesisi, Newcastle upon Tyne, Birleşik Krallık

³ Newcastle Doğum Merkezi, The Newcastle upon Tyne Hastanesi, NHS Foundation Trust, Newcastle upon Tyne, Birleşik Krallık

⁴ Radboudumc Üniversite Hastanesi, Kadın Hastalıkları ve Jinekoloji Anabilim Dalı, Nijmegen, Hollanda

⁵ Radboudumc Üniversite Hastanesi, Donders Beyin, Biliş ve Davranış Enstitüsü, İnsan Genetiği Bölümü, Nijmegen, Hollanda

ABSTRACT

Objectives: Chromosomal abnormalities are a common cause of male infertility (MI), but little is known about the role of copy number variants (CNVs). Here, we identify CNVs affecting known and novel MI candidate genes by performing whole genome sequencing (WGS) in 200 patient-parent trios with quantitative sperm defects.

Materials and Methods: WGS of 200 patients with idiopathic azoospermia and their parents was performed. CNVs were detected and annotated using CNVRobot, dysgu-SV, and AnnotSV. Rare de novo and maternally inherited CNVs were prioritized for further studies. Whole exome sequencing (WES) was performed on 235 additional patients with azoospermia to be used as a replication cohort.

Results: Approximately 212.000 CNVs were detected in 200 trios with an average of 1083 per proband. A total of 12 rare de novo CNVs were identified ranging from 2 kb to 256 kb, mostly affecting non-coding DNA. One of these de novo deletions affects the CTNNA3 gene, which was also affected by a duplication CNV in an unrelated azoospermic patient. Disruption of CTNNA3 has been previously described in an azoospermic patient (PMID: 35017386). In addition, 1863 rare maternally inherited CNVs were identified. A STRING analysis revealed significant enrichment in the number of interacting genes (p -value < 4.99e-10) affected by these CNVs, many of which are known to play a prominent role in spermatogenesis.

Conclusion: In this WGS study, we discovered many CNVs affecting genes involved in spermatogenesis, potentially explaining male infertility. Large-scale replication studies are ongoing to further determine their relevance and improve our understanding of male infertility genetics.

Keywords: Genomics, structural variations, male infertility

GenNext: Yeni Nesil Dizileme Verileri için Yeni Nesil Analiz Platformu

Berk Ergun², Hasan Ali Turan¹, Eylül Aydın^{1,3}, Özkan Özdemir^{1,2}

¹ Geniva Bilişim Sağlık Hizmetleri A.Ş., İstanbul Türkiye

² Acıbadem Mehmet Ali Aydınlar Üniversitesi Tıp Fakültesi, Tıbbi Biyoloji Anabilim Dalı, İstanbul, Türkiye

³ Acıbadem Mehmet Ali Aydınlar Üniversitesi Sağlık Bilimleri Enstitüsü, Genom Çalışmaları, İstanbul, Türkiye

ÖZET

Amaç: GenNext, yeni nesil dizileme (YND) verilerinin analizi için geliştirilmiş, uçtan uca (P2P) ve merkezi-olmayan bir analiz platformudur. GenNext ile veri işleme, anotasyon, varyant filtreleme ve raporlama gibi süreçler kullanıcı dostu bir arayüz ile sunulmaktadır. Platform, FASTQ, BAM ve VCF verilerini işleyebilir, germ-hattı ve somatik varyant analizleri için kullanılabilir. Yapısal varyant analizinde ise yapay zekâ ile desteklenmektedir. GenNext'in KVKK-uyumlu özgün mimarisi, ulusal genom projeleri için ideal bir çözümdür. Bu çalışmada GenNext'in merkezi-olmayan yapısının KVKK uyumluluğu, YND veri analizindeki etkinliği ve kullanım kolaylığı değerlendirilmiştir.

Gereç ve Yöntem: GenNext, tasarımı gereği merkez işlem düğümleri (nodları) ile kullanıcı ve hesaplama düğümlerini birbirine bağlar. Bu işlem için platform en kısa yolu seçer, işlemler hesaplama düğümünde gerçekleştirildikten sonra merkez düğümlerinde analiz edilir. Düğümler Kubernetes ile yönetilir ve veriler WebRTC protokolü ile transfer edilir. Argo Events ve Workflows ise platformdaki iş yükünün yönetimi ve delegasyonu için kullanılır.

Bulgular: Beta-test aşamasında olan GenNext, Türkiye'de dört, Malezya ve Almanya'da ise birer merkezde kullanılmaktadır. Bugüne kadar özellikle kanser ve tanısız ve nadir hastalıklar kapsamında, hedefli-panel, tüm-ekzom, ve tüm-genom-dizilemeden elde edilen yaklaşık 3000 kadar verinin başarıyla işlendiği ve analiz edildiği tespit edilmiştir.

Sonuç: GenNext, YND verisi analizinde güçlü bir araç olup merkezi-olmayan ve KVKK-uyumlu özgün tasarımı ile ulusal ve uluslararası düzeyde güvenli ve etkili bir kullanım sunmaktadır. Yapay zekâ ile desteklenen yapısı sayesinde analiz kapasitesini arttırmakta ve genetik araştırmalarda yeni ufuklar açmaktadır. Meta veri paylaşımı ve entegrasyonunu da mümkün kılan GenNext'in, ulusal genom projelerini destekleme potansiyelinde olduğu öngörülmektedir.

Anahtar kelimeler: Yeni nesil dizileme, genomik, biyoinformatik

GenNext: Next-Generation Analysis Platform for Next-Generation Sequencing Data

Berk Ergun², Hasan Ali Turan¹, Eylül Aydın^{1,3}, Özkan Özdemir^{1,2}

¹ Geniva Informatics and Health Services Inc., İstanbul, Türkiye

² Department of Medical Biology, Acıbadem Mehmet Ali Aydınlar University, Faculty of Medicine, İstanbul, Türkiye

³ Institute of Health Sciences, Genome Studies, Acıbadem Mehmet Ali Aydınlar University, İstanbul, Türkiye

ABSTRACT

Objectives: GenNext is an end-to-end, decentralized analysis platform developed for next-generation sequencing (NGS) data analysis. It provides user-friendly processes such as data processing, annotation, variant filtering, and reporting. The platform can handle FASTQ, BAM, and VCF files and is used for germline and somatic variant analyses, with artificial intelligence support for structural variant analysis. GenNext's unique, Personal Data Protection Law-compliant design is ideal for national genome projects. This study evaluates the compliance of its decentralized architecture with data protection laws, assesses its efficacy in NGS data analysis, and examines its usability.

Materials and Methods: GenNext connects central processing nodes to user and computation nodes, selecting the shortest path for operations, including post-computation node processing. Data is then analyzed at the central nodes. Nodes are managed with Kubernetes, and data is transferred using the WebRTC protocol. Argo Events and Workflows manage and delegate platform workload.

Results: Currently in beta testing, GenNext is used in four centers in Türkiye, one in Malaysia, and one in Germany. Approximately 3000 datasets, encompassing targeted-panel, whole-exome, and whole-genome sequencing data, primarily focused on cancer and undiagnosed/rare diseases, have been successfully processed and analyzed.

Conclusion: With its decentralized and data protection-compliant design, GenNext proves to be a robust tool for NGS data analysis, ensuring secure and effective usage both nationally and internationally. Its AI-supported structure increases analysis capacity and opens new horizons in genetic research. GenNext is also expected to support national genome projects through its metadata sharing and integration capabilities.

Keywords: Next-generation sequencing, genomics, bioinformatics

Akciğer Kanserinde Kritik RNA Hedeflerinin Belirlenmesi için Değişen MikroRNA ve Yarışan Endojen RNA Etkileşimlerinin Kapsamlı Analizi

Selcen Arı Yuka^{1,2}, Alper Yılmaz¹

¹ Yıldız Teknik Üniversitesi Kimya-Metalürji Fakültesi, Biyomühendislik Bölümü, İstanbul, Türkiye

² Sağlık Biyoteknolojisi Mükemmeliyet Ortak Uygulama ve Araştırma Merkezi, İstanbul, Türkiye

ÖZET

Amaç: Ortak baskılayıcı miRNA'lara karşı miktar bağımlı olarak yarışmalı profil sergileyen kodlanmayan RNA ve mRNA'lar yarışan endojen RNA'lar olarak nitelendirilmektedir. Güncel yaklaşımlar bu etkileşimleri küçük endojen RNA ağları olarak ele alır ancak sağlıklı ve kanser dokularındaki değişen endojen RNA dinamiklerini geniş kapsamlı inceleyen çalışmalar sınırlıdır. Bu çalışmada dünyada en yüksek insidansa sahip türlerden biri olan akciğer kanserinde terapötik hedef olarak kullanılabilir RNA'lar, sağlıklı ve tümör dokuları arasında değişen yarışan endojen RNA etkileşimlerinin kapsamlı analizi ile belirlenmiştir.

Gereç ve Yöntem: Akciğer için normal ve tümör dokularının RNA ve miRNA dizileme verileri TCGA veri tabanından R (4.2.1) TCGAbiolinks (2.29.3) aracılığıyla elde edilmiş. miRNA:hedef etkileşimleri Encyclopedia of RNA Interactomes veri tabanından indirilmiştir. RNA, miRNA ifade ve miRNA:hedef matriksleri kullanarak ayrık kısmi korelasyon metodu uygulanmış, normal ve tümör dokularına özgü potansiyel yarışan endojen gen çiftleri elde edilmiştir. Normal ve tümör dokularında geniş kapsamlı dokuya özgü miRNA:ceRNA eksenleri kurulmuş, ağ merkezilik ölçütleri kullanılarak alt ağ yapıları olarak da incelenmiştir.

Bulgular: Akciğer sağlıklı ve tümör dokularında sırasıyla 393 ve 3374 yarışan endojen RNA tespit edilmiştir. Yarışan endojen RNA'ların ifade profillerindeki değişikliklerden bağımsız şekilde akciğer dokusundaki miRNA:ceRNA etkileşimlerinde katılabileceği ve akciğer kanserine özgü 3204 endojen RNA'nın özellikle kanser ilişkili yollarda görev aldığı görülmüştür. Ayrıca, 3204 ceRNA'dan elde edilen geniş kapsamlı miRNA:ceRNA etkileşiminde yedi alt ağ yapısı öne çıkmıştır.

Sonuç: Akciğer kanserinde değişen miRNA:ceRNA etkileşimlerinin analizi sonucunda yedi ceRNA ve altı miRNA'dan oluşan kritik bir ağ yapısının hedef olarak kullanılmasının terapötik potansiyelinin olabileceği saptanmıştır.

Anahtar kelimeler: Kodlanmayan RNA'lar, yarışmalı endojen RNA, ağ biyolojisi

Comprehensive Analysis of Fluctuating MicroRNA and Competing Endogenous RNA Interactions for Identification of Key RNA Targets in Lung Cancer

Selcen Arı Yuka^{1,2}, Alper Yılmaz¹

¹ Department of Bioengineering, Yıldız Technical University Faculty of Chemistry-Metallurgy, İstanbul, Türkiye

² Health Biotechnology Joint Research and Application Center of Excellence, İstanbul, Türkiye

ABSTRACT

Objectives: Non-coding RNAs and mRNAs that exhibit an abundance-dependent competitive profile against shared repressor miRNAs are characterized as competing endogenous RNAs. Current approaches consider these interactions as small endogenous RNA networks, but comprehensive studies on changed endogenous RNA dynamics in healthy and cancer tissues are limited. We identified RNAs that could serve as therapeutic targets in lung cancer, one of the highest incidence types in the world, through a comprehensive analysis of competing endogenous RNA(s) interactions that fluctuate between healthy and tumor tissues.

Materials and Methods: RNA and miRNA sequencing data of normal and tumor tissues of the lung were obtained from the TCGA database R (4.2.1) via the TCGAbiolinks (2.29.3). miRNA:target interactions were downloaded from the Encyclopedia of RNA Interactomes database. Using the RNA, miRNA expression, and miRNA:target matrices, the sparse partial correlation method was applied and potential competing endogenous gene pairs were curated. Comprehensive tissue-specific miRNA:ceRNA axes were established from normal and tumor tissues and examined as sub-networks using network centrality measures.

Results: We identified 393 competing endogenous RNAs in healthy lung tissues and 3374 in tumor tissues. We found that competing endogenous RNAs may participate in miRNA:ceRNA interactions in lung cancer regardless of expression change, and 3204 lung cancer-specific endogenous RNAs are specifically implicated in cancer-related pathways. Seven sub-networks emerged in the comprehensive miRNA:ceRNA interaction derived from 3204 ceRNAs.

Conclusion: As a result, it was determined that the use of a key network consisting of seven ceRNAs and six miRNAs as a target may have therapeutic potential for lung cancer.

Keywords: Non-coding RNAs, competitive endogenous RNA, network biology

Büyük Ölçekli Prenatal Test Genomik Verilerinden Türk Popülasyonunun Genetik Profilinin Ortaya Konulması

Selim Can Kuralay¹, Eren Akdeniz², Gülperi Özdoğan², İsmihan Merve Tekin², Özgecan Kayalar³,
Gökmen Zararsız⁴, Gökhan Yıldız⁵, Vahap Eldem¹

¹ İstanbul Üniversitesi Fen Fakültesi, Biyoloji Bölümü, İstanbul, Türkiye

² Genoks Genetik Hastalıklar Değerlendirme Merkezi, Ankara, Türkiye

³ Koç Üniversitesi Translasyonel Tıp Araştırma Merkezi (KUTTAM), Koç Üniversitesi Tıp Fakültesi, İstanbul, Türkiye

⁴ Erciyes Üniversitesi Tıp Fakültesi, Biyoistatistik Bölümü, Kayseri, Türkiye

⁵ Karadeniz Teknik Üniversitesi Tıp Fakültesi, Tıbbi Biyoloji Bölümü, Trabzon, Türkiye

ÖZET

Amaç: Her ne kadar tüm genom ve ekzom dizileme verileri büyük ölçekli genomik projelerde ana genomik kaynak olarak görülsede düşük kapsamlı WGS temelli invaziv olmayan doğum öncesi test (NIPT) verileri günümüzde, genetik çeşitliliği profileleme amacı bakımından henüz değerlendirilmemiş kaynak olarak kabul edilmektedir. Bu çalışma kapsamında, büyük çoğunluğu NIPT testi için üretilen 30,000 bireye ait düşük kapsamlı WGS verisini kullanarak Türk popülasyonunun varyant profilinin analiz edilmesi amaçlanmıştır.

Gereç ve Yöntem: Bu çalışma İstanbul Üniversitesi İstanbul Tıp Fakültesi Etik Kurulu onayı ile gerçekleştirilmiştir. Düşük kaliteli okumaların filtrelenmesini takiben temiz okumalar bwa (v0.7.17) ile insan genomuna (GRCh38) hizalanmıştır, hizalama verileri ile GATK (v4.2.6.1) kullanılarak varyant çağırma işlemi gerçekleştirilmiştir. Tüm bireylere ait varyant verisi popülasyona özgü varyant seti oluşturulması amacıyla birleştirilmiştir. Varyant seti SnpEff ve SnpSift (v5.2) kullanılarak anotasyonu gerçekleştirilmiştir, popülasyona özgü anote edilmiş Türk varyasyonu daha sonra dbSNP ve gnomAD veri tabanlarında bulunan diğer popülasyonlarla karşılaştırılmıştır.

Bulgular: 30,000 kişiden elde edilen yaklaşık 9,8 terabayt düşük kapsamlı WGS verisinin işlenmesi sonucunda 0,15x kapsamda ortalama %96 hizalanma oranı tespit edilmiştir. Türk varyomunda 34,400,478 SNP, 395,688 delesyon, 47,672 insersiyon ve beş MNP olmak üzere 34,843,843 genomik varyant bulunmaktadır. Klinik verileri elde edilebilen 49,857 varyantın çoğu benign olmakla birlikte, 1096 tanesinin muhtemel patojenik olduğu belirlenmiştir. Popülasyon yapısı analizleri sonucunda Türk popülasyonunun Büyük Orta Doğu, Kafkaslar, Balkan ve Avrupa popülasyonlarıyla yakın genetik ilişki içinde olduğu saptanmıştır.

Sonuç: WGS tabanlı NIPT'den elde edilen genomik veriler, Türk genom yapısına ve Türk popülasyonuna özgü katalog varyantlara ilişkin bilgimizi genişletebilmek için alternatif bir kaynak olarak kullanılabilirliği ortaya konulmuştur.

Anahtar kelimeler: WGS, düşük kapsama, NIPT, varyom, Türk varyom

Large-Scale Analysis of Low-Coverage WGS-Based Nipt Data Contributes to Genetic Variant Profile of the Turkish Population

Selim Can Kuralay¹, Eren Akdeniz², Gülperi Özdoğan², İsmihan Merve Tekin², Özgecan Kayalar³, Gökmen Zararsız⁴, Gökhan Yıldız⁵, Vahap Eldem¹

¹ Department of Biology, İstanbul University Faculty of Science, İstanbul, Türkiye

² GENOKS Genetic Disease Diagnostic Center, Ankara, Türkiye

³ Koç University Research Center for Translational Medicine (KUTTAM), Koç University Faculty of Medicine, İstanbul, Türkiye

⁴ Department of Biostatistics, Erciyes University Faculty of Medicine, Kayseri, Türkiye

⁵ Department of Medical Biology, Karadeniz Technical University Faculty of Medicine, Trabzon, Türkiye

ABSTRACT

Objectives: While whole-genome and exome sequencing data are commonly utilized as the primary genomic sources in large-scale genomic projects, low-coverage whole-genome sequencing-based non-invasive prenatal tests (NIPT) data is currently regarded as an untapped resource for genetic diversity profiling. This study aimed to analyze the variant profile of the Turkish population using low-coverage WGS data of 30.000 individuals mainly generated for the NIPT test.

Materials and Methods: The study received approval from the İstanbul University İstanbul Faculty of Medicine Ethics Committee. After filtering low-quality reads, the clean reads were aligned to the human genome (GRCh38) using bwa (v0.7.17), and the alignment data were employed to call variants with GATK (v4.2.6.1). Variant data for all individuals were merged to generate a population-specific variant set. The variant set was annotated using SnpEff and SnpSift (v5.2). Subsequently, the population-specific annotated Turkish variome data was compared with other populations deposited in the dbSNP and gnomAD databases.

Results: Approximately 9.8 terabases of low-coverage WGS data from 30.000 individuals were processed. The average sequence alignment rate was >96% with coverage at 0.15x. Turkish variome harbors 34.843.843 genomic variants, including 34.400.478 SNPs, 395.688 deletions, 47.672 insertions, and five MNPs. Clinical data were available for 49.857 variants; most were benign, while 1096 were likely pathogenic. Population structure analyses revealed that the Turkish population has close genetic relationships with the Greater Middle East, Caucasus, Balkan, and European populations.

Conclusion: Genomic data from WGS-based NIPT can serve as an alternative source to expand our understanding of the Turkish genome compositions and catalog variants that are unique to the Turkish population.

Keywords: WGS, low-coverage, NIPT, variome, Turkish variome

HISAT2 ve Star Hizalama Araçlarının Çölyak Kohortundan Elde Edilen RNA Dizileme Verilerinde Karşılaştırılması

Yeliz Ekici^{1,2}, Selçuk Candan^{3,4}, Billur Canbakan⁴, Feyza Nur Tuncer¹

¹ İstanbul Üniversitesi, Aziz Sançar Deneysel Tıp Araştırma Enstitüsü, Genetik Anabilim Dalı, İstanbul, Türkiye

² İstanbul Üniversitesi, Sağlık Bilimleri Enstitüsü, İstanbul, Türkiye

³ Sağlık Bilimleri Üniversitesi, Şehit Prof. Dr. İlhan Varank Sancaktepe Eğitim ve Araştırma Hastanesi, Gastroenteroloji Bölümü, İstanbul, Türkiye

⁴ İstanbul Üniversitesi-Cerrahpaşa, Cerrahpaşa Tıp Fakültesi, Gastroenteroloji Bölümü, İstanbul, Türkiye

ÖZET

Amaç: Yeni nesil dizileme yöntemlerinin hızla gelişmesi ile RNA dizileme klinik araştırmalarda daha yaygın kullanılmaya başlanmıştır. Buna paralel olarak gen anlatım miktarının belirlenmesini kolaylaştırmak için birçok biyoformatik araç geliştirilmiştir. RNA dizileme ile yüksek miktarda ham gen anlatım verisi üretilmektedir. Veri miktarının fazla olması, çalışmanın amacına uygun şekilde değerlendirilmesini zorlaştırmaktadır. Veri analizinde okumaların kırılması ve referans genoma hizalanması en temel adımları oluşturmaktadır. Bu çalışma, HISAT2 ve STAR hizalama araçlarının RNA-dizi analizi üzerindeki etkisini, analiz sonuçlarının biyolojik ilgisini ve hesaplama sürelerinin değerlendirilmesini amaçlamaktadır.

Gereç ve Yöntem: Çalışma kapsamında, çölyak hastaları (n=6) ile çölyak tanısı dışlanmış gönüllü kontrollerin (n=6) duodenal biyopsilerinden elde edilen RNA-dizi verileri kullanılarak HISAT2 ve STAR hizalama araçları ile gen ifadesini ölçmek için popüler programlardan biri olan DESeq2 arasındaki uyum test edilmiştir.

Bulgular: Her iki iş akışında p değeri < 0,05 olan sonuçlar için filtrelendiğinde, HISAT2 analizinden 3431, STAR analizinden ise 3324 gen kaydedilmiştir. İki liste karşılaştırıldığında ise ortak gen sayısı 3221'di. Ortak genler çölyak hastalık mekanizması ile bağlantılıydı. Ortak olmayanlar ise her iki iş akışında da bağışıklık sistemi yollarını işaret etmekteydi. Ancak STAR sonuçları bağışıklık sistemi hastalıkları ile ilişkililikten, HISAT2 sonuçları enfeksiyon hastalıklarıyla ilişkiliydi.

Sonuç: Çalışmamızda Türk çölyak hasta grubunda ilk kez RNA-dizi analizi gerçekleştirilmiştir. Elde edilen veriler belirli bir araç veya yöntemin tahmin hatalarını azaltmak için birden fazla iş akışından yararlanılarak sonuçların değerlendirilmesini önermektedir. Bu çalışmadan elde edilen veriler gelecekte genişletilmiş hasta grubu ile gerçekleştirmeyi planladığımız RNA-dizi analizi için ön çalışma niteliği taşımaktadır.

Anahtar kelimeler: RNA dizileme, çölyak hastalığı, STAR, HISAT2, hizalama araçları

Comparison of HISAT2 and Star Alignment Tools Utilizing RNA Sequencing Data Obtained from a Celiac Cohort

Yeliz Ekici^{1,2}, Selçuk Candan^{3,4}, Billur Canbakan⁴, Feyza Nur Tuncer¹

¹ Department of Genetics, İstanbul University Aziz Sançar Institute of Experimental Medicine, İstanbul, Türkiye

² İstanbul University, Graduate School of Health Sciences, İstanbul, Türkiye

³ Clinic of Gastroenterology, Sancaktepe Şehit Prof. Dr. İlhan Varank Training and Research Hospital, İstanbul, Türkiye

⁴ Department of Gastroenterology, İstanbul University-Cerrahpaşa, İstanbul, Türkiye

ABSTRACT

Objectives: RNA sequencing is now extensively employed in clinical research, leading to the development of numerous bioinformatics tools for improved gene expression quantification. However, the generation of large RNA-seq datasets has introduced challenges in data analysis itself. Read trimming and alignment are the most essential steps in data analysis. We aimed to evaluate HISAT2 and STAR alignment tools in terms of their impacts on RNA-seq analysis. For this purpose, we compared the computational times and biological relevance of the results.

Materials and Methods: The concordance between HISAT2 and STAR alignment tools and DESeq2 which measures gene expression, was tested using RNA-seq data obtained from duodenal biopsies of celiac patients (n= 6) and non-celiac controls (n= 6).

Results: When data were filtered for $p < 0.05$ in both workflows, there were 3431 genes in the HISAT2 analysis, while STAR had 3324. When the two lists were compared, the number of common genes was 3221. Common genes were linked to the celiac mechanism, whereas the others were linked to immune-system pathways. STAR results were associated with immune-system diseases while HISAT2 results were associated with infectious diseases.

Conclusion: RNA-seq analysis was performed for the first time in a Turkish celiac patient cohort. The data obtained suggest that evaluating the results using multiple workflows is necessary to reduce the prediction errors of a particular tool or method. These preliminary findings lay the groundwork for future RNA-seq research involving a larger celiac cohort planned by our research group.

Keywords: RNA sequencing, celiac disease, STAR, HISAT2, alignment tools