



Protein – Protein Etkileşimi Tespit Yöntemleri, Veri Tabanları ve Veri Güvenilirliği

Volkan Altuntaş^{1*}, Murat Gök²

¹ Bursa Teknik Üniversitesi, Rektörlük, Bursa, Türkiye (ORCID: 0000-0003-3144-8724)

² Yalova Üniversitesi, Mühendislik Fakültesi, Bilgisayar Mühendisliği Bölümü, Yalova, Türkiye (ORCID: 0000-0003-2261-9288)

(İlk Geliş Tarihi 21 Nisan 2020 ve Kabul Tarihi 6 Temmuz 2020)

(DOI: 10.31590/ejosat.724390)

ATIF/REFERENCE: Altuntaş, V. & Gök, M. (2020). Protein – Protein Etkileşimi Tespit Yöntemleri, Veri Tabanları ve Veri Güvenilirliği. *Avrupa Bilim ve Teknoloji Dergisi*, (19), 722-733.

Öz

Önemli biyolojik aktiviteler tek bir molekülün sonucu değil, birbirleriyle etkileşime giren çoklu moleküllerin etkilerinin ürünü olarak ortaya çıkmaktadır. Protein-protein etkileşimlerinin belirlenmesi, ilgili proteinlere ait fonksiyonların tespit edilmesi için önemli bilgi sağlamaktadır. Genlerin ve proteinlerin büyük bir çoğunluğu işlevlerini birbirleriyle etkileşimleri sonucunda oluşturmaktadırlar. Protein-protein etkileşimlerini incelemek için çok sayıda yöntem geliştirilmiştir. Etkileşimlerin tespitinde in vitro, in vivo ve in silico olarak adlandırılan 3 temel yaklaşım bulunmaktadır. In vitro ve in vivo yöntemlerin maliyet, zaman gibi sınırlamaları bulunur. İn silico yöntemler deneysel yönlendirme ile maliyet ve zaman kazancı için geliştirilmiştir. Yöntemler sonucunda oluşan veri setleri gürültülüdür, çok sayıda yanlış pozitif ve yanlış negatif değerler içermektedirler. Protein etkileşim tespit yöntemlerindeki gelişmeler hastalıkların tespit edilmesi, model organizmalara ait yolların ve protein komplekslerinin belirlenmesi gibi birçok alana doğrudan etki etmektedir. Yapılan çalışmalar sonucunda tespit edilen etkileşimler veri tabanlarında saklanmakta ve ücretsiz olarak erişilebilmektedir. Metotların hızlanması ile tespit edilen etkileşim sayısındaki artış, elde edilen bu verilerin analiz edilmesini, bir veya birden fazla metot ile sağlanmasını ve doğruluğunun belirlenmesini önemli hale getirmektedir. Bu çalışmada protein-protein etkileşim tespitinde kullanılan in vitro, in vivo ve in silico yöntemler ve protein-protein etkileşim veri tabanları incelenmektedir. Tespit yöntemlerinin artıları ve eksileri araştırılmış ve yöntemlerin avantaj ve dezavantajları paylaşılmıştır. Veri tabanlarının içerdiği bilgiler karşılaştırılmış, benzerlik oranları ve sebepleri araştırılmıştır.

Anahtar Kelimeler: Protein, Protein-protein etkileşimi, PPI tespit yöntemleri, PPI veri tabanları.

Protein - Protein Interaction Detection Methods, Databases and Data Reliability

Abstract

Important biological activities do not result from a single molecule but as a result of the effects of multiple molecules interacting with each other. The determination of protein-protein interactions provides important information for determining the functions of the respective proteins. The most majority of genes and proteins function as a result of interactions with each other. Numerous methods have been developed to study protein-protein interactions. In the determination of interactions, there are three basic approaches called in vitro, in vivo, and in silico. In vitro and in vivo methods have limitations such as cost and time. In silico methods have been developed for cost and time savings with experimental guidance. The data sets generated by the methods are noisy and contain a large number of false-positive and false-negative values. Advances in protein interaction detection methods have a direct impact on many areas such as the detection of diseases, pathways of model organisms, and protein complexes. The interactions identified as a result of the studies are stored in the databases and can be accessed free of charge. With the increase in the number of interactions detected by

* Sorumlu Yazar: Bursa Teknik Üniversitesi, Rektörlük, Bursa, Türkiye, ORCID: 0000-0003-3144-8724, valtuntas@gmail.com

accelerated methods, it became important to analyze the obtained data, verify it with one or more methods, and determine its accuracy. In this study, in vitro, in vivo and in silico methods and protein-protein interaction databases used for determination of protein-protein interaction are examined. The pros and cons of detection methods were investigated and the advantages and disadvantages of the methods were shared. The information contained in the databases was compared, investigated the similarity rates and reasons.

Keywords: Protein, Protein-protein interaction, PPI detection methods, PPI databases.

1. Giriş

1.1. Protein Nedir?

Bütün canlı organizmalar yaşayan hücrelerden oluşur ve temel hücresel mekanizmaları paylaşırlar. Bu hücre yapıları, bakteri veya insana ait olsa da, karmaşıklıkları bakımından farklılık içerseler de aynı yapı bloklarını (DNA, RNA ve protein) içermektedirler [1]. Protein sekansları DNA'da kodlanır ve moleküler biyolojinin santral dogması olarak anılan bir yolla sentezlenir. Her hücreye ait plan, DNA diziliminde kodlanmaktadır. DNA kopyalanır ve bir kısmı haberci RNA'ya (mRNA) aktarılır ve mRNA da proteinlere dönüştürülür. Bir mRNA'nın proteine tercümesinin ardından, proteinin işlevsel hale gelmeden önce geçmesi gereken birçok süreç bulunur. Protein önce katlanarak doğru üç boyutlu yapıyı oluşturmalı, daha sonra belirli bir hücresel bölgeye taşınmalı ve çoğu zaman ilgili proteine özgü değişiklikler yapılmalıdır [2].

Bir hücrede birçok farklı tipte protein bulunmaktadır. Örneğin, model mikroorganizmalardan *Saccharomyces cerevisiae*'de 6000 farklı protein türü bulunmaktadır [3]. Her bir protein bir kopya veya binlerce kopya halinde sentezlenebilir. Protein ekspresyon seviyeleri sıkı bir şekilde düzenlenmektedir. Bu yüzden benzer hücrelerden, farklı koşullar altında sentezlenen proteinler belirgin farklılıklar içerebilir. Proteinler, sinyal yollarının oluşturulması, hücresel süreçlerin düzenlenmesi, hücre zarı üzerinde seçici taşıyıcılık, kimyasal reaksiyonların hızlandırılması gibi hücresel süreçlerde önemli rol oynarlar [2]. Bu görevlerin çoğunda, proteinler, farklı boyutlarda kompleksler oluşturarak, birbirlerini değiştirerek ve diğer proteinleri transfer ederek birlikte çalışırlar [4]. Bu protein kompleksleri, hücrede çeşitli görevleri yerine getiren küçük makineler gibi davranmaktadırlar. Hücrenin metabolizmasına, sinyal iletimine, DNA transkripsiyon ve kopyalanmasına, DNA hasar onarımına ve daha birçok mekanizmaya katılırlar. Bu hücresel makinelerin tam fonksiyonlarını tanımlamak, karakterize etmek ve aralarındaki etkileşimleri belirlemek, canlı bir hücrenin işlevselliğini anlamada önemlidir.

1.2. Protein –protein etkileşimi nedir ve önemi

Protein-protein etkileşimleri (PPI) hücrenin metaboliksel gelişimi, kontrolü, hücreler arası iletişim, gibi birçok biyolojik süreci temsil etmektedir [5]. Bu önemli süreçler sebebi ile PPI'lar sistem biyolojisinin temel konularından biridir. PPI'lar zincirler arasındaki kovalent olmayan temaslar ile oluşmaktadır [6]. Bu temasların sonucu olarak proteinler arası etkileşim ve ilişkiler meydana gelir. Yapısal ve işlevsel özelliklerine göre PPI'lar çeşitli şekillerde sınıflandırılabilir [7]. Etkileşim yüzeyleri temelinde homooligomerik veya heterooligomerik olarak, istikrarına göre mecburi veya mecburi olmayan ve kalıcılıklarına göre geçici veya kalıcı olarak sınıflandırılabilirler [8]. Belirli bir PPI, bu sınıfların bir kombinasyonu olabilir. Proteinler arasındaki geçici etkileşimler sinyallerin iletimini sağlayan sinyal yollarını oluşturulmakta ve kalıcı etkileşimler kararlı bir protein kompleksine dönüşmektedir [9].

In vivo ortamda yapılan deneysel çalışmalarda doku ekstraktı veya ölü organizma kullanımı yerine canlı organizmalar veya hücreler kullanılır [10]. Proteinler in vivo ortamlarda fonksiyonlarını yerine getirirken izole edilmiş türler gibi davranmazlar [11]. Proteinlerin büyük bir kısmının yalnız çalışmadığı, yalnızca protein komplekslerinde faaliyet gösterebildiği ortaya çıkmıştır [12]. Doğrulanmış proteinlerin analizi, aynı hücresel süreçlerde yer alan proteinlerin art arda birbirleri ile etkileşime girdiğini ortaya koymaktadır [13]. PPI çalışmaları, proteinlerin hücre içindeki fonksiyonlarını anlamak için önemlidir. Tanımlanamayan bir proteinin fonksiyonları, fonksiyonları daha önce tespit edilmiş bir protein ile olan etkileşimleri tespit edilerek tahmin edilebilir. PPI'ların ayrıntılı bir şekilde incelenmesi, hücresel süreçlerin anlaşılması için model geliştirme sürecini hızlandırmıştır [8]. Belli bir proteomdaki proteinlerin etkileşimlerinin modellenmesi, hücrenin biyokimyasını anlamak için fayda sağlayacaktır [8]. İki veya daha fazla proteinin belli bir fonksiyonu yerine getirmek amacı ile etkileşmesi çeşitli yollardan gerçekleşebilir. PPI'lar enzimlerin kinetik özelliklerini değiştirebilir, efektör molekülleri için yeni bir bağlanma yeri oluşturabilir, bir proteini inaktive edebilir, hücresel düzenleyici olarak rol alabilirler. PPI'lar heterojen süreçler içerir ve oluşmaları için gerekli koşulların kapsamı geniştir. Hücredeki önemlerinin anlaşılabilmesi için, yapmış oldukları etkileşimler tanımlanmalı ve etkileşimlerin sonuçları belirlenmelidir [8]. Farklı deneysel kaynaklar kullanılarak kapsamlı PPI ağları inşa edilebilmesi mümkündür. PPI deneylerinin yüksek çıktılı verileri laboratuvar doğrulamalarını zorlaştırmaktadır. Keşfedilmemiş proteinlerin işlevlerini anlamak için PPI ağlarının hesaplamalı analizi ile verilerin hacminin azaltılması ve deneysel yönlendirme artık zorunlu bir araç haline gelmiştir. Günümüzde, protein-protein etkileşimi modern sistem biyolojisinin gelişimi ve ilerlemesi için en önemli konulardan biridir.

2. PPI Tespit Yöntemleri

Protein-protein etkileşim tespit yöntemleri kategorik olarak in vivo, in vitro ve in silico yöntemler olarak sınıflandırılabilir. In vivo tekniklerinde belirli bir prosedür canlı organizma üzerinde [10], in vitro tekniklerde ise canlı organizmanın dışında kontrollü bir ortamda gerçekleştirilir [14]. In silico teknikler, bilgisayar simülasyonları aracılığı ile bilgisayar ortamında gerçekleştirilir [15]. Her yöntemin farklı avantaj ve dezavantajları bulunmaktadır. In vivo ve in vitro yöntemler maliyet zaman gibi sınırlamalara sahip iken in silico yöntemlerde de deneysel doğrulama gereksinimi bulunmaktadır.

2.1. In Vivo Yöntemler

In vivo çalışmalarda, çeşitli biyolojik varlıkların etkileri, bir doku ekstraktı veya ölü organizma kullanımı yerine, canlı organizmalar veya hücreler (insanlar, hayvanlar ve bitkiler) üzerinde test edilir [10]. Hayvanlarda yapılan testler ve klinik çalışmalar in vivo araştırmanın temel unsurlarıdır. In vivo testler genellikle in vitro testlerden sonra uygulanır çünkü bir deneyin canlı organizma üzerindeki genel etkilerini gözlemlemek için daha uygundur. In vitro testler bazen yanıltıcı sonuçlar ortaya çıkarabileceği için yapılan deneylerin in vivo etkinliğinin doğrulanması önemlidir. Literatürde in vitro ortamda etkin olan ve in vivo ortamda etkinlik kazanamayan moleküllere [16,17], in vitro ortamda etkisiz iken in vivo ortamda ölümcül olan organizma [16,19] raporlanmıştır. In vivo çalışmaların kesin görüş sunma potansiyeline sahip olduğu düşünülse de, bu sonuçların yanıltıcı olabileceği çeşitli durumlarda olabilir. Bir tedavinin kısa vadede fayda sağlarken uzun vadede zarar vermesi buna örnek olabilir.

2.1.1. Y2H (Yeast two-hybrid)

Y2H yöntemi PPI'ların saptanmasında uygulanan in vivo bir yöntemdir. Y2H yönteminde, spesifik fonksiyonlara sahip olan iki protein bölgesi gereklidir:

- DNA'ya bağlanmaya yardımcı olan bir DNA bağlama alanı (DBD).
- DNA'nın transkripsiyonunu aktive etmekten sorumlu aktivasyon bölgesi (AD).

Her iki alan bir haberci genin transkripsiyonu için gereklidir. Y2H analizi, protein çiftleri arasında PPI'nın doğrudan tanınmasına olanak tanır. Ancak, yöntem çok sayıda yanlış pozitif etkileşimli sonuca neden olabilir. İlave olarak, Y2H yönteminde yanlış negatif sonuç olarak elde edilen pek çok gerçek etkileşim tespit edilemeyebilir. Y2H analizinde, etkileşimli proteinler çekirdeğe yerleştirilir. Proteinler işlevlerini yerine getirmek için post-translasyonel modifikasyonlara ihtiyaç duyar bu yüzden Y2H deneyinde proteinlerin normal davranması ve etkileşimde olması olasılığı düşer. Proteinler doğal fizyolojik ortamlarında olmadığında etkileşime uygun şekilde katlanamazlar [20] – [22].

2.1.2. Sentetik letalite (Synthetic lethality)

Sentetik letalite fiziksel etkileşimden ziyade fonksiyonel etkileşimlere dayanır. Sentetik letalite, varyasyon, çevresel değişimler ve mutasyonlar gibi rastgele olaylara rağmen fenotipik stabiliteye izin veren mekanizmaları anlamaya çalışan bir in vivo genetik tarama türüdür. Bu yöntemde, tek başına uygulandığında zarar vermeyen fakat belirli koşullar altında birleştirildiğinde öldürücü olan iki veya daha fazla gende mutasyonlar üretilir. Sentetik letalite ile tespit edilen sonuçlar proteinler arasında fiziksel etkileşim gerektirmez. Bu tür ilişkiler fonksiyonel etkileşim olarak adlandırılır [23] – [25].

2.1.3. Afinite etiketleri (Affinity tags)

Rekombinant proteinlerin ekspresyonu ve saflaştırılması, proteinlerin yapısını ve işlevini karakterize etmek için kullanılmaktadır. Yüksek saflıkta yeterli konsantrasyon elde etmek için ilgili proteini saflaştırmaya ihtiyaç vardır. Afinite bazlı protein saflaştırması, genetik olarak kaynaştırılmış bir afinite etiketinin seçici bağlanmasından yararlanır. İlk olarak, hücreler, etikete kaynaşmış bir yem proteinini kodlayan bir plazmid ile transfekte edilir. Uygun bir ekspresyon döneminden sonra, hücreler parçalanır ve bağlı proteinler ile birlikte etiketli yem, katı bir desteğe bağlı spesifik bir kimyasal veya biyolojik ligand kullanılarak izole edilir. Ayrıştırılan proteinler daha sonra jel-elektroforez ile ayrılır ve spesifik olarak bağlı proteinler kütle spektrometresi ile tanımlanır. Kullanılacak etiketin nihai seçiminden önce dikkate alınması gereken kendine özgü avantajları ve dezavantajları vardır. Bu uygulama, özgüllük, çözünürlük, bağlama ve elüsyon koşulları için şartlara bağlıdır. Bu yöntemler, yüksek afinite ve yavaş kinetikler ile etkileşime giren proteinlere karşı yanlıdır. Afinite kromatografisine dayalı yöntemler, geçici protein etkileşimlerinin saptanması için uygun olmayabilir. Protein aktivitesinin translasyon sonrası kontrolünde yer alan geçici kompleksler saptanamazlar. Hücre içi ortam ile seyreltme tüpü arasındaki konsantrasyon farkları hatalı sonuçların sebebi olarak gösterilmektedir [26] – [28].

2.1.4. Kimyasal çapraz bağlama (Chemical crosslinking)

Protein-protein etkileşimlerini yapısal olarak çözmek için alternatif bir strateji kimyasal çapraz bağlama, enzimatik sindirim ve reaksiyon ürünlerinin kütle spektrometresi ile analizidir. Çapraz bağlama, bir protein içindeki fonksiyonel grup çiftlerini kovalent olarak bağlayarak, bir dizi yapısal olarak tanımlanmış etkileşimin oluşturulmasını sağlar. Kimyasal çapraz bağların sağladığı kısıtlamalardan yararlanarak, düşük çözünürlüklü 3 boyutlu yapıların çıkarılmasında temel oluşturacak bir protein veya protein kompleksi içinde mesafe haritaları oluşturulabilir. Bu deneylerde kimyasal sabitleştirme gereklidir. Doğal durumunda hücrelerden kromatin yakalanamaz. Çapraz bağlı kromatin, bazı antikörlerin epitoplarını antikör-kromatin bağlanmasını etkileyerek maskeleyebilir [29] – [32].

2.2. In Vitro Yöntemler

In vitro çalışmalar, normal biyolojik ortam dışındaki (canlı organizma dışı) mikroorganizmalar, hücreler veya biyolojik moleküller ile yapılır [10]. Test tüpü deneyleri olarak adlandırılan bu çalışmalar, test tüpleri, şişeler, petri tabakları ve mikrotiter plakalar gibi ekipmanlar ile laboratuvar ortamında yapılır [33]. Biyolojik çevreden izole edilmiş bir organizma kullanılarak yapılan çalışmalar, tüm organizmalarla yapılabilecekten daha detaylı ve daha uygun bir analiz yapmaya izin verir. Bununla birlikte, in vitro deneylerden elde edilen sonuçlar, deneyin bütün bir organizma üzerindeki etkilerini tam olarak ve doğru bir şekilde ortaya koymayabilir [17,19,16,18]. In vitro deneylerin zayıf yönlerinden biri, organizmanın hassas hücresel koşullarını oluşturamamalarıdır. In vitro deneylerin aksine in vivo çalışmalar biyolojik ortamda yapılan araştırmalardır. In vitro çalışmaların doğrulanması için in vivo çalışmalara başvurulur [14, 34].

2.2.1. TAP-MS (Tandem affinity purification-mass spectroscopy)

Tandem afinite arıtma-kütle spektroskopisi TAP-MS, kromozomal lokusu üzerindeki ilgilenilen proteinin çift etiketlenmesine dayanır. Bu işlemi iki aşamalı bir saflaştırma ve kütle spektroskopik analizi takip eder. Kütle spektrometrik protein tanımlaması ile birleştirilmiş tandem-afinite saflaştırması, biyolojik olarak ilgili protein komplekslerinin bileşimini belirlemek için güçlü bir yaklaşımdır. TAP etiketi kaynaştırılan bir protein, hücrelerde ifade edilir ve TAP-etiketli proteini birleştiren protein komplekslerini saflaştırmak için bir yem olarak kullanılır. TAP-etiketlemenin önemi, çok çeşitli protein komplekslerini tanımlama ve in vivo olarak var olan monomerik veya multimerik protein komplekslerinin aktivitesini test etme yeteneğidir. TAP'ın kütle spektroskopisi ile kullanımı, protein etkileşimlerini ve protein komplekslerini tanımda güç kazandırmaktadır [35] – [37].

2.2.2. Afinite kromatografi (Affinity chromatography)

Afinite kromatografi yüksek oranda duyarlıdır ve proteinlerdeki en zayıf etkileşimleri bile tespit edebilir. Etkileşim genellikle tersinirdir. Saflaştırma, bir yüzeyine sabitlenmiş olan moleküllerden (ligand) biri ile bifazik bir etkileşim yoluyla elde edilirken, eşi (hedef), karmaşık bir karışımın parçası olarak hareketli bir fazdadır. Yakalama adımı genellikle yıkama ve elüsyon ile takip edilir. Sonuçta yüksek oranda saflaştırılmış protein elde edilir. Hücresel sisteme dâhil olmasalar bile proteinler arasındaki yüksek belirlilik nedeniyle yanlış pozitif sonuçlar ortaya çıkar. Bu nedenle elde edilen sonuçları çapraz kontrol etmek ve doğrulamak için başka yöntemler kullanılması gerekir [38].

2.2.3. Eş-immüno-çökeltme (Co-immunoprecipitation)

Eş-immüno-çökeltme (co-IP), spesifik bir hedef proteine bağlı olan proteinleri dolaylı olarak yakalamak için kullanılır. Hedef protein-spesifik antikorların kullanımı ile fizyolojik olarak ilgili protein-protein etkileşimlerini tanımlayan bir tekniktir. Bilinen protein (antijen) yem proteini olarak adlandırılır ve etkileştiği protein av protein olarak adlandırılır. Hücreler denatüre edici olmayan koşullar altında tamamen parçalandıktan sonra, birbirine bağlanan proteinler korunur. Proteinlerin, doğal etkileşimlerinde gerekli olabilecek karmaşık hücresel bileşenler karışımında mevcut olan bütün bir hücre ekstresi kullanılır. Ökaryotik hücrelerin kullanımı post translasyonel modifikasyonu mümkün kılar. [39] – [41]

2.2.4. Mikrodizi (Microarrays)

Bir protein mikrodizisi, üzerinde çeşitli şekillerde, proteinlerin ayrı yerlere düzenli şekilde bağlandığı bir cam parçasıdır. Protein mikrodizi gelişiminin ardındaki amaç, otomatik işlemlerle paralel olarak çok sayıda belirleme gerçekleştirerek verimli protein analizleri gerçekleştirmektir. Protein mikrodizileri proteinleri tespit etmek, ekspresyon seviyelerini izlemek ve protein etkileşimlerini ve fonksiyonlarını hızlı bir şekilde araştırmak için yaygın olarak kullanılmaktadır. İlk olarak hedef protein dizilimin fragmanlarına floresan işaretleyiciler eklenir. Bunların çipin problemleri ile reaksiyona girmelerine izin verilir. Etkileşen dizilimlere ait fragmanlar çipin problemlerine bağlanır. Çip yıkanarak temizlenir. Etkileşim kuran dizilimler bir lazer ışını ile tespit edilir. Bu teknik, çok sayıda dizilimin hızlı ve eş zamanlı olarak tanımlanmasını sağlar [42] – [44].

2.2.5. Parça tamamlama (Fragment complementation)

Protein-parça tamamlama deneyleri protein-protein etkileşimlerini saptamak için oluşturulmuş analiz ailesidir. Basit ve doğrudan yollar sağlar. Herhangi bir moleküler ağırlığa sahip proteinler arasındaki PPI'yı tespit etmek ve endojen seviyelerinde ifade etmek için kullanılabilir. Kütle spektroskopisi kullanarak protein tanımlaması için peptit parmak izi ve geniş yelpaze (shotgun) proteomik yöntemleri kullanılır. Peptit parmak izi ile az protein içeren saflaştırılmış örneklerde çalışılabilir. Geniş yelpaze proteomik de ise birçok protein içeren örnekler kullanılabilir. Karmaşık karışımları analiz etmek için güçlü bir yöntemdir. İncelenen fragmanların bağlayıcı proteinlerin yokluğunda kendiliğinden birleşmemesi metodun doğruluğu için önemlidir. Bu durum, yanlış pozitiflere ve yöntemin yararsız hale gelmesine neden olabilir [45] – [47].

2.2.6. Faj Ekranı (Phase display)

Faj ekranı eski ve güçlü yöntemlerinden biridir. Teknoloji bakteriyofaj kaplama proteinlerine kaynaşmış polipeptidlerin, kodlama geni de içeren faj partikülleri üzerinde ifade edilebilmesine dayanmaktadır. Bu şekilde, genotip ve fenotip arasında fiziksel bir bağlantı kurulur. Basit moleküler biyoloji yöntemleri ile DNA kodlu peptidlerin veya proteinlerin çeşitli kütüphaneleri üretilebilir. Faj gösterimli kütüphaneler, bir bakteri konakçı kullanılarak çoğaltılabilir. Herhangi bir amino asit dizisi, DNA'nın faj partiküllerinin içinde dizilmesiyle kolayca tespit edilebilir [48] – [50].

2.2.7. X-ışını (X-ray)

X-ışını kristalografisi protein yapılarının atomik düzeyde görselleştirilmesini ve bu şekilde protein fonksiyonunun anlaşılmasını sağlayan, yüksek çözünürlüklü bir mikroskopidir. Proteinlerin diğer moleküller ile nasıl etkileştiğini ve enzimlerin durumlarındaki değişimler izlenebilir. Bir kristal yapısından, protein-protein etkileşimini ayırt etmek genellikle kolay değildir. Bunları ayırt etmek için bir dizi hesaplama gereklidir ve hata oranı yüksektir [51] – [53].

2.2.8. NMR (Nükleer manyetik rezonans)

NMR spektroskopisi ile protein-protein etkileşiminin analizinde, bağlanma arayüzünün yeri, protein etkileşim tayini için önemli bir özelliktir. NMR spektroskopisinin temeli, güçlü bir manyetik kaynağın yönlendirdiği manyetik olarak aktif çekirdeklerin, kimyasal çevreleri tarafından yönetilen karakteristik frekanslarda elektromanyetik radyasyonu absorbe etmesine dayanmaktadır. Yüksek alanlı spektrometre ve kriyojenik problemler ile tekniğin duyarlılığı artmıştır. Büyük ve biyolojik olarak ilişkili biyomoleküler kompleksleri yüksek doğrulukla karakterize etmek mümkündür. [71] – [73]

2.3. In Siliko Yöntemler

In siliko çalışmalar, bilgisayar veya simülasyonlar ile gerçekleştirilen çalışmaları ifade etmektedir [15]. Adını bilgisayar teknolojilerinde yoğun kullanılan silikona atfen almıştır. In siliko çalışmaları, pahalı laboratuvar deneyleri ve klinik araştırmalara olan ihtiyacı azaltılması için avantaj sağlamaktadır. In siliko teknikler, protein tasarımı, hücre davranışlarının modellenmesi, biyoproses optimizasyonu ve benzeri alanlarda kullanılabilir [57, 58]. Deneysel yaklaşımla etkileşim tespitine destek sağlama ve alternatif oluşturmak için çeşitli in siliko yöntemler geliştirilmiştir. In siliko tespit, yapı, dizilim, gen komşuluğu, gen birleşimi, filogenetik ağaç ve gen ifadesine dayalı yaklaşımları içerir.

2.3.1. Yapı tabanlı yaklaşımlar (Structure based approaches)

Yapıya dayalı yaklaşımda iki protein benzer bir yapıya sahipse etkileşim kurulan proteinler arasında da aynı benzerliğin olacağı kabul edilmektedir. Etkileşim tahmini için ilk olarak proteinlerin yapıları çeşitli yöntemler ile oluşturulmalıdır. Yapıları ve aralarındaki etkileşimleri bilinen proteinler ile bir havuz oluşturulur. Etkileşimi tahmin edilmek istenen hedef proteinin yapısal olarak benzediği proteinler belirlenir. Belirlenen proteinlerin sahip olduğu protein-protein etkileşimleri hedef protein içinde geçerli olduğu kabul edilir. Protein yapısı ve yapılar arası benzerliklerin skorlanması için çeşitli algoritmalar geliştirilmiştir. Yapılan işlem protein yapı benzerliği ile protein – protein etkileşim tahminidir [59] – [61].

2.3.2. Dizilim tabanlı yaklaşımlar (Sequence based approaches)

Protein sekansı verilerindeki hızlı artış, sekans bilgisi kullanılarak tahmin yapılması için gerekli olan verilerin oluşmasını sağlamıştır. Bilinen protein etkileşimleri protein homoloji bilgileri ile birleştirilerek bilinmeyen etkileşimler tahmin edilir. Bu yaklaşım, bir türün içinde bulunan bir etkileşimin diğer türlerdeki etkileşimi anlamak için kullanılabilir kavramına dayanmaktadır. Protein birincil yapıları, ikincil yapıları, homoloji, gen ontolojileri gibi bilgilerin farklı teknikler ile birleştirilmesi ve daha güçlü tahminler yapılması için çeşitli algoritmalar bulunmaktadır [62] – [64].

2.3.3. Gen Komşuluğu (Gene neighbourhood)

Operonlar gibi işlevsel olarak ilişkili proteinler, genomlardaki yakın bölgelerde organize olma eğilimindedirler. Komşuluk ilişkisi birden fazla genom boyunca korunduğunda ilgili genler tarafından kodlanan proteinler arasındaki fonksiyonel bağlantının olma olasılığı daha yüksek olacaktır. Bu yöntemde genlerin komşuluk ilişkileri bilinen protein-protein etkileşim ilişkileri ile birleştirilerek bilinmeyen etkileşimler tahmin edilmeye çalışılır. Komşuluk ilişkilerinin varlığı ve göreceli gen oryantasyonundan bağımsız olduğu deneysel sonuçlarla doğrulanmıştır. Komşuluk ilişkileri, transkripsiyon düzenleyicileri için hedef süreçleri ve düzenleyici özellikleri tahmin etmek için kullanılmaktadır. Gen komşulukları, bakteriyel genomlarda korunduğu için doğrudan tespit edilebilmesine karşın diğer genomlarda zor bir problem olma özelliğini korumaktadır [65, 66].

2.3.4. Gen birleşimi (Gene fusion)

Füzyon gen, çeşitli sebeplerle, iki ayrı gen veya parçalarının birleşiminden oluşan hibrid bir gendir. Bu değişim sonucu üretilen füzyon proteinleri, bazı kanser türlerinin gelişmesine yol açmaktadır. Füzyon genleri ve proteinleri, kanserin tanı ve tedavisinde önemli rol oynamaktadır. Yöntem füzyon proteinlerinin temel özelliği olan etkileşime dayanmaktadır. Bir organizmada tekil olan proteinler başka bir organizmada etkileşim ile protein kompleksleri oluşturabilir. Bu durumun tespit edilmesi, ilgili proteinler için fonksiyonel birlikteliklerinde belirlenmesini sağlayacaktır. Metabolik yollara katılan proteinlerde birleşim daha sık rastlanmaktadır. Bu yöntemde farklı genomlara ait olan düzenleme bilgileri kullanılarak, protein-protein etkileşimleri tahmin edilir. Yöntem sadece alan düzenlemesinin mevcut olduğu füzyon proteinlerine uygulanabilir [67] – [69].

2.3.5. Filogenetik ağaçlar (Phylogenetic trees)

Protein tahmini evrim sürecini hakkında bilgi verir. Bu yöntemde temel olarak bu bilgi kullanılarak proteinler arası etkileşimler tahmin edilir. Yöntemdeki temel kabul, etkileşimli proteinlerin birlikte evrimleşmesi sebebi ile benzer filogenetik ağaçlar oluşturacağıdır. Benzer filogenetik ağaçlar protein-protein etkileşiminin göstergesidir. Araştırılan proteinleri içeren organizmalar için filogenetik ağaçlar ve benzerlik matrisleri oluşturulur. Benzerlikler arasındaki ilişkiler ile proteinler arasındaki etkileşim tahmin edilir. Yüksek benzerlik oluşturulan filogenetik ağaç benzerliğini dolayısıyla proteinlerin etkileşimli olduğunu göstermektedir [70] – [72].

2.3.6. Gen ifadesi (Gene expression)

Bu yöntemde genlerin ifade modelleri ile kodladıkları proteinlerin etkileşimleri arasındaki ilişki kullanılarak protein-protein etkileşim tahmini yapılır. Gen ifadesi, belirli bir genin, farklı deney koşulları altında ve zaman aralıklarında bir hücre, doku veya organizma içinde ifade edildiği seviyenin belirlenmesidir. Büyük ölçekli gen ifadesi ve protein-protein etkileşim verilerinin istatistiksel analizi ile birlikte, ifade edilen genler tarafından kodlanan protein çiftlerinin, birbirleriyle rastgele proteinlerden daha sık etkileştiği gösterilmiştir. Bu kabul ile gen ifade ve protein-protein etkileşim bilgileri birleştirilerek yapılan analizler ile bilinmeyen etkileşimler tespit edilmeye çalışılır. Bu yöntem protein etkileşimini ortaya çıkarmanın dolaylı bir yolu olduğu için tespit edilen protein etkileşimlerinin doğruluğu hakkında çeşitli görüşler bulunmaktadır. Diğer deneysel yöntemlerle birlikte doğrulama amacı ile de kullanılabilir [73] – [75].

2.3.7. Metin Madenciliği (Text Mining)

Metin madenciliğine dayalı hesaplama metodolojileri, proteinler ve protein-protein etkileşimlerinin tahmini için literatür ve biyolojik veri tabanlarını metin madenciliği metotları ile analiz etmektedir [76]. Literatürdeki üstel artış göz önüne alındığında metin

madenciliği tekniklerinin daha fazla önem kazanacağı görülmektedir. Tahmin edilen protein ve etkileşimlerin hatalarının azaltılması ile doğruluğun artırılması üzerine olan çalışmalar yöntemin uygulanabilirliği açısından önem arz etmektedir [77].

3. PPI Veri Tabanları ve Erişim Kaynakları

3.1. PUBMED

Yaşam bilimleri ve biyomedikal konularla ilgili MEDLINE veri tabanına ve özetlere erişim sağlayan ücretsiz bir arama motorudur. Ulusal Sağlık Enstitüleri'ndeki Amerika Birleşik Devletleri Ulusal Tıp Kütüphanesi (NLM) Ulusal Biyoteknoloji Bilgi Merkezi (NCBI) tarafından sağlanan bir hizmettir. Bu veri tabanı, Entrez sisteminin bilgi alma parçası olarak kullanılmaktadır. PubMed, tüketiciler için kolay okunan özetler ve tam teknik raporlar ile klinik etkinlik araştırması incelemelerinde uzmanlaşmıştır. [78] – [80]

3.2. DIP (Database of Interacting Proteins)

Etkileşen proteinlerin veri tabanı (DIP) deneysel olarak belirlenen protein-protein ikili etkileşimlerini içeren bir veri tabanıdır. Etkileşimler hakemli literatürden incelenir ve uzman sorumlular tarafından veri tabanına manuel olarak girilir. Manuel kayıtlara ek olarak, DIP'deki etkileşimler hesaplama değerlendirmesine tabi tutulur. Değerlendirme algoritmaları, bildirilen etkileşimin özelliklerini, DIP etkileşimlerinin en güvenilir özellikleriyle karşılaştırarak biyolojik olarak anlamlı olup olmadıklarını denetler. Değerlendirme sonuçları etkileşimin bilgi sayfasında görüntülenir. Protein-protein etkileşim detaylarının kataloglanmasının ötesinde, DIP protein fonksiyonu ve protein-protein ilişkilerini anlamak, etkileşen proteinlerin ağlarının özelliklerini incelemek, protein-protein etkileşimlerinin kestirimlerini ölçmek ve protein-protein etkileşimlerinin evrimini incelemek için yararlıdır. Etkileşimli verilerin ana kamu sağlayıcılarından oluşan bir grup olan Uluslararası Moleküler Etkileşim Değişim Konsorsiyumunun (IMEx) [81] bir üyesidir. [82] – [84]

3.3. IntAct (An Open Source Molecular Interaction Database)

IntAct protein etkileşimlerinin depolanması, sunulması ve analizi için açık kaynak veri tabanı ve analiz araçları sunmaktadır. Web ara yüzü, protein etkileşimlerinin hem metinsel hem de grafiksel gösterimlerini sunar ve proteinlerinin GO(Gene Ontology) [85] ek açıklamalarına ulaşılabilir. Tüm etkileşimler literatür madenciliği veya doğrudan kullanıcı gönderimlerinden türetilmiştir. [86,87]

3.4. MINT (The Molecular Interaction Database)

Moleküler Etkileşim veri tabanı (MINT), deneysel olarak elde edilen literatürden oluşturulan PPI verilerinin veri tabanıdır. Etkileşimlere ilave olarak, her etkileşim için kanıt ağırlığı bilgisi de sunulur. Web ara yüzü, kullanıcılara etkileşimli verileri arama, görselleştirme ve indirme olanağı sağlar. MINT, Uluslararası Moleküler Değişim Konsorsiyumunun (IMEx) [81] üyelerinden biridir. İyileştirme ve veri alışverişi için Proteomik Standart Girişimi (PSI-MI) standartlarının Moleküler Etkileşim Ontolojisini benimsemektedir [88, 89].

3.5. MIPS (Munich Information Center for Protein Sequences)

Münih Bilgi Merkezi Protein Sekansları (MIPS), Neuherberg, Almanya'daki Biyoenformatik Enstitüsü'nde (IBI) bulunan bir araştırma merkezidir. Genom yönelimli biyoinformatik üzerine odaklanılmıştır. Bakteri, mantar ve bitki genomlarının sistematik karşılaştırmalı analizleri de sunulmaktadır. Yüksek ilişki düzeyine sahip seçilmiş veri kümeleri manuel kontrollere tabi tutulur, yüksek verimli veri kümeleri ise otomatik araçlarla etiketlenir. Alternatif olarak hesaplama araçlarının kullanılması ve ikincil bilgiler veri tabanlarının oluşturulması, bilginin daha verimli bir şekilde kullanılmasını sağlamaktadır. [90, 91]

3.6. BIOGRID (The Biological General Repository for Interaction Datasets)

Yüksek çıktılı veri kümelerinden ve birincil literatürden alınan ve kapsamlı iyileştirme çalışmaları ile derlenen etkileşimleri içerir. Tüm veriler arama dizini aracılığıyla ücretsiz olarak temin edilmekte ve standart formatlarda indirilebilmektedir. Bünyesinde protein-protein etkileşimleri, genetik etkileşimler, kimyasal etkileşimler ve post-translasyonel modifikasyon bilgileri bulunur. Kanada ve Amerika Birleşik Devletleri'nde barındırılmaktadır. BBSRC, NIH ve CIHR tarafından finanse edilmektedir. Uluslararası Moleküler Değişim Konsorsiyumunun (IMEx) [81] üyesidir. İyi bir ara birime sahiptir. Kullanıcılar ilgi alanlarına giren proteinleri veya yayınları araştırıp, ek açıklamaların yanı sıra, rapor edilen, literatürde bulunan ve büyük ölçekli kontrol çalışmaları ile derlenen verilere kolaylıkla ulaşabilirler. Dinamik etkileşim ağ görüntüleyicisi, tüm genetik ve protein etkileşim verilerinin yanı sıra biyoaktif bileşikler ve bunların yerleşik hedefleri için kolay gezinme ve filtreleme sağlar [92] – [94].

3.7. STRING (Search Tool for the Retrieval of Interacting Genes)

Protein etkileşimlerinin hem fiziksel hem de fonksiyonel ilişkilerini içerir. Deneysel veriler, hesaplamalı tahmin yöntemleri ve genel metin koleksiyonları dâhil olmak üzere çok sayıda kaynaktan bilgi bulunmaktadır. Veriler ağırlıklandırılmış ve tüm protein etkileşimleri için bir güven skoru ile bütünleştirilmiştir. GO [85], Pfam [95] ve KEGG [96] gibi fonksiyonel sınıflandırma sistemleri kullanarak protein listelerine fonksiyon bilgileri eklenmektedir. Novo Nordisk Foundation Center for Protein Research (CPR), European Molecular Biology Laboratory (EMBL), University of Copenhagen (KU), Swiss Institute of Bioinformatics (SIB), Technische Universität Dresden (TUD) ve University of Zurich (UZH) gibi akademik kurumların bir konsorsiyumu tarafından geliştirilmiştir. Protein etkileşimlerinin hızlı ve kolay bir şekilde sunulduğu web ara birimi bulunmaktadır [97] – [99].

3.8. HPID (Human Protein Interaction Database)

HPID, insan proteomasındaki her bir protein için domain mimarisi, translasyon sonrası değişiklikler, etkileşim ağları ve hastalık birliği ile ilgili bilgileri görsel olarak göstermek ve entegre etmek için merkezi bir platformdur. Mevcut yapısal ve deneysel verilerden hesaplanan insan protein etkileşimlerine, proteinler arasındaki tahmin edilmiş etkileşimlere ve kullanıcılardan gelen yeni insan protein etkileşimi verilerine bir depo sağlamak üzere tasarlanmıştır. Tüm bilgiler, yayınlanmış verileri okuyan, yorumlayan ve analiz eden uzman biyologlar tarafından incelenmektedir. Web tabanlı ara yüzü sayesinde kullanıcılar protein etkileşimi ağlarını görselleştirip analiz edebilmektedir [100] – [102].

3.9. BIND (The Biomolecular Interaction Network Database)

Proteomik kaynak oluşturma, diğer kaynaklardan veri madenciliğini sağlayacak bir platform oluşturma ve karmaşık moleküler etkileşimlerin görselleştirmelerini sunabilen bir platform oluşturma amaçları ile tasarlanmıştır. Biomoleküler etkileşim, kompleks ve yolak bilgilerini içerir. Kayıtları sorgulamak ve görüntülemek için web tabanlı bir arayüz bulunmaktadır. İlgilenilen bölgelere odaklanmaya yardımcı olacak bir etkileşim ağı kümeleme aracına sahiptir [103, 104].

3.10. HINT (High-Quality Interactomes)

Farklı organizmalar için yüksek kaliteli protein-protein etkileşimlerinden oluşan bir veri tabanıdır. Veriler farklı kaynaklardan derlenir ve hatalı veya düşük kaliteli etkileşimleri gidermek için hem sistematik hem de manuel olarak filtrelenir. Projenin amacı, hücresel ağların global özelliklerini analiz etmenin yanı sıra, belirli proteinler veya yolaklar hakkında spesifik hipotezler üretmek için kullanılabilir veri kümesi oluşturmaktır [105].

3.11. HitPredict (A Database of Quality Assessed Protein-Protein Interactions)

Deneysel olarak tanımlanmış fiziksel protein-protein etkileşimlerinin güven puanlarıyla birlikte yayınlandığı bir kaynaktır. Veri entegrasyonu sırasında karşılaşılabilecek sorunları çözmek için manuel doğrulama kullanılmıştır. Etkileşime giren proteinlerin sekansı, yapısı ve işlevsel ek açıklamalarından elde edilen kanıtlar kullanılarak hesaplanan güvenilirlik puanına dayalı olarak etkileşimlere bir güven düzeyi atanır. Protein aramak ve etkileşimlerini görselleştirmek için web ara yüzüne sahiptir, veriler çevrimdışı kullanım için indirilebilir [106] – [108].

3.12. GeneMANIA

GeneMANIA genleri analiz etmek ve gen fonksiyonlarını önceliklendirmek için geliştirilmiş web tabanlı bir kaynaktır. İşlevsel olarak benzer genler tespit edilebilir ve gen fonksiyonları tahmin edilebilir. Veriler BioGRID, gibi farklı kaynaklardan derlenir. Cytoscape [109] yazılımı ile entegre çalışabilmektedir [110].

3.13. CORUM

CORUM memeli organizmalarından deneysel olarak karakterize edilmiş protein komplekslerine ait veriler içerir. Veriler elle küratörlüğe tabi tutulmaktadır. Birimler arasında bilinen etkileşimleri gösteren Cytoscape [109] tabanlı grafiksel bir araç içermektedir. İleri çalışmalar için tüm veri setlerini farklı formatlarda indirilebilir [111].

3.14. iRefIndex

Proteinlere ait etkileşim verileri farklı veri tabanlarında bulunabilmektedir. iRefIndex bu tür verileri aramayı kolaylaştırmak için tasarlanmıştır. iRefIndex veri kaynağı olarak BIND, BioGRID, CORUM, DIP, HPRD, IntAct, MINT gibi birçok birincil etkileşim veri tabanında mevcut olan protein etkileşimlerini kullanır. Her bir etkileşim ve etkileşime katılan proteinler için benzersiz bir anahtar üretilmektedir. Anahtar üretim esnasında hatalı biçimlendirilmiş, kullanımdan kaldırılmış, belirsiz veya sorunlu referansların belirlenebilmesi için bir puanlama sunulmaktadır. İleri çalışmalar için tüm veri setlerini indirilebilir [112].

4. PPI Veri Tabanları Arasındaki Farklar

Veri tabanları arasında kullanım kolaylığı, veri görselleştirme ve kapsayıcılık gibi konularda farklılıklar bulunmaktadır. Deneysel ve hesaplamalı etkileşimleri sunan STRING veri tabanı mükemmel grafikler sağlar, kullanımı kolaydır, başarılı grafiksel etkileşim haritasına sahiptir, kanıtlar ayrıntılı olarak açıklanır. Literatür küratörlüğünden veya doğrudan kullanıcı sunumlarından türetilmiş tablolardan oluşan IntAct, sonuçların sunumunda anlaşılması zor kodlamalar kullanmaktadır. Kolay tür seçimi yoktur ve grafiksel arabirime sahip değildir. Literatürden ve yüksek verimli veri setlerinden elde edilen verileri paylaşan BIOGRID veri tabanı mayalar için zengin veri kaynağına sahiptir. İyi bir ara yüze sahiptir, genellikle memeli genleri için diğer veri tabanlarına kıyasla daha az veri içerir. Referans ve açıklamalar iyi bir şekilde sunulmaktadır. Kullanıcı arabirimi ve sonuçların kullanım kolaylığı konularında yaşanan farklılaşmaya veri tabanlarının içeriklerinde de karşılaşılmaktadır. Tablo 1 de yer alan bilgilerde görüldüğü üzere, her veri tabanında yer alan canlı tür sayısı, protein sayısı, etkileşim sayısı farklı olmakla beraber aynı kaynaktan türetilen bilgi sayısında farklılıklarda gözlenmektedir. En az iki veri tabanı tarafından paylaşılan 14.899 yayının, 5.782'si (yüzde 39) farklı veri tabanlarında farklı sayıda etkileşimle bildirilmiştir [113]. Veri tabanlarında bulunan düşük oranlı farklılıklar, bu veriyi kullanan algoritmaların sonuçlarını verideki farklılıktan daha fazla oranda etkilemektedir [114, 115]. Veri tabanlarının aynı yayından aldığı bilgilerde de önemli ölçüde farklılıklar olabilmektedir. Bu durumun tanımlayıcı eşlemesi kaynaklı olduğu düşünülmektedir. Farklı veri tabanlarının aynı yayından küratörlük ile elde ettiği bilgide, etkileşimlerin %42 ve proteinler %62 oranında aynı olduğu görülmüştür [116]. Yapılan bir diğer deneysel çalışmada Y2H kullanılarak insan proteinleri arasında 2,754 etkileşimi tespit edilip veri tabanları ile kıyaslanmıştır. Yapılan

kıyaslamaya göre, HPRD 2,371 etkileşim, IntAct 2,671 etkileşim ve MINT 2,463 etkileşim bildirildiği tespit edilmiştir [117]. Elde edilen sonuçların aynı olmaması, farklı güven kümesi veya eşik kullanımı gibi deneysel verinin farklı yorumlanmasını sağlayan parametrelerin aynı şekilde kullanılmamasına bağlanmaktadır.

Tablo 1: Veri Tabanlarında Yer Alan Protein, Etkileşim ve Organizma Sayıları

Veri Tabanı	Protein	Etkileşim	Organizma	Ara Yüz
BIOGRID	73.159	1.706.694	69	Var
MINT	26.100	130.733	645	Var
BIND	23.643	1.673.783	80	Var
DIP	21.167	53.431	134	Var
IntAct	98.289	720.711	131	Var
HPRD	30.047	41.327	1	Var
HINT	-	387.615	12	Var
MIPS	982	1.859	3	Yok
STRING	9.643.763	1.380.838.440	2.031	Var
GeneMania	163.599	597.392.998	9	Var
CORUM	4.274	257.941	3	Var
iRefIndex	77.827	1.076.405	13	Yok

*2020.04.20 tarihinde derlenmiştir

4. Sonuçlar

Gelişmelere bağlı olarak, protein-protein etkileşimlerinin saptanması için birçok yöntem geliştirilmiştir ve geliştirilmeye devam etmektedir. Yöntemlerde kullanılan prensipler ve performansları çeşitlilik göstermektedir. Bazı durumlarda yöntemler birbirlerini tamamlamakta veya birlikte kullanılmaktadır. Eski yöntemlerin performanslarının artırılmasının yanında yeni yöntemler de geliştirilmiştir. Yeni yöntemler yüksek miktarda veri üretmekte ve bu verilerin istatistiksel olarak işlenmesi gerekmektedir. Deneylerin yürütülme şekli ve yorumlanması da yeni yöntemler için önemli kriterler arasında yer almaktadır. Yeni veya eski kullanılabilir tüm yöntemler yüzde yüz doğrulukla etkileşimleri tahmin edememektedir. Birçoğu yüksek hata oranlarına sahiptir. Düşük hata oranlı yöntemler genellikle uzun zaman gereksinimi ve düşük verimli yöntemlerdir. Hesaplamalı yöntemler olası etkileşim kümelerini tahmin etmektedir. Tahmin edilen bu etkileşimler yapılması planlanan laboratuvar deneyleri için yol gösterici olarak kullanılmaktadır. Farklı kaynaklardan elde edilen bilgiler birlikte kullanılarak protein-protein etkileşimlerinin ve PPI ağlarının doğruluğu artırılmaya çalışılmaktadır. Teknolojik gelişmeler deneysel süreçleri hızlandırmakla beraber deneysel doğruluğun sorgulanmasını zorunlu hale getirmiştir. Hesaplamalı yöntemler olası deneylerin sayısını azaltarak hedef deneylerin belirlenmesinde önemli rol üstlenmiş duruma gelmiştir. Her grup tespit yöntemi içinde doğruluk analizi, tespiti ve iyileştirilmesi üzerine çalışmalar oldukça azdır. Eksikliği fark edilen ve iyileştirmeye açık olan bu alanlarda yapılacak çalışmaların sayısında yaşanacak artış ve hesaplamalı yöntemlerin tahmin doğruluklarının iyileştirilmesi ile yöntemlerin sonuç üretme doğruluğu ve sonuç üretme hızlarında önemli gelişmeler yaşanacaktır.

Kaynakça

1. Ingber, D. E. (2000). The origin of cellular life. *Bioessays*, 22(12), 1160-1170.
2. Lodish, H., Berk, A., Zipursky, S. L., Matsudaira, P., Baltimore, D., & Darnell, J. (2000). *Molecular cell biology* 4th edition. National Center for Biotechnology Information, Bookshelf.
3. Lu, L., Arakaki, A. K., Lu, H., & Skolnick, J. (2003). Multimeric threading-based prediction of protein-protein interactions on a genomic scale: Application to the *Saccharomyces cerevisiae* proteome. *Genome Research*, 13(6a), 1146-1154.
4. Ghaemmaghami, S., Huh, W. K., Bower, K., Howson, R. W., Belle, A., Dephoure, N., ... & Weissman, J. S. (2003). Global analysis of protein expression in yeast. *Nature*, 425(6959), 737-741.
5. Braun, P., & Gingras, A. C. (2012). History of protein-protein interactions: From egg-white to complex networks. *Proteomics*, 12(10), 1478-1498.

6. Yan, C., Wu, F., Jernigan, R. L., Dobbs, D., & Honavar, V. (2008). Characterization of protein–protein interfaces. *The protein journal*, 27(1), 59-70.
7. Nooren, I. M., & Thornton, J. M. (2003). Diversity of protein–protein interactions. *The EMBO journal*, 22(14), 3486-3492.
8. Zhang, A. (2009). *Protein interaction networks: computational analysis*. Cambridge University Press.
9. Iqbal, M. (2018). Introductory Chapter: Protein-Protein Interactions and Assays. *Protein-Protein Interaction Assays*, 1.
10. Klein, S. (2010). The use of biorelevant dissolution media to forecast the in vivo performance of a drug. *The AAPS journal*, 12(3), 397-406.
11. Yanagida, M. (2002). Functional proteomics; current achievements. *Journal of Chromatography B*, 771(1-2), 89-106.
12. Berggård, T., Linse, S., & James, P. (2007). Methods for the detection and analysis of protein–protein interactions. *Proteomics*, 7(16), 2833-2842.
13. Von Mering, C., Krause, R., Snel, B., Cornell, M., Oliver, S. G., Fields, S., & Bork, P. (2002). Comparative assessment of large-scale data sets of protein–protein interactions. *Nature*, 417(6887), 399-403.
14. Rishton, G. M. (1997). Reactive compounds and in vitro false positives in HTS. *Drug discovery today*, 2(9), 382-384.
15. Vivona, S., Gardy, J. L., Ramachandran, S., Brinkman, F. S., Raghava, G. P. S., Flower, D. R., & Filippini, F. (2008). Computer-aided biotechnology: from immuno-informatics to reverse vaccinology. *Trends in biotechnology*, 26(4), 190-200.
16. Atanasov, A. G., Waltenberger, B., Pferschy-Wenzig, E. M., Linder, T., Wawrosch, C., Uhrin, P., ... & Rollinger, J. M. (2015). Discovery and resupply of pharmacologically active plant-derived natural products: A review. *Biotechnology advances*, 33(8), 1582-1614.
17. Lipinski, C., & Hopkins, A. (2004). Navigating chemical space for biology and medicine. *Nature*, 432(7019), 855-861.
18. Smith, H. (2001). Discovery of the anthrax toxin: the beginning of studies of virulence determinants regulated in vivo. *International journal of medical microbiology*, 291(6-7), 411-417.
19. Relman, D. A. (1998). Detection and identification of previously unrecognized microbial pathogens. *Emerging infectious diseases*, 4(3), 382.
20. Uetz, P., Giot, L., Cagney, G., Mansfield, T. A., Judson, R. S., Knight, J. R., ... & Qureshi-Emili, A. (2000). A comprehensive analysis of protein–protein interactions in *Saccharomyces cerevisiae*. *Nature*, 403(6770), 623-627.
21. Ito, T., Chiba, T., Ozawa, R., Yoshida, M., Hattori, M., & Sakaki, Y. (2001). A comprehensive two-hybrid analysis to explore the yeast protein interactome. *Proceedings of the National Academy of Sciences*, 98(8), 4569-4574.
22. Oliver, S. (2000). Guilt-by-association goes global. *Nature*, 403(6770), 601-602.
23. Rutherford, S. L. (2000). From genotype to phenotype: buffering mechanisms and the storage of genetic information. *Bioessays*, 22(12), 1095-1105.
24. Hartman, J. L., Garvik, B., & Hartwell, L. (2001). Principles for the buffering of genetic variation. *Science*, 291(5506), 1001-1004.
25. Bender, A. L. A. N., & Pringle, J. R. (1991). Use of a screen for synthetic lethal and multicopy suppressor mutants to identify two new genes involved in morphogenesis in *Saccharomyces cerevisiae*. *Molecular and cellular biology*, 11(3), 1295-1305.
26. Arnau, J., Lauritzen, C., Petersen, G. E., & Pedersen, J. (2006). Current strategies for the use of affinity tags and tag removal for the purification of recombinant proteins. *Protein expression and purification*, 48(1), 1-13.
27. Waugh, D. S. (2005). Making the most of affinity tags. *Trends in biotechnology*, 23(6), 316-320.
28. Von Mering, C., Krause, R., Snel, B., Cornell, M., Oliver, S. G., Fields, S., & Bork, P. (2002). Comparative assessment of large-scale data sets of protein–protein interactions. *Nature*, 417(6887), 399-403.
29. Sinz, A. (2014). The advancement of chemical cross-linking and mass spectrometry for structural proteomics: from single proteins to protein interaction networks. *Expert review of proteomics*, 11(6), 733-743.
30. Rappsilber, J. (2011). The beginning of a beautiful friendship: cross-linking/mass spectrometry and modelling of proteins and multi-protein complexes. *Journal of structural biology*, 173(3), 530-540.
31. Bruce, J. E. (2012). In vivo protein complex topologies: Sights through a cross-linking lens. *Proteomics*, 12(10), 1565-1575.
32. Serpa, J. J., Parker, C. E., Petrotchenko, E. V., Han, J., Pan, J., & Borchers, C. H. (2012). Mass spectrometry-based structural proteomics. *European Journal of Mass Spectrometry*, 18(2), 251-267.
33. Kuramochi, J., & Sakakibara, Y. (2005, June). Intensive in vitro experiments of implementing and executing finite automata in test tube. In *International Workshop on DNA-Based Computers* (pp. 193-202). Springer, Berlin, Heidelberg.
34. Dehghan, B. (2016). Synergistic Modeling of in-vitro and in-vivo data via Stochastic Kriging with Qualitative Factors (SKQ).
35. Rigaut, G., Shevchenko, A., Rutz, B., Wilm, M., Mann, M., & Séraphin, B. (1999). A generic protein purification method for protein complex characterization and proteome exploration. *Nature biotechnology*, 17(10), 1030-1032.
36. Denison, C., Rudner, A. D., Gerber, S. A., Bakalarski, C. E., Moazed, D., & Gygi, S. P. (2005). A proteomic strategy for gaining insights into protein sumoylation in yeast. *Molecular & Cellular Proteomics*, 4(3), 246-254.
37. Graumann, J., Dunipace, L. A., Seol, J. H., McDonald, W. H., Yates, J. R., Wold, B. J., & Deshaies, R. J. (2004). Applicability of tandem affinity purification MudPIT to pathway proteomics in yeast. *Molecular & Cellular Proteomics*, 3(3), 226-237.
38. Urh, M., Simpson, D., & Zhao, K. (2009). Affinity chromatography: general methods. In *Methods in enzymology* (Vol. 463, pp. 417-438). Academic Press.
39. Phizicky, E. M., & Fields, S. (1995). Protein-protein interactions: methods for detection and analysis. *Microbiol. Mol. Biol. Rev.*, 59(1), 94-123.
40. Golemis, E., & Adams, P. D. (Eds.). (2002). *Protein-protein interactions: a molecular cloning manual* (p. 3). Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.

41. Ohh, M., Yauch, R. L., Lonergan, K. M., Whaley, J. M., Stemmer-Rachamimov, A. O., Louis, D. N., ... & Iliopoulos, O. (1998). The von Hippel-Lindau tumor suppressor protein is required for proper assembly of an extracellular fibronectin matrix. *Molecular cell*, 1(7), 959-968.
42. MacBeath, G., & Schreiber, S. L. (2000). Printing proteins as microarrays for high-throughput function determination. *Science*, 289(5485), 1760-1763.
43. Brown, P. O., & Botstein, D. (1999). Exploring the new world of the genome with DNA microarrays. *Nature genetics*, 21(1), 33-37.
44. Pinkel, D., Seagraves, R., Sudar, D., Clark, S., Poole, I., Kowbel, D., ... & Dairkee, S. H. (1998). High resolution analysis of DNA copy number variation using comparative genomic hybridization to microarrays. *Nature genetics*, 20(2), 207-211.
45. Michnick, S. W., Ear, P. H., Landry, C., Malleshiah, M. K., & Messier, V. (2011). Protein-fragment complementation assays for large-scale analysis, functional dissection and dynamic studies of protein-protein interactions in living cells. In *Signal Transduction Protocols* (pp. 395-425). Humana Press, Totowa, NJ.
46. Moresco, J. J., Carvalho, P. C., & Yates III, J. R. (2010). Identifying components of protein complexes in *C. elegans* using co-immunoprecipitation and mass spectrometry. *Journal of proteomics*, 73(11), 2198-2204.
47. Morell, M., Espargaró, A., Avilés, F. X., & Ventura, S. (2007). Detection of transient protein-protein interactions by bimolecular fluorescence complementation: The Abl-SH3 case. *Proteomics*, 7(7), 1023-1036.
48. Scott, J. K., & Smith, G. P. (1990). Searching for peptide ligands with an epitope library. *Science*, 249(4967), 386-390.
49. Bass, S., Greene, R., & Wells, J. A. (1990). Hormone phage: an enrichment method for variant proteins with altered binding properties. *Proteins: Structure, Function, and Bioinformatics*, 8(4), 309-314.
50. Sidhu, S. S., Lowman, H. B., Cunningham, B. C., & Wells, J. A. (2000). [21] Phage display for selection of novel binding peptides. In *Methods in enzymology* (Vol. 328, pp. 333-IN5). Academic Press.
51. Tong, A. H. Y., Evangelista, M., Parsons, A. B., Xu, H., Bader, G. D., Pagé, N., ... & Andrews, B. (2001). Systematic genetic analysis with ordered arrays of yeast deletion mutants. *Science*, 294(5550), 2364-2368.
52. Kobe, B., Guncar, G., Buchholz, R., Huber, T., Maco, B., Cowieson, N., ... & Forwood, J. K. (2008). Crystallography and protein-protein interactions: biological interfaces and crystal contacts.
53. Urakubo, Y., Ikura, T., & Ito, N. (2008). Crystal structural analysis of protein-protein interactions drastically destabilized by a single mutation. *Protein Science*, 17(6), 1055-1065.
54. Scott, D. E., Bayly, A. R., Abell, C., & Skidmore, J. (2016). Small molecules, big targets: drug discovery faces the protein-protein interaction challenge. *Nature Reviews Drug Discovery*, 15(8), 533.
55. Barbieri, L., Luchinat, E., & Banci, L. (2015). Protein interaction patterns in different cellular environments are revealed by in-cell NMR. *Scientific reports*, 5, 14456.
56. Gao, G., Williams, J. G., & Campbell, S. L. (2004). Protein-protein interaction analysis by nuclear magnetic resonance spectroscopy. In *Protein-Protein Interactions* (pp. 79-91). Humana Press.
57. Lescot, M., Déhais, P., Thijs, G., Marchal, K., Moreau, Y., Van de Peer, Y., ... & Rombauts, S. (2002). PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. *Nucleic acids research*, 30(1), 325-327.
58. Wishart, D. S., Knox, C., Guo, A. C., Shrivastava, S., Hassanali, M., Stothard, P., ... & Woolsey, J. (2006). DrugBank: a comprehensive resource for in silico drug discovery and exploration. *Nucleic acids research*, 34(suppl_1), D668-D672.
59. Shin, W. H., Christoffer, C. W., & Kihara, D. (2017). In silico structure-based approaches to discover protein-protein interaction-targeting drugs. *Methods*, 131, 22-32.
60. Shortridge, M. D., & Varani, G. (2015). Structure based approaches for targeting non-coding RNAs with small molecules. *Current opinion in structural biology*, 30, 79-88.
61. Blom, N., Gammeltoft, S., & Brunak, S. (1999). Sequence and structure-based prediction of eukaryotic protein phosphorylation sites. *Journal of molecular biology*, 294(5), 1351-1362.
62. Chen, X. W., & Jeong, J. C. (2009). Sequence-based prediction of protein interaction sites with an integrative method. *Bioinformatics*, 25(5), 585-591.
63. Huang, Y. A., You, Z. H., Chen, X., Chan, K., & Luo, X. (2016). Sequence-based prediction of protein-protein interactions using weighted sparse representation model combined with global encoding. *BMC bioinformatics*, 17(1), 184.
64. Sun, T., Zhou, B., Lai, L., & Pei, J. (2017). Sequence-based prediction of protein protein interaction using a deep-learning algorithm. *BMC bioinformatics*, 18(1), 277.
65. De, S., & Babu, M. M. (2010). Genomic neighbourhood and the regulation of gene expression. *Current opinion in cell biology*, 22(3), 326-333.
66. Oliver, B., Parisi, M., & Clark, D. (2002). Gene expression neighborhoods. *Journal of biology*, 1(1), 4.
67. Enright, A. J., Iliopoulos, I., Kyripides, N. C., & Ouzounis, C. A. (1999). Protein interaction maps for complete genomes based on gene fusion events. *Nature*, 402(6757), 86-90.
68. Thieme, S., & Groth, P. (2013). Genome fusion detection: a novel method to detect fusion genes from SNP-array data. *Bioinformatics*, 29(6), 671-677.
69. Latysheva, N. S., Oates, M. E., Maddox, L., Flock, T., Gough, J., Buljan, M., ... & Babu, M. M. (2016). Molecular principles of gene fusion mediated rewiring of protein interaction networks in cancer. *Molecular cell*, 63(4), 579-592.
70. Pazos, F., & Valencia, A. (2001). Similarity of phylogenetic trees as indicator of protein-protein interaction. *Protein engineering*, 14(9), 609-614.
71. Erten, S., Li, X., Bebek, G., Li, J., & Koyutürk, M. (2009). Phylogenetic analysis of modularity in protein interaction networks. *BMC bioinformatics*, 10(1), 333..

72. Pazos, F., Juan, D., Izarzugaza, J. M., Leon, E., & Valencia, A. (2008). Prediction of protein interaction based on similarity of phylogenetic trees. In *Functional Proteomics* (pp. 523-535). Humana Press.
73. Grigoriev, A. (2001). A relationship between gene expression and protein interactions on the proteome scale: analysis of the bacteriophage T7 and the yeast *Saccharomyces cerevisiae*. *Nucleic acids research*, 29(17), 3513-3519.
74. Wu, C., Zhu, J., & Zhang, X. (2012). Integrating gene expression and protein-protein interaction network to prioritize cancer-associated genes. *BMC bioinformatics*, 13(1), 182.
75. Chen, R., Zhang, Z., Xue, Z., Wang, L., Fu, M., Lu, Y., ... & Fan, Z. (2015). Protein-protein interaction network of gene expression in the hydrocortisone-treated keloid. *International journal of dermatology*, 54(5), 549-554.
76. Papanikolaou, N., Pavlopoulos, G. A., Theodosiou, T., & Iliopoulos, I. (2015). Protein-protein interaction predictions using text mining methods. *Methods*, 74, 47-53.
77. Badal, V. D., Kundrotas, P. J., & Vakser, I. A. (2018). Natural language processing in text mining for structural modeling of protein complexes. *BMC bioinformatics*, 19(1), 84.
78. Fleischer Jr, A. B. (2016). Increasing Incidence within PubMed of the Use of the Misspelling. *Acta dermato-venereologica*, 96(6), 826-827.
79. Garofalo, R., & Schilling, J. L. (2017). Transgender Health Accepted for Indexing in PubMed Central and Inclusion in PubMed.
80. Reyes-Aldasoro, C. C. (2017). The proportion of cancer-related entries in PubMed has increased considerably; is cancer truly "The Emperor of All Maladies"? *PloS one*, 12(3).
81. Orchard, S., Kerrien, S., Abbani, S., Aranda, B., Bhate, J., Bidwell, S., ... & Chatr-Aryamontri, A. (2012). Protein interaction data curation: the International Molecular Exchange (IMEx) consortium. *Nature methods*, 9(4), 345-350.
82. Xenarios, I., Rice, D. W., Salwinski, L., Baron, M. K., Marcotte, E. M., & Eisenberg, D. (2000). DIP: the database of interacting proteins. *Nucleic acids research*, 28(1), 289-291.
83. Xenarios, I., Salwinski, L., Duan, X. J., Higney, P., Kim, S. M., & Eisenberg, D. (2002). DIP, the Database of Interacting Proteins: a research tool for studying cellular networks of protein interactions. *Nucleic acids research*, 30(1), 303-305.
84. Salwinski, L., Miller, C. S., Smith, A. J., Pettit, F. K., Bowie, J. U., & Eisenberg, D. (2004). The database of interacting proteins: 2004 update. *Nucleic acids research*, 32(suppl_1), D449-D451.
85. Gene Ontology Consortium. (2006). The gene ontology (GO) project in 2006. *Nucleic acids research*, 34(suppl_1), D322-D326.
86. Hermjakob, H., Montecchi-Palazzi, L., Lewington, C., Mudali, S., Kerrien, S., Orchard, S., ... & Margalit, H. (2004). IntAct: an open source molecular interaction database. *Nucleic acids research*, 32(suppl_1), D452-D455.
87. Kerrien, S., Aranda, B., Breuza, L., Bridge, A., Broackes-Carter, F., Chen, C., ... & Jandrasits, C. (2012). The IntAct molecular interaction database in 2012. *Nucleic acids research*, 40(D1), D841-D846.
88. Licata, L., Briganti, L., Peluso, D., Perfetto, L., Iannuccelli, M., Galeota, E., ... & Castagnoli, L. (2012). MINT, the molecular interaction database: 2012 update. *Nucleic acids research*, 40(D1), D857-D861.
89. Chatr-Aryamontri, A., Ceol, A., Palazzi, L. M., Nardelli, G., Schneider, M. V., Castagnoli, L., & Cesareni, G. (2007). MINT: the Molecular INTeraction database. *Nucleic acids research*, 35(suppl_1), D572-D574.
90. Mewes, H. W., Ruepp, A., Theis, F., Rattei, T., Walter, M., Frishman, D., ... & Antonov, A. (2011). MIPS: curated databases and comprehensive secondary data resources in 2010. *Nucleic acids research*, 39(suppl_1), D220-D224.
91. Mewes, H. W., Frishman, D., Gruber, C., Geier, B., Haase, D., Kaps, A., ... & Stocker, S. (2000). MIPS: a database for genomes and protein sequences. *Nucleic acids research*, 28(1), 37-40.
92. Chatr-Aryamontri, A., Oughtred, R., Boucher, L., Rust, J., Chang, C., Kolas, N. K., ... & Stark, C. (2017). The BioGRID interaction database: 2017 update. *Nucleic acids research*, 45(D1), D369-D379.
93. Stark, C., Breitkreutz, B. J., Reguly, T., Boucher, L., Breitkreutz, A., & Tyers, M. (2006). BioGRID: a general repository for interaction datasets. *Nucleic acids research*, 34(suppl_1), D535-D539.
94. Winter, A. G., Wildenhain, J., & Tyers, M. (2011). BioGRID REST Service, BiogridPlugin2 and BioGRID WebGraph: new tools for access to interaction data at BioGRID. *Bioinformatics*, 27(7), 1043-1044.
95. Finn, R. D., Bateman, A., Clements, J., Coggill, P., Eberhardt, R. Y., Eddy, S. R., ... & Sonnhammer, E. L. (2014). Pfam: the protein families database. *Nucleic acids research*, 42(D1), D222-D230.
96. Ogata, H., Goto, S., Sato, K., Fujibuchi, W., Bono, H., & Kanehisa, M. (1999). KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic acids research*, 27(1), 29-34.
97. Szklarczyk, D., Morris, J. H., Cook, H., Kuhn, M., Wyder, S., Simonovic, M., ... & Jensen, L. J. (2016). The STRING database in 2017: quality-controlled protein-protein association networks, made broadly accessible. *Nucleic acids research*, gkw937.
98. Szklarczyk, D., Franceschini, A., Wyder, S., Forslund, K., Heller, D., Huerta-Cepas, J., ... & Kuhn, M. (2015). STRING v10: protein-protein interaction networks, integrated over the tree of life. *Nucleic acids research*, 43(D1), D447-D452.
99. Mering, C. V., Huynen, M., Jaeggi, D., Schmidt, S., Bork, P., & Snel, B. (2003). STRING: a database of predicted functional associations between proteins. *Nucleic acids research*, 31(1), 258-261.
100. Han, K., Park, B., Kim, H., Hong, J., & Park, J. (2004). HPID: the human protein interaction database. *Bioinformatics*, 20(15), 2466-2470.
101. Peri, S., Navarro, J. D., Amanchy, R., Kristiansen, T. Z., Jonnalagadda, C. K., Surendranath, V., ... & Ibarrola, N. (2003). Development of human protein reference database as an initial platform for approaching systems biology in humans. *Genome research*, 13(10), 2363-2371.
102. Keshava Prasad, T. S., Goel, R., Kandasamy, K., Keerthikumar, S., Kumar, S., Mathivanan, S., ... & Balakrishnan, L. (2009). Human protein reference database—2009 update. *Nucleic acids research*, 37(suppl_1), D767-D772.
103. Bader, G. D., Betel, D., & Hogue, C. W. (2003). BIND: the biomolecular interaction network database. *Nucleic acids research*, 31(1), 248-250.

104. Bader, G. D., Donaldson, I., Wolting, C., Ouellette, B. F., Pawson, T., & Hogue, C. W. (2001). BIND—the biomolecular interaction network database. *Nucleic acids research*, 29(1), 242-245.
105. Das, J., & Yu, H. (2012). HINT: High-quality protein interactomes and their applications in understanding human disease. *BMC systems biology*, 6(1), 92.
106. López, Y., Nakai, K., & Patil, A. (2015). HitPredict version 4: comprehensive reliability scoring of physical protein–protein interactions from more than 100 species. *Database*, 2015.
107. Patil, A., Nakai, K., & Nakamura, H. (2011). HitPredict: a database of quality assessed protein–protein interactions in nine species. *Nucleic acids research*, 39(suppl_1), D744-D749.
108. Patil, A., & Nakamura, H. (2005). Filtering high-throughput protein-protein interaction data using a combination of genomic features. *BMC bioinformatics*, 6(1), 100.
109. Smoot, M. E., Ono, K., Ruschinski, J., Wang, P. L., & Ideker, T. (2011). Cytoscape 2.8: new features for data integration and network visualization. *Bioinformatics*, 27(3), 431-432.
110. Franz, M., Rodriguez, H., Lopes, C., Zuberi, K., Montojo, J., Bader, G. D., & Morris, Q. (2018). GeneMANIA update 2018. *Nucleic acids research*, 46(W1), W60-W64.
111. Giurgiu, M., Reinhard, J., Brauner, B., Dunger-Kaltenbach, I., Fobo, G., Frishman, G., ... & Ruepp, A. (2019). CORUM: the comprehensive resource of mammalian protein complexes—2019. *Nucleic acids research*, 47(D1), D559-D563.
112. Razick, S., Magklaras, G., & Donaldson, I. M. (2008). iRefIndex: a consolidated protein interaction database with provenance. *BMC bioinformatics*, 9(1), 405.
113. Lehne, B., & Schlitt, T. (2009). Protein-protein interaction databases: keeping up with growing interactomes. *Human genomics*, 3(3), 291.
114. Altuntas, V., Gök, M., Kahveci, T. "Stability Analysis of Biological Networks' Diffusion State." *IEEE/ACM transactions on computational biology and bioinformatics* (2018).
115. Altuntaş, V., & Gök, M. (2017, October). The stability and fragility of biological networks: Eukaryotic model organism *Saccharomyces cerevisiae*. In 2017 International Conference on Computer Science and Engineering (UBMK) (pp. 116-118). IEEE.
116. Turinsky, A. L., Razick, S., Turner, B., Donaldson, I. M., & Wodak, S. J. (2010). Literature curation of protein interactions: measuring agreement across major public databases. *Database*, 2010.
117. Bhardwaj, N., & Lu, H. (2005). Correlation between gene expression profiles and protein–protein interactions within and across genomes. *Bioinformatics*, 21(11), 2730-2738.