

Filogenetik Aęaęlandırma Metotları

Seyyide SARIÇAM, H. Kaan MÜŞTAK

Ankara Üniversitesi Veteriner Fakóltesi Mikrobiyoloji Anabilim Dalı

Geliş Tarihi / Received: 02.09.2015, Kabul Tarihi / Accepted: 19.11.2015

Özet: Bu derlemede; filogenetik bilimi ve esasları, temel terminolojisi, aęaę gösterim şekilleri hakkında bilgi verilip, bir filogenetik analizde izlenmesi gereken adımlar anlatıldı. Analizde kullanılacak verinin taşınması gereken başlıca kriterler, veri tabanlarından referans birimlerin elde edilmesi ve genlerin analiz süreci ele alındı. Filogenetik aęaę oluřturmada kullanılan bilgisayar tabanlı metotların başlıca özellikleri ve üstünlükleri belirtildi. Analiz sonucu elde edilen aęaęın güvenilirlięi ve bu deęerin yorumlanması üzerinde duruldu. Özetle bu derlemede temel filogenetik bilgisinin verilmesi amaçlandı.

Anahtar kelimeler: Aęaęlandırma metotları, filogenetik, takson

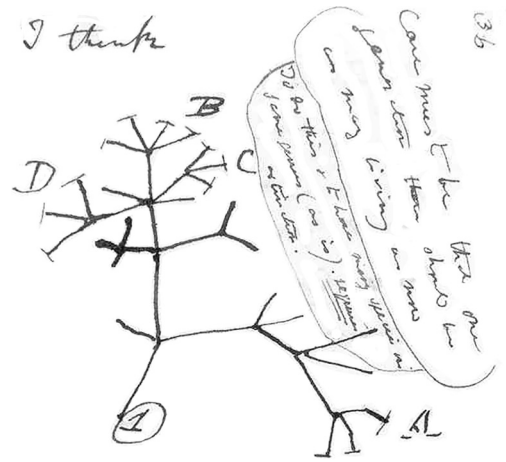
Phylogenetic Tree Construction Methods

Abstract: In this review; the knowledge about the phylogenetic science and its principles, basic phylogeny terminology, type of phylogenetic tree presentations were given and also the steps that must be followed in a phylogenetic analysis was mentioned. The main criteria for the data to be used in analysis, obtaining the reference units from databanks and the process of data analysis were described. The main features and advantages of the computer-based methods that were used to create phylogenetic trees were also showed. The reliability of the phylogenetic tree obtained from the results of analysis and the interpretation of these values were described. In summary it was aimed to give the basic phylogenetic knowledge.

Key words: Tree construction methods, phylogenetic, taxa

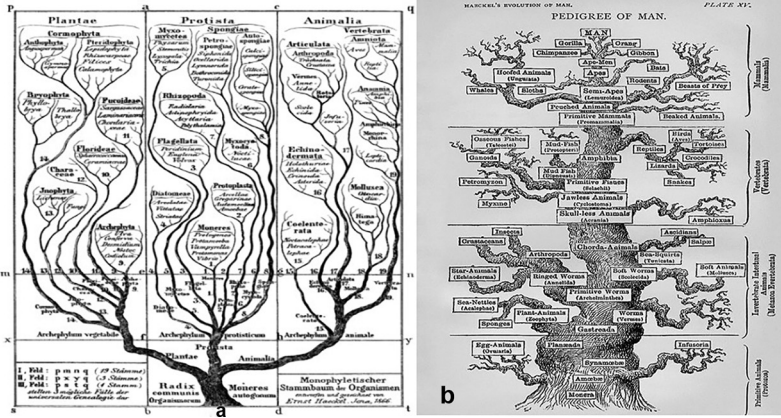
Giriş

“Filogenetik”, tüm organizma grupları arasındaki evrimsel iliřkiyi ata-soy iliřkileri řeklinde ortaya çıkarmayı amaçlar. Organizmaların sahip olduęu moleküler mekanizmalar, tek bir ataya sahip olduklarını göstermektedir. Ortak atadan evrimleşmeleri sayesinde türler, birbirleriyle iliřkilendirilebilirler [6,7]. Filogenetik sistematilerinin kurucularından Alman biyolog Emil Hans Willi Hennig, bu iliřkilendirmenin; türler arası morfolojik, fizyolojik, genetik, coęrafik ve ekolojik farklılıklar dikkate alınarak gerçekleştirilebileceğini ortaya koymuřtur [12]. Saptanan filogenetik iliřkinin, grafiksel olarak gösterimi ise “filogenetik aęaęlar” aracılıęıyla olur. Organizmalar arası evrimsel iliřkileri gösteren bu filogeniler, “evrim aęacı” ya da “yařam aęacı” olarak da bilinmektedir. Yařam aęacı kavramı, tek atadan köken almıř ve dallanarak farklılařmıř türleri tek bir konsept halinde göstermek için, ilk kez İngiliz biyolog Charles Darwin tarafından (1809-1882) evrim teorisinde kullanılmıřtır (Şekil 1) [1].



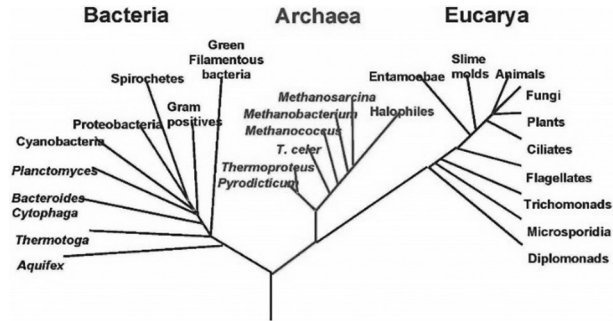
Şekil 1. Charles Darwin'in “Türlerin kökeni (1859)” bölüm IV’de kullandıęı yařam aęacı.

Alman biyolog Ernst Haeckel, Darwin’in görüşlerini benimseyerek biyogenetiğin temelini attıęı çalışmalarında (1834-1919) daha kapsamlı ve ayrıntılı yařam aęaęları oluřturmuřtur (Şekil 2).



Şekil 2. a) Ernst Haeckel; 1866'da bitki, protista ve hayvanları kapsayan yaşam ağacını, b) 1879'da ise insan yaşam ağacını oluşturmuştur [18].

1970'lerin sonunda ise Carl Woese ve George Fox, prokaryot ve ökaryot ayrımını 16S rRNA analizleriyle ortaya koymuş ve filogenetik ağacın üç parçalı evrensel halini önermiştir (Şekil 3) [1].



Şekil 3. Carl Woese tarafından ortaya konan "Yaşamın filogenetik ağacı"

Filogenetik'te Kullanılan Temel Terminoloji

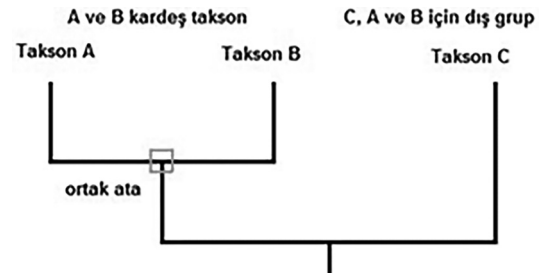
Takson

Canlıların sınıflandırılmasında, alemde türe kadar bir hiyerarşi içinde düzenlenmiş birimlerin (şube, sınıf, takım, familya, cins) her birine "takson" denir. Bir gruba takson denilebilmesi için, belirli bir kategoriye girebilecek derecede ayırt edici farklılıklara sahip olması gerekir. En üst kategorilerde yer alan taksonlar, daha aşağıdaki taksonları içine alan basamaklar şeklinde gösterilirler. Filogenetik ağaçlarda dallanma noktalarında bulunurlar (Şekil 5) [20]. Filogenetik ağaçta ortak bir atadan köken alan, birbiriyle yakın ilişkili taksonlara da "kardeş taksonlar" denir (Şekil 4) [6].

Dış Grup

Filogenetik ağaçta diğer tüm taksonlardan evrimsel süreçte açık bir biçimde en erken ayrılan taksona

denir. Diğer taksonlar, dış gruba kıyaslanınca birbiriyle daha yakından ilişkilidir. Bu nedenle ağacın köküne yakın bulunur (Şekil 4).



Şekil 4. Filogenetik ağaç birimleri

Klan

Ortak bir atadan gelen takson grubuna (Tür ya da popülasyonlar) "klan" ya da "monofiletik" denir. Bir monofiletik gruptaki taksonların paylaştığı bir ata, başka herhangi bir takson tarafından paylaşılmaz. Bu nedenle "tek kökenlilik" olarak da bilinmektedir (Şekil 5) [7].

Polifiletik ve Parafiletik grup

Birden fazla atadan oluşan türlere "polifiletik grup" ya da "çok kökenlilik" denir. Ortak atanın bütün türlerini içermeyen bazı gruplara ise "parafiletik" grup denir (Şekil 5) [3].

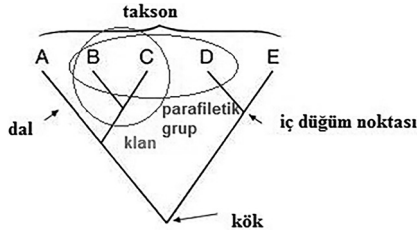
Soy

Bir filogenetik ağaçta ata-torun ilişkisini tanımlayan ve monofiletik gruba giden dala denir.

Düğüm noktası

Filogenetik ağaçta komşu iki taksonun kesiştiği noktadır. Her düğüm bir taksonomik birimdir. "Hi-

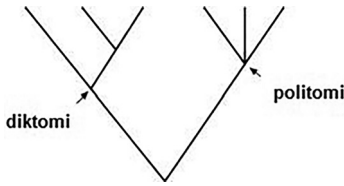
potetik taksonlar “ya da “çatallanma noktası” olarak da adlandırılabilirler (Şekil 5).



Şekil 5. Filogenetik ağaç birimleri

Diktomi ve Politomi

Filogenetik ağacın dallanma modeline ya da desenine denir. Bu model ayrıık iki ana dal üzerinde oluştuğunda “diktomi”; tek bir düğüm noktasından köken alan ikiden fazla ana dal üzerinde oluştuğunda “politomi” adını alır (Şekil 6) [7].



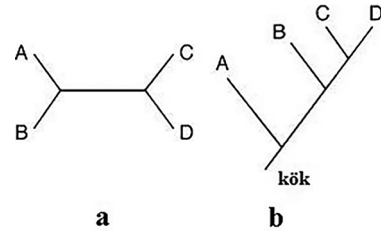
Şekil 6. Filogenetik ağaç birimleri

1962 yılında Emile Zuckerkandl ve Linus Pauling tarafından ortaya konulan “Moleküler Saat Hipotezi” (MSH); jeolojik geçmişte iki türün veya taksonların birbirinden ne zaman ayrıldıklarını tespit etmek için, moleküler değişim oranlarının kullanıldığı bir tekniktir. Moleküler saat belirlemede kullanılan moleküler veriler, nükleik asitlerde nükleotid dizileri veya proteinlerdeki aminoasit dizileridir [21]. “Gen saati, genetik saat ya da evrimsel saat” dendiği de olur [1].

Filogenetik Ağaç Gösterim Şekilleri

Köklü ve Köksüz Ağaçlar

Filogenetik ağaçlar, köklü ya da köksüz formatlarda hazırlanabilir. Köksüz ağaçta, her bir taksonun diğeriyle ilişkisi görülür ancak ortak ata tahmin edilemediği için evrimsel yönü yoktur. Bu evrimsel ilişkiyi belirlemek için kullanılan köklü ağaçlarda ise taksonlar, ortak bir atadan köken alarak yerleştirilir. Bundan dolayı köklü ağaç, köksüz filogenetik ağaçlara oranla daha fazla bilgi sağlamaktadır (Şekil 7) [7].



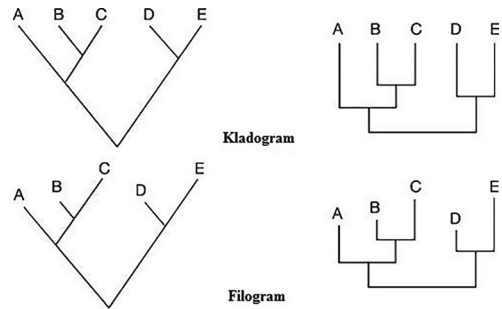
Şekil 7. a) Köksüz, b) Köklü filogenetik ağaç

Dendrogram, Kladogram ve Filogram

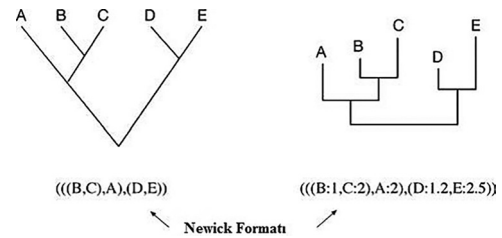
Filogenetik ağaçlarda, taksonlar arası ilişkiler tanımlanır. Bu gösterim, dendrogram, kladogram ve filogram gibi farklı yollarla gerçekleştirilebilir.

“Dendrogram”, yunanca ağaç (dendro) ve çizim (gramma) kelimelerinden köken almaktadır. Kümeleme ile hiyerarşik düzen halinde ele alınan verileri bir araya getiren en temel gösterimdir. Öbek ağacı olarak da bilinmektedir [23].

“Kladogram”, dal uzunluklarının önemli olmadığı filogenetik ağaç şekilleridir. Bu nedenle dal uzunlukları farklı değildir ve evrimsel yakınlık için bir anlamı yoktur. “Filogram”, ise dal uzunluklarının önemli olduğu filogenetik ağaç şekilleridir. Dal uzunlukları ile evrimsel uzaklıkların doğru orantılı olarak ifade edildiği gösterimdir (Şekil 8) [9].



Şekil 8. Kladogram ve Filogram



Şekil 9. Newick formatı.

Kladogram ve filogram çizimleri için geliştirilmiş bir format olan “Newick Formatı” ile taksonlar,

iç içe parantezler şeklinde yazılır. Bir monofilogenetik gruba ait taksonlar virgül ile ayrılarak parantez içinde belirtilir. Filogramlar ölçekli ağaçlar olduğu için, uzaklık önemini vurgulamak için de sayılardan yararlanılır (Şekil 9) [7].

Filogenetik Analizde İzlenecek Adımlar

Gen seçimi ve gene ait özellikler

Karşılaştırma yapmak için analiz edilecek örnekler ait genomik ya da proteomik veriler (DNA, RNA ya da aminoasit dizileri) toplanır. Bu aşamada referans birimin (Örn; belirli genler) seçilmesinde dikkat edilmesi gereken noktalar vardır. Bazı diziler diğerlerinden birkaç noktayla öne çıkar ve filogenide daha kullanışlı hale gelir.

Ele alınacak dizi tüm organizmalarda yapı ve fonksiyonunu korumuş olmalıdır. Örneğin; 16S rRNA, ribozomal küçük alt birimin yapısına katılan 16S ribozomal RNA molekülünün sentezinde kalıp olarak kullanılarak, tüm hücrelerde aynı fonksiyonda bulunur. Aynı zamanda kompleks sekonder yapısını da koruduğundan filogenetik ağaçlandırmada sıklıkla tercih edilir. Ökaryotik hücrelerde ise fonksiyonel olmayan diziler yani intronlar, çok yakın akrabaların filogenetik analizini saptırdığı için kullanılmamaktadır.

Elde edilen sekans verisi, yeteri kadar varyasyon göstermesinin yanında kolay sıralamanın mümkün olabilmesi için benzer diziler de -homolog kalıntılar- içermelidir. Örneğin; 16S rRNA, diğer iki ribozomal RNA molekülü ve pek çok şaperon protein ile katlanmış kendine özgü bir kuaternar yapıya sahip olduğundan; yapısında tüm bakterilerde korunmuş dizileri bulundururken, yakın türler arasında bile değişebilen ve değişken bölgeler olarak bilinen dizileri de içerir. Bu da, hem dizileri benzerliklerine göre sıralamanın hem de farklılıklarına göre tür ye da başka bir taksonomik birim düzeyinde ayırt edebilmenin mümkün olmasını sağlar.

Soylar arası ilişkilendirmede kriter olabilmesi için ağaçlandırmada kullanılacak dizi, horizontal olarak aktarılmamalı yani yalnızca ata bireyden yeni nesle vertikal olarak aktarılmalıdır. Organizmalar arasında horizontal gen alışverişinin olması, onların evrimsel açıdan ata-soy ilişkisinde gösterilmesini engeller [7,24].

Seçilen genin karşılaştırılması için kullanılan veri tabanları

Seçilen sekans verisinin karşılaştırılması ve filogenetik ağaçlandırma için kullanılacak olan “referans diziler” çeşitli veri tabanları aracılığıyla elde edilir.

Amerikan Ulusal Kütüphanesinin bir parçası olarak oluşturulmuş NCBI (National Center for Biotechnology Information); yerel veritabanları, bilgisayar destekli araştırmalar, dizi analizi yapan yazılımlara odaklanmış bir kuruluştur. Yapısına BLAST (Basic Local Alignment Search Tool), GenBank ve Pubmed gibi uygulamaları ekleyerek daha kapsamlı hale gelmiştir. INSDC'nin (International Nucleotide Sequence Databases) bir parçası olarak DNA dizilerini toplayan GenBank; DDBJ ve EMBL ile işbirliği içerisinde [13].

EMBL (European Molecular Biology Laboratory), Avrupa Birliği üyesi devletlerin bilimsel topluluklarına eğitim ve hizmet vermektedir. Avrupa'ya bir uçtan diğer uca kapsayan; ana laboratuvar, Avrupa Biyoinformatik Enstitüsü (EMBL-EBI), yapısal biyoloji araştırmaları üzerine ve fare biyolojisi üzerine çalışan 5 enstitüden oluşturmaktadır [14].

DDBJ (DNA Data Bank of Japan), nükleotid dizi verilerini toplayarak; yaşam bilimleri araştırma aktivitelerine, süper bilgisayar sistemleri ve serbestçe kullanılabilir verileri ile destek sağlamaktadır [15].

PDB (Protein Data Bank), nükleik asitler ve proteinleri kapsayan biyolojik moleküllerin üç boyutlu yapılarını araştırmakta ve bunları toplamaktadır [13].

Genlerin analizi

Hizalama: Ele alınan dizileri karşılaştırmak için ilk olarak hizalama yapılır. Dizilerin hizalanması, filogenetik analiz sürecindeki en kritik adımdır. Çünkü evrimsel süreçte dizinin uygun pozisyonlarını saptamak, yalnızca doğru hizalama ile doğru filogenetik çıkarım elde edildiği için oldukça önemlidir. Yanlış hizalama, final ağacında sistematik hatalara ya da tamamen yanlış bir ağaca neden olmaktadır. Bu nedenle diziler doğru bir şekilde hizalanmalıdır. Genel olarak doğru bir hizalama, benzer kalıntıların eşleşmesini ve benzer fizikokimyasal özelliklere uymasını sağlamalıdır. Sekonder yapısal elementler bilinir ya da tahmin edilirse, bunlar hizalamaya rehberlik edebilir.

Manuel düzenleme, hizalama kalitesinin yükseltilmesinde önemlidir. Belirsiz hizalanmış bölgelerin, filogenetik analizin öncesinde kaldırılması gerekir. Hizalamanın hangi bölümünde bunların kaldırılacağını, araştırmacının taktığı belirler. Bu nedenle oldukça sübjektif bir süreçtir.

Bunlara ek olarak, hizalama hatalarının düzeltilmesini ve ilişkisiz ya da yüksek derecede bağımsız dizilerin ortadan kaldırılmasını sağlayarak hizalama kalitesinin iyileştirilmesinde önemli olan otomatik yaklaşımlar vardır. Çeşitli programlar ile doğru hizalanmış diziler, filogenetik analize hazır hale gelir [7].

Çoklu Yer Değiştirmeler: İki dizi arası farklılığın basit bir ölçümü, hizalamadaki yer değiştirmelerin sayılmasıyla belirlenir. İki dizi arasındaki gözlenen uzaklığı, bu yer değiştirme oranı belirler. Ancak gözlenen bu sayı, gerçek evrimsel olayları yansıtmayıp rastgele bir mutasyon ile oluşmuş olabilir. Çoklu yer değiştirmeler ve bireysel pozisyonlarda değişimler diziler arası evrimsel uzaklığın saptanmasını zorlaştırır. Bu etki "homoplazi" olarak bilinir ve yanlış homoplazi, yanlış filogenetik ağaç oluşturmaya neden olur. Doğru homoplazi, yani diziler arası uzaklığı evrimsel olarak doğru şekilde bulmak için çeşitli istatistiksel modellere ihtiyaç duyar.

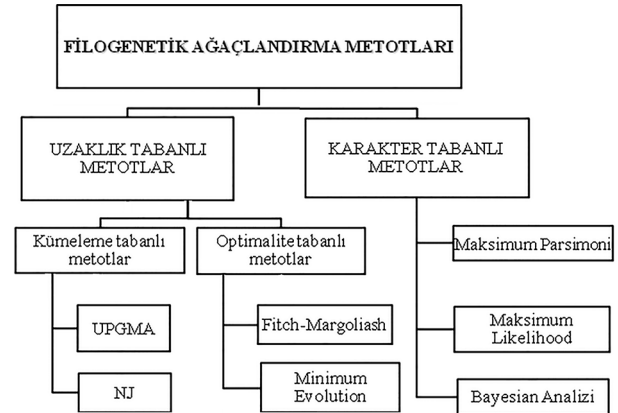
Yer Değiştirme Modellerinin Seçimi: Doğru homoplazi için, yer değiştirme modelleri veya evrimsel modeller olarak da bilinen istatistiksel modeller kullanılır. DNA filogenileri oluşturmak için, çok sayıda yer değiştirme modeli vardır. Bu modellerde, her bir nükleotid için uygulanan çoklu yer değiştirmeler yani mutasyonlar farklıdır. Birbirinden çok uzaklaşmış dizilerde, çok sayıda çoklu yer değiştirme tespit edilir. Yer değiştirme modelleri ile bu pozisyonlar doğru olarak saptanır. Bu, evrimsel uzaklığın istatistiksel metodun doğruluğuna bağlı olduğu anlamına gelir. Bu nedenle, yalnızca uygun diziler filogenetik ağaçlandırmada kullanılabilir [7].

Jukes-Cantor Modeli (JC Modeli): 1969 yılında oluşturulan ve tek parametrelili model olarak adlandırılan bu yer değiştirme modelinde; DNA dizilerinin, eşit olasılıkla birbirine dönüştüğü varsayılır. Bu nedenle eşit olasılıkla nükleotidlerin yerleştirilmesine dayanır. En basit nükleotid yer değiştirme modelidir. Evrimsel uzaklıkları kullanan bir formülle çalışır.

Kimura Modeli (K2P Modeli): 1980 yılında oluşturulan ve iki parametrelili model olarak adlandırılan bu yer değiştirme modelinde; DNA dizilerinin, transisyonlar ve transversiyonlar için farklı olasılıklarda birbirine dönüştüğünü yani yer değiştirdiğini varsayar. Bu nedenle daha karmaşık ama daha gerçekçi bir modeldir. Bu modele göre, transisyonlar transversiyonlardan daha sık meydana gelir ve bu da evrimsel uzaklıkların saptanmasında önemli bir kriterdir. Model, daha kapsamlı bir formüle dayanmaktadır [7,16].

Filogenetik Ağaçlandırma Metotları

Filogenetik ağaçlandırma metotları, dayandıkları esaslara göre "uzaklık tabanlı" ve "karakter tabanlı" olmak üzere ikiye ayrılırlar. Uzaklık tabanlı metotlar, taksonların aralarındaki uzaklıklara göre kümelendirilerek yerleştirilmesi ya da oluşturulan birden fazla ağaçtan optimal olanın seçilmesine göre kendi içerisinde ikiye ayrılırken; karakter tabanlı metotlar tek başlık altında toplanabilen metotlardır (Şekil 10).



Şekil 10. Filogenetik ağaçlandırma metotları

Uzaklık Tabanlı Metotlar

Tüm taksonların, aralarındaki uzaklıklar dikkate alınarak yerleştirildiği filogenetik ağaçlardır. Genetik uzaklık yöntemi, filogenetik ağacı oluşturmak için dizi grubunda her bir çift arasında değişikliklerin sayısını temel alır. Birbirlerine genetik uzaklığı en az olan türler birleştirilerek bir ağaç oluşturulur. Uzaklık metotları ile hizalanan diziler arasındaki farklılıkların miktarına göre ağaç oluşturulur [10]. Uzaklık metotları, diğer yöntemlerden daha kolay ve hızlıdır [3].

Kümeleme tabanlı metotlar (Unweighted Pair Group Method Using Arithmetic Average

(UPGMA): “Aritmetik Ortalama ile Ağırlıksız Çift Grup Metodu” olarak çevrilir. Kümelemenin en basit ve en hızlı metodudur. Köklü ağaçlar oluşturan ultrametrik (kökten tüm uçlara eşit uzunlukta dalları olan) bir ağaçlandırma bir metottur. Verileri uzaklık bakımından algoritmik olarak düzenleyerek taksonları kümeleyen bu metot, bu uzaklığı elde ederken bir formül kullanır [7]. En yakın iki taksonun gruplandırılmasından başlar. Artan uzaklık dikkate alınarak, tüm taksonları gruplamaya dayanır. Kademeli olarak uzaklık arttıkça taksonlar yeni gruplara girmeye ve birbirinden farklılaşmaya başlar [2]. Neighbor Joining (NJ) “Komşu Birleştirme Metodu” olarak çevrilen bu metot, 1987’de Saitou ve Nei tarafından önerilmiştir [11]. Verileri, genetik uzaklık bakımından kümeleyerek analiz eden bu kümelenme metodu, evrimsel saate dayandırılmadığı için köksüz ağaçlar oluşturulur. Köksüz filogenetik ağaç oluşturan en basit metottur [7].

Sonuç olarak algoritmik uzaklık metotları; optimal ağacı bulamaz çünkü alternatif ağaçlar oluşturamaz tek tip ağaç oluşturur, oluşturulan ağaçta her hangi bir dal değiştirilemez, hızlı ve kolay olduğu için başlangıç analizleri için kullanışlıdır. UPGMA köklü ve dal uzunlukları eşit ağaçlar oluşturan en hızlı ve basit yöntemken, bu metodun köksüz ve dal uzunlukları farklı ağaçlar oluşturan formatı ise Neighbor Joining metodudur [22].

Tablo 1. Kümeleme tabanlı metotların karşılaştırılması

	UPGMA	NJ
Esas	Uzaklık, kümeleme	Uzaklık, kümeleme
Ağaç sayısı	Tek	Tek
Formatı	Köklü, kladogram	Köksüz, filogram
Neden bu metot?	Başlangıç analizi yapılacaktır köklü basit bir ağaç elde etmek istenirse	Başlangıç analizi yapılacaktır köksüz basit bir ağaç elde etmek istenirse

Optimalite tabanlı metotlar: Özellikle geniş veri setleri kullanılarak filogenetik ağaçlandırma yapılacağına tercih edilen metotlardır. Bu nedenle kapsamlı ve daha yavaş bir hesaplama ile ayrıntılı araştırma gerektirir. Fitch-Margoliash ve Minimum Evolution olmak üzere iki temel metottan oluşur.

Fitch-Margoliash: Orijinal veri kümesi içinde uzaklıkları ve ağacın tüm dal uzaklıklarında minimum sapmaya dayanarak muhtemel ağaçlardan en iyisini seçer. İki taksonun rastgele bir düğüm noktasında kümelendirilmesiyle başlar ve uzaklıkları

tanımlamak için eşitlikler oluşturur. Kümelenen iki takson, yeni bir matris oluşturmaya yardım eder. Bu işlem, ağaç tamamlanıncaya kadar devam eder. Tüm ağaç topolojilerini araştırır ve gerçek uzaklıklardan en küçük sapmayla dal uzunluklarını hesaplar.

Minimum Evolution: Minimum Evrim yöntemi, benzer bir prosedür ile ağaç oluşturur ama muhtemel ağaçlar arasından en uygun olanı seçerken farklı bir optimalite kriteri kullanır. Farklı dal uzunluğunu bulurken kullandığı cebirsel eşitliğe dayanır [19].

Sonuç olarak Fitch-Margoliash ve Minimum evolution metodu, farklı aritmetiklerle uzaklığa dayalı birden fazla ağaç oluşturan optimalite metotlarıdır.

Tablo 2. Optimalite tabanlı metotların karşılaştırılması,

	Fitch-Margoliash	Minimum evolution
Esas	Uzaklık, optimalite	Uzaklık, optimalite
Ağaç sayısı	Çok	Çok
Formatı	İlaveli	İlaveli
Neden bu metot?	Daha geniş veri setlerinde, daha yavaş bir hesaplama	Daha geniş veri setlerinde, daha yavaş bir hesaplama
Farkları		Cebirsel eşitlik

Karakter Dayalı Metotlar

Karakter dayalı metotlar, daha karmaşık olduğu için daha fazla emek ve zaman isteyen metotlardır. Bunlar Maksimum parsimoni (MP), Maksimum likelihood (ML) ve Bayesian analizidir.

Maksimum Parsimoni: Bir gözlemin en az karmaşık olarak açıklanması “parsimoni” şeklinde tanımlanır. Bu nedenle bu ağaçlandırma yöntemi de incelenen diziler ile uyumlu bir ağaç elde etmek için gerekli en az mutasyonların saptanmasını esas alır. Başka bir deyişle “tutumluluk” olarak tanımlanabilir yani, biyolojik değişim süreci boyunca karmaşıklık yerine basit bir açıklama yaparak verilerin yorumlanmasıdır. MP yöntemi uygulanırken, dizi pozisyonlarının farklı puanlamaları tercih edilebilir. Örneğin; korunmuş bölgede gerçekleşen bazı mutasyonlar, değişken bölgedeki mutasyonlardan daha çok vurgulanmak istenebilir. Ya da transversiyonlar transisyonlardan daha önemli olarak vurgulanabilir. MP ile ağaçların oluşturulmasında ‘kesin’ ve ‘tahmini’ yaklaşımlar söz konusu olmaktadır. Çok zaman alıcıdır ve çok sayıda örnek ele alındığında kullanıma uygun değildir [3].

Maksimum Likelihood: Joseph Felsenstein tarafından 1981 yılında MP'ye alternatif olarak ortaya konulmuş bir yöntemdir [4]. Birçok olası ağaç içerisinde en iyi ağacı seçmede istatistiksel testler kullanma olanağı ortaya konmuştur. Bu nedenle, maksimum likelihood metodu, her ağaç topolojisini değerlendirir veya gözlenen verinin oluşturulması olasılığını hesaplar. Eğer ağaç doğruysa her dalın oluşturulma olasılıkları toplamı, gözlenen verinin oluşturulması olasılığını temsil eder. Bu olasılık, ağaçların olasılığı olarak kabul edilir. [3].

Bayesian Analizi: İngiliz istatistikçi Thomas Bayes (1702-1761) tarafından oluşturulduğu için

“Bayes Teoremi” olarak da bilinmektedir. Olasılık kuramı içinde incelenir ve bütün istatistikçiler için kabul edilir bir ilişkiyi açıklar. Bu teorem bir rasal (rastgele) değişken için olasılık dağılımı içinde koşullu olasılıklar ile arasındaki ilişkiyi gösterir. Olasılık teorisi içinde incelenen bir olay olarak B olayına, koşullu bir A olayı (yani B olayı bilindiği durumdaki A olayı) için olasılık değeri; A olayına koşullu olarak B olayı (yani A olayı bilindiği durumdaki B olayı) için olasılık değerinden farklıdır. Ayrıca Bayes teoremi, olasılık değeri hakkındaki subjektif inanışların güncelleştirilip değiştirilmesini sağlayan temel bir gereçtir [17].

Tablo 3. Karakter tabanlı metodların karşılaştırılması

	Maksimum Parsimoni	Minimum Evolution	Bayesian Analizi
Esası	Karakter	Karakter	Karakter
Ağaç sayısı	Çok	Çok	Çok
Formatı	İlaveli	İlaveli	İlaveli
Neden bu metod?	Puanlama ile vurgulanmak istenen nokta öne çıkarılır	Ağaçtaki her dalın olasılığı görülür.	Daha geniş veri setlerinde objektivite maksimum

Filogenetik Ağaç Doğruluk Ölçümü

Filogenetik ağaç oluşturduktan sonraki adım, istatistiksel olarak filogeninin güvenilirliğini değerlendirmektir. Bu amaca hizmet eden çeşitli analizler mevcuttur. En çok kullanılan “Bootstrapping”, bilgisayar tabanlı tekniklerin doğruluklarını yansıtan istatistiksel bir değer bulmaya yarayan bir analiz olarak tanımlanır. Bootstrap, elde edilen ağaçların dalları üzerinden parsimoni kriterini kullanarak istatistiksel yönden en güvenilir olan dalları belirlemede kullanılır [5]. Bootstrap değeri %0 ile %100 arasında değişir. Kress ve arkadaşlarının karakterize ettiği bootstrap destek değerlerine göre, %85’den büyükse çok güçlü, %70-85 arası güçlü, %50-70 arası zayıf ve %50’den küçükse çok zayıf şekilde tanımlanmıştır. Bootstrap desteğinin %70 ya da daha büyük oluşu genellikle doğru filogeninin tanımlanması ile ilişkilendirilir. Eğer belli bir dal için bootstrap desteği % 50’nin altında ise, dalı oluşturulan taksonların birbiriyle doğru olarak ilişkilendirilemediği düşünülür [8].

Kaynaklar

1. Woese C.R, Stackebrandt E, Macke T.J, Fox G.E. (1985). *A Phylogenetic Definition of the Major Eubacterial Taxa*. *System. Appl. Microbiol* 6: 143-151.
2. Morrison David A. (1996). *Phylogenetic Tree-building*. *International Journal for Parasitology*, Vol 26, No 6.
3. Freeman S, Herron J. C. (1999). *Evrimsel Analiz*. Çıplak, B, Başbüyük H, Karaytuğ, S, Gündüz, İ. (eds). Palme Yayıncılık. Ankara. 28-29: 438-708.
4. Felsenstein J. (1987). *Estimation of hominoid phylogeny from a DNA hybridization data set*. *Molecular Evolution*. 26: 123-31.
5. Felsenstein J. (1985). *“Confidence Limits on Phylogenies: an Approach using the Bootstrap”*, *Evolution*. 39: 783-791.

6. Singh Gautam B. (2015). *Fundamentals of Bioinformatics and Computational Biology, Methods and Exercises in MATLAB*. Vol 6: 235-270.
7. Xiong Jin. (2006). *Essential Bioinformatics*: 127-169
8. Kress W.J. (2005). *“The Molecular Phylogeny of Alpinia (Zingiberaceae): A Complex and Polyphyletic Genus of Gingers”*, *American Journal of Botany*. 92/1: 167-178.
9. Podsiadlo L., Polz-Dacewicz M. (2013). *Molecular evolution and phylogenetic implications in clinical research*. *Annals of agricultural and Environmental Medicine*. Vol 20.No 3: 455-459.
10. Mount D.W. (2001). *Bioinformatics Cold Spring Harbor Laboratory Press*. *Cold Spring Harbor, New York*. Chapter 3. *Alignment of pairs of sequences*: 52-137.
11. Saitou N, Nei M. (1987). *The neighbor-joining method: a new method for reconstruction phylogenetic trees*. *Molecular Biology and Evolution* 4: 406- 425.
12. Onursal Sila Suzan, Uğur Avni. (1997). *Böceklerin Filogenisi*. *Türk.entomol.derg*. 21(1): 65-80.
13. Anonim. (2015). *RCSB Protein Data Bank*. Erişim adresi: <http://www.rcsb.org/pdb/home/home.do>, Erişim tarihi: 30.05.2015.
14. Anonim. (2015). *European Molecular Biology Laboratory*. Erişim adresi: <http://www.embl.org/>, Erişim tarihi: 30.05.2015
15. Anonim. (2015). *DNA Data Bank of Japan*. Erişim adresi: <http://www.ddbj.nig.ac.jp/>, Erişim tarihi: 30.05.2015
16. Anonim. (2015). *Distance Methods*. Erişim adresi: <http://www.cs.rice.edu/~nakhleh/COMP571/Slides/Phylogenetics-DistanceMethods.pdf>, Erişim tarihi: 03.06.2015.
17. Anonim. (2015). *Bayes Teoremi*. Erişim adresi: http://tr.wikipedia.org/wiki/Bayes_teoremi, Erişim tarihi: 07.06.2015
18. Anonim. (2015). *Tree of Life*. Erişim adresi: http://en.wikipedia.org/wiki/Tree_of_life_%28biology%29, Erişim tarihi: 04.06.2015.
19. Anonim. (2015). *Phylogenetic Fundamentals*. Erişim adresi: <http://artedi.ebc.uu.se/course/X3-2004/Phylogeny/Phylogeny-Criteria/Phylogeny-Criteria.html>, Erişim tarihi: 27.05.2015.
20. Anonim. (2015). *Taksonomi*. Erişim adresi: <http://tr.wikipedia.org/wiki/Taksonomi>, Erişim tarihi: 04.06.2015.
21. Anonim. (2015). *Molecular Clock*. Erişim adresi: http://en.wikipedia.org/wiki/Molecular_clock, Erişim tarihi: 04.06.2015.
22. Anonim. (2015). *Model Based Distance Methods*. Erişim adresi: <http://www.life.umd.edu/labs/delwiche/MSyst/lec/Distance1.html>, Erişim tarihi: 06.06.2015.
23. Anonim. (2015). *Dendrogram*. Erişim adresi: <http://en.wikipedia.org/wiki/Dendrogram>, Erişim tarihi: 02.06.2015.
24. Anonim. (2015). *Molecular phylogenetic analysis using ribosomal RNA (rRNA)*. Erişim adresi: http://eebweb.arizona.edu/blast/rna_lecture.pdf, Erişim tarihi: 06.06.2015.
25. Anonim. (2015) Erişim adresi: http://www.cs.rice.edu/~nakhleh/COMP571/Slides/Phylogenetics_BuildingPhyloTrees.pdf Erişim Tarihi: 03.06.2015.