

Filogenetik Analiz Metotları

FİLOGENİ

- Filogeni , en kısa deyimle ile evrimsel şecere ilişkisi olarak tanımlanabilir. Tür ve tür üstü kategoriler jeolojik dönemlerde türleşme süreçleri ile oluşmuşlardır.
- Bu türleşme süreçlerinin açıklanması ile taksonlar arasındaki akrabalık ilişkide açıklanmış olur.

FENETİK EKOL VEYA NUMERİK FENETİK

- Taksonlar arasındaki benzerliğin derecesini esas almaktadır. Bu yol ile de bir sınıflandırmanın oluşacağı savunulmaktadır.
- Çünkü en fazla benzerlik ortak bir ataya sahip bireyler arasında beklenir.
- Numerik fenetik: Bireylerin eşit ölçüde değer verilmiş hemen bütün benzerliklerinin değerlendirilmesidir.

FENETİK EKOL VEYA NUMERİK FENETİK

- Numerik fenetik evrimsel ve filogenetikler tarafından homoplastik karakterlerin (paralel evrim veya bağımsız olarak kazanılan benzerlikler) ve eski atalardan kalan benzerliklerin eleştirilmesidir.
- Her bir karaktere eşit değer biçer ve olası ölçüde tüm benzerlikleri dikkate aldığı için doğal sınıflandırmaya yakın bir sınıflandımadır.

FİLOGENETİK AĞAÇ OLUŞTURMA METOTLARI

- Bu ağaçlar genellikle oldukça karmaşık yöntemler ile elde edilirler. Günümüzde kullanılmakta olan yöntemler temel olarak ini başlık altında toplanmaktadır.
- **1) Mesafe Temelli Yöntemler:** Dizi hizalanması (alignment) temeline dayanarak hesaplanan dizi çiftleri arasındaki farklılıkların miktarına (mesafeye) dayanır.

FİLOGENETİK AĞAÇ OLUŞTURMA METOTLARI

2) Karakter Temelli Yöntemler: Farklı karakterlere dayanmaktadır ki bu karakterler her bir taksondaki canlıya ait olan moleküler dizilerdir.

FİLOGENETİK AĞAÇ OLUŞTURMA METOTLARI

- **1) Mesafe Temelli Yöntemler(Distance-based)**
 - a) Kümelenme Temelli Yöntem
- UPGMA(Unweighted Pair Group Method Algorithm)
- Neighbor-Joining (NJ) b) En iyi
durum(optimum) Temelli Yöntem

FİLOGENETİK AĞAÇ OLUŞTURMA METOTLARI

2) Karakter Temelli Yöntemler(Character-based)

- Maximum Parsimony(MP)
- Maximum Likelihood(ML)
- Bayesian inference

MESAFE TEMELLİ YÖNTEMLER

- Dizi hizalanması (alignment) temeline dayanarak hesaplanan dizi çiftleri arasındaki farklılıkların miktarına (mesafeye) dayanır. Dizi hizalanması sonucu hesaplanan evrimsel mesafeler, her bir takson çifti arasındaki mesafelerin bir matrisinin oluşturulmasında kullanılabilirler.

MESAFE TEMELLİ YÖNTEMLER

Matristeki bu çiftli mesafe skorlarına dayanarak tüm taksonlar için bir filogenetik ağaç oluşturulabilir. Bu yöntemde kullanılan algoritmalar, kümelenme temelli veya en iyi durum (optimum durum) temelli olarak iki kola ayrılırlar

MESAFE TEMELLİ YÖNTEMLER

Kümelenme temelli algoritmalar, en benzer dizi çiftlerinden başlayan bir mesafe matrisine dayanarak filogenetik ağacı hesap ederler. Bu algoritmalar, aritmetik ortalamayı kullanarak ağırlıklı olmayan çift grup yöntemini (UPGMA (unweighted pair group mean arithmetic) ve komşu birleştirme yöntemini içerirler.

MESAFE TEMELLİ YÖNTEMLER

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. Aralarında az sayıda nükleotid değişikliği olan bu dizi çiftleri komsu (neighbours) olarak adlandırılır.

MESAFE TEMELLİ YÖNTEMLER

a) Kümelenme temelli algoritmalar : En benzer dizi çiftlerinden başlayan bir mesafe matrisine dayanarak filogenetik ağacı hesap ederler. Bu algoritmalar, aritmetik ortalamayı kullanırlar

UPGMA

- UPGMA yönteminde tahmin edilen mesafeler ile gerçek evrimsel mesafeler tam olarak uyuşmamaktadır. Bu da yöntemin yeterince başarılı olmadığını göstermektedir.
- Ağacın dalları boyunca değişiklik hızının sabit olduğunu varsayar
- Bu nedenle hesaplamaları yaparken ağacın kökünü (root – ortak ata) de hesaplar.

UPGMA

Avantajları:

- İndirekt ölçümlü uzaklıklarda kullanılabilir (immünojenik, hibridizasyon).
- Uzaklıklar görülmeyen eventslarla düzeltilebilir.
- En hızlı metotlardandır.
- Bundan dolayı çok geniş data setleri çabukça analiz edebilir.
- Bazı orantı analizlerinde kullanılabilirler.

UPGMA

- **Dezavantajları:**
- Benzerlik ve ilişkiler aynı karakterde kullanılmaz, benzerliğe göre sıralama, evrimsel ağacın verilişinde kullanılmaz.
- Karakter analizinde ***kullanılamazlar.***

UPGMA

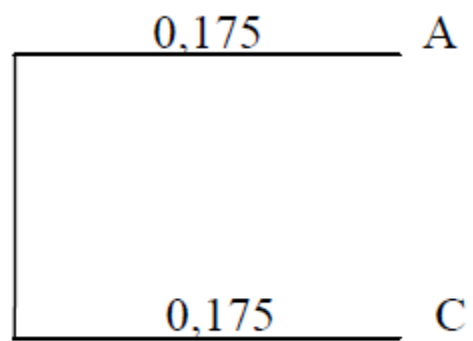
- UPGMA yöntemi ile filogenetik ağaç oluşturmaya bir örnek :

	A	B	C
B	0.40		
C	0,35	0,45	
D	0,60	0,70	0,55

UPGMA

4 takson (A,B,C,D) içeren bir mesafe matrisi kullanılarak UPGMA yöntemi ile öncelikle en yakın iki takson (A ve C, aralarındaki mesafe 0,35) birleştirilir. Tüm taksonlar boğum noktasından eşit uzaklıkta olduklarından (Bu, UPGMA yönteminde bir varsayımdır.) A'dan ve C'den düğüm noktasında olan dal uzunluğu $0,35/2= 0,175$ 'tir.

UPGMA



NEIGHBOR-JOINING

- Komşu Birleştirme Yöntemi'nde ise kümelenme temelli algoritmada olduğu gibi taksonların kökten eşit uzaklıkta olduğu varsayılmaz.
- Bu yöntem ile sadece bir tane ağaç oluşturulur ve diğer olası ağaç topolojileri test edilmez. Bu sorunun giderilmesi için genelleştirilmiş komşu birleştirme yöntemi geliştirilmiştir.

NEIGHBOR-JOINING

Avantajları :

- Çok verimlidir.
- Geniş veri kümelerini analiz edebilir.

Dezavantajları :

- Tüm olası topolojileri inceleyemez

MESAFE TEMELLİ YÖNTEMLER

- **b) En iyi durum temelli algoritmalar:**

Pek çok farklı ağaç topolojisini (pek çok noktadan ve dallardan oluşan bir ağ yapısı) karşılaştırırlar ve ağaçtaki önceden tahmin edilen mesafeler ile gerçek evrimsel mesafeler arasındaki en iyi uyumu göstereni seçerler.

KARAKTER TEMELLİ YÖNTEMLER

- Farklı karakterlere dayanmaktadır ki bu karakterler her bir taksondaki canlıya ait olan moleküler dizilerdir.
- ***MAKSİMUM PARSIMONY***
- ***MAKSİMUM LİKELİHOOD***
- ***BAYESİAN METODU***

MAKSİMUM PARSIMONY METODU

- İki birim arasında en az deęişiklik gerektiren ağaç en makul olan ağaçtır. Yani en az farklı olan birimler, birbirine en yakın ve benzeyen birimlerdir. Ağacın belirlenmesinde işlemin daha kısa zamanda tamamlanması için en zengin filogenetik bilgiye sahip olan az sayıdaki bölgeler kullanılır.

MAKSİMUM PARSIMONY METODU

- Bilgi verici bölgeler adını alan bu bölgelerde en azından iki tane farklı türde karakter bulunur ve bunların her biri en az iki kere bölgelerde yer alır.
- Bilgi verici olmayan bölgeler ya sabit bölgelerdir ya da sadece bir kere görülen değişiklikler içerirler.

MAKSİMUM PARSIMONY METODU

- Sabit bölgeler, her taksonda aynı halde bulunurlar ve bilgi verici olmayan bölgeler gibi filogenetik ağaç çizilmesinde yarasızdırlar. tabloda 2,5,8 numaralı bölgeler bilgi verici bölgelerdir.

MAKSİMUM PARSIMONY METODU

Bölgeler Taksonlar	1	2	3	4	5	6	7	8
I	A	A	T	T	A	G	C	T
II	G	G	T	C	G	T	A	G
III	A	A	T	G	C	G	C	T
IV	A	G	T	A	A	G	C	A
V	A	C	T	A	C	G	C	G
VI	A	C	A	T	G	G	C	A

MAKSİMUM PARSIMONY METODU

Avantajları:

- Basittirler, sezgi yolu ve yaratıcılıkla çıkarılabilirler (kağıt-ve-kalem ile yapılabilirler).
- Moleküler ve morfolojik datalarda kullanılabilirler.
- Benzerlik çeşitlerini ayırt edebiliriz.(sinapomorfik, plesiomorfik, homoplasi)
- Karakter ve orantı analizlerinde kullanılabilirler

MAKSİMUM PARSIMONY METODU

- **Dezavantajı:**
- Yüksek seviye homoplasilerde yanılabirler.
- İncelenen diziler ya da genetik uzaklıklar ile uyumlu bir ağaç elde etmek için gerekli en az mutasyonların saptanmasına dayanan bir yöntemdir. Maksimum Parsimoni (MP), minimum evrimsel metot (parsimoni = tutumluluk) olarak tanımlanabilir .

MAKSİMUM PARSIMONY METODU

- MP analizi ile en iyi sonuçlar dizi çiftleri arasındaki benzerliklerin çok güçlü olduğu ve az sayıda dizinin olduğu durumlarda alınır. MP ile ağaçların oluşturulmasında kesin ve tahmini yaklaşımlar söz konusudur.
- Kesin yaklaşımda olası tüm ağaçlar gözden geçirilir ve kullanılan optimalite ölçütüne en uygun ağaç belirlenir.

MAXİMUM LİKELİHOOD (ML) METODU

- Joseph Felsenstein tarafından 1981 yılında MP“ye alternatif olarak ortaya konulmuş bir yöntemdir.
- Bu yöntem, farklı tipteki nükleotit deęişiklięinin aıęa ıkma olasılıklarını tanımlayan bir matematiksel formül ve dal uzunlukları bilinen belli bir aęaç verildięinde, bu belli DNA dizisi setini elde etme olasılıęı nedir sorusunu sormaktadır.

MAXIMUM LIKELIHOOD (ML) METODU

- Bu yöntem için bir bilgisayar programı, her ağaç topolojisinin değerlendirir veya gözlenen verinin oluşturulması olasılığını hesaplar.
- Eğer ağaç doğruysa her dalın oluşturulma olasılığı toplamı, gözlenen verinin oluşturulması olasılığını temsil eder.
- Çok iyi temellendirilmiş istatistiğe dayandığından matematiksel olarak daha zahmetli bir yöntemdir.

MAXIMUM LIKELIHOOD (ML) METODU

- **Avantajları :**

- Mevcut metodların içinde genelde en tutarlı olanıdır.
- Karakter ve oran analizlerinde kullanılabilirler.
- Sönmüş (hipotetik) ataların sekanslarını tahmin etmede kullanılabilir.
- Nükleotid, aminoasit sekansları ve diğer data tiplerine uygulanabilir.

MAXİMUM LIKELİHOOD (ML) METODU

- **Dezavantajları :**
- Diğer pek çok metotda olduğu gibi basit ve sezgisel değildirler.
- Parsimonide olduğu gibi yüksek seviye homoplasilerde yanılabirler.

BAYES METODU

- Filogenetiğin en popüler metodudur. Temelde **Likelihood metoduna benzer, ancak sonraki (posterior) olasılık kullanımı ile bu yöntemden ayrılır.**
- Bu yöntemde ağaç seçiminde, “önceki olasılık”, analiz öncesinde tüm olası ağaç topolojileri için geçerli olan olasılıktır.

BAYES METODU

- Ağacın oluşturulmasından önce her bir topolojinin olasılığı birbirine eşittir. “Şarta bağlı olasılık”, dizi hizalanmasında gözlenen karakterlerin değişikliğe uğrama frekansıdır.
- Amaç tek bir “doğru” filogeniyi bulmayı değil, bütün muhtemel filogenilerin sonraki (posterior) olasılık dağılımlarını hesaplamaktır.