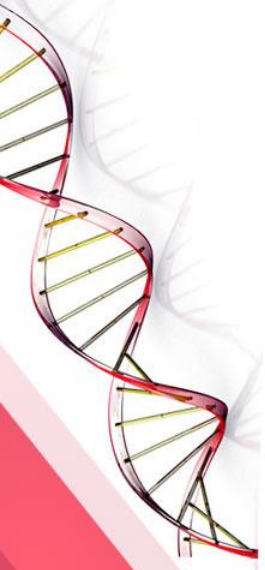


Biyoinformatik





Dr. Öğr. Üyesi

Mehmet MİLLİ

Oda: 330

Mail: mehmetmilli@ibu.edu.tr



Ders Saati ve Yeri

Ders Saati: Cuma 11:00

Ders Yeri 226



Değerlendirme

Yöntem	Oran
Vize	20
Performans	20
Laboratuvar	20
Final	40



Vize ve Final Sınavı

- Vize ve Final Sınavlarınız çoktan seçmeli, doldurma ve klasik sorulardan oluşacaktır.
- Sınavlarda sorulan sorular kesinlikle ders sunumlarında olan ve derste anlatılan konulardan olacaktır.
- Sınavlardan bir hafta önce şimdiye kadar gördüğünüz konulardan ve algortimalardan tekrar yapılacaktır.



Performans

Dönemin son 3 haftası Öğrenci sunumları olacaktır.

Bu sunumlarda sizden beklenenler

- Bir konu seçip onu sınıf ortamında arkadaşları ile paylaşmak
- Laboratuvarlarda görmediğimiz bir algoritmayı yazıp sınıfta anlatmak
- **NCBI gibi bir veri tabanını inceleyip arkadaşlarına sınıfta tanıtmak**
- Literatür araştırması yapıp sınıfta paylaşmak.
(Son yıllara ait en az 5 adet makaleyi okuyup sınıfta tartışmak.)



Performans

- Yeni Nesil Dizileme
- Farklı bir Alignment Algoritması
- Filogenetik Ağaç Oluşturma (MEGA-7) ***
- .
- .
- .



Laboratuvar Uygulama Puanı

Dönem boyunca laboratuvarda 3 veya 4 algoritmanın kodunu yazacağız.
Bunların ortalamasından alacağınız puanlar laboratuvar puanınızı oluşturacaktır.

Needelman Wunsch Algoritması
Yarı Global Alignment Algoritması
Genetik Algortimalar
Filogenetik Ağaç Oluşturma***



Biyoenformatik

- **Dersin amacı**
- Genetik ve Biyoloji bilmi ile bilgisayar bilimleri arasındaki ilişkinin incelenmesi.

Tanımı:

Özellikle 70 yılların sonlarına doğru artan ve insan gözü ile iredelenemeyecek boyutlara ulaşan genetik ve biyolojik bilginin bilgisayar bilimleri yardım ile inceleyen bilim alanına Biyoenformatik denir.



Öğrenim Kazancı

Öğrencilerin biyoenformatiğin ve algoritmanın temelleri konusunda bilgi sahibi olması, veritabanlarının nasıl kullanıldığının öğrenilmesi, biyoinformatikte sıkça kullanılan bazı algoritmaların kavratılması ve uygulama sahası konusunda geniş bir bakış açısı kazandırılması amaçlanmaktadır.

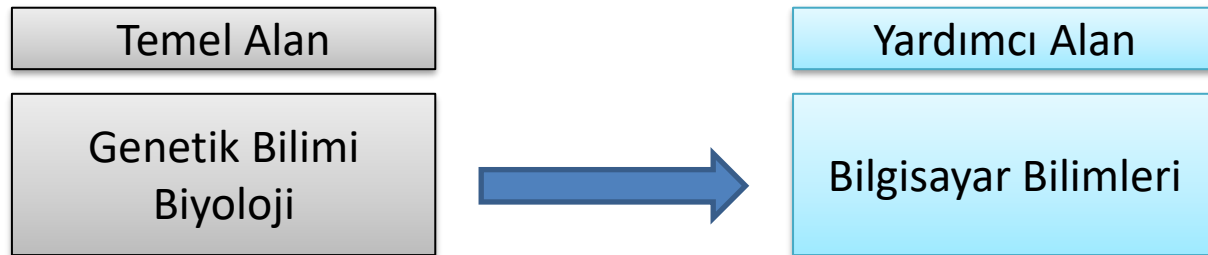


Biyoenformatik Dersini Kimler Almalı

- Biyoenformatiğe ilgi duyan yazılımcılar
 - İleride Lisans üstü (Yüksek Lisans ve Doktora) eğitim almak isteyenler
 - Mezun olduktan sonra bir araştırma merkezinde (Genetik ve Biyoloji) çalışmak
 - Genel kültür olsun isteyenler
-
- Biyoenformatik dersinde öğrendikleri ile iş bulma hevesinde olan arkadaşlarada almamaları gerekir.

Dersin amacı

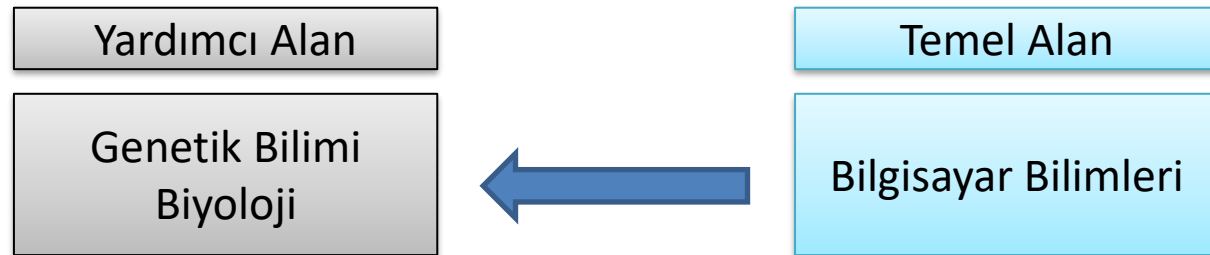
- Genetik ve Biyoloji bilmi ile bilgisayar bilimleri arasındaki ilişkinin incelenmesi.



- Genetik Bilginin Depolanması
- Veri Tabalarının Oluşturulması
- Bilginin Biçimlendirilmesi
- Genetik Bilginin Paylaşılması
- Uygulamalar
- Alignement Algoritmaları
 - Global Alignment
 - Local Alignment
 - Hibrit Alignment
- Primer Dizayn
- Filogenetik Ağaç oluşturma

Dersin amacı

- Bilgisayar bilimlerinde genetikten faydalandığı konular ve zamanlar vardır.



- Genetik Algoritmalar
- Karınca Kolonisi
- Yapay arı kolonisi
- Parçacık sürü optimizasyonu
- Evrimsel Programlama Algoritması
- Benzetimli Tavlama



Genetik Algoritmalar

- Arama ve eniyileme yöntemidir.
- Bilgisayar Bilimlerinin Genetikten esinlenilerek oluşturduğu bir yöntemdir.
- Başlangıç Populasyonu
- Ebeveyn Seçimi
- Çaprazlama
- Mutasyon
- Tek nokta veya çoklu mutasyon
- Rank Tekniği
- Rulet tekniği gibi kavramları mevcuttur.



Boşluk Doldurma-2

- Nükleotidler bulundurdukları baz'a göre isimlendirilirler.
Adenin, Guanin, Sitozin, Timin, Urasil
- DNA da bulunan anlamlı gen bölgelerine exon denir.



Doğru Yanlış-2

- (D) Nükleotidler şeker, fosfat ve organik bazlardan oluşur.
- (D) DNA sarmalında Adenin ve Timin arasında 2 adet zayıf hidrojen bağı varken, Guanin ile sitozin arasında 3 adet zayıf hidrojen bağı bulunmaktadır.
- (D) Değerli bir gıda ürününün içersine daha değersiz bir gıda ürününün bilerek karıştırılmasına tahşiş denir.
- (Y) Calının karmaşıklığı arttıkça kromozom sayısında artar.



Klasik-2

1- Biyoenformatik nedir?

Cevap: Özellikle 70 yılların sonlarına doğru artan ve insan gözü ile iredelenemeyecek boyutlara ulaşan genetik ve biyolojik bilginin bilgisayar bilimleri yardım ile inceleyen bilim alanına Biyoenformatik denir.

Sorular



Teşekkür



**thanks for
listening!**

Biyoinformatik





Dr. Öğr. Üyesi

Mehmet MİLLİ

Oda: 330

Mail: mehmetmilli@ibu.edu.tr



1 Hafta Genel Kavramlar

- Biyoinformatik nedir?
- Biyoinformatiğin Amacı
- Biyoinformatiği oluşturan bilim dalları
- Biyoinformatiğin Tarihçesi
- Neden Biyoinformatiğe ihtiyaç duyulur?
- Genel Tanımlar
 - DNA, RNA, Protein, Gen, Genom, Kromozom...

Biyoinformatik/Biyoenformatik?-1



Biyoinformatik mi yoksa **Biyoenformatik** mi?

- **Biyoloji**, Canlı bilimi.....
- **İnformatik**, Bilişim ve bilgisayar bilimleri
- **Enformatik**, her yönüyle insan yaşamının vazgeçilmez bir parçası olan bilginin, nasıl üretildiği, iletildiği ve kullanıldığını inceleyen bir bilim dalıdır.

Biyoinformatik/Biyoenformatik?-2



- TÜBİTAK İleri Genom ve **Biyoenformatik** Araştırmaları Merkezi (İGBAM)
- ODTU Enformatik Enstitüsü, **Biyoenformatik** Programı
- Boğaziçi Üniversitesi Hesapsal Biyoloji ve **Biyoinformatik** Laboratuvarı
- Bahçeşehir Üniversitesi, Genetik ve **Biyoinformatik** Bölümü

Eş Anlamlar



- **Literatürde Biyoinformatik yerine kullanılan bazı eş anlamlar**

- Molecular Bioinformatics → • Moleküler Biyoinformatik
- Computational Biology → • Hesaplamalı Biyoloji
- Biocomputing → • Biyolojik bilgi işlem

Biyoinformatik Nedir?



Biyoinformatik nedir?

- Biyoinformatik; büyük miktarda biyolojik verinin yorumlanması, depolanması ve elde edilmesi için enformatik tekniklerinin uygulanmasıdır.
- Biyoenformatik, biyolojinin çeşitli dalları, ancak özellikle moleküler biyoloji ile bilgisayar teknolojisini ve bununla ilişkili veri işleme aygıtlarını bünyesinde barındıran bilimsel disiplin.
- Biyoinformatik, kabaca söyleyecek olursak, biyolojik bilginin bilgisayar yardımı ile incelenmesi ve işlenmesidir.

Biyoinformatiğin Amacı-1



GENEL AMAÇ

- Biyoinformatiğin amacı; hızla gelişen ve artan biyolojik bilgiyi toplamak, yönetmek, dağıtmak, bilgiye en hızlı ve kolay biçimde ulaşılmasını sağlamak, çok karmaşık yapıda olan biyolojik sistemleri incelemek ve çözebilmektir.

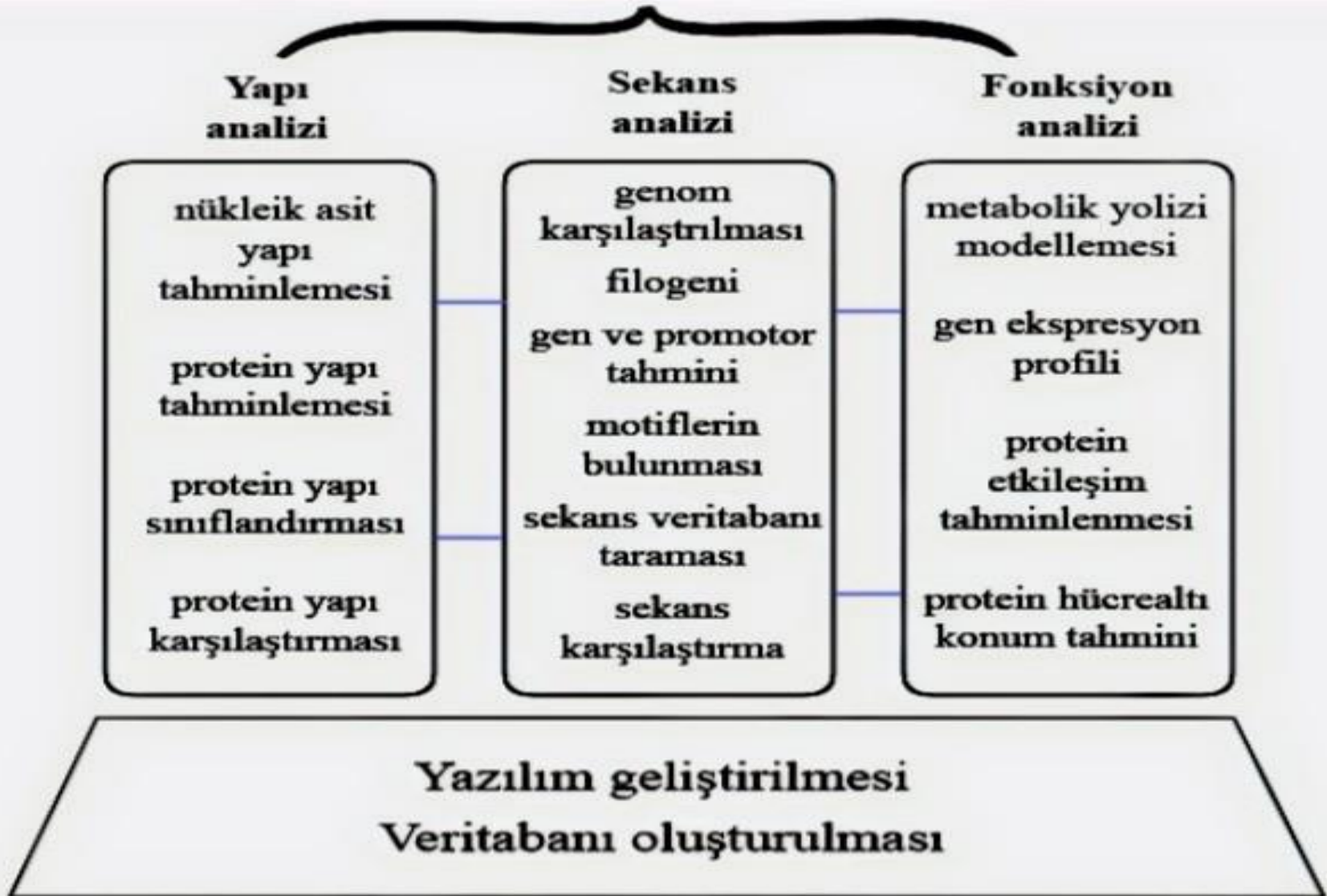
Biyoinformatiğin Amacı -2



Amaçlar

- DNA, RNA ve protein dizilerinin yapılarını ve fonksiyonlarını araştırmak
- Hastalıklara ve genetik bozukluklara çare üretebilmek
- Genetik hastalıkları tedavi edebilecek ilaçları üretebilmek
- Anamolilerin tespiti ve tedavi edilmesi

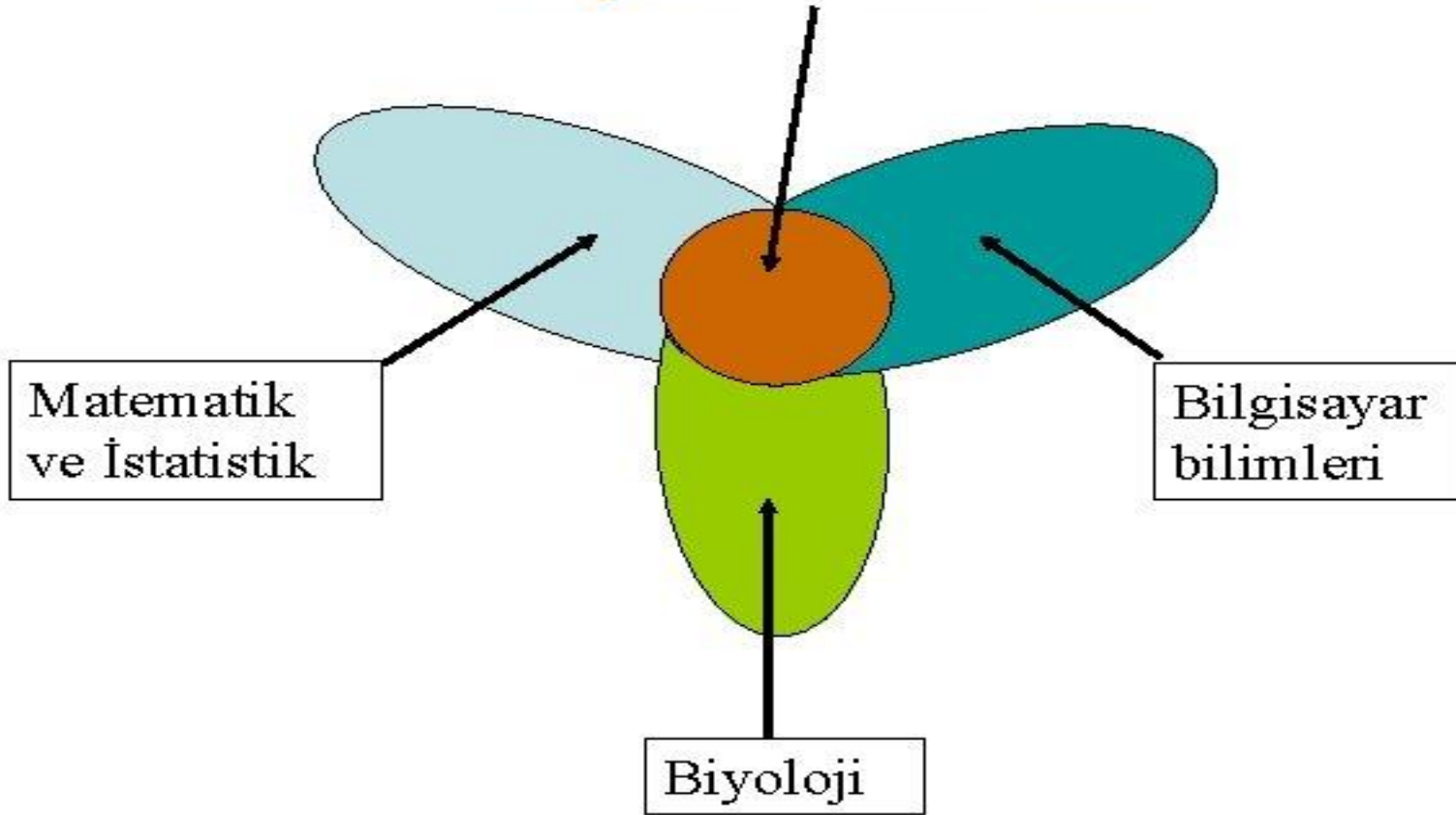
Biyoinformatiğin Amacı -3



Biyoinformatik Alanları



Biyoinformatik



Biyoinformatiği Oluşturan Temel Alanlar

Biyoinformatiğin Tarihi-1



Biyoinformatiğin Ortaya Çıkışı?

- Pauling ve Corey'in 1951 yılında proteinlerin sekonder yapılarının doğru tahmini için geliştirdikleri yaklaşım biyoinformatik için başlangıç kabul edilebilir.

Biyoinformatiğin Tarihi-2



Biyoinformatiğin Ortaya Çıkışı?

- Bir kaynağa göre;
- Günümüzde birçok araştırmacı asıl anlamda biyoinformatik biliminin başlangıcı olarak, 1966 yılında bilgisayarla moleküler grafiklerin çizimine ait ilk makalenin Scientific American dergisinde yayımlanması olarak kabul edilebilir.

Biyoinformatiğin Tarihi-3



Biyoinformatiğin Ortaya Çıkışı?

- Başka bir kaynağa göre
- Biyoinformatik adı ilk defa Paulien Hogeweg tarafından 1970 yılında, canlı sistemlerin bilgisinin incelenmesinde kullanıldı.

Biyoinformatiğin Tarihi-4



Biyoinformatiğin Ortaya Çıkışı?

- İlk biyolojik veritabanı - Protein Identification Resource, Margaret Dayhoff tarafından 1972'de kuruldu.
- İlk DNA veritabanı 1979'da kuruldu.

Biyoinformatiğin Tarihi-5



Ne zamandan beri bu isimle anılmaktadır?

- Bununla beraber “Biyoinformatik” terimi 1980’li yılların ortalarından sonra kullanılmaya başlanmıştır
- Hızla yükselen veri miktarının doğurduğu ihtiyaç ve bir o kadar hızla yükselen bilgisayar teknolojisinin sunduğu imkanlar ile Biyoinformatiğin 1990’larda ayrı bir bilim dalı olarak kabul edilmiştir.

Biyoinformatiğin Tarihi-6



National Center for Biotechnology Information

- Temel moleküler ve genetik süreçlerin anlaşılmasında ve karmaşık verilerin analizi ve yorumlanması için yeni yöntemler geliştirilmesinde Dünya çapında en etkin kurum olan “National Center for Biotechnology Information (NCBI)” 1988’de kurulmuştur.

Biyoinformatiğin Tarihi-7



Bu alandaki en önemli proje

- Biyoinformatik alanındaki en önemli projelerden biri olan İnsan Genom Projesi (İGP) Ekim 1990 da başlamıştır.
- İGP çalışmaları biyoinformatiğin gelişiminde çok önemli bir yere sahiptir. ABD Enerji Bakanlığı ve Ulusal Sağlık Enstitüsü tarafından koordine edilen 13 yıllık bu proje 2003 yılında tamamlanmıştır. Projeye İngiltere, Japonya, Fransa, Almanya gibi ülkeler de katkıda bulunmuşlardır.

Biyoinformatiğin Tarihi-8



Biyoinformatikte 2000'li yıllar

- 2000'li yıllardan sonra gerçekleştirilen diğer genom projeleri ile 1000'den fazla organizmanın gen dizileri elde edilmiştir. Böylece, farklı canlı türlerinin gen dizilimlerini karşılaştıran çalışmalar da yapılmaya başlanmıştır.
- Örnek: Sığır genomu, Domuz genomu, At Genomu, Eşşek Genomu....

Neden Biyoinformatik?



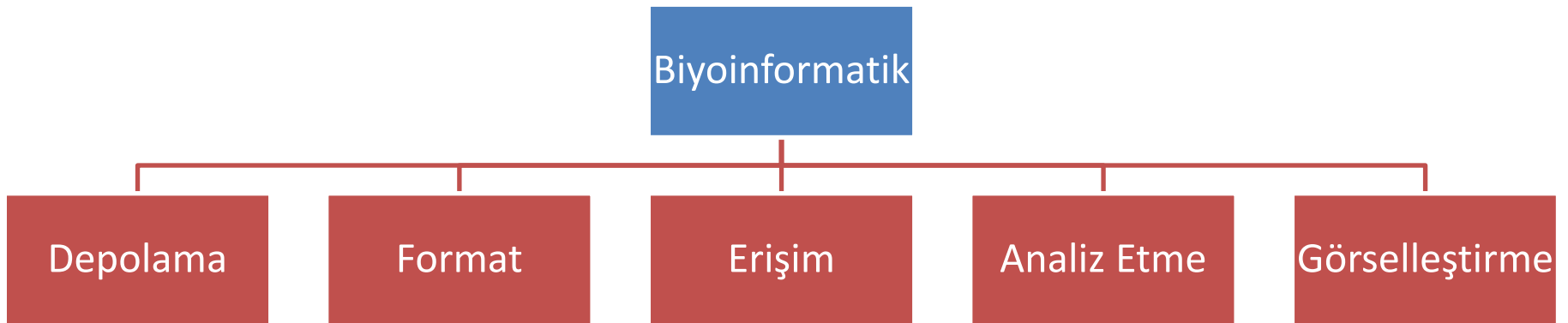
Biyoinformatiğe neden ihtiyaç duyulmuştur?

- Biyoloji, Moleküler ve Genetik konu alanlarında artan veri
- Muazzam büyüklükteki bu verinin insan gücü ile irdelenememesi
- Bu verilerin uygun koşullarda saklanma ihtiyacı
- Verilerin uygun formatlarda temsil edilmesi
- Gerektiğinde verilerden çıkarım yapılabilmesi için biyoinformatik alanına ihtiyaç duyulmuştur.

Çalışma Alanları



Biyoinformatikte ki çalışma alanları



Bilişim sektörünün biyoinformatik konu alanındaki çalışma alanları



Biyoinformatik Bilgi Akışları

- Biyoinformatikte İki temel bilgi akışı vardır.
 1. Genetik bilgi akışı
 2. Deneysel bilgi akışı



Genetik bilgi akışı

- **Genetik bilgi akışı:** bir organizmanın genetik materyalinin en küçük biriminden o organizma türünün oluşturduğu toplulukların karakteristik özelliklerine kadar olan bilgi akışıdır.
- Genetik Materyal= DNA ve RNA



Deneysel bilgi akışı

- **Deneysel bilgi akışı:** Biyolojik olaylar gözlenerek elde edilen bilgi açıklayıcı modeller ile tarif edilir, daha sonra bu modellerin doğruluğu yeni deneylerle test edilir.

Temel Kavramlar-4

Nükleotid



- Nükleotidlerin 3 bileşeni vardır.

1. Azotlu baz

2. Pentoz şekeri (5-karbonlu şeker)

3. Fosfat grubu

Fosfat

+

Şeker

+

Organik baz

=

Nükleotit

- Deoksiriboz**
- Riboz**

- Adenin (A)**
- Timin (T)**
- Sitozin (C)**
- Guanin (G)**
- Urasil (U)**

Temel Kavramlar-4

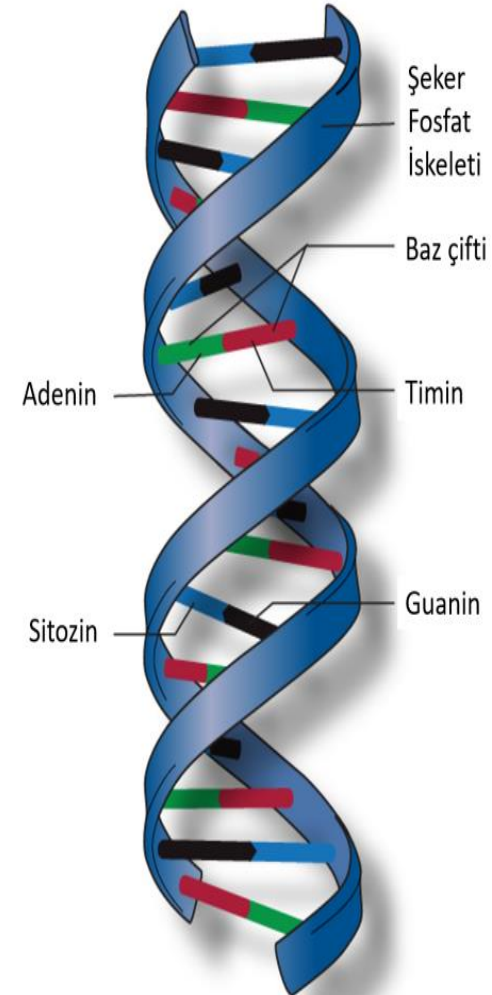


Nükleotid

- Nükleotidler bulundurdukları organik baza göre adlandırılırlar.

Organik Bazlar

- Adenin
- Guainin
- Sitozin
- Timin
- Urasil



Temel Kavramlar-1



DNA

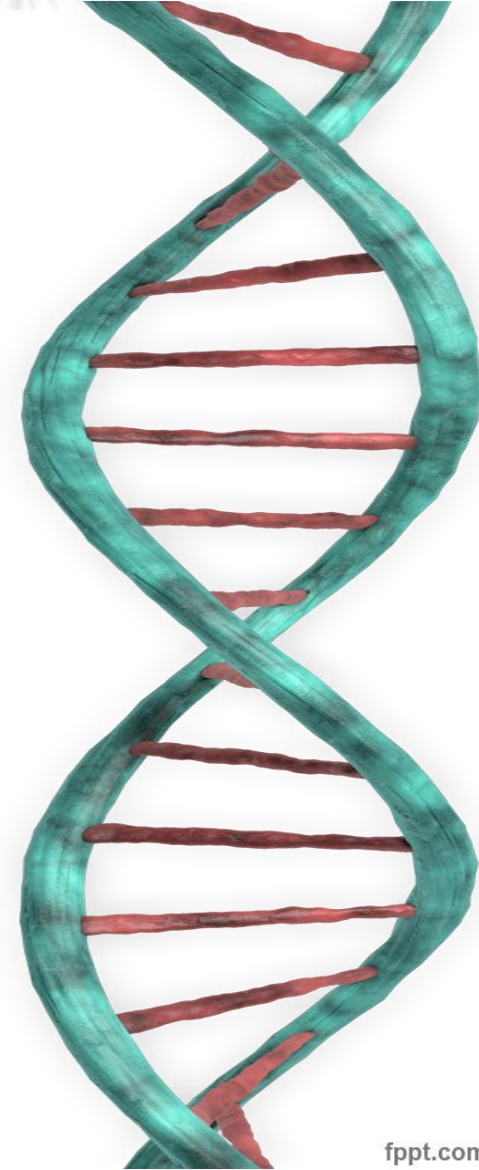
- DNA'nın açılımı deoksiribonükleik asittir.
- Birbiri etrafında dolanarak sarmal oluşturan bir asit molekülüdür.
- Tüm organizmalar ve bazı virüslerin canlılık işlevleri ve biyolojik gelişmeleri için gerekli olan genetik talimatları taşıyan bir nükleik asittir.

Temel Kavramlar-1



DNA

- DNA'nın başlıca rolü bilgiyi uzun süre saklamasıdır.
- Protein ve RNA gibi hücrenin diğer bileşenlerinin inşası için gerekli olan bilgileri taşır.
- Bir başka deyişle DNA organizmaların canlılık fonksiyonlarını icra etmesi için gereken kodu taşır.



Temel Kavramlar-1



DNA'nın yapısı

- Kimyasal olarak DNA, nükleotit olarak adlandırılan basit birimlerden oluşan iki uzun polimerden oluşur.
- Bu polimerlerin omurgaları, ester bağları ile birbirine bağlanmış şeker ve fosfat gruplarından meydana gelir.
- Bu iki iplik birbirine ters yönde uzanır.

Temel Kavramlar-1



DNA'nın yapısı

- Her bir şeker grubuna baz olarak adlandırılan dört tip molekülden biri bağlıdır.
- DNA'nın omurgası boyunca bu bazların oluşturduğu dizi, genetik bilgiyi kodlar.
- Protein sentezi sırasında bu bilgi, genetik kod aracılığıyla okununca proteinlerin amino asit dizisini belirler.
- Bu süreç sırasında DNA'daki bilgi, DNA'ya benzer yapıya sahip başka bir nükleik asit olan RNA'ya kopyalanır. Bu işleme transkripsiyon denir.

Temel Kavramlar-1

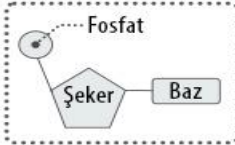
DNA'nın yapısı

DNA'nın Kimyasal Yapısı

DNA ikili sarmal yapıdadır. Yani DNA'nın yapısında birbirine dolanmış şekilde iki zincir bulunur.

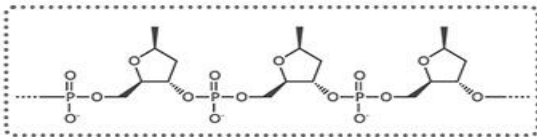
Nükleotid Birimleri

DNA'nın ikili sarmal yapısındaki her bir zincir nükleotid birimlerinden oluşur. Nükleotidlerin yapısında beş tane karbon içeren şeker (deoksiriboz), fosfat grubu (PO_4^{3-}) ve azotlu organik bazlar bulunur.



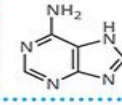
Şeker-Fosfat İskeleti

DNA'yı oluşturan nükleotidler şeker ve fosfat grupları arasındaki kimyasal bağlar ile birbirine bağlanır.

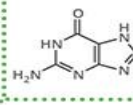


Azotlu Bazlar

A Adenin

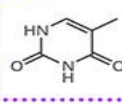


G Guanin

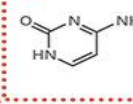


Adenin ve Guanin
İki halkaya sahip pürin sınıfı azotlu bazlardır.

T Timin



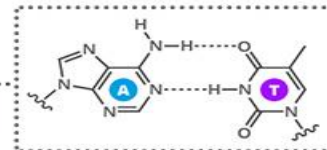
S Sitozin



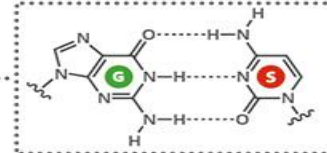
Timin ve Sitozin
Tek halkaya sahip pirimidin sınıfı azotlu bazlardır.

DNA Zincirini Bir Arada Tutan Etkileşimler

DNA'nın yapısındaki iki zincir, bazlar arasındaki hidrojen bağları ile birbirine tutunur.



Adenin ve Timin
Aralarında iki hidrojen bağı vardır.

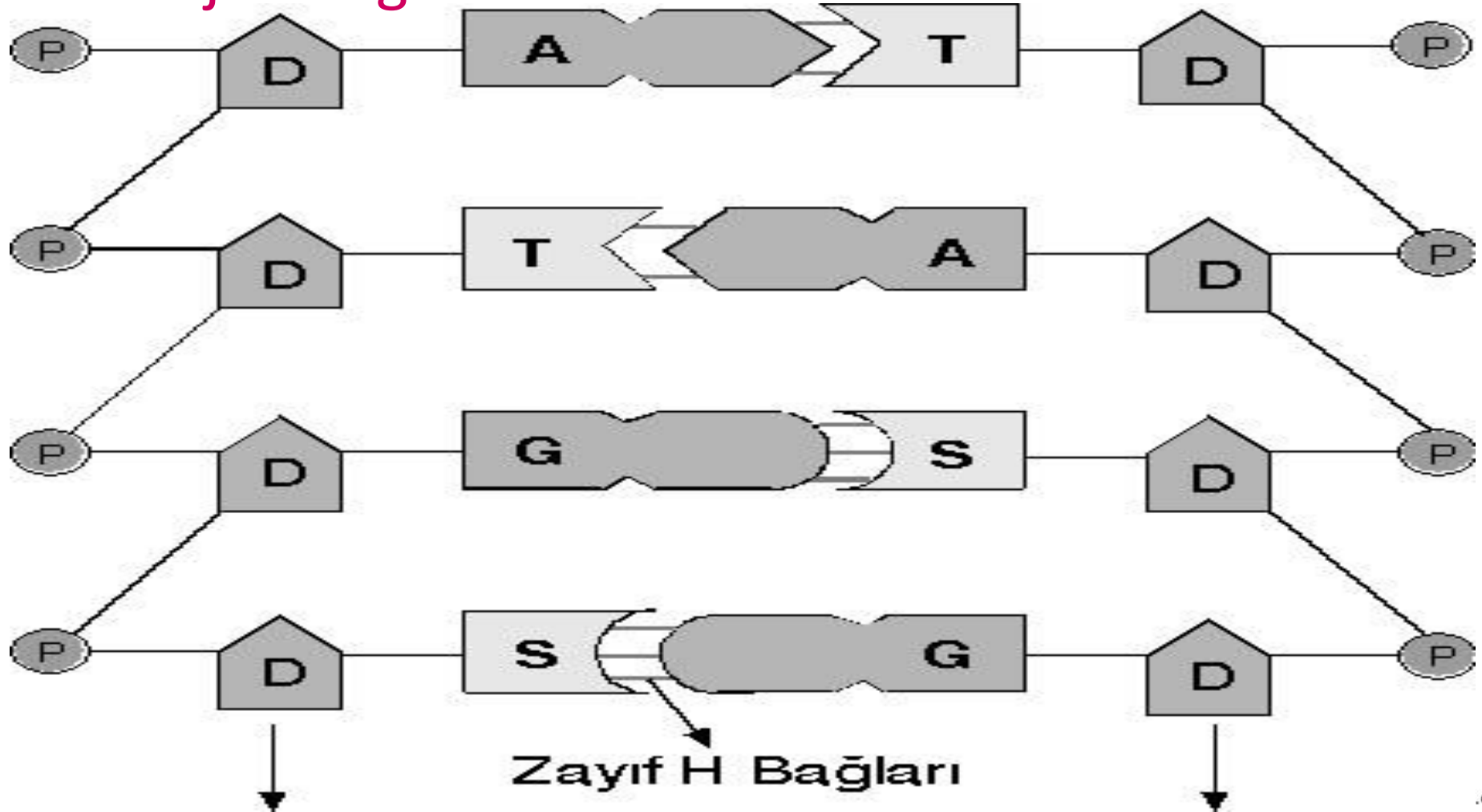


Guanin ve Sitozin
Aralarında üç hidrojen bağı vardır.

Temel Kavramlar-1



Hidrojen Bağları



Temel Kavramlar-1



RNA

- RNA'nın açılımı Ribonükleik asittir.
- Nükleotitlerden oluşan bir polimerdir.
- Her nükleotit bir azotlu baz, bir riboz şeker ve bir fosfattan oluşur.

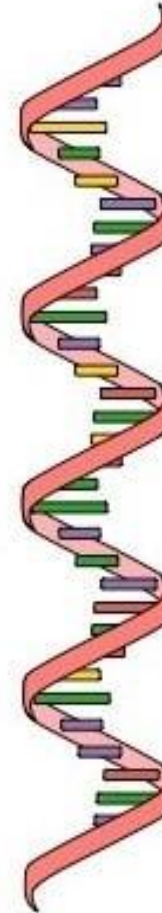


Temel Kavramlar-1



RNA

- RNA pek çok önemli biyolojik rol oynar.
- DNA'da taşınan genetik bilginin proteine çevirisi (translasyon) ile ilişkili çeşitli süreçlerde de yer alır.



RNA

Adenine



Guanine



Cytosine



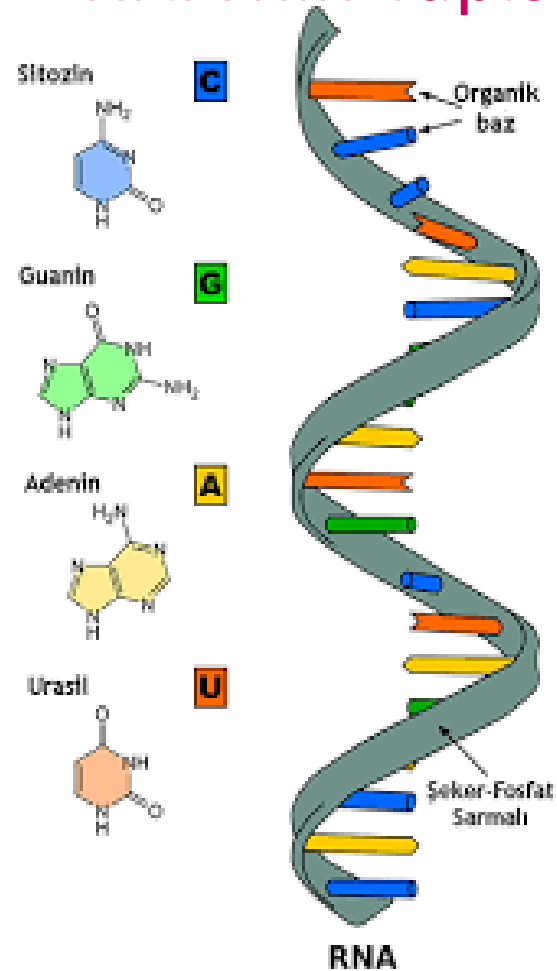
Uracil



Temel Kavramlar-1

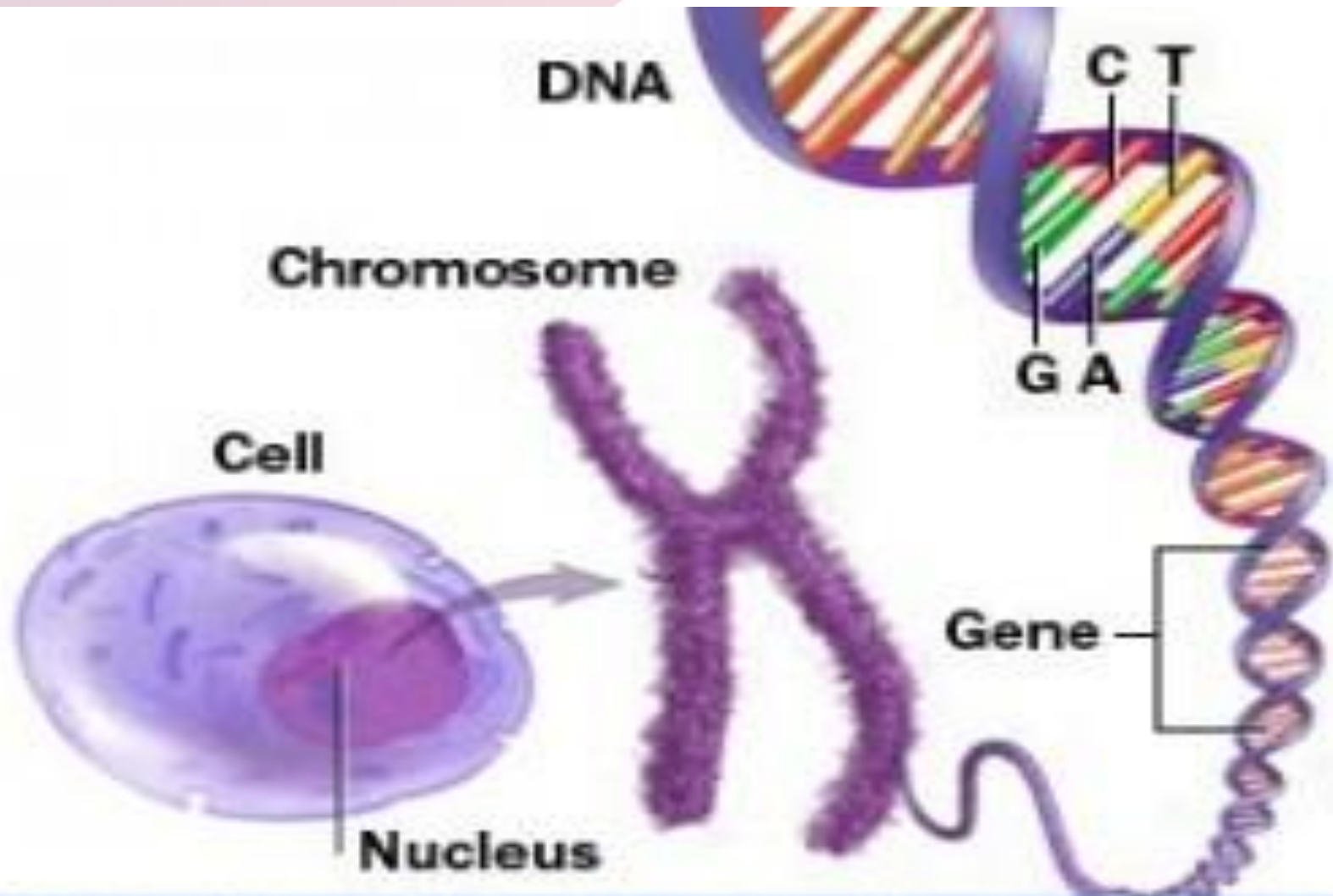


RNA'nın Yapısı

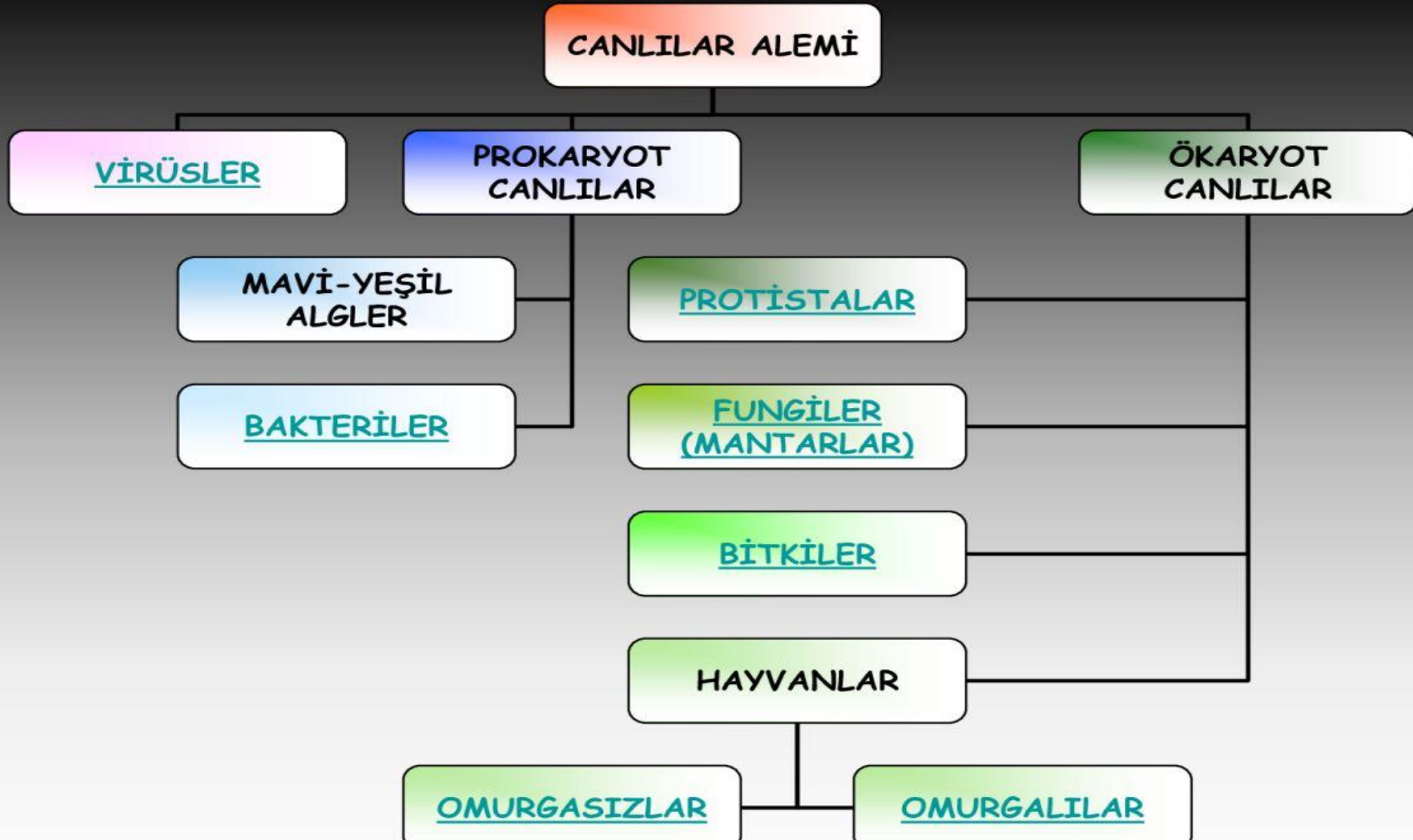


- RNA, DNA'ya çok benzer olmakla beraber bazı yapısal ayrıntılarında farklılık gösterir.
- Hücre içinde RNA genelde tek zincirli, DNA ise genelde çift zincirlidir.
- RNA nükleotitleri riboz içerirler, DNA ise deoksiriboz vardır.
- DNA'da bulunan timin bazı yerine RNA'da urasil vardır ve genelde RNA'daki bazlar ayrıca kimyasal modifikasyona uğrar.

Temel Kavramlar-1



Temel Kavramlar-1



Temel Kavramlar-1



Prokaryotlar Canlılar

- Prokaryotlar ökaryotlara göre daha az evrimleşmiş canlılardır.
- Zarla çevrili bir çekirdeği yoktur.
- Ribozom dışında bir organeli bulunmamaktadır.
- Hem DNA hem RNA bu canlılarda sitoplazmaya dağılmış haldedir.

Temel Kavramlar-1



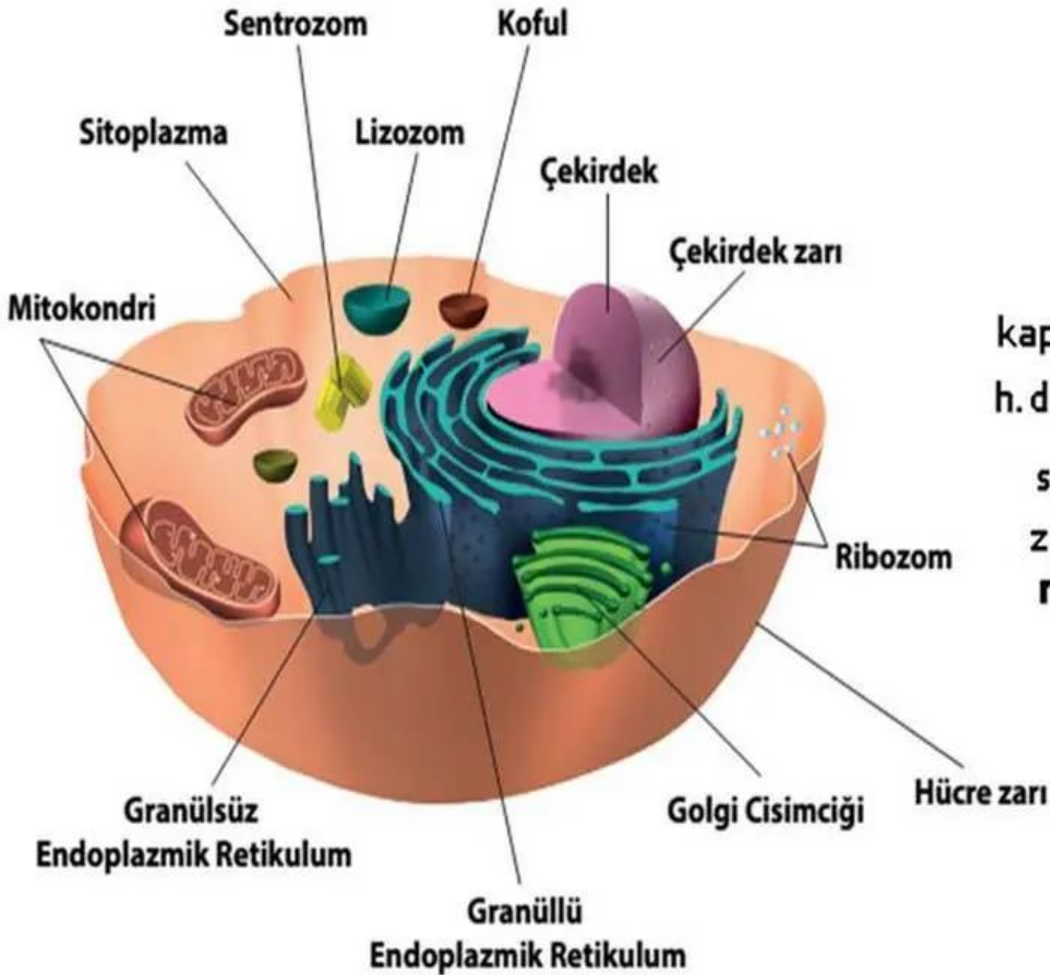
Ökaryot Canlılar

- Ökaryotlar ise evrimin daha üst seviyesinde yer alan canlılardır.
- Hücre zarı, zarla çevrili çekirdek ve pek çok organelle sahiptir.
- DNA, bu hücrelerde çekirdeğin içinde bulunurken RNA hem çekirdek hem sitoplazmada bulunabilir.
- Protein sentezinin gerçekleşebilmesi için RNA'nın taşınabilir olması gerekmektedir.

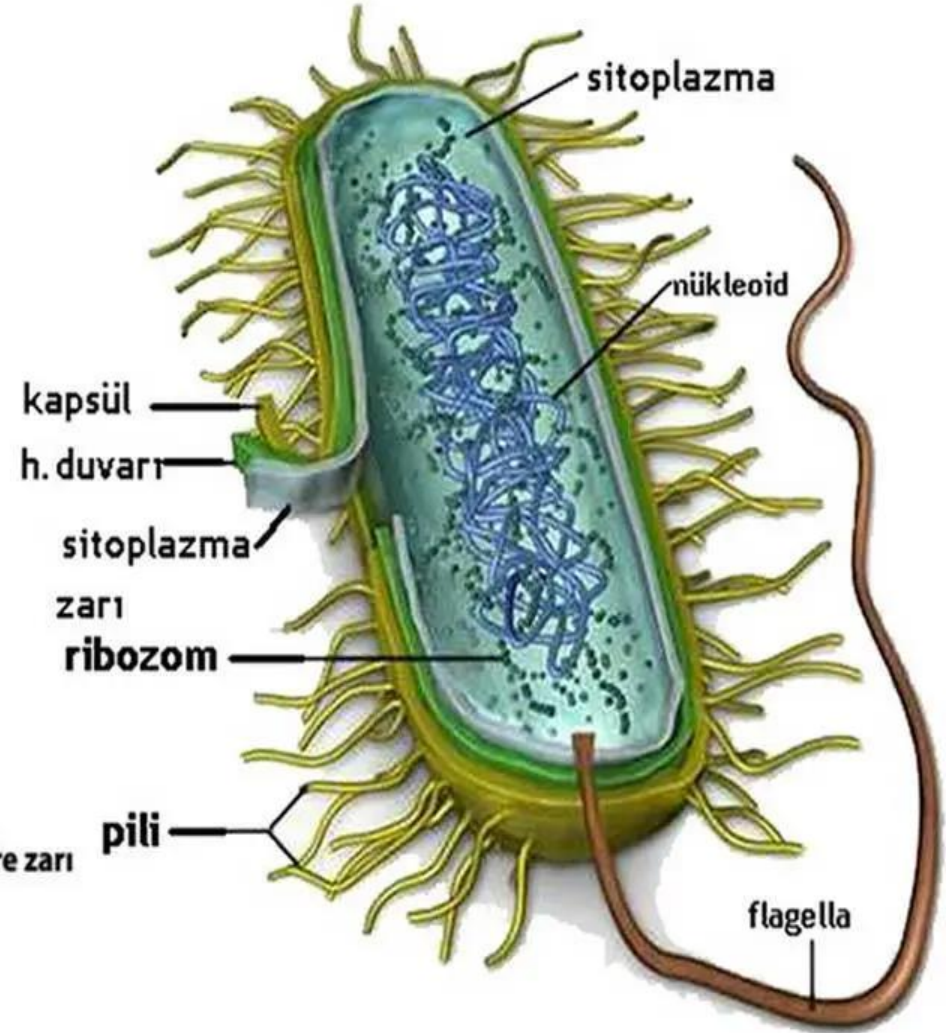
Temel Kavramlar-1



ÖKARYOT HÜCRE



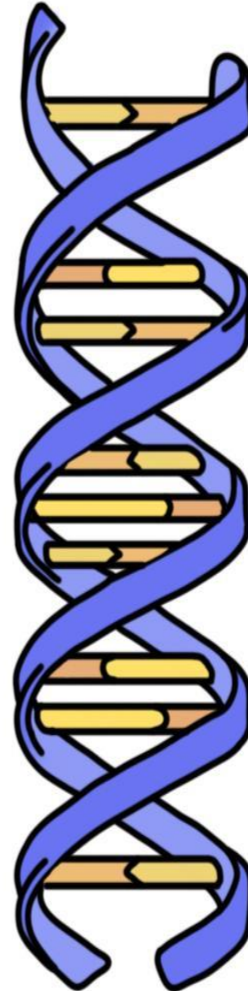
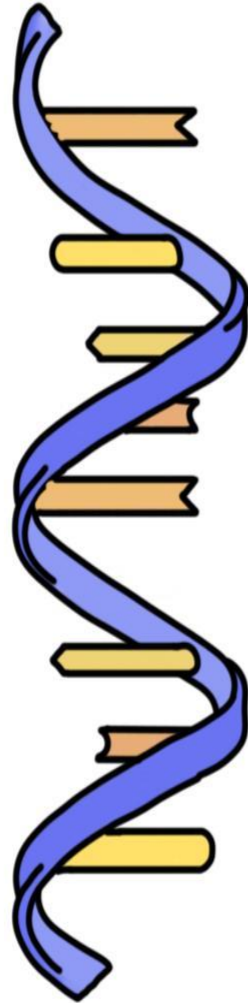
PROKARYOT HÜCRE



Temel Kavramlar-2

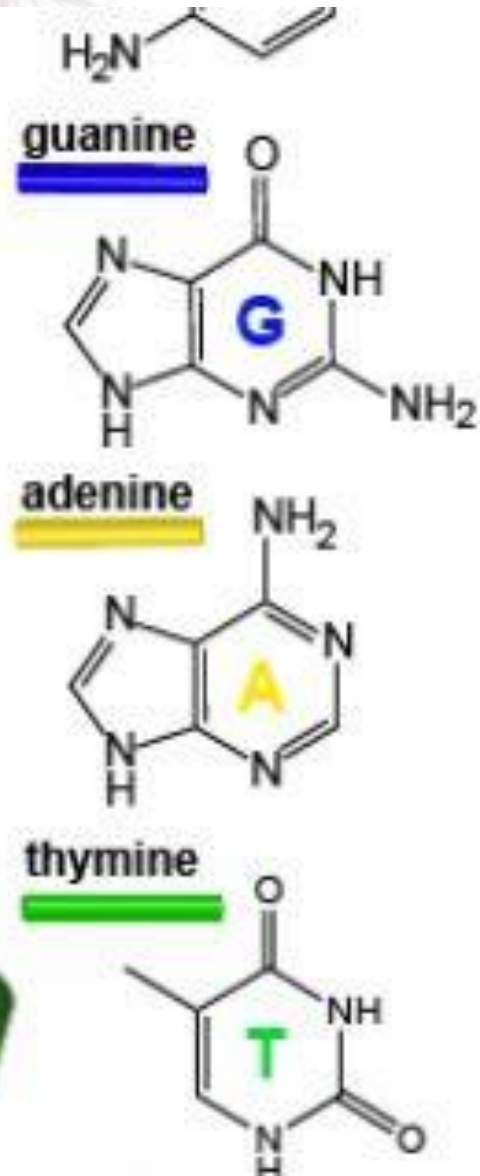
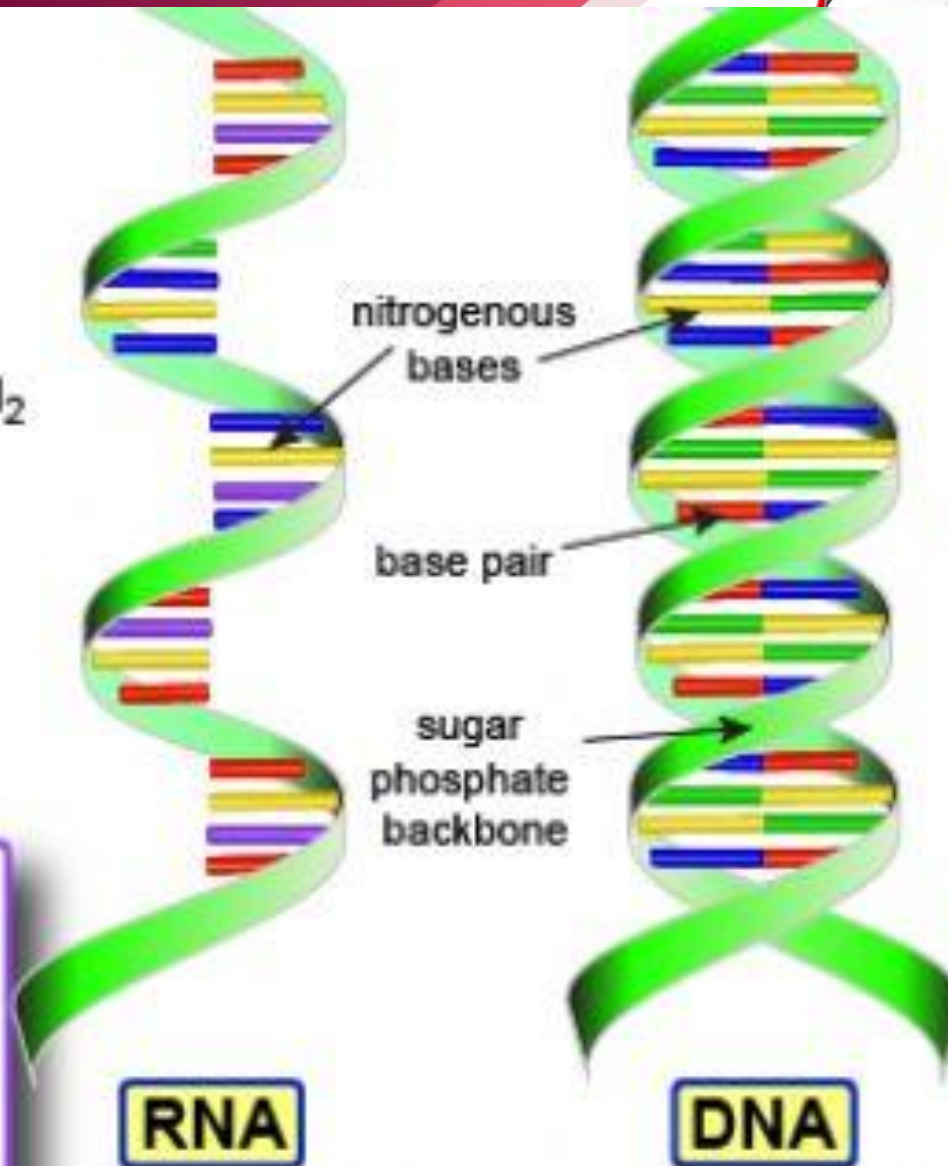
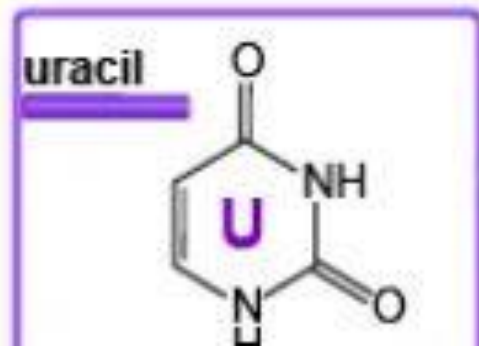


RNA



DNA

Temel Kavramlar-2



Temel Kavramlar-2



DNA ve RNA arasındaki farklar

1	Bazları A,G,C,T	1	Bazları A,G,C,U
2	Kendine özgü bazı timindir.	2	Kendine özgü bazı urasildir.
3	Genetik bilgiyi taşırlar	3	Protein sentezini yaparlar.
4	Çift ipliklidir.	4	Tek tek ipliklidir.
5	5C'lu şekeri deoksiribozdur.	5	5C'lu şekeri ribozdur.
6	DNA polimeraz ile sentezlenir.	6	RNA polimeraz ile sentezlenir.
7	Zayıf hidrojen bağı bulunur.	7	Zayıf hidrojen bağı bulunmaz. (tRNA hariç)
8	Deoksiribonükleaz (DNAaz) ile hidrolize olur.	8	Ribonükleaz (RNAaz) ile hidroliz olur.
9	DNA, çift iplikli olduğu için yapısındaki hataların bazılarını onarabilir.	9	RNA, tek iplikli olduğu için yapısındaki hatalar onarılamaz.

Temel Kavramlar-2



DNA ve RNA arasındaki farklar

10	A=T, G=C eşitliği vardır.	10	A=U, G=C eşitliği zorunlu değildir
11	Kendini eşleyebilir.	11	Eşleyemez.(Bazı virüsler hariç). DNA tarafından sentezlenir.
12	Tek çeşittir.	12	Üç çeşidi vardır. (mRNA, tRNA, rRNA, 16 S rRNA, 23S rRNA, snRNA, hnRNA)
13	Prokaryot hücrede sitoplazmada bulunur.	13	Prokaryot hücrede sitoplazmada ribozomda bulunur.
14	Ökaryot hücrede çekirdek, mitokondri ve kloroplastta bulunur.	14	Ökaryot hücrede çekirdek, çekirdekçik, sitoplazma, ribozom, mitokondri ve kloroplastta bulunur.

Temel Kavramlar-2



DNA ve RNA arasındaki farklar

15	Her türün diploit hücresinde miktarı sabittir.	15	Hücreden hücreye miktarı değişir.
16	DNA'da gen bulunur.	16	RNA'da gen bulunmaz.
17	Yöneticidir. Emir verir.	17	DNA'nın emirlerini uygular.
18	<ul style="list-style-type: none">- DNA eşlenmesi sentezi (replikasyonu) sırasında;- Nükleotidler hazır olarak kullanılırsa, $n-2$ su oluşur.- DNA'nın en küçük bileşenlerinden (fosfat, şeker ve bazlardan) sentezlenmesi sırasında toplam $3n-2$ molekül su oluşur. (n=nükleotid sayısıdır.)	18	RNA sentezinde $3n-1$ tane su molekülü açığa çıkar. (n =nükleotid sayısıdır.)
19	Ph değeri 8'dir	19	Ph değeri 4.7'dir
20	Uzun süreli depolama sıcaklığı -20°C 'dir	20	Uzun süreli depolama sıcaklığı -80°C 'dir
21	Sürekli dir.	21	Geçici dir.

Temel Kavramlar-3



mRNA, tRNA, rRNA, tmRNA

Hücre içindeki RNA tipleri;

1.mRNA (haberci RNA): DNA'daki genetik bilginin proteine çevrilmesinde aracılık eden RNA'lardır.

2.rRNA (ribozomal RNA): Ribozom alt birimlerini oluşturan ve protein sentezinde görev alan RNA'lardır.

Temel Kavramlar-3



mRNA, tRNA, rRNA, tmRNA

3.t-RNA (taşıyıcı RNA) : Protein sentezi sırasında aktifleşmiş aa (amino asit)'leri mRNA'daki kodon dizisine uygun olarak ribozom-mRNA kompleksine taşır.

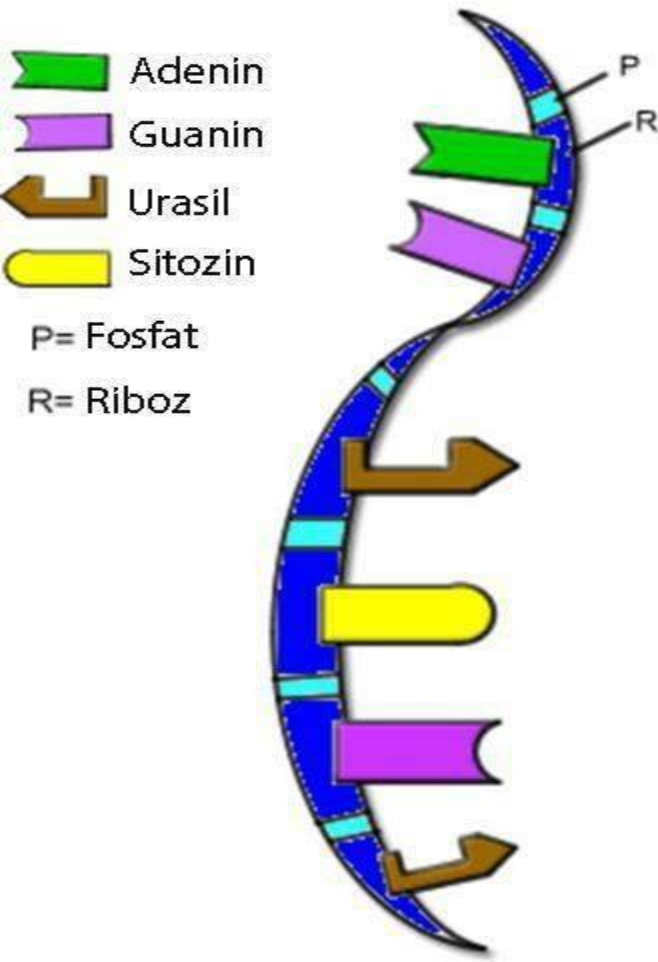
4.snRNA (küçük nükleer RNA): Küçük nükleer RNA'lar yeni sentezlenen pre- mRNA'ların işlenmesi sırasında kesip-ekleme işlemlerinde görevli olan RNA'lardır.

5.snoRNA (küçük nükleolar RNA): Küçük nükleolar RNA'lar çekirdekçikte rRNA'ların işlenmesinde görev alan RNA'lardır

Temel Kavramlar-3

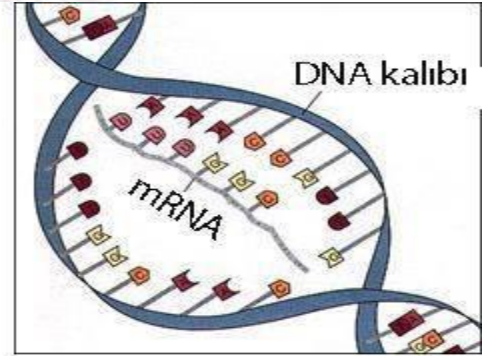


RNA



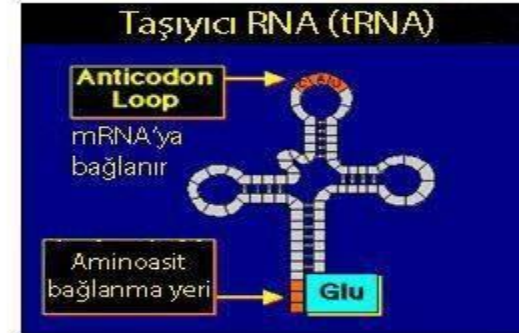
mRNA:

Protein sentezi için gerekli genetik bilgiyi nükleustan sitoplazmaya taşır.



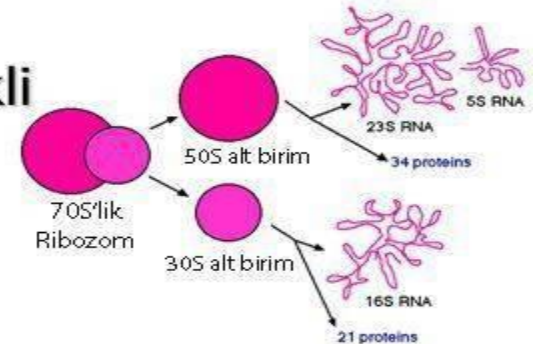
tRNA:

mRNA'daki bilgiyi çözer.



rRNA:

Protein sentezi için gerekli hücresel yapı olan ribozomun bileşenidir. (%50sini oluşturur.)



Temel Kavramlar-5



Kromozom

Kromozom: Hayvan ve bitki hücrelerinde çekirdek içinde konumlanmış, ipliksi bir yapıdır. Kromozomlar çok küçük boyutlarda olup sadece mikroskop yardımıyla görüntülenebilir.

Kromatid: Kromatit, DNA'nın hücre bölünmesi esnasında sentezlenen, protein bir zarf ile paketlenerek sentromer ile birbirine tutturulmuş iki kopyasından her biridir. Kromatitlere aynı zamanda eş kromozomlar denir.

Sentromer: Sentromer, bir çift kardeş kromatidi birbirine bağlayan bir kromozomun DNA dizisidir.

Temel Kavramlar-5

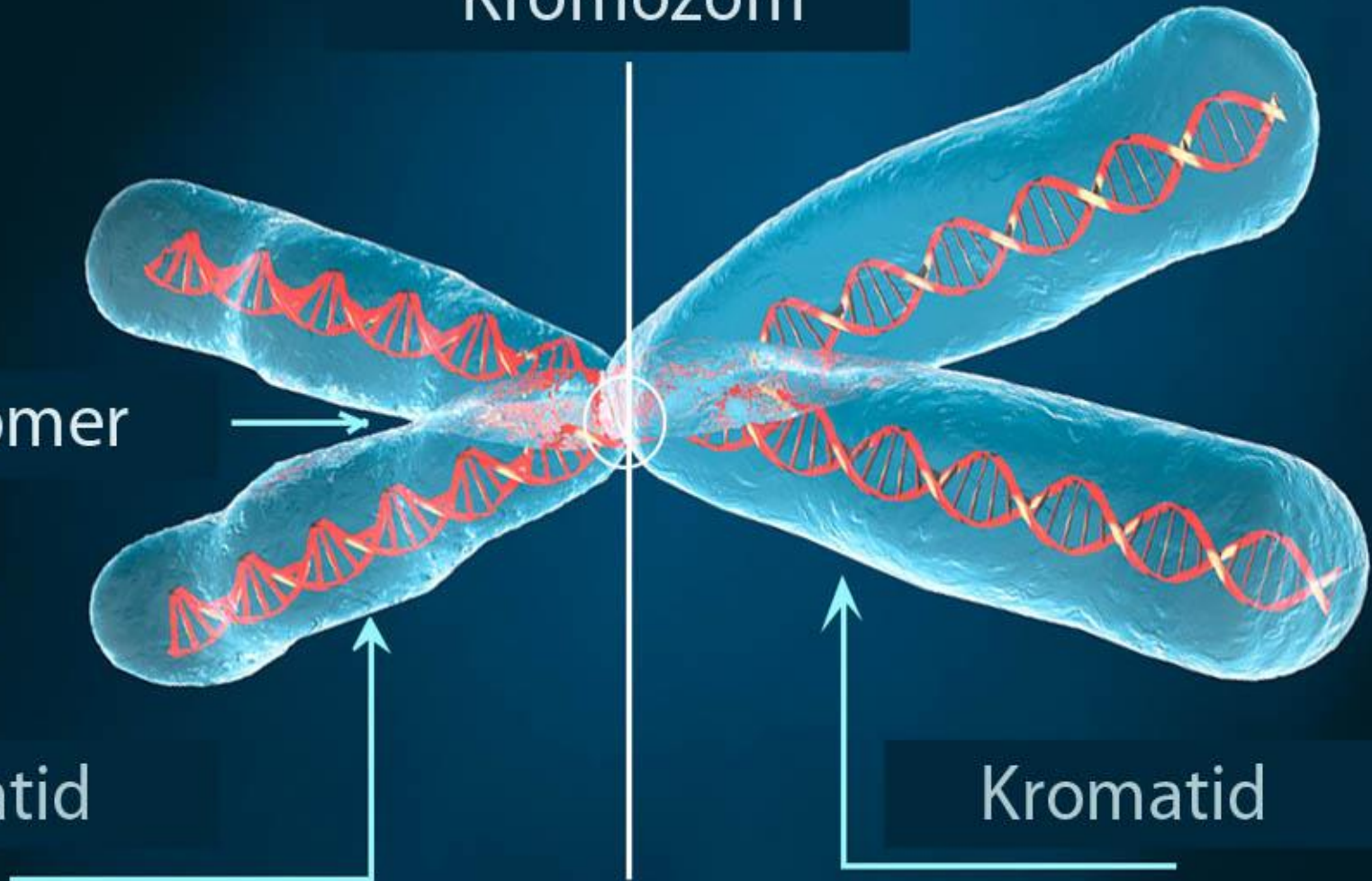


Kromozom

Sentromer

Kromatid

Kromatid



Temel Kavramlar-5



Bazı Türlerin Kromozom Sayıları

- İnsan 46 Kromozom
- Domuz 38 Kromozom
- Sığır 60 Kromozom
- Eşşek 62 Kromozom
- At 64 Kromozom
- Kromozom sayısının fazlalığı türün gelişmişliğini ifade etmez.
- Neden bu türler? **Tağşiş**
- Hindi 62 Kromozom
- Deve 74 Kromozom
- Koyun 54 Kromozom
- Tavşan 44 Kromozom
- Tavuk 80 Kromozom



Tağışış- Bulaşma

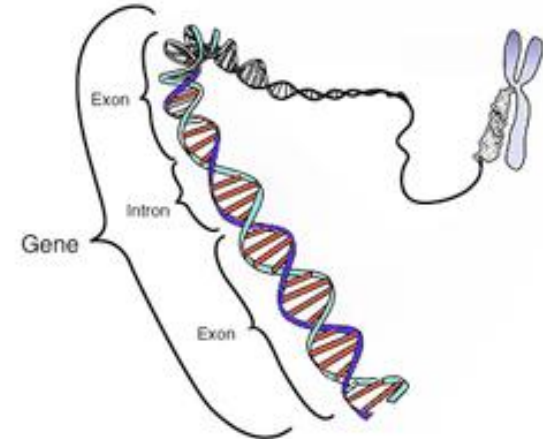
- **Tağışış (Genel):** Değerli bir madenin içersine daha değersiz bir madenin karıştırılmasıdır.
- **Tağışış (Gıda):** Değerli bir gıda ürününün içersine daha değersiz bir gıda ürününün bilerek karıştırılmasıdır.
- **Contamination (Bulaşma):** Gıda ürünlerinin taşıma, üretim ve satışı esnasında bilinçsiz bir şekilde bir birine karışması.

Temel Kavramlar-6

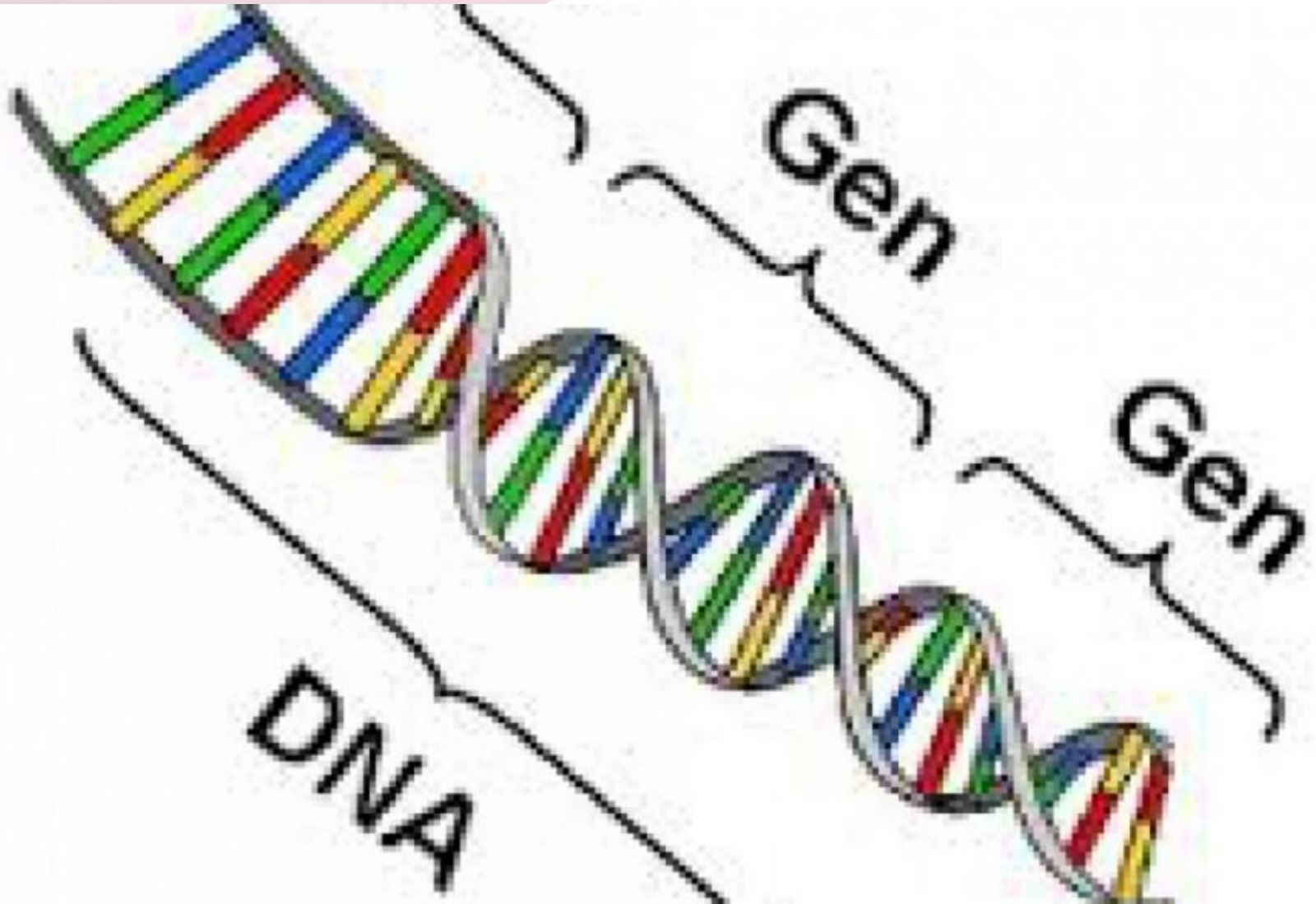


Gen

- Gen, genom dizisinde yeri tanımlanabilen, transkripsiyonu yapılan, düzenleyici ve/veya fonksiyonel bölgeleri olan bir bölgedir.
- Gen, bir kalıtım birimidir. Bir DNA'nın belirli bir kısmını oluşturan nükleotid dizisidir.
- Kromozomun kesitleri olan genler birbirinden çok farklı işlevlerde ve büyüklüklerde (uzunluklarda) olabilirler.
- Genlerin büyüklükleri ve işlevleri her zaman doğru orantılı değildir.



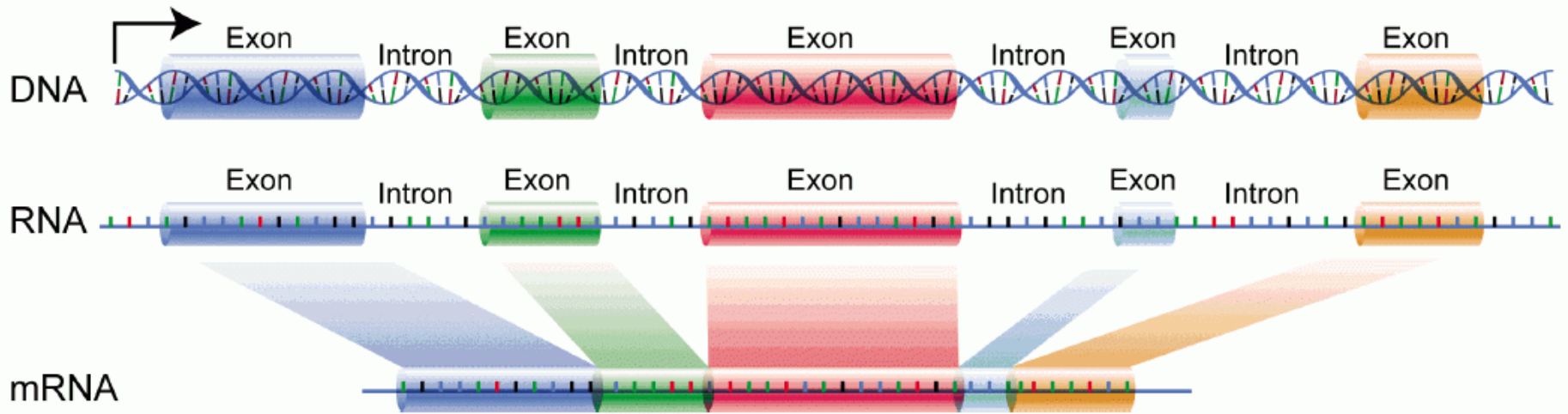
Temel Kavramlar-2



Temel Kavramlar-6

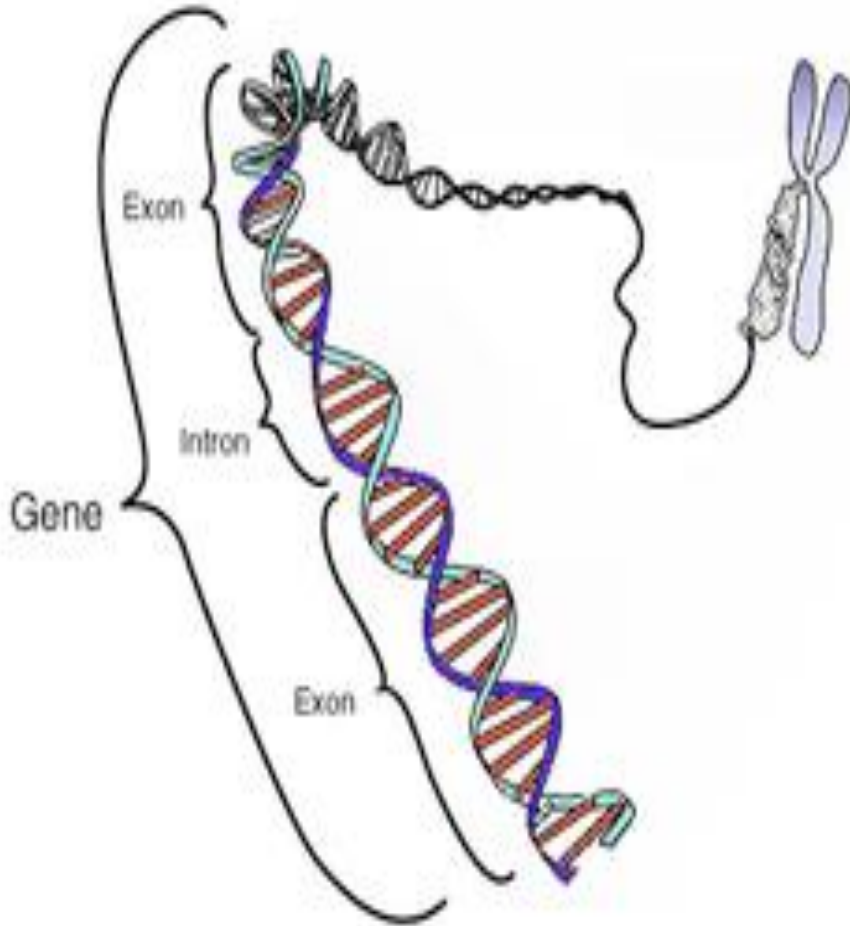


- Protein ve enzimler üretilirken, DNA üzerindeki genlerin harf dizilimi örnek alınarak bu genlere karşılık gelen RNA kopya dizilimleri çıkarılır.



- Kopyalanan bu RNA'lar yapılırken, genin harf dizilimi baştan sona tümüyle okunmaz.
- Bir kısım dizilim okunup kopyası çıkarıldıktan sonra, uzun bir bölüm okunmadan atlanıp başka bir bölüme geçilir ve oradan devam edilir.

Temel Kavramlar-6



Kısaca

Exon: DNA da bulunan anlamlı gen bölgeleridir.

Intron: DNA'nın anlamsız bölgeleri

Genlerin kodlamaya katılmayan bu bölümü, toplam insan genomunun yaklaşık %97'lik bir kısmını oluşturur.

Temel Kavramlar-7



Genom

- Bir organizmanın kalıtım materyalinde bulunan genetik şifrelerin tamamını simgeler.
- Her canlının hücrelerinin içine yerleştirilmiş genetik program, o canlının "genom"unu oluşturur.

Replikasyon

- Bir genetik materyalin kendi kendini eşlemesidir.

Temel Kavramlar-8



Kodon

- Protein sentezi için gerekli bilgiyi taşıyan mRNA da bir amino asidi temsil eden peş peşe üç nükleotitten (triplet) oluşan grup.
- Bilinen 64 kodondan 61 tanesi amino asitlerin her biri için özel, üç tanesi ise polipeptit sentezini bitirmek içindir.



Aminoasitler ve Proteinler

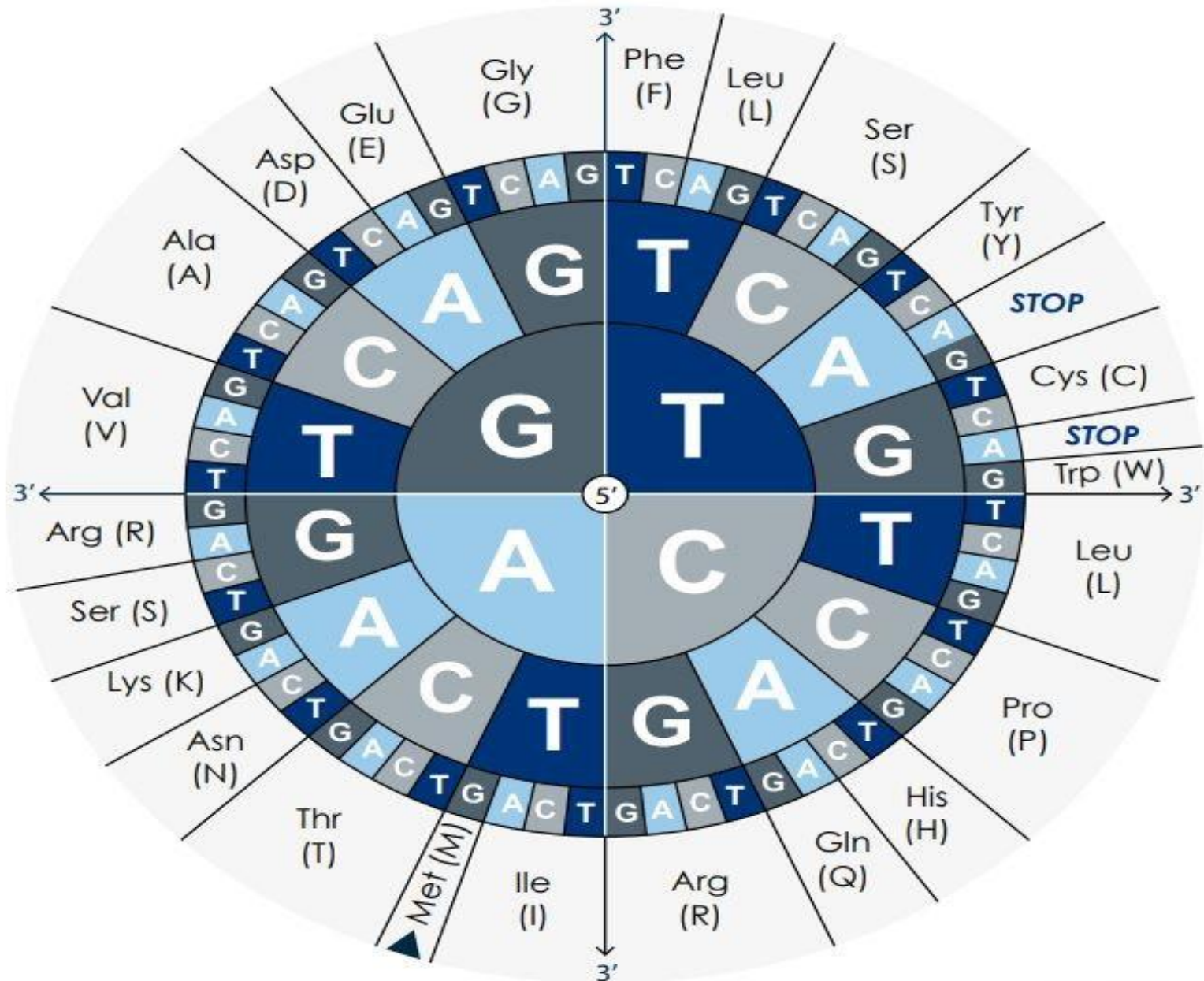
- Amino asitler proteinleri oluşturan temel yapı taşlarıdır.
- Aminoasitlerin peptit bağlarıyla uç uca eklenmesiyle oluşturdukları kısa polimer zincirler "peptit", uzun polimer zincirler ise "polipeptit" veya "protein" olarak adlandırılırlar.
- Hücre içerisinde ribozomlar, mRNA moleküllerini kalıp olarak kullanarak amino asitleri uç uca ekleyerek proteinleri sentezlerler.
- Bu işleme translasyon (okuma) denir.



Aminoasitler ve Proteinler

- Doğada keşfedilmiş 500 çeşit aminoasit mevcuttur.
- İnsan doğası bunlardan yalnız 20 tanesini kodlayabilir.
- Toplam 64 farklı kodon vardır.
- Farklı kodonlar aynı amino asiti oluşturabiliyorken, bir kodon yalnız tek bir amino asiti oluşturur.

cDNA Genetic Code



cDNA Codon Table

Second Position

First Position

	T	C	A	G	
T	TTT } Phe TTC } TTA } Leu TTG }	TCT } Ser TCC } TCA } TCG }	TAT } Tyr TAC } TAA STOP TAG STOP	TGT } Cys TGC } TGA STOP TGG Trp	T C A G
C	CTT } Leu CTC } CTA } CTG }	CCT } Pro CCC } CCA } CCG }	CAT } His CAC } CAA } Gln CAG }	CGT } Arg CGC } CGA } CGG }	T C A G
A	ATT } Ile ATC } ATA } Met ATG }	ACT } Thr ACC } ACA } ACG }	AAT } Asn AAC } AAA } Lys AAG }	AGT } Ser AGC } AGA } Arg AGG }	T C A G
G	GTT } Val GTC } GTA } GTG }	GCT } Ala GCC } GCA } GCG }	GAT } Asp GAC } GAA } Glu GAG }	GGT } Gly GGC } GGA } GGG }	T C A G

Third Position



Boşluk Doldurma-1

- Genetik bilgiyi analiz etmek ve anlamak için bilgisayar bilimleri, bilgi teknolojileri ve genetik bilim dallarından oluşan kombinasyona **biyoinformatik** denir.
- Organizma türünün oluşturduğu toplulukların karakteristik özelliklerine kadar olan bilgi akışı **genetik bilgi akışı** dır.
- İlk biyolojik veritabanı **Protein Identification Resource (PIR)** 'dır.



Boşluk Doldurma-2

- Nükleotidler bulundurdukları **baz'a** göre isimlendirilirler.
Adenin, Guanin, Sitozin, Timin, Urasil
- DNA da bulunan anlamlı gen bölgelerine **Exon** denir.



Doğru Yanlış-1

- (Y) Biyoinformatik matematik ve istatistik gibi bilim dalları ile ilişkili değildir.
- (D) İlk DNA veritabanı 1979'da kuruldu
- (D) İlk biyolojik veritabanı Protein Identification Resource Margaret Dayhoff tarafından 1972'de kuruldu.
- (Y) Bir canlının genetik materyali DNA, RNA ve Proteinden oluşmaktadır.
- (D) Biyolojik bilgiler ile ilgili veritabanı oluşturulup yönetilmesi biyoinformatiğin kapsam alanındadır.



Doğru Yanlış-2

- (D) Nükleotidler şeker, fosfat ve organik bazlardan oluşur.
- (Y) DNA sarmalında Adenin ve Timin arasında 2 adet zayıf hidrojen bağı varken, Guanin ile sitozin arasında 3 adet zayıf hidrojen bağı bulunmaktadır.
- (D) Değerli bir gıda ürününün içersine daha değersiz bir gıda ürününün bilerek karıştırılmasına taşış denir.
- (Y) Calının karmaşıklığı arttıkça kromozom sayısında artar.

Sorular



Klasik-1

- Biyoinformatiğin amaçları nelerdir ?
- Canlıların genetik materyalinin küçükten büyüğe sıralınışı nasıldır?

Sorular



Teşekkür



**thanks for
listening!**

Biyoinformatik





Dr. Öğr. Üyesi

Mehmet MİLLİ

Oda: 330

Mail: mehmetmilli@ibu.edu.tr



Alignment

- Alignment Nedir?
- Neden İhtiyaç Vardır?
- Alignment Türleri
 - Global Alignment
 - Local Alignment
- Alignment Algoritmaları
- Global Hizalama Algoritmaları
 - Needleman Wunsch Algoritması
- Local Hizalama Algoritmaları
 - Smith-Waterman algoritması

Alignment Nedir?



Alignment Nedir?

- Biyoenformatikte dizi hizalaması, **DNA, RNA veya protein dizilerini** düzenleyerek benzer bölgelerin tespit edilmesidir.
- Diziler arasında benzer bölgelerin bulunması, diziler arasında işlevsel, yapısal veya evrimsel bir ilişki olduğu anlamına gelir.

Alignment Nedir?



- İki dizi ortak bir ataya sahipse, hizalamada eşleşmeyen harfler, bu dizilerin birbirinden ıraksadığı zamandan beri meydana gelmiş noktasal mutasyonlar olarak yorumlanır.
- Hizalamadaki boşluklar ise insersiyon veya delesyonlar (indel) olarak yorumlanır.
- İnsersioyon=Ekleme
- Delesyon =Silme
- İndel

Alignment Neden Yapılır?



Alignment Neden Yapılır?

1. DNA/Protein Sekanslarının benzerlik ve homoloji tespiti amacıyla kıyaslanması
2. Gen/Protein fonksiyon tahmini
3. Filogenetik (Evrimsel ilişki) ağaçlarının oluşturulması
4. Protein motiflerinin bulunması
5. Shotgun Sekans analizi için birleştirme işlemleri

Alignment Türleri?



A. Alignment Türleri

- a) Global Alignment
- b) Local Alignment

Alignment Türleri?



a) Global Alignment:

- iki dizinin gerekirse boşluklar kullanılarak eşit uzunlukta olacak biçimde hizalanmaya zorlanmasına Global Alignment denmektedir.

Global Alignment

Target Sequence

5' ACTACTAGATTACTTACGGATCAGGTACTTTAGAGGCTTGCAACCA 3'

||||| | | | | | | | | | | | | | | | | | | | | | | | | | | |

5' ACTACTAGATT---ACGGATC--GTACTTTAGAGGCTAGCAACCA 3'

Query Sequence

Alignment Türleri?



a) Global Alignment:

- Global hizalamalarda her dizideki her harfin hizalanması amaçlanır.
- Sorgu kümesindeki diziler birbirine benzer ve yaklaşık aynı uzunlukta olursa global hizalama çok daha verimli olur.
- Ama bu, global hizalamaların boşluklarla sonlanamayacağı anlamına gelmez.

Alignment Türleri?



a) Global Alignment:

- Birbirine çok benzeyen ve ortak ataları olma olasılığı yüksek olan türler arasında global hizalama yapmak çok daha doğru bir yaklaşımdır.
- O halde kromozomların veya türlerin homolog olup olmadığına karar vermek için global hizalama teknikleri kullanılmalıdır.

Alignment Türleri?



b) Local Alignment:

- Gen dizilerinin bire bir tamamını değilde sadece belirli bir kısmını hizalama yöntemine yerel hizalama (local alignment) denir.

Local Alignment

Target Sequence

5' ACTACTAGATTACTTACGGATCAGGTACTTTAGAGGCTTGCAACCA 3'

||||| ||||| ||||| ||||| |||||

Query Sequence

5' TACTCACGGATGAGGTACTTTAGAGGC 3'

Alignment Türleri?



b) Local Alignment:

- Birbirine benzemeyen ama benzer bölgeler içerdiği tahmin edilen diziler için lokal hizalamalar daha yararlıdır.
- Bunun yanında nükleotid sayıları (eleman sayıları) arasında büyük farklar olmsdı durumunda da yine local hizalama kullanılmalıdır.



En Uzun Ortak Altdizi (Longest Common Subsequences)

Benzerlik analizinin en basit halidir.

Yalnızca baz ya da amino asitler için eklemelere (insertions) ve silmelere (deletion) izin verilir.

Bir v dizisinin alt dizisi (subsequence) basitçe (sıralı) bir v 'den gelen bir karakter dizisidir.



En Uzun Ortak Altdizi (Longest Common Subsequences)

Örneğin, $v = \mathbf{ATTGCTA}$ şeklinde bir nükleotid dizimiz olsun.

AGCA ve **ATTA** dizilerine v 'nin alt dizileri denebilir.

Fakat **TGTT** ve **TCG** diziler v 'nin alt dizisi değildir.

Burada önemli olan alt dizinin ana diziye soldan sağa baz silinmesi veya eklenmesiyle uyumlu olmasıdır.



En Uzun Ortak Altdizi (Longest Common Subsequences)

İki dizinin hizalanmasında amaç mümkün olan **en fazla** veya mümkün olan **en büyük ortak** alt dizinlerin bulunmasıdır.

Yani bulunacak olan **common subsequences (ortak alt dizin)** iki ayrı dizininde alt dizisi olmalıdır.

Örneğin ATCTGAT ve TGCATA şeklinde iki nükleotid dizisi olsun.

TCTA alt dizisi bu diziler için ortak alt dizilerden biridir. Hatta bu diziler için LCS'dir denebilir.

İndeller ve Skorlama



Boşlukların Önemi ve Skorlama

Hizalama sırasında en uygun eşleşmelerin bulunabilmesi için indellerin kullanılabileceğini söylemiştik.

O zaman eşleşme matrisimde ki bir nükleotide karşılık gelen olası ihtimaller;

- Doğru Eşleşmeler
- Yanlış Eşleşmeler
- Boşluklar

O zaman bir sorun çıkıyor karşımıza.....

Olası değişik ihtimallerde hangi alt kümeleri en iyi eşleşme olarak seçilmesi gerekmektedir?

İndeller ve Skorlama



Boşlukların Önemi ve Skorlama

İki basit nükleotit dizisini hizalamaya çalışalım

Bu iki dizi aşağıdaki gibi olsun;

1. Dizi=AGGCTAGTT
2. Dizi=AGCGAAGTTT .

Bu iki diziye hizaladığımızda hangi ihtimaller ortaya çıkar veya bu iki diziye nasıl hizalar sınız?

İndeller ve Skoring



Boşlukların Önemi ve Skoring

Birinci Olasılık;

AGGCTA-GTT-
AG-CGAAGTTT

Burada 7 doğru eşleşme, 1 hatalı eşleşme ve 3 boşluk bulunmaktadır.

İkinci Olasılık;

AGGC-TA-GTT-
AG-CG-AAGTTT

Bu hizalamada ise 7 doğru eşleşme, 0 hatalı eşleşme ve 5 boşluk bulunmaktadır.

İndeller ve Skoring



Boşlukların Önemi ve Skoring

Hangisi türleri kıyaslamak için daha faydalı? Veya hangisini tercih etmeliyiz.

Yazacağımız algoritma bize çözüm olarak

7 doğru eşleşme, 1 hatalı eşleşme ve 3 boşluk

yoksa

7 doğru eşleşme, 0 hatalı eşleşme ve 5 boşluk

Vermeli.

İndeller ve Skorlama



Boşlukların Önemi ve Skorlama

Açıkça görülüyorki tercih yapabilmek için bir skarlama sistemine ihtiyacımız var.

Bu skarlama sistemi her bir, doğru eşleşme, yanlış eşleşme ve boşluk için bir skor vermeli ve bunların toplamı ile toplam skoru elde etmeli.

Örneğin, bir skarlama algoritması, **m** doğru eşleşmeler için **s** yanlış eşleşmeler için ve **d** boşluklar için verilen skoru temsil etmek üzere aşağıdaki biçimde yazılabilir.

$$F = (\text{doğru eşleme sayısı}) \times m - (\text{yanlış eşleme sayısı}) \times s - (\text{boşluk sayısı}) \times d$$

İndeller ve Skorlama



Boşlukların Önemi ve Skorlama

Bu noktada sadece matematik değil biyoloji ile de bu probleme yaklaşmalıyız.

Örneğin aşağıdaki gibi ardışık bir delesyon işlemi,

AATCTATA
AA- -GATA

pek çok defa oluşan delesyonlardan,

AATCTATA
AA-G-ATA

çok daha olasıdır.

İndeller ve Skarlama



Boşlukların Önemi ve Skarlama

Bu durumu matematiksel olarak yeni açılan boşluklar için -3 puan verirken, devam eden boşluklar için -1 puan vererek gösterebiliriz. Skarlama algoritmaları geliştirilirken bu ve benzeri pek çok durum dikkate alınmalıdır.

Skarlama Puanları

Doğru Eşleşme=1 Puan

Yanlış Eşleşme=-1 Puan

İndel Puanı->Eğer Yeni Açılan Bir Boşluk ise -3

->Eğer Ardaşık Devam Eden Bir Boşluk ise -1 olarak belirlediğimizde.

A A T C T A T A

A A - - G A T A

1 1 -3-1-11 1 1

Toplam Puan=0

A A T C T A T A

A A - G - A T A

1 1 -3-1-31 1 1

Toplam Puan=-2

Global Alignment Algoritması



Needleman-Wunsch Algoritması

İki diziye birbiri ile hizalarken izlenebilecek pekçok farklı alternatif bulunur.

Matematiksel olarak yaklaşık (N hizalanacak dizilerin toplam uzunluğu olmak üzere) yaklaşık 2^N alternatif bulunduğunu söyleyebiliriz.

Needleman & Wunsch algoritması benzer uzunluktaki dizilerin hizalanması için kullanılır ve genelde benzer genler ile büyük bölgelerin hizalanmasında fayda sağlar.

Global Alignment Algoritması



Needleman-Wunsch Algoritması

Bütün alternatifleri denemek dizi uzunluğu arttığında çok vakit almakta ve bazı durumlarda imkansız hale gelmektedir.

Dolayısı ile kalabalık çözüm uzayı içersinden en iyi çözümlerden birini bulmak için bilgisayar bilimlerinde kullanılan dinamik algoritma mantığından yararlanılması gerekmektedir.

Global Alignment Algoritması



Dinamik Programlama Nedir?

Dinamik programlama temel olarak bir problemi aynı biçimde çözülebilecek parçalara bölüp bu parçaları çözmeyi ifade eder.

Bu çözüm sırasında dinamik programlama bir önceki adımlarda yapılan hesaplamayı da kullanır.

Global Alignment Algoritması



Dinamik Programlama Nedir?

Dinamik programlamada hesaplanan değerler bir tabloda/matrisde tutulur.

Daha detaylı bir anlatımla önce tablonun ilk satır ve sütununu hesaplar ve buradan yola çıkarak diğer satır sütunları hesaplayarak devam eder.

Dolayısı ile b işlemi gerçekleştirmek için n ve m dizi uzunlukları olmak üzere $n*m$ boyutunda bir tablo oluşturulması gerekmektedir.

Global Alignment Algoritması



1. Adım Tabloların Oluşturulması

Örnek: İki dizi üzerinde Needleman Wunsch algoritmasını uygulayalım.

1. Dizi=SEND
2. Dizi=AND

Needelman Wunsch algoritmasını uygularken ilk adım 5x4 büyüklüğünde iki matris oluşturmaktır.

Burada 5 birinci dizisinin elaman sayısının bir fazlası ve 4 ise ikinci dizisinin elaman sayısının bir fazlasını temsil etmektedir.

Global Alignment Algoritması



1. Adım Tabloların Oluşturulması

	S	E	N	D	
A	$C(1,1)$	$C(1,2)$	$C(1,3)$	$C(1,4)$	$C(1,5)$
	$C(2,1)$	$C(2,2)$	$C(2,3)$	$C(2,4)$	$C(2,5)$
	$C(3,1)$	$C(3,2)$	$C(3,3)$	$C(3,4)$	$C(3,5)$
	$C(4,1)$	$C(4,2)$	$C(4,3)$	$C(4,4)$	$C(4,5)$

	S	E	N	D

Burada $i = 1, 2, \dots, N$ ve $j = 1, 2, \dots, M$ olmak üzere $C(i, j)$ ler için tablo;



2. Skorlama Kurallarının Belirlenmesi

Doğru Eşleşme=+6 Puan

Yanlış Eşleşme= +1 Puan

İndel Eşleşme =-10 Puan

Global Alignment Algoritması



3. Tablonun Birinci satır ve sütunun doldurulması

Bu durumda her açılan boşluk için -10 puan verecek olursak tablomuzun ilk satır ve sütununu aşağıdaki biçimde doldurmuş oluruz.

	S	E	N	D	
	0	-10	-20	-30	-40
A	-10				
N	-20				
D	-30				



4. Tablonun Diğer satır ve sütunun doldurulması

Tablonun ilk satır ve sütunu ceza puanı baz alınarak doldurulduktan sonra diğer hücrelerin doldurulması işlemine geçilebilir.

Tabloda bulunan her bir hücre için üç alternatif bulunmaktadır.

1. Birinci dizinin sıradaki elemanını, ikinci dizinin sıradaki elemanı ile hizalamak.
2. Birinci dizinin sıradaki elemanını, boşluk ile hizalamak, yani ikinci dizide bir boşluk bırakmak.
3. İkinci dizinin sıradaki elemanını, boşluk ile hizalamak, yani birinci dizide bir boşluk bırakmak.

Global Alignment Algoritması



4. Tablonun Diğer satır ve sütunun doldurulması

İlk Pozisyon	Skor	Kalan Dizi
S A	1	END ND
– A	-10	SEND ND
S –	-10	END AND



4. Tablonun Diğer satır ve sütunun doldurulması

1- Sol üst hücreden diyagonal (çapraz) olarak ilerleyerek dizinin hedef hücrede kesişen harfleri için bir skor hesaplar ($s(a,b)$) ve bu skoru geldiğimiz hücrenin skoru ile toplayarak hedef hücrenin skorunu bulur ve geçici olarak kaydederiz.



4. Tablonun Diğer satır ve sütunun doldurulması

2- Soldaki hücreden sağa doğru yatay olarak ilerleyerek dizinin hedef hücrede kesişen harf ve boşluk için bir skor hesaplar ($s(a,-)$) ve bu skoru geldiğimiz hücrenin skoru ile toplayarak hedef hücrenin skorunu bulur ve geçici olarak kaydederiz.



4. Tablonun Diğer satır ve sütunun doldurulması

3- Üstteki hücreden aşağı doğru dikey olarak ilerleyerek dizinin hedef hücrede kesişen harf ve boşluk için bir skor hesaplar ($s(-,b)$) ve bu skoru geldiğimiz hücrenin skoru ile toplayarak hedef hücrenin skorunu bulur ve geçici olarak kaydederiz.

Global Alignment Algoritması

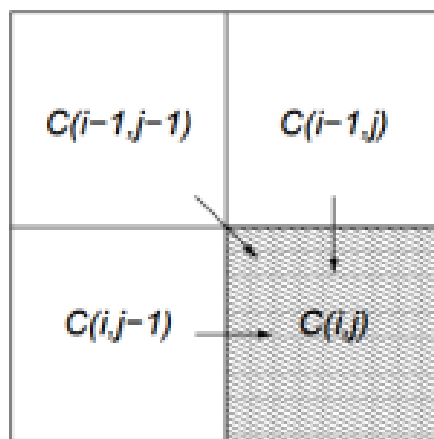


4. Tablonun Diğer satır ve sütunun doldurulması

$$q_{diag} = C(i-1, j-1) + S(i, j)$$

$$q_{up} = C(i-1, j) + g$$

$$q_{left} = C(i, j-1) + g$$





4. Tablonun Diğer satır ve sütunun doldurulması

$$q_{diag} = C(1, 1) + S(S, A) = 0 + 1 = 1$$

$$q_{up} = C(1, 2) + g = -10 + (-10) = -20$$

$$q_{left} = C(2, 1) + g = -10 + (-10) = -20$$

Global Alignment Algoritması



4. Tablonun Diğer satır ve sütunun doldurulması

Her hücre için aynı hesapları yapar ve en yüksek skoru kaydederiz ve bir sonraki hücreye ilerleriz

		S	E	N	D
A	0	-10	-20	-30	-40
	-10	1	?		
	-20				
	-30				

	S	E	N	D
done	left	left	left	left
up	diag	?		
up				
up				

Global Alignment Algoritması



4. Tablonun Diğer satır ve sütunun doldurulması

		S	E	N	D
A	0	-10	-20	-30	-40
	-10	1	-9	-19	-29
	-20	-9	-1	-3	-13
	-30	-19	-11	2	3

	S	E	N	D
done	left	left	left	left
up	diag	left	left	left
up	diag	diag	diag	left
up	up	diag	diag	diag

Global Alignment Algoritması



5. Traceback yaparak hizalanmış dizilerin oluşturulması

Bu şekilde tüm hücreleri doldurduktan sonra son hesapladığımız hücreden (sol alt köşe) başlayarak sol üst köşeye doğru geriye ilerleriz. Geriye doğru ilerlerken skor hesaplama sırasında geldiğimiz hücreleri kullanırız.

		S	E	N	D	
		done	left	left	left	left
A	up	diag	left	left	left	
N	up	diag	diag	diag	left	
D	up	up	diag	diag	diag	

Traceback starts here

Needleman-Wunsch

match = 1

mismatch = -1

gap = -1

		G	C	A	T	G	C	U
	0	-1	-2	-3	-4	-5	-6	-7
G	-1	1	0	-1	-2	-3	-4	-5
A	-2	0	0	1	0	-1	-2	-3
T	-3	-1	-1	0	2	1	0	-1
T	-4	-2	-2	-1	1	1	0	-1
A	-5	-3	-3	-1	0	0	0	-1
C	-6	-4	-2	-2	-1	-1	1	0
A	-7	-5	-3	-1	-2	-2	0	0

The diagram illustrates the Needleman-Wunsch algorithm for sequence alignment. The matrix shows the scores for aligning the sequence GGCATGCU (top) with GTTACUA (left). The optimal alignment path is highlighted with blue arrows, showing a match between G and G, a mismatch between C and T, a match between A and A, a match between T and T, a mismatch between G and A, a match between C and C, and a match between U and U. The final score is 0.



5. Traceback yaparak hizalanmış dizilerin oluşturulması

- 1- Eğer çapraz olarak ilerlediyssek bulunduğumuz hücrede kesişen hücreleri karşılıklı olarak hizalarız.
- 2- Eğer yatay olarak ilerlediyssek, tabloya yatay olarak yazdığımız dizide hücrenin dikey ekseninde karşısına gelen dizi elemanın karşısına bir boşluk açarız.
- 3- Eğer dikey olarak ilerlediyssek, tabloya dikey olarak yazdığımız dizide hücrenin yatay ekseninde karşısına gelen dizi elemanın karşısına bir boşluk açarız.

SEND

A-ND

Sonuç

		S	E	N	D
		done	left	left	left
A		up	diag	left	left
N		up	diag	diag	left
D		up	up	diag	diag

Traceback starts here

Local Alignment Algoritması



Smith-Waterman algoritması

Yerel hizalama için kullanılan Smith & Waterman algoritması ise farklı uzunluklardaki ve birbirine benzemeyen bölgelerin bulunması için fayda sağlar.



Klasik-1

- Neden Alignment yapılır?
- Alignment hangi konularda araştırmacılara faydalı olur?
- Alignment türleri nelerdir?
- Global Alignment ve local Alignment hakkında bilgi veriniz.
- Global Alignment ve local Alignment hakkında bilgi veriniz.

Sorular



Teşekkür



**thanks for
listening!**