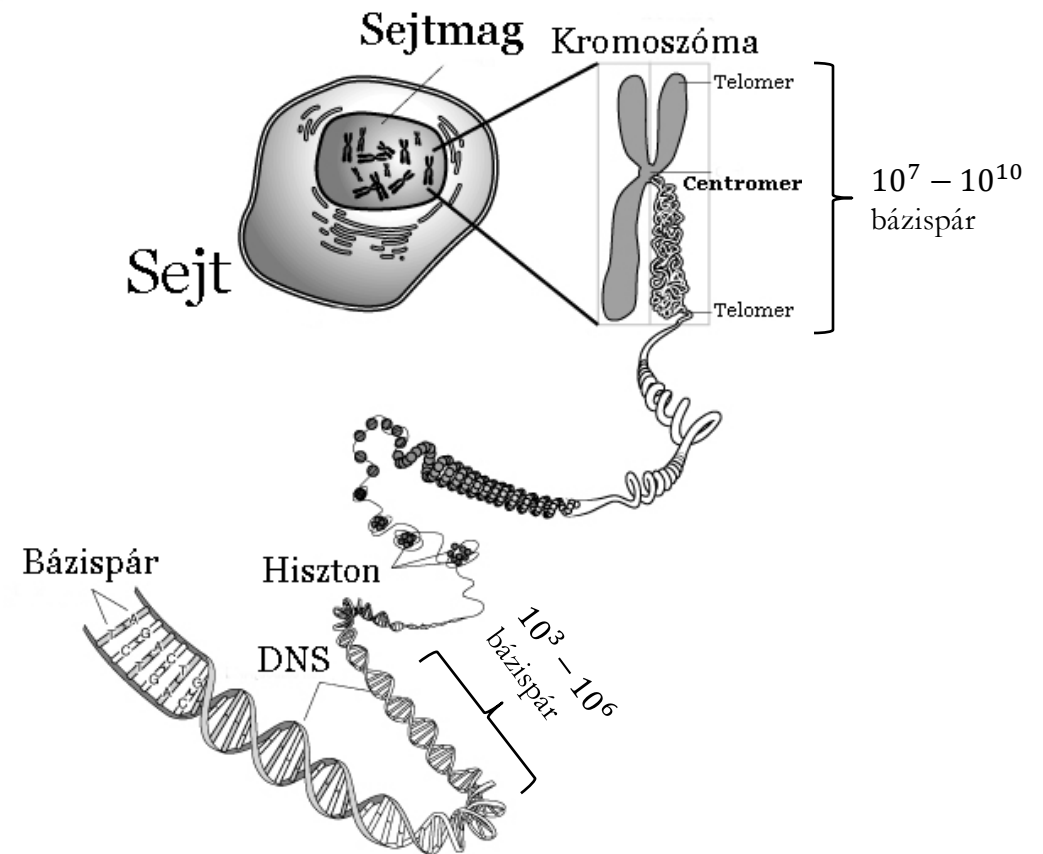


DNS számítások

- **Miért érdekes** a DNS az informatikában?
- **Adattárolás:** rendkívül nagy adatsűrűség
 - A bázisok közötti 0.35nm távolságot alapul véve, a DNS adatsűrűsége nagyobb, mint 1 millió Gbit/inch (a HDD-ké kb. 7 Gbit/inch)
 - Másképpen: 1 gramm DNS = 4 milliárd Tbit
- **Nagyfokú párhuzamosság:** egy kémcsőnyi DNS spirálok trillióit tartalmazhatja és a kémcső tartalmán elvégzett műveletek az összes szálon párhuzamosan végrehajthatók
- Rendkívül **kis energiafelhasználás**
 - az Adleman által végrehajtott kísérletben (lásd később) a rendszer mindössze 1 Joule energiát használt fel $2 \cdot 10^{19}$ számításához

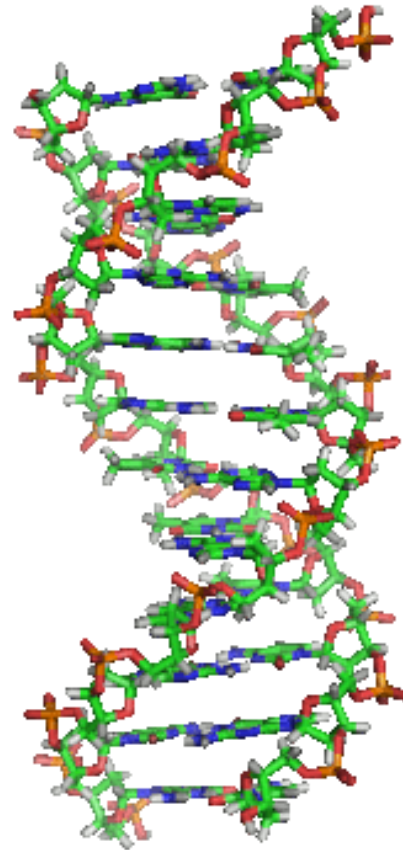
Mi is az a DNS?

- Minden élőlény sejtjében megtalálható a **DNS** (dezoxiribonukleinsav), ami egy **nukleotidokból** felépülő kettős hélix (csavar vagy spirál)
- A DNS tárolja magában a sejt **genetikai információját**
- A gének kódolják az, hogyan kell a **szervezet működését és növekedését** befolyásoló **fehérjéket** előállítani

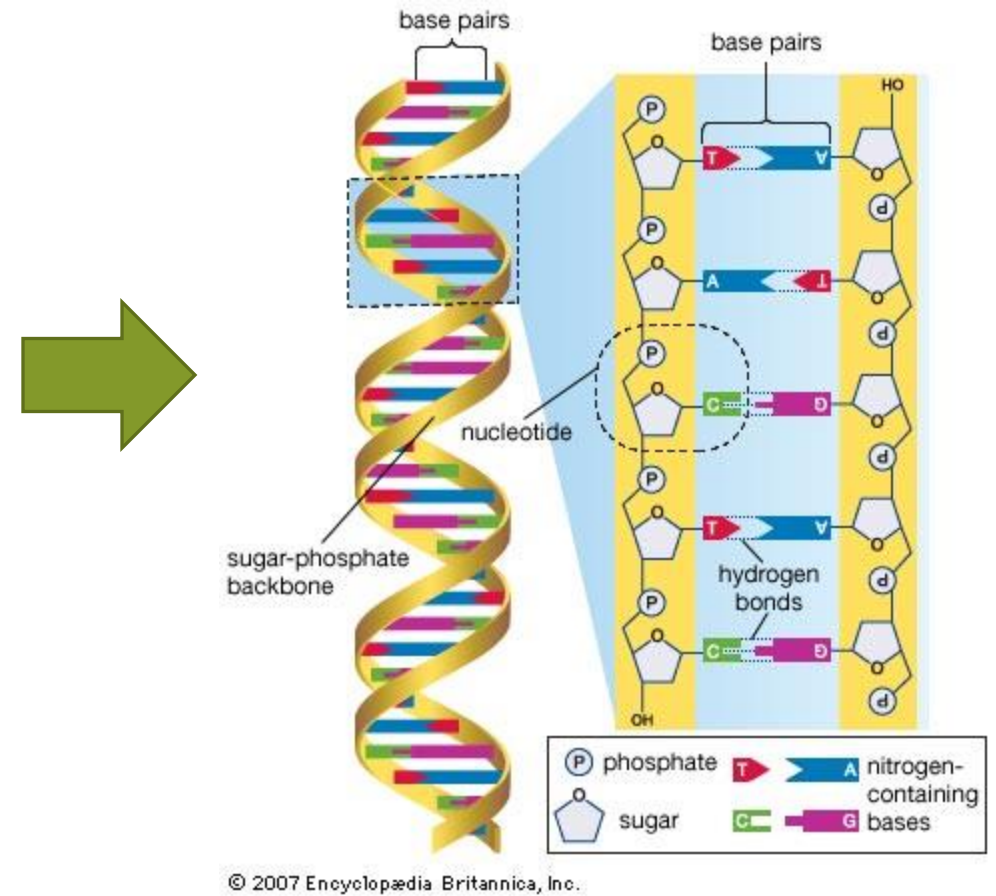


A DNS felépítése

- A nukleinsavak **nukleotidokból** felépülő szerves makromolekulák
- A nukleotidok **komponensei**:
 - **bázisok**
 - **pentózcukor** (dezoxiribóz, 5 szénatomos cukor)
 - **foszforsav**
- A pentózcukor és a foszforsav alkotják a DNS **cukor-foszfát gerincét**
- A bázisok: **adenin** (A), **guanin** (G), **citozin** (C), **timin** (T)

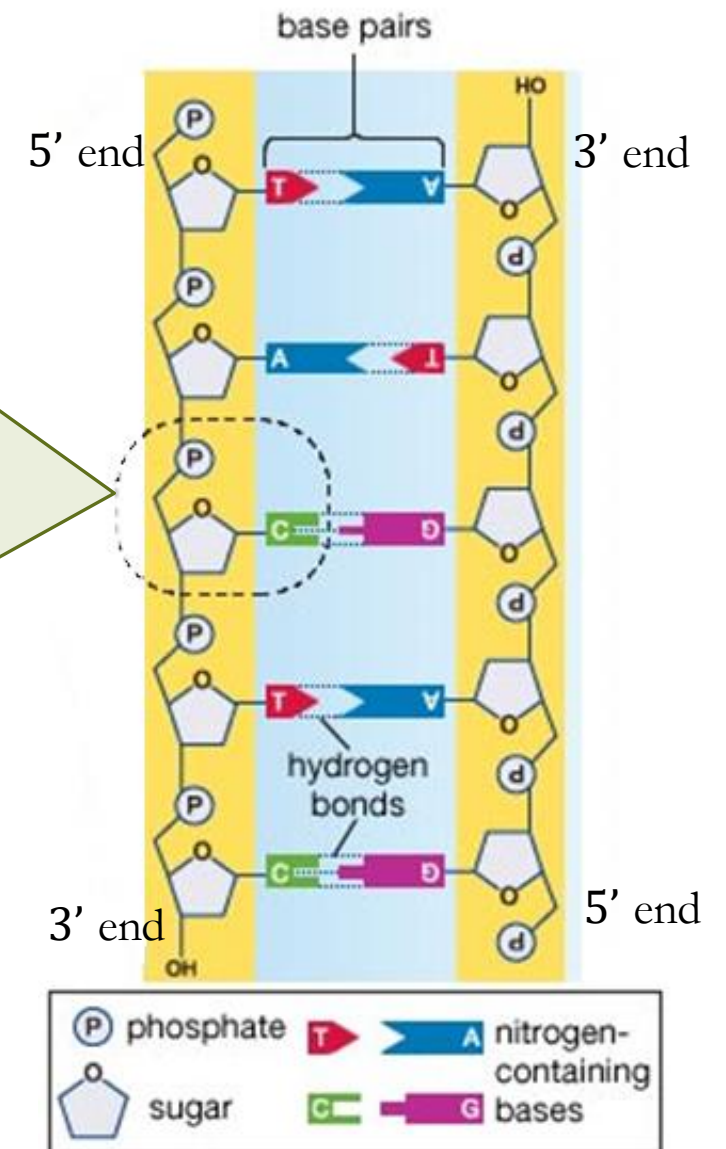
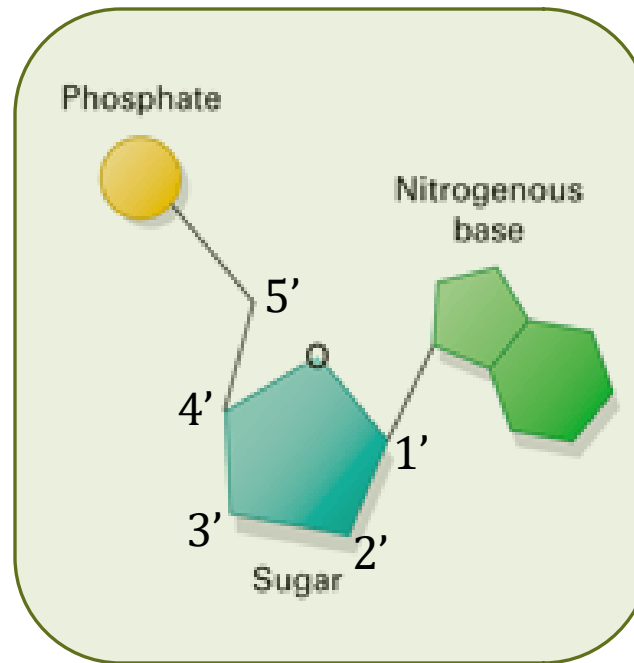


Forrás: Wikipedia



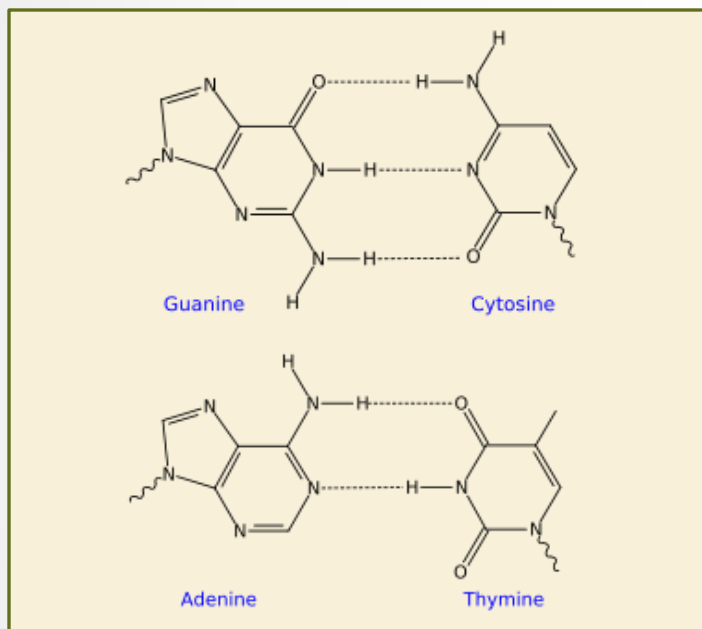
A DNS felépítése

- Szokás a szénatomokat a bázisoktól kezdve **számozni**
- Pl. az oxigén az 1'-es és 4'-est szénatomot köti össze
- A foszfátcsoport pedig az 5'-ik atomhoz kapcsolódik
- Ezért a spirálok szükségszerűen az **ellentétes irányokba** haladnak



Watson-Crick komplementaritás

Csak az alábbi típusú bázispárok lehetségesek:



Ezeket a felfedezőik után **Watson-Crick bázispároknak** hívjuk

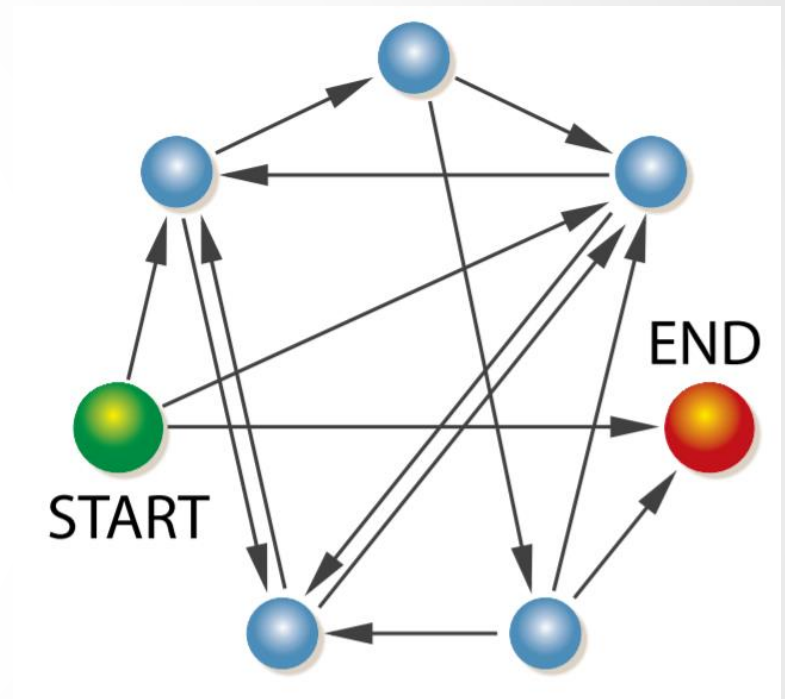
- Mindegyik bázispár szimmetriája lehetővé teszi, hogy a DNS kettős spiráljába kétféle módon illeszkedjen be:
 - $A = T, T = A, G \equiv C, C \equiv G$
- Így mindegyik szálon minden bázis **tetszőleges kombinációban** fordulhat elő
- Ha ismerjük az egyik szál egy **részszekvenciáját** (pl. TCGCAT), akkor meg tudjuk határozni a másik **megfelelő részszekvenciáját** (AGCGTA)
- Ezek egymásnak **komplementerei** lesznek, ezt hívjuk **Watson-Crick komplementaritásnak**
- Fontos a szekvencia iránya is (5'-től halad 3'-ig vagy fordítva)

Hogyan számol a DNS?

- A DNS-t a **Watson-Crick komplementaritást** felhasználva lehet számítások elvégzésére bírni:
 - Ha egy kémcsőbe egyszeres DNS szálak szekvenciáit rakjuk
 - és minden szekvenciából nagyon sok van
 - akkor a bázisok nagy valószínűséggel **minden lehetséges módon kapcsolódni** fognak egymáshoz
 - Ezek a kapcsolódások egyszerre, **párhuzamosan** történnek
 - Fontos, hogy a szálak közötti kötés gyengébb, mint a szálon belüli, tehát egy szál nem szakad szét azért, hogy kötődjön egy másik szálhoz
- Ezt megfigyelve adott meg **Leonard Adleman** a HAMILTON-ÚT probléma eldöntésére egy DNS-eken alapuló megoldást

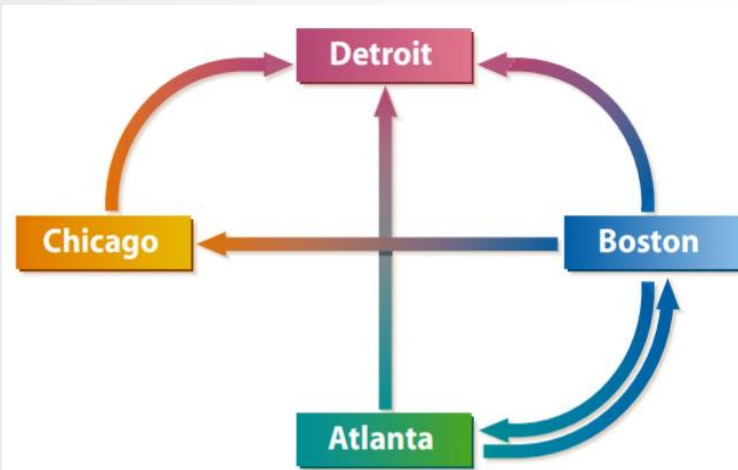
A HAMILTON-ÚT/ELÉRHETŐSÉG probléma

- Emlékeztető: A **HAMILTON-ÚT / ELÉRHETŐSÉG** probléma
 - Adott egy irányított $G = (V, E)$ gráf és $s, t \in V$ csúcsok
 - Azt kell eldönteni, hogy van-e *Hamilton-út / tetszőleges út* G -ben s -ből t -be
- **Gyakorlati vonatkozás:** van-e *Hamilton út / tetszőleges út* bizonyos városok gráfján?
- Például a jobboldali gráfban van Hamilton-út START-ból END-be:



Hogyan keressünk egy gráfban utakat (a DNS-sel)?

- Az egyszerűség az alábbi gráfban keresünk Hamilton utat DNS számítással Atlantából Detroitba:



- Kódoljuk el megfelelő bázis szekvenciákkal a csúcsokat (a DNS név komplementere lesz a csúcsot kódoló szekvencia)

Város	DNS név	Komplementer
Atlanta	ACTTGCAG	TGAACGTC
Boston	TCGGACTG	AGCCTGAC
Chicago	GGCTATGT	CCGATACA
Detroit	CCGAGCAA	GGCTCGTT

- és az éleket:

Rep.járat	DNS kód	Rep.járat	DNS kód
ATL-BOS	GCAGTCGG	BOS-DET	ACTGCCGA
ATL-DET	GCAGCCGA	BOS-ATL	ACTGACTT
BOS-CHI	ACTGGGCT	CHI-DET	ATGTCCGA

Hogyan keressünk egy gráfban utakat (a DNS-sel)?

- Ha ezekből a szekvenciákból nagyon sok van a kémcsőben, akkor a kialakult gráfbeli séták nagy valószínűséggel tartalmazni fogják az összes gráfbeli utat

- Például

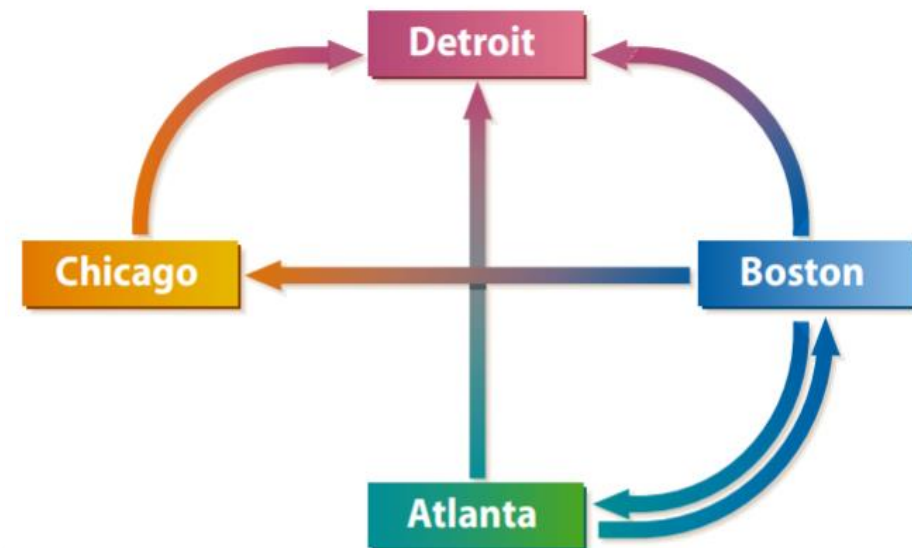
Boston	Atlanta	Boston	Detroit
AGCCTGACTGAACGTCAGCCTGACCCGAGCAA			
ACTGACTTGCAGTCGGACTGGGCT			
BOS-ATL	ATL-BOS	BOS-DET	

- vagy

Atlanta	Boston	Chicago	Detroit
TGAACGTCAGCCTGACCCGATACACCGAGCAA			
GCAGTCGGACTGGGCTATGTCCGA			
ATL-BOS	BOS-ATL	CHI-DET	

- vagy

Boston	Atlanta	Detroit
AGCCTGACTGAACGTCCCGAGCAA		
ACTGACTTGCAGCCGA		
BOS-ATL	ATL-DET	



Hogyan keressünk egy gráfban utakat (a DNS-sel)?

- Ha csak az **ELÉRHETŐSÉG** problémát szeretnénk megoldani, azaz azt, hogy van-e A és B város között tetszőleges (átszállásos) légiút, akkor a következőt kell tennünk:
 - Készítsünk a városokat és a repülőjáratokat kódoló szekvenciákból **nagyon sokat** (10^{18} nagyságrend)
 - Tegyük őket **egy kémcsőbe**
 - Kis idő múlva nézzük meg, hogy **van-e a kémcsőben megfelelő DNS szekvencia**, azaz olyan ami tartalmazza A és B kódját
- Ha a **HAMILTON-ÚT** problémát szeretnénk megoldani, akkor kicsit nehezebb kiválogatni a megfelelő DNS láncokat
- A következőkben összefoglaljuk, hogy **milyen eszközeink** vannak a fenti lépések elvégzéséhez

A DNS láncok manipulálásának eszközei

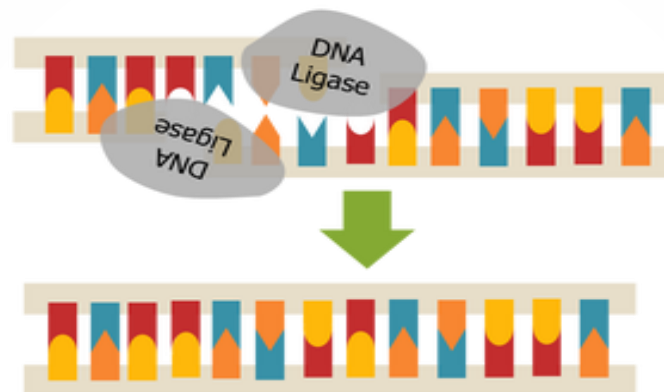
Elnevezés: A kémcsőben lévő anyagok összességét **levesnek** nevezzük

A DNS láncok **manipulálásának** eszközei:

- Watson-Crick **komplementaritás**
- **Helikázok: enzimek**, melyek a hidrogén kötések bontásával képesek a DNS-ből egyszeres DNS szálakat létrehozni
 - A DNS szétbontása elérhető a leves hőmérsékletének növelésével is (lásd PCR később)
- **Nukleázok: enzimek**, képesek a DNS láncot egy adott ponton elvágni
 - Azt sejtik, hogy bizonyos típusaik **védelmi mechanizmusként** alakultak ki

A DNS láncok manipulálásának eszközei

- **Ligázok: enzimek**, kettős szálak összekapcsolódását teszik lehetővé a cukor-foszfát gerinc mentén:



"DNA Ligase" by Heather Tsai Licensed under CC BY-SA 3.0

- Alkalmas a DNS-en lévő **hibák kijavítására** is, pl. UV sugárzás hatására létrejövő hibákat a borsejtben
- **Polimerázok:** enzimek, megyek az egyszeres DNS láncok végétől kezdve képesek felépíteni a lánc komplementerét (PCR-ben a P=Polimeráz)

A DNS láncok manipulálásának eszközei

Polimeráz-láncreakció (PCR)

- **Melegítéssel** elérjük, hogy a kettős szálak szétváljanak
- **Hűtéssel** elérjük, hogy úgynevezett primérek kapcsolódjanak az egyszeres láncok 3' végéhez
- A **polimeráz a primérből kiindulva**, a levesben található nukleotidok segítségével egy új **komplementis láncot** kapcsol az egyszeres lánchoz
- Ezt iterálva lehet lineáris időben exponenciálisan megsokszorozni bizonyos láncokat
- **Megjegyzés:** A PCR során a sokszorozítandó szál és annak a Watson-Crick komplementere is másolásra kerül

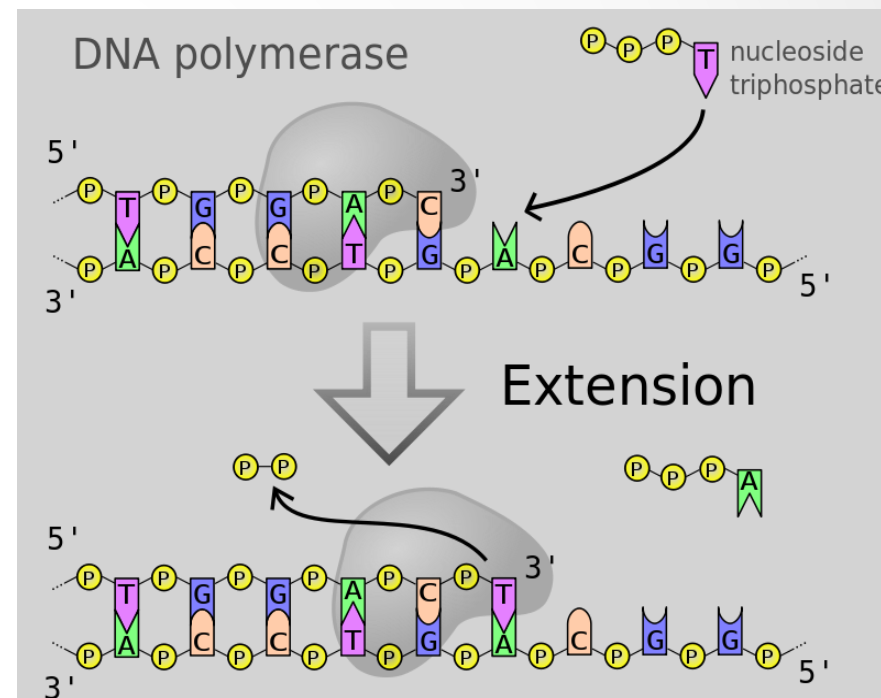


www.dnalc.org

A DNS láncok manipulálásának eszközei

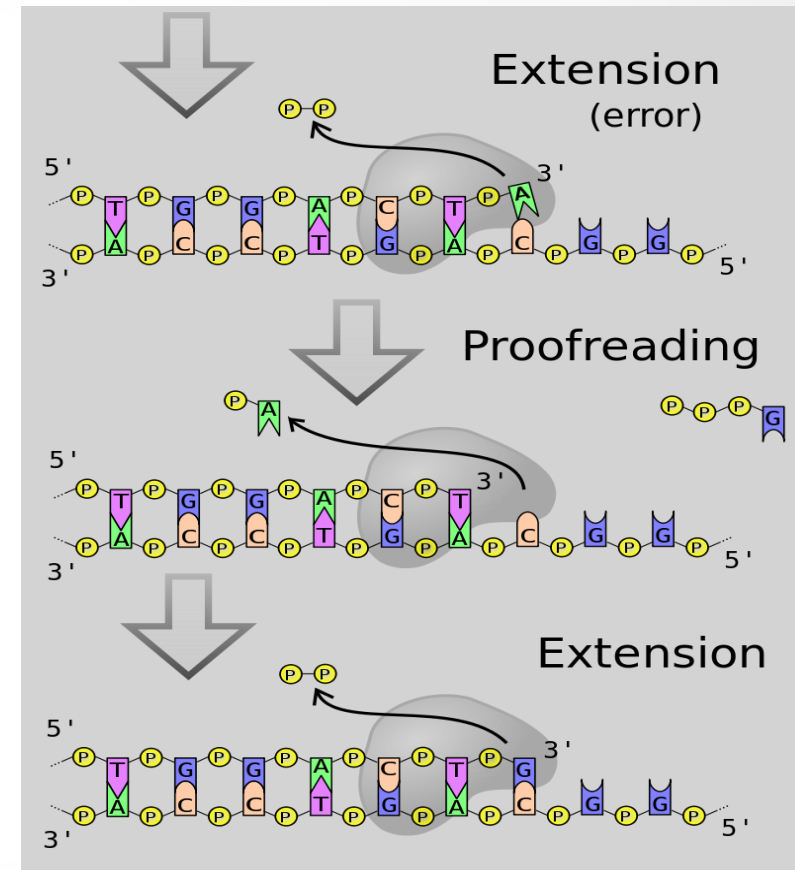
Polimeráz-láncreakció

- **Miért kell** a polimeráz enzim a nukleotidok kapcsolódásához?
- A levesben lévő nukleotidok az **enzimek nélkül nem** tudnak kapcsolódni a lánchoz mert
- két további **foszfor atom** van bennük
- Az enzim nagyon **pontos**, 1 milliárd bázispáronként vét egy hibát



A DNS láncok manipulálásának eszközei

- **Hibajavításra** is képes az exonukleáz segítségével
- Általában legfeljebb 10000 bázispár hosszú szekvenciák sokszorosítására alkalmazható
- A reakció sebessége kb. 1000 bázispár per perc



A DNS láncok manipulálásának eszközei

- **Gélelektroforézis**
 - A DNS láncok töltéssel rendelkeznek, melynek erős elektromos mező hatására elmozdulnak
 - A nagyobb láncok nehezebben mozognak
 - Így a DNS láncok a méretük alapján szeparálhatók
- **Megjegyzés:** ma már kereskedelmileg megrendelhető olyan kémcső, melyben a leves meghatározott - rövid - DNS szekvenciákat tartalmaz akár 10^{18} nagyságrendben

Agarose Gel Electrophoresis

Biotechnology Explorer™

BIO-RAD