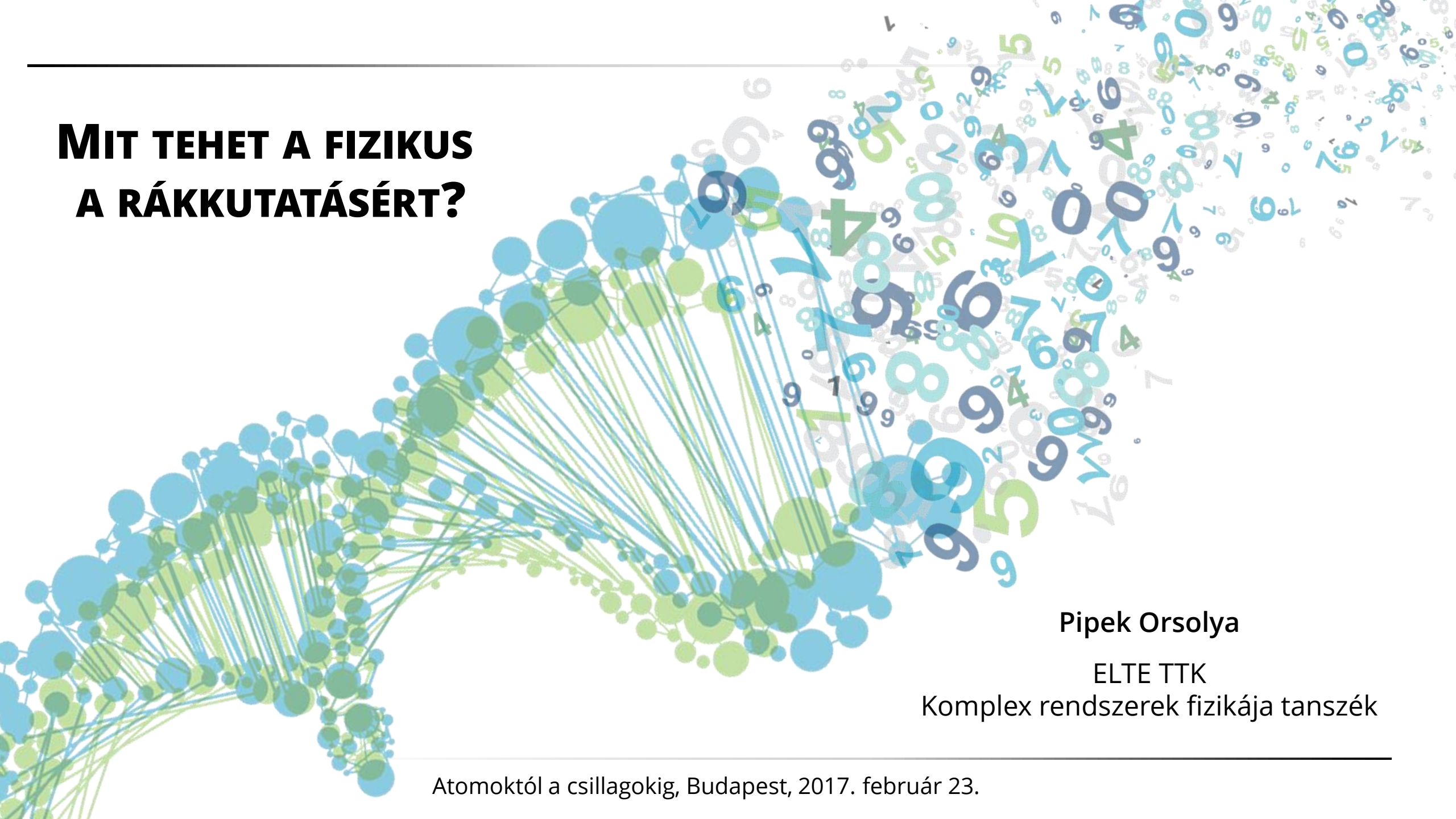


---

# MIT TEHET A FIZIKUS A RÁKKUTATÁSÉRT?



Pipek Orsolya

ELTE TTK

Komplex rendszerek fizikája tanszék

# ÁTTEKINTÉS



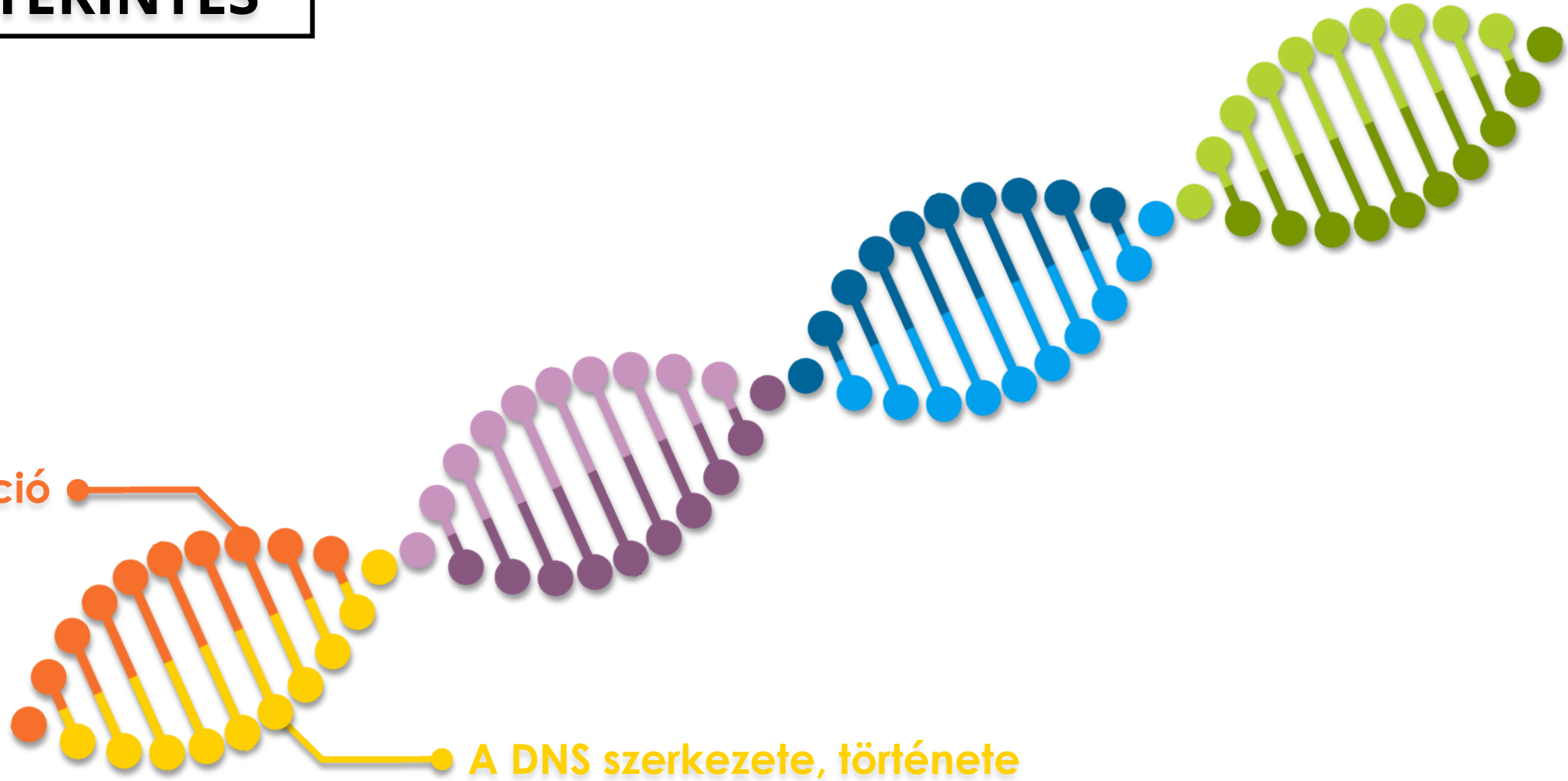
# ÁTTEKINTÉS

Motiváció



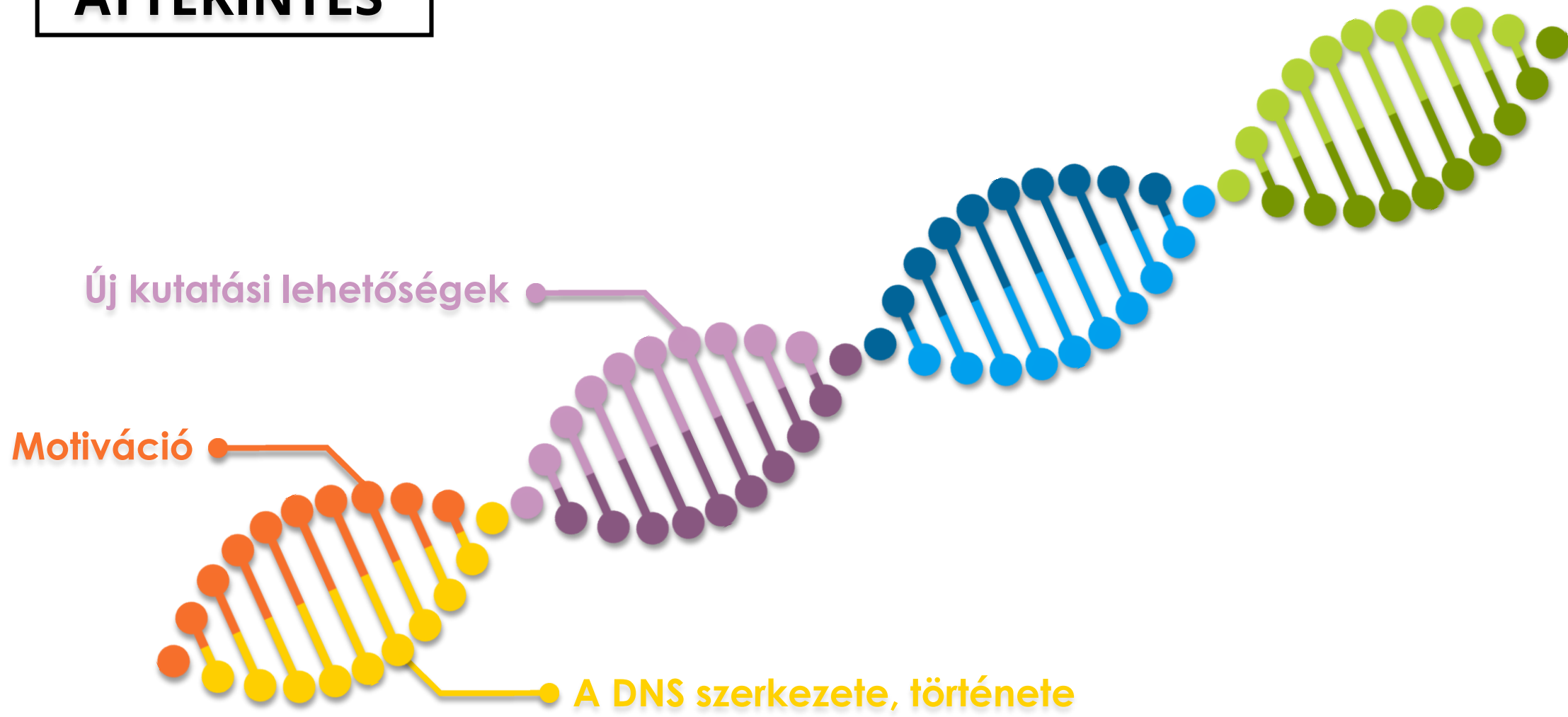
# ÁTTEKINTÉS

Motiváció

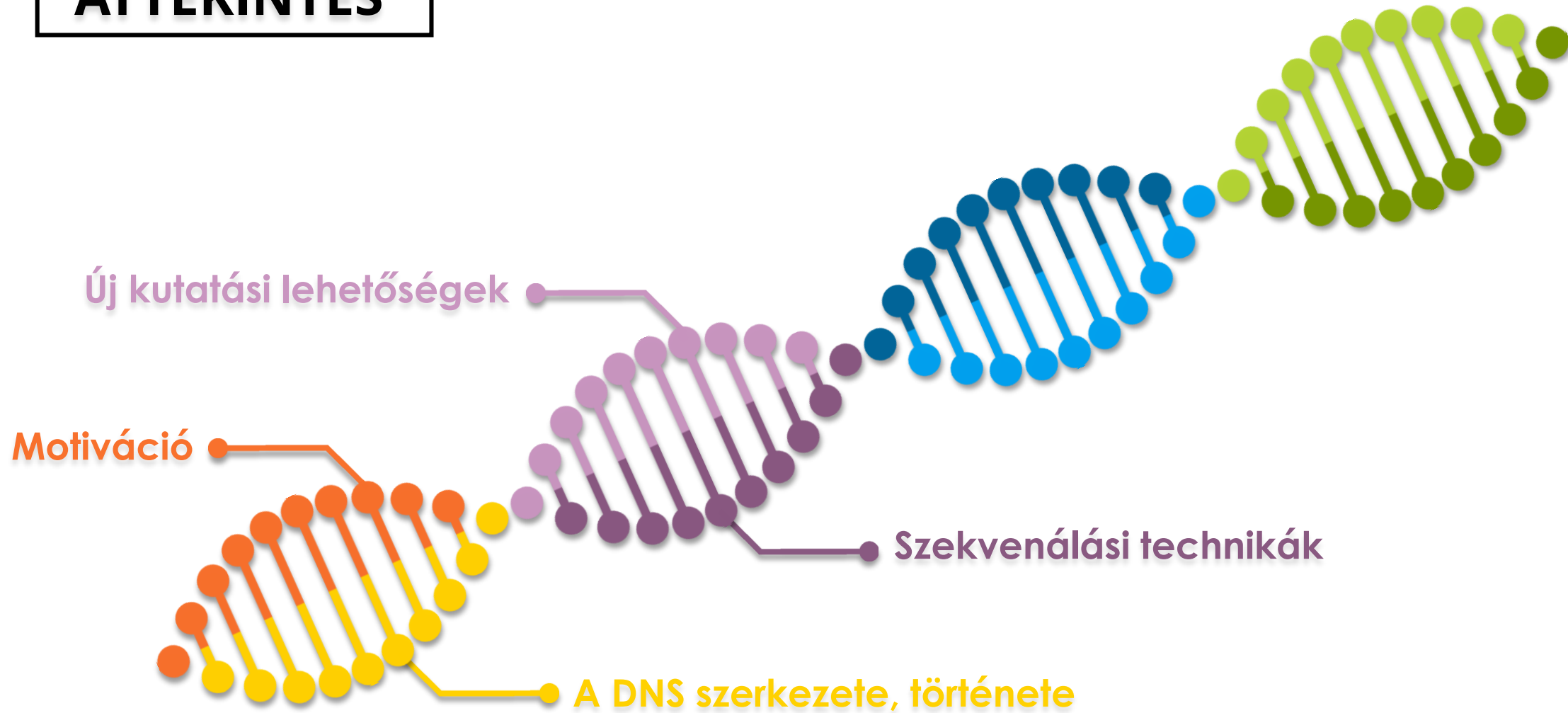


A DNS szerkezete, története

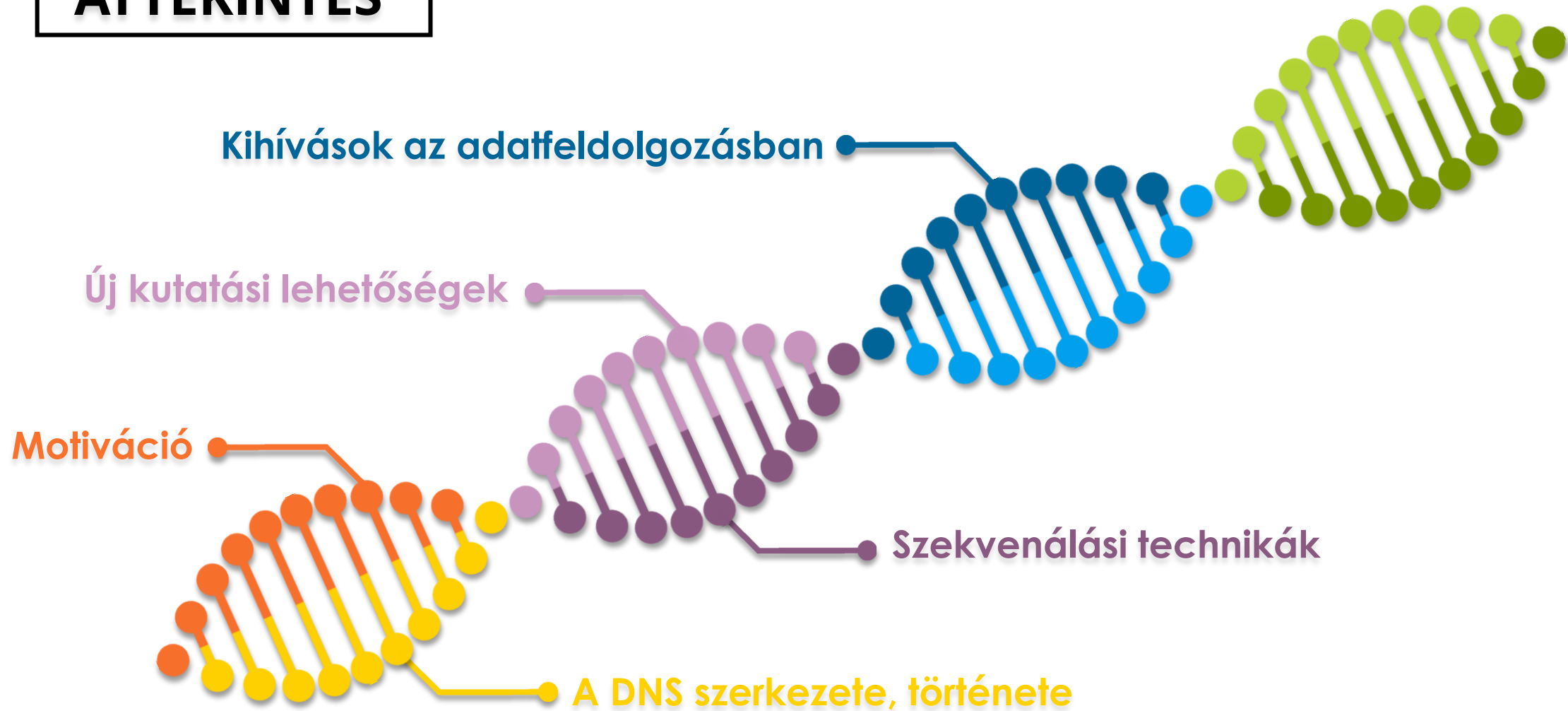
# ÁTTEKINTÉS



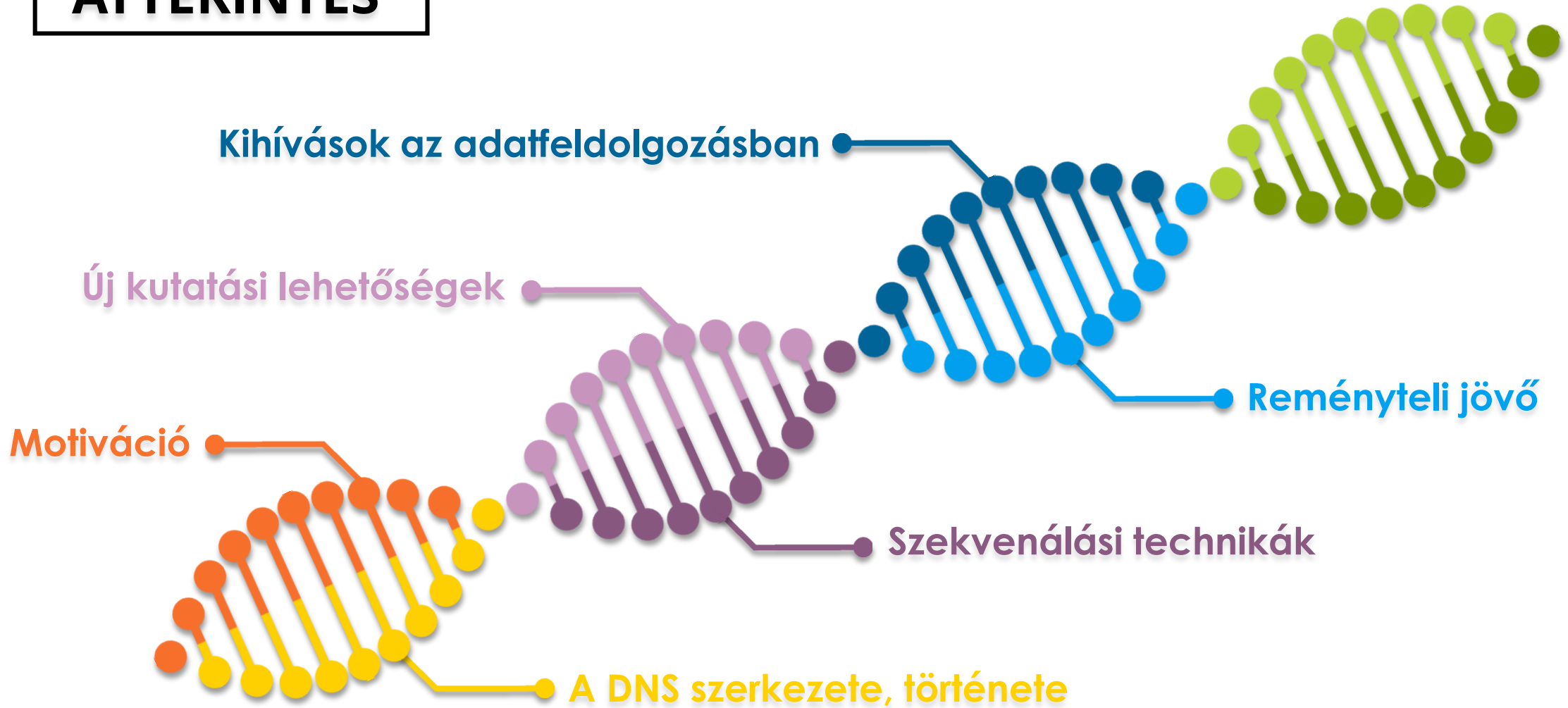
# ÁTTEKINTÉS



# ÁTTEKINTÉS

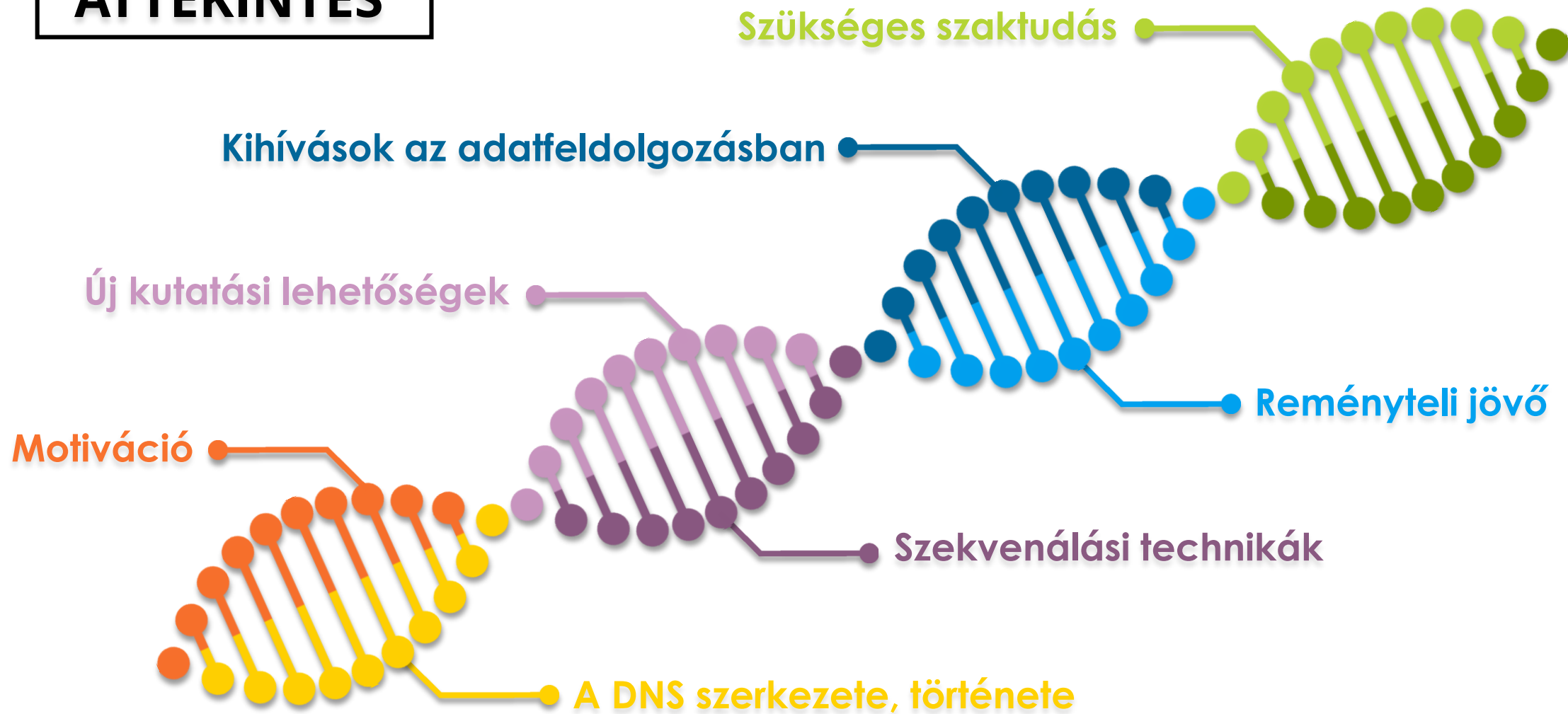


# ÁTTEKINTÉS

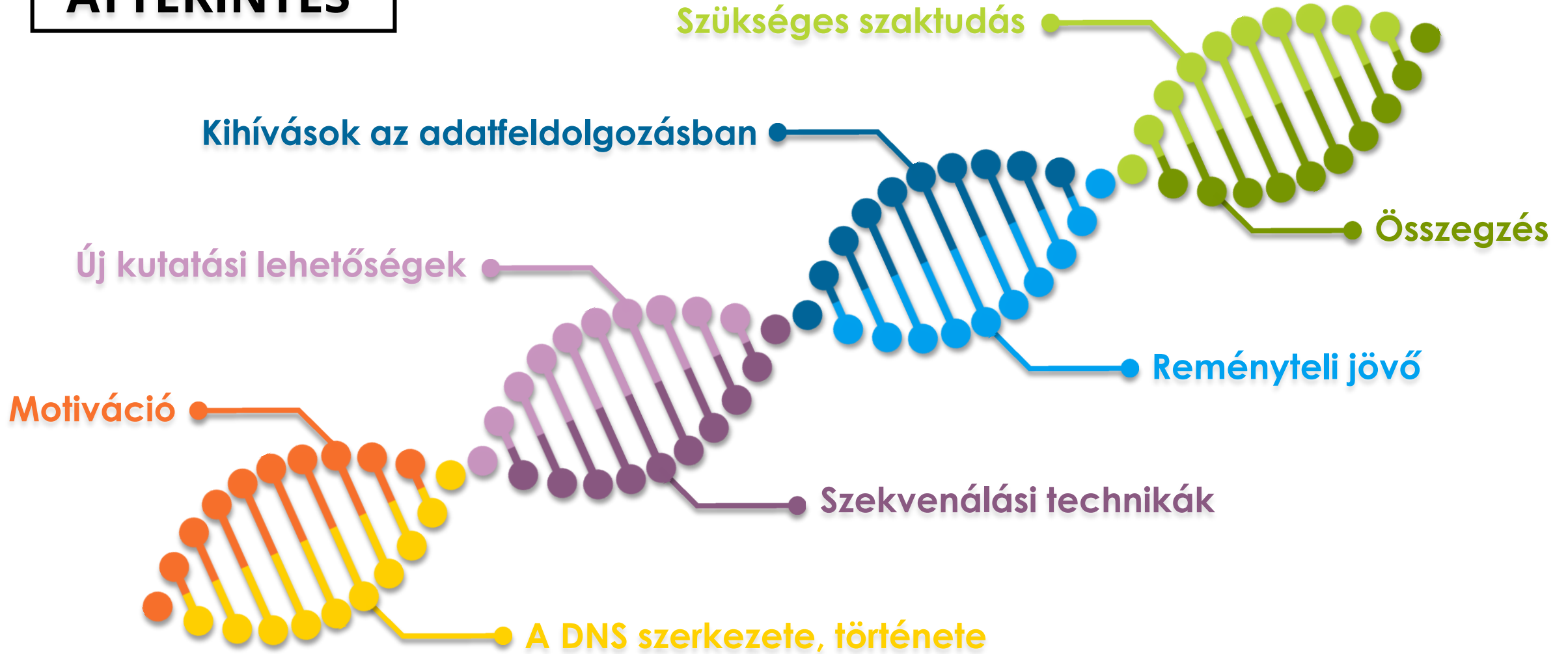




# ÁTTEKINTÉS



# ÁTTEKINTÉS



# **MOTIVÁCIÓ: MIÉRT FOGLALKOZUNK GENETIKÁVAL?**

# MOTIVÁCIÓ: MIÉRT FOGLALKOZUNK GENETIKÁVAL?

↑ főleg fizikusként...

## MOTIVÁCIÓ: MIÉRT FOGLALKOZUNK GENETIKÁVAL?

*↑ főleg fizikusként...*



**hasznos**

genetikai  
betegségek

## MOTIVÁCIÓ: MIÉRT FOGLALKOZUNK GENETIKÁVAL?

*↑ főleg fizikusként...*



**hasznos**

genetikai  
betegségek



**szükséges**

óriási  
feldolgozatlan  
adatmennyiség

## MOTIVÁCIÓ: MIÉRT FOGLALKOZUNK GENETIKÁVAL?

*↑ főleg fizikusként...*



**hasznos**

genetikai  
betegségek



**szükséges**

óriási  
feldolgozatlan  
adatmennyiség



**érdekes**

új, felfedezetlen  
területek,  
kihívások

## MOTIVÁCIÓ: MIÉRT FOGLALKOZUNK GENETIKÁVAL?

*↑ főleg fizikusként...*



**hasznos**

genetikai  
betegségek



**szükséges**

óriási  
feldolgozatlan  
adatmennyiség



**érdekes**

új, felfedezetlen  
területek,  
kihívások



**képesek  
vagyunk rá**

fizikusok jó  
probléma-  
megoldók

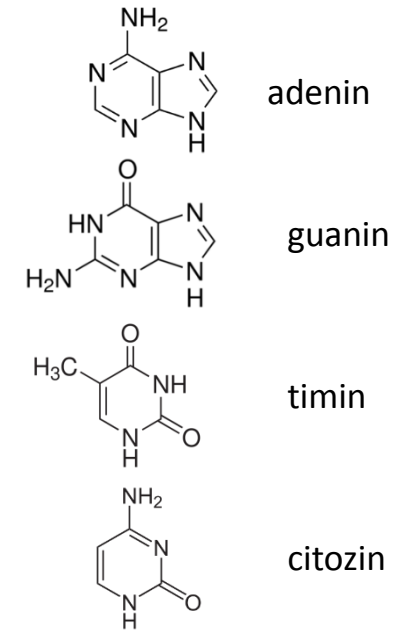
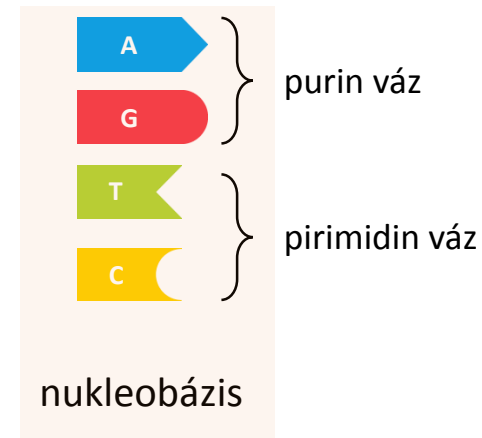
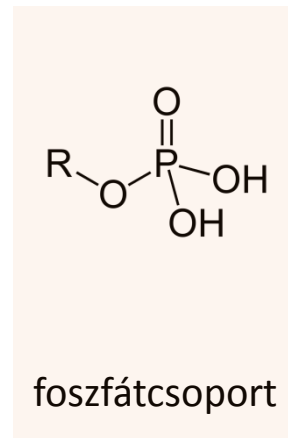
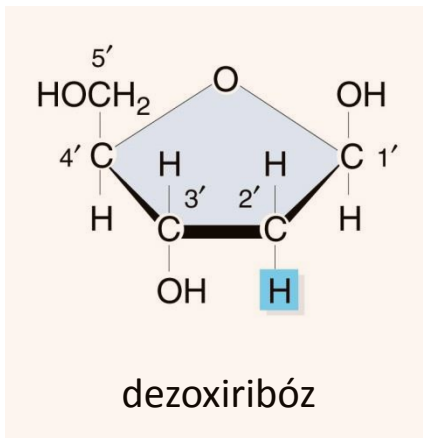


# MI A DNS?

# MI A DNS?

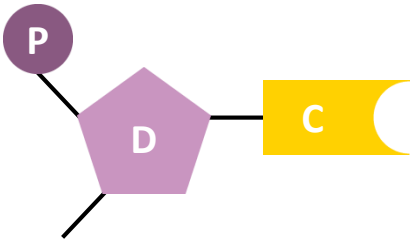
*deoxiribonukleinsav*

- Feladata: **genetikai utasítások összességének kódolása**
  - „örökítőanyag”, „genom” (vírusoknál RNS)
    - utód megőröklí az apai és anyai DNS felét → öröklött tulajdonságok
- Felépítése: hatalmas **makromolekula**, két egymás köré csavarodott **polinukleotid szál**
  - monomer egysége: nukleotid

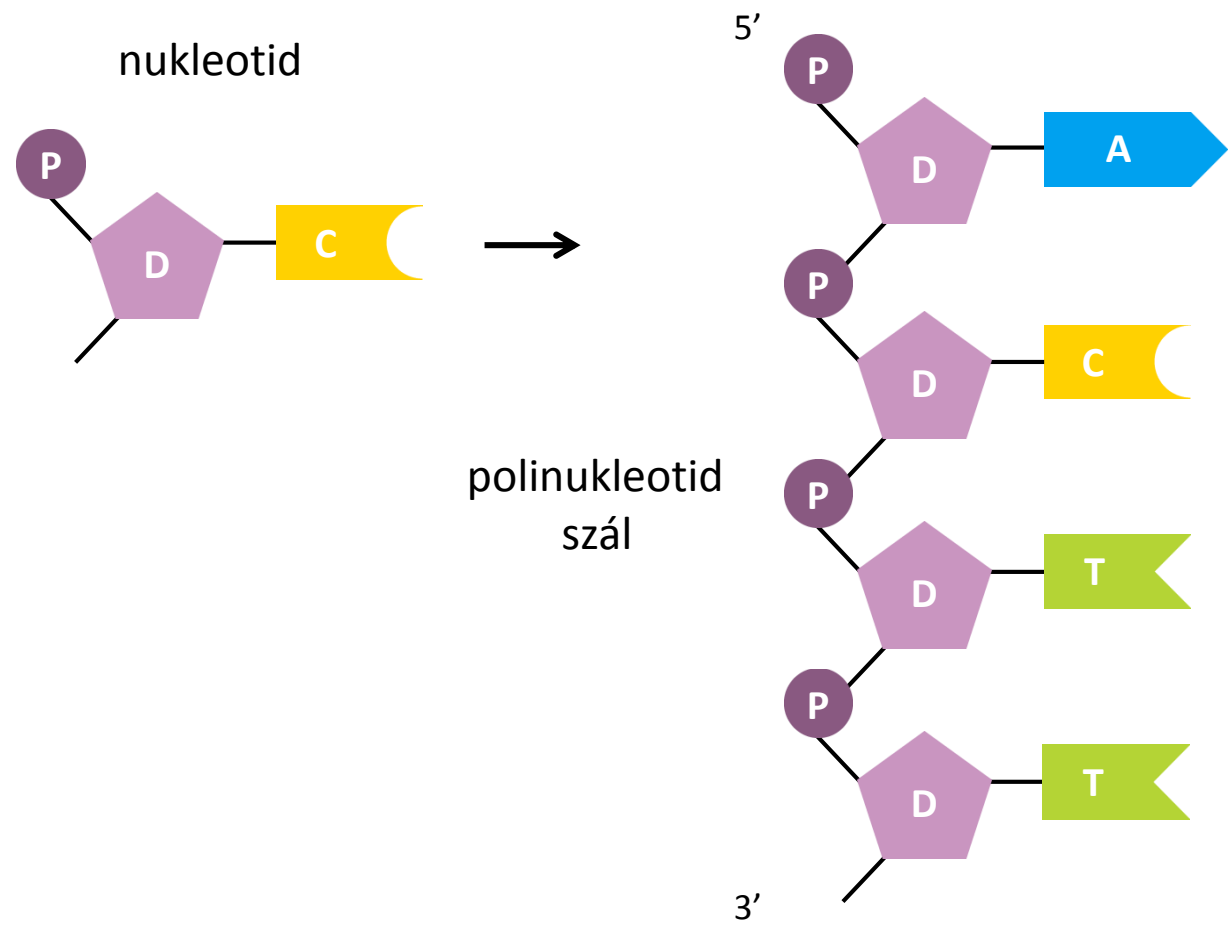


# MI A DNS?

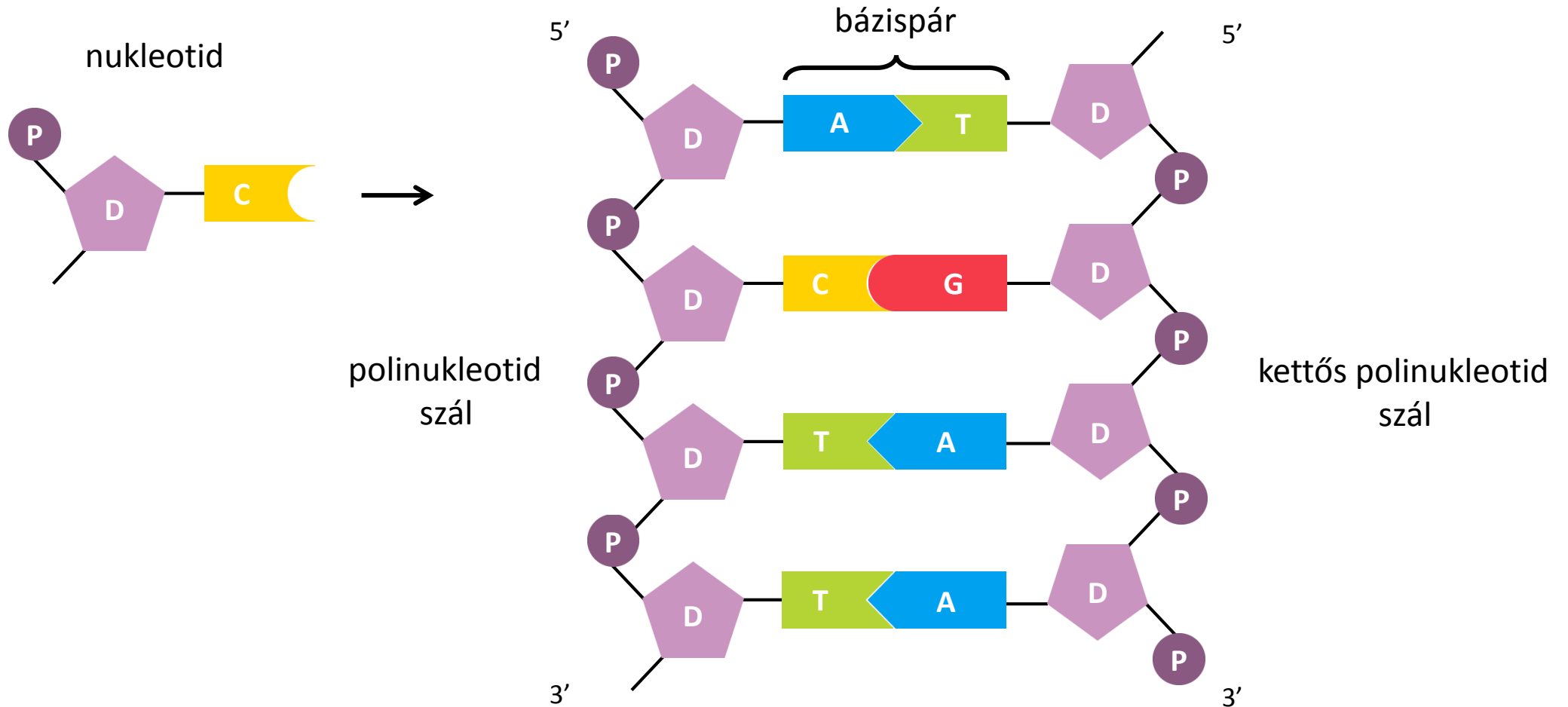
nukleotid



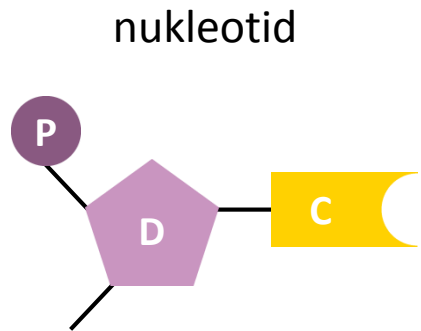
# MI A DNS?



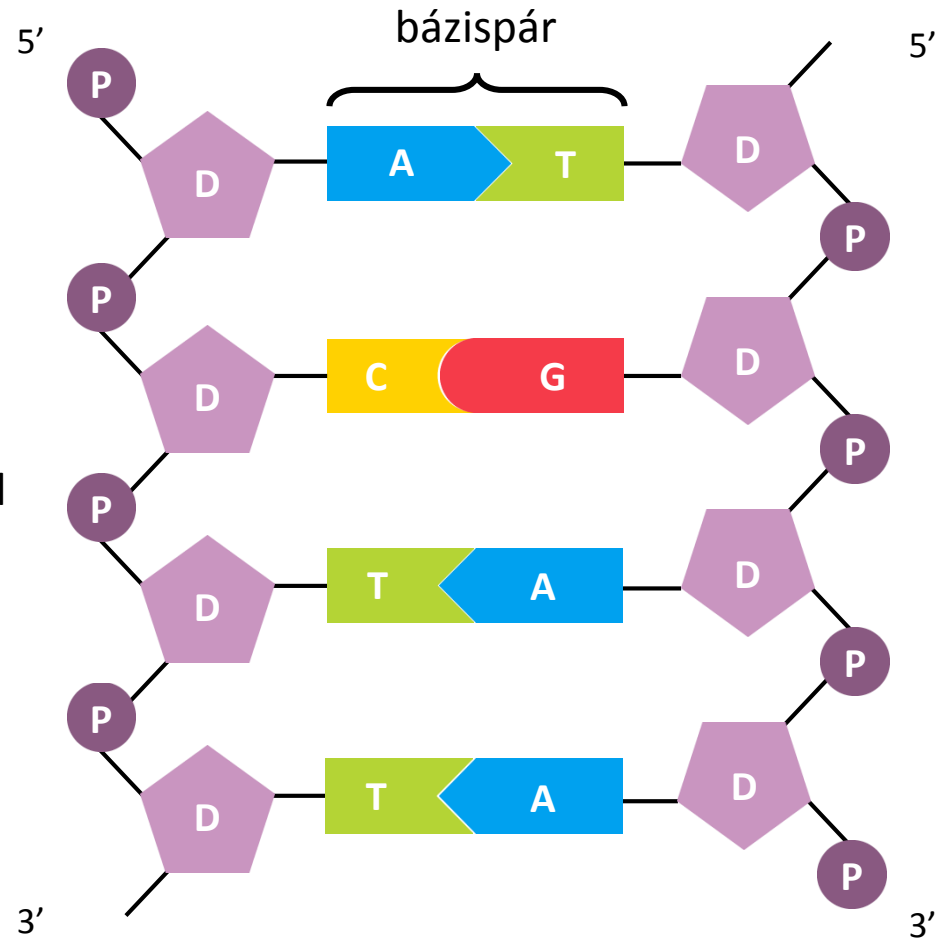
# MI A DNS?



# MI A DNS?



polinukleotid  
szál



kettős polinukleotid  
szál

Humán genom:  
3 milliárd ( $3 \cdot 10^9$ ) bázispár  
minden testi sejtben!

## MI A DNS?

Humán genom:  
3 milliárd ( $3 \cdot 10^9$ ) bázispár  
minden testi sejtben!

## MI A DNS?

Bázispárok átlagos távolsága:  
0,34 nm ( $0,34 \cdot 10^{-9}$  m)



Humán genom:  
3 milliárd ( $3 \cdot 10^9$ ) bázispár  
minden testi sejtben!



## MI A DNS?

Emberi test:

**37,2 billió ( $3,72 \cdot 10^{13}$ ) sejt!**

(2013, DOI: 10.3109/03014460.2013.807878)



Bázispárok átlagos távolsága:

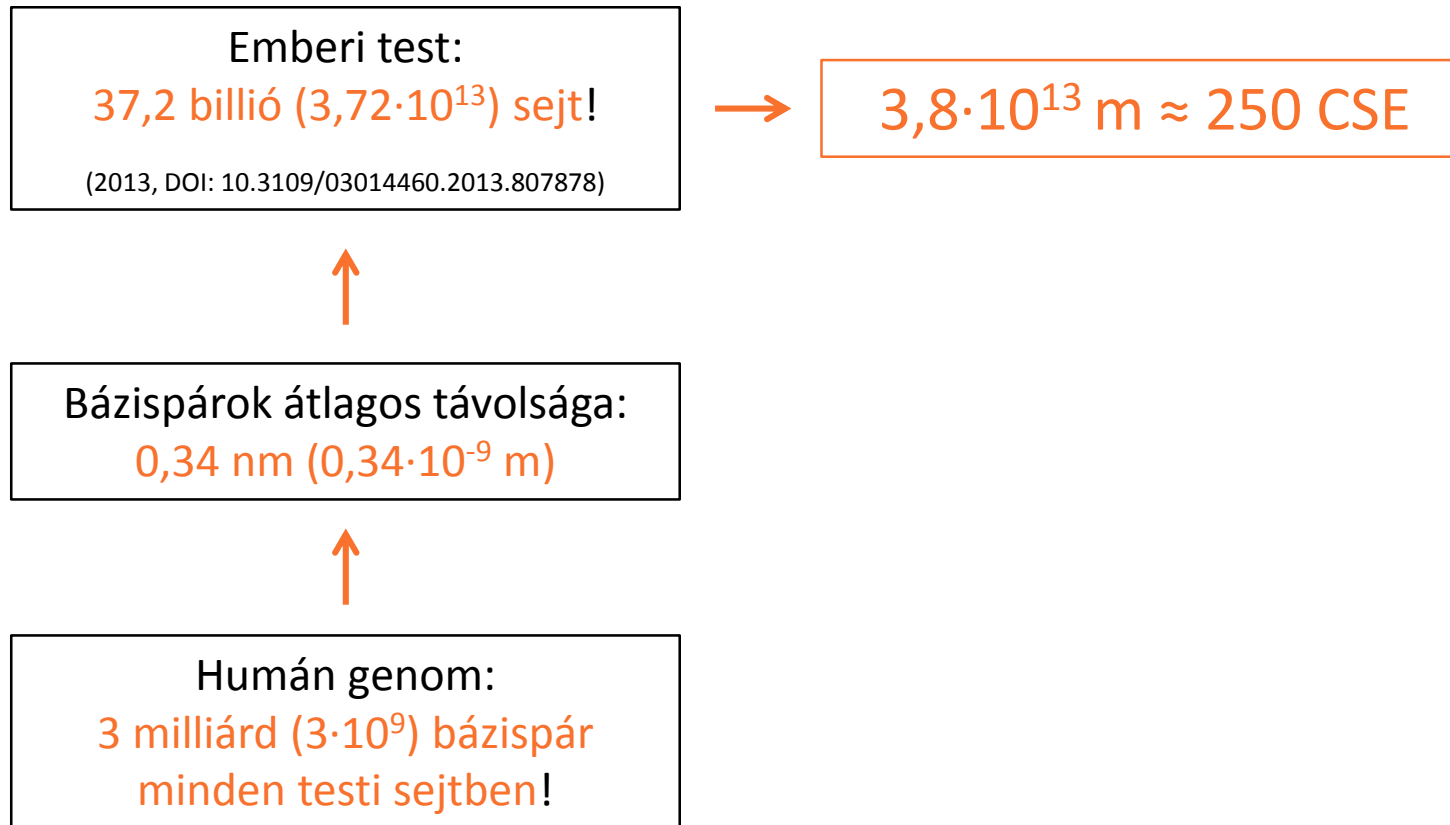
**0,34 nm ( $0,34 \cdot 10^{-9}$  m)**



Humán genom:

**3 milliárd ( $3 \cdot 10^9$ ) bázispár  
minden testi sejtben!**

## MI A DNS?



## MI A DNS?

Emberi test:  
37,2 billió ( $3,72 \cdot 10^{13}$ ) sejt!  
(2013, DOI: 10.3109/03014460.2013.807878)



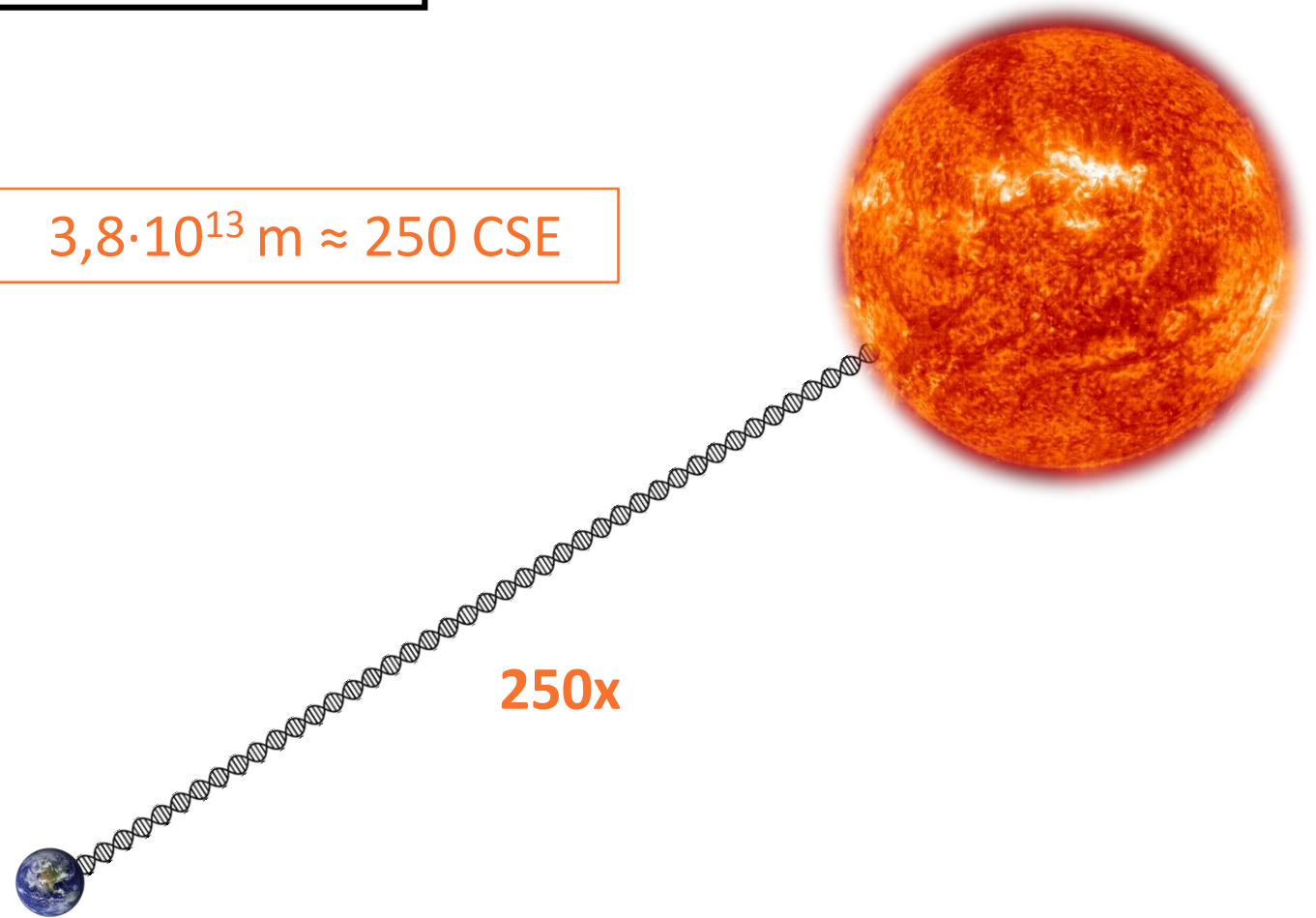
Bázispárok átlagos távolsága:  
0,34 nm ( $0,34 \cdot 10^{-9}$  m)



Humán genom:  
3 milliárd ( $3 \cdot 10^9$ ) bázispár  
minden testi sejtben!



$3,8 \cdot 10^{13}$  m  $\approx$  250 CSE



## MI A DNS?

Emberi test:  
37,2 billió ( $3,72 \cdot 10^{13}$ ) sejt!  
(2013, DOI: 10.3109/03014460.2013.807878)



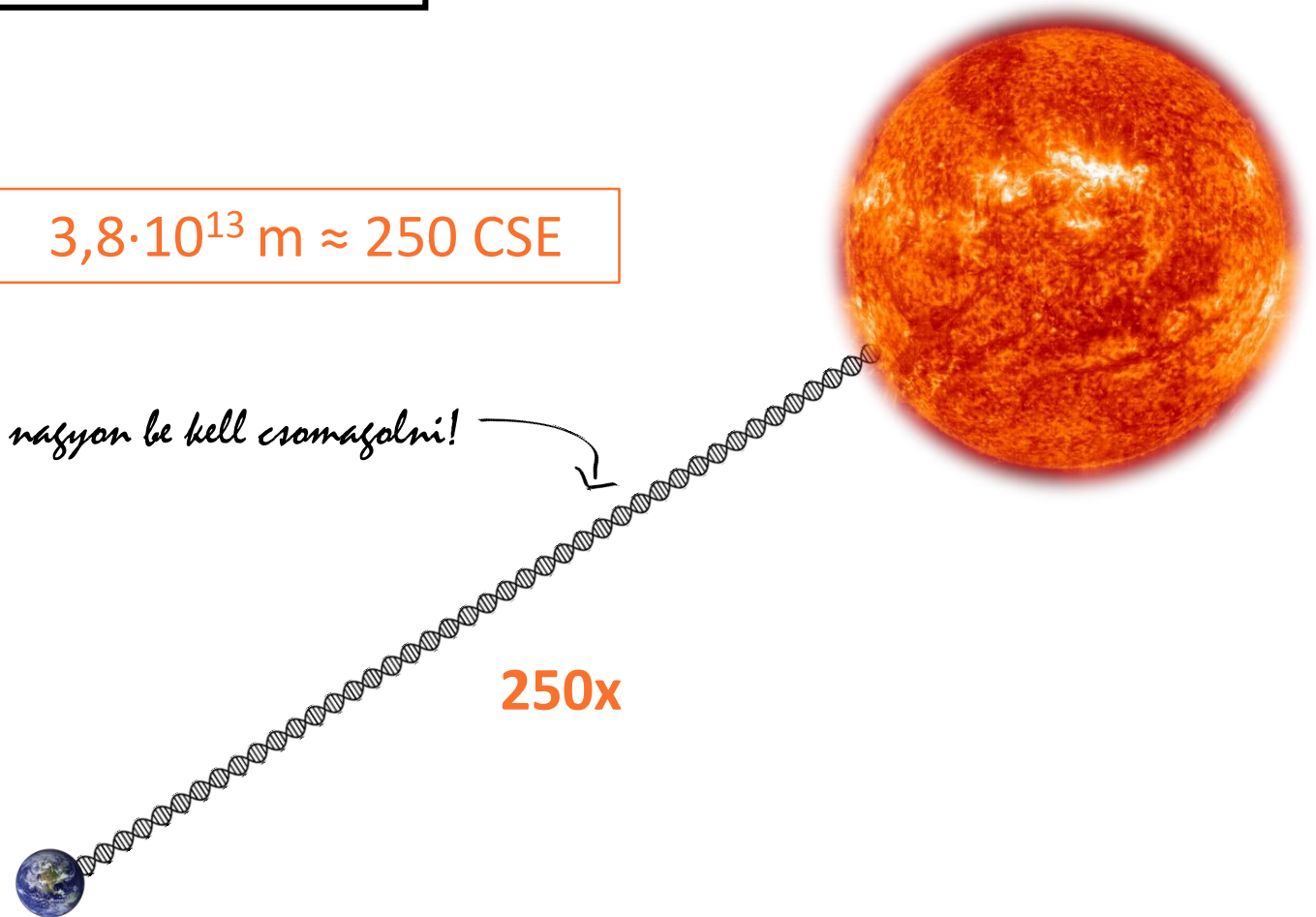
Bázispárok átlagos távolsága:  
0,34 nm ( $0,34 \cdot 10^{-9}$  m)



Humán genom:  
3 milliárd ( $3 \cdot 10^9$ ) bázispár  
minden testi sejtben!

→  $3,8 \cdot 10^{13}$  m  $\approx$  250 CSE

*Ezt nagyon be kell csomagolni!*



250x

# MI A DNS?

Emberi test:

37,2 billió ( $3,72 \cdot 10^{13}$ ) sejt!

(2013, DOI: 10.3109/03014460.2013.807878)



Bázispárok átlagos távolsága:

0,34 nm ( $0,34 \cdot 10^{-9}$  m)



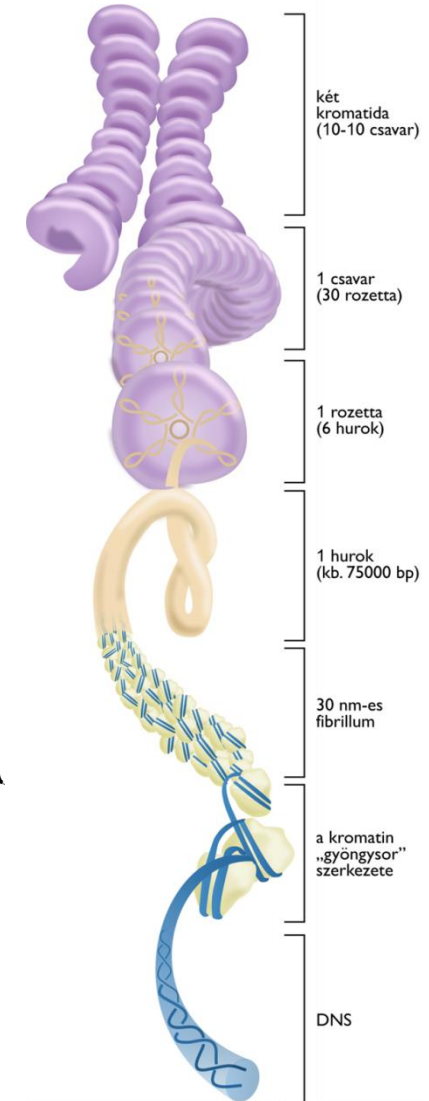
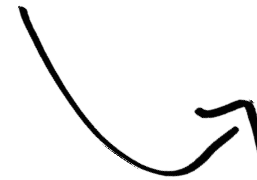
Humán genom:

3 milliárd ( $3 \cdot 10^9$ ) bázispár  
minden testi sejtben!



$3,8 \cdot 10^{13}$  m  $\approx$  250 CSE

*Ezt nagyon be kell csomagolni!*



# MI A DNS?

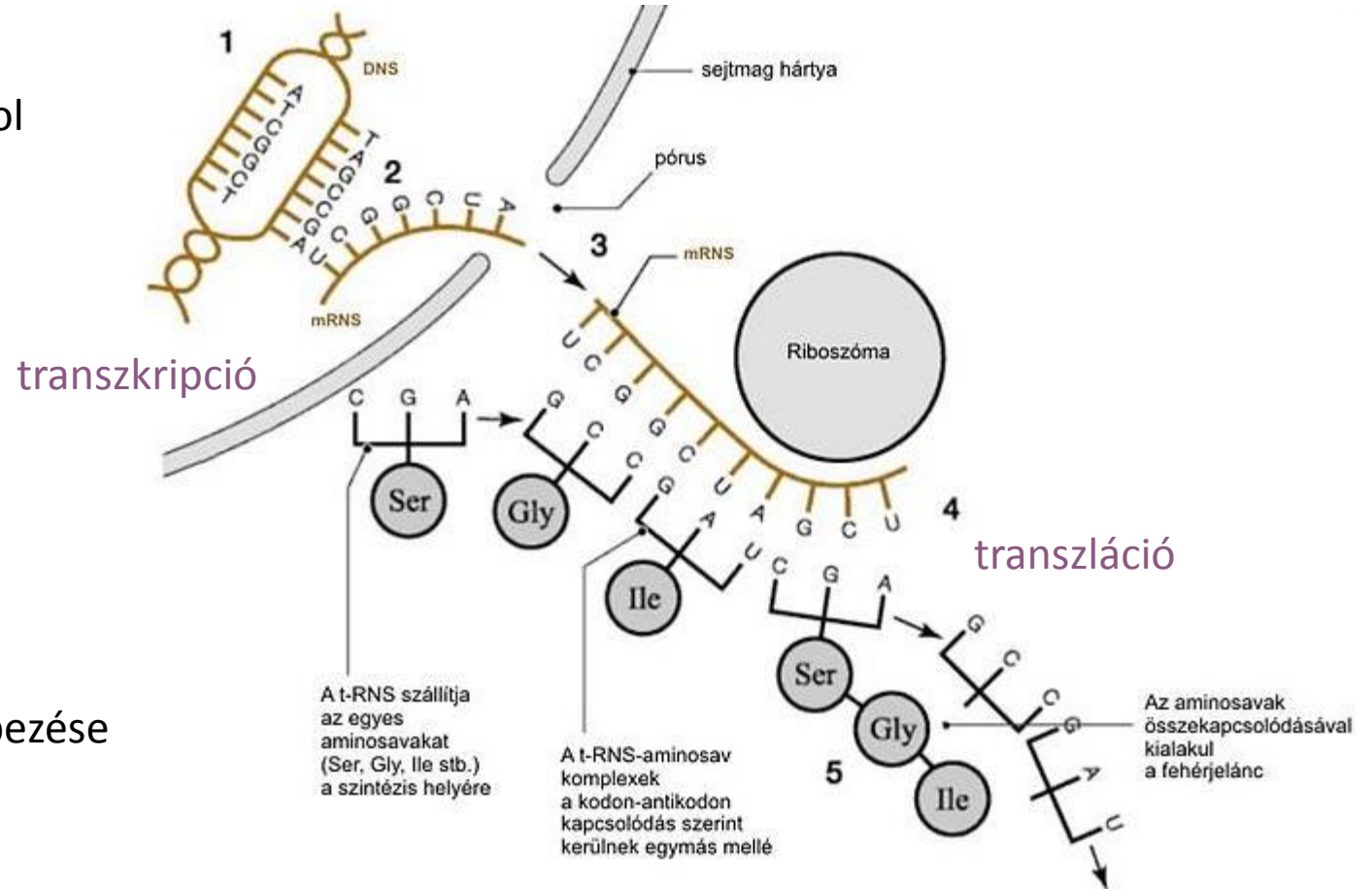
- Információ kódolása: **bázissorendben**
  - 3 egymás melletti bázis („kodon”) kódol egy aminosavat
  - aminosavak → **fehérjék**



A szervezet működésében a konkrét bázissorendnek meghatározó szerepe van.



Cél: a bázissorend (és következményeinek) feltérképezése  
**SZEKVENÁLÁS**



# A DNS TÖRTÉNETE

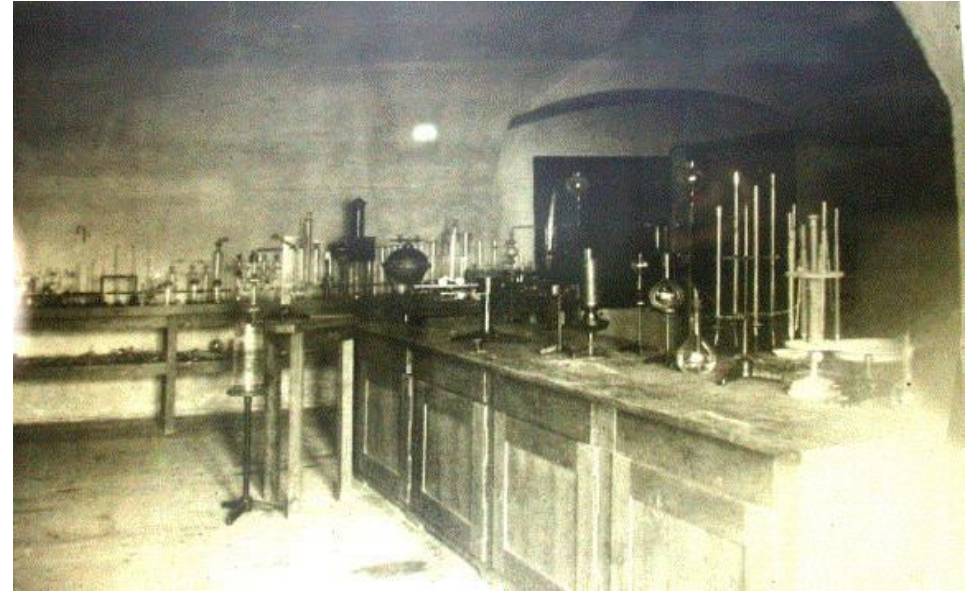
## A DNS TÖRTÉNETE

1879  
izzólámpa

1869

Friedrich Miescher izolálja a „nukleint”

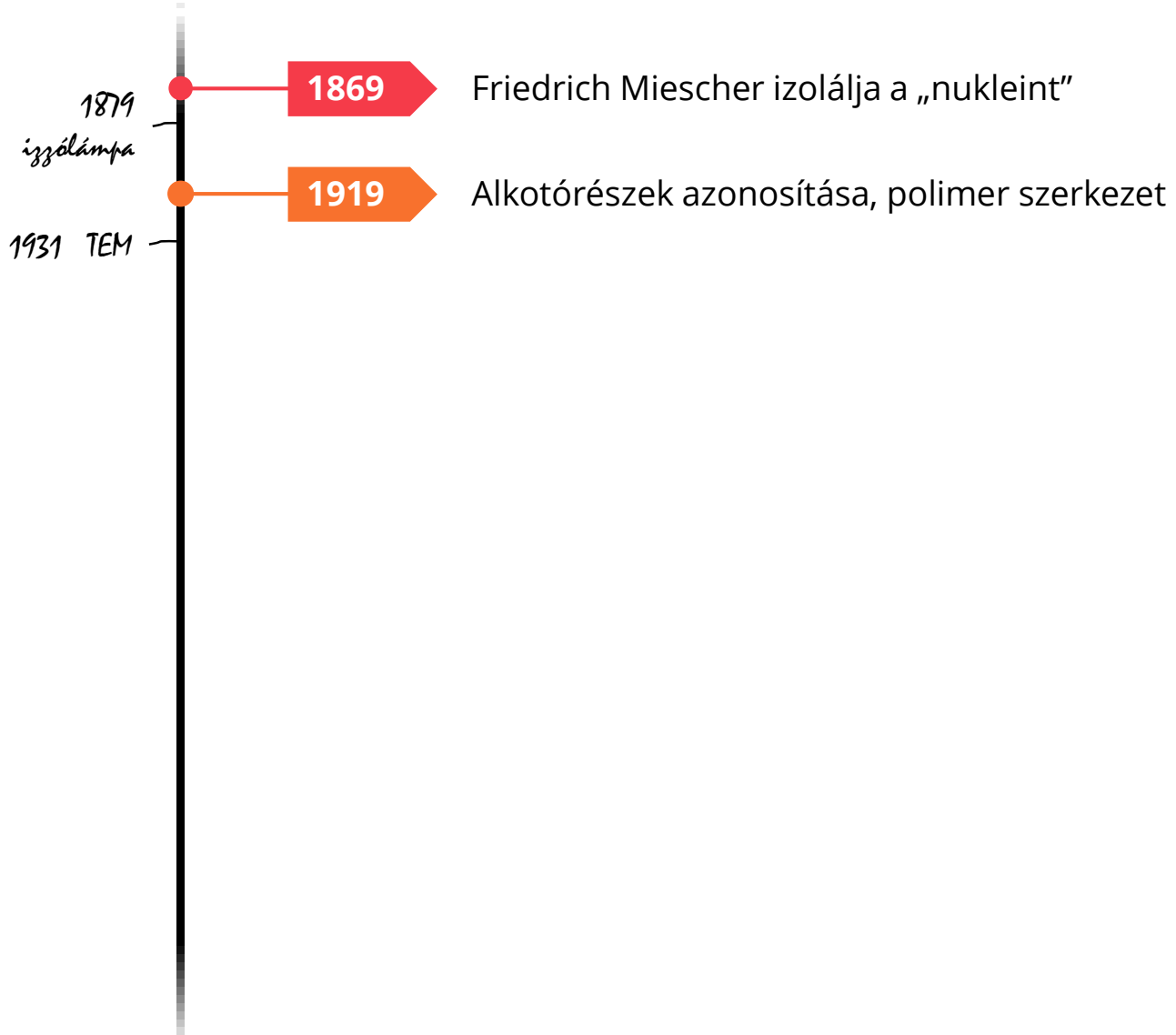
- eddig ismeretlen, a sejtmagban („nukleusz”) elhelyezkedő „anyag”
- nem fehérje



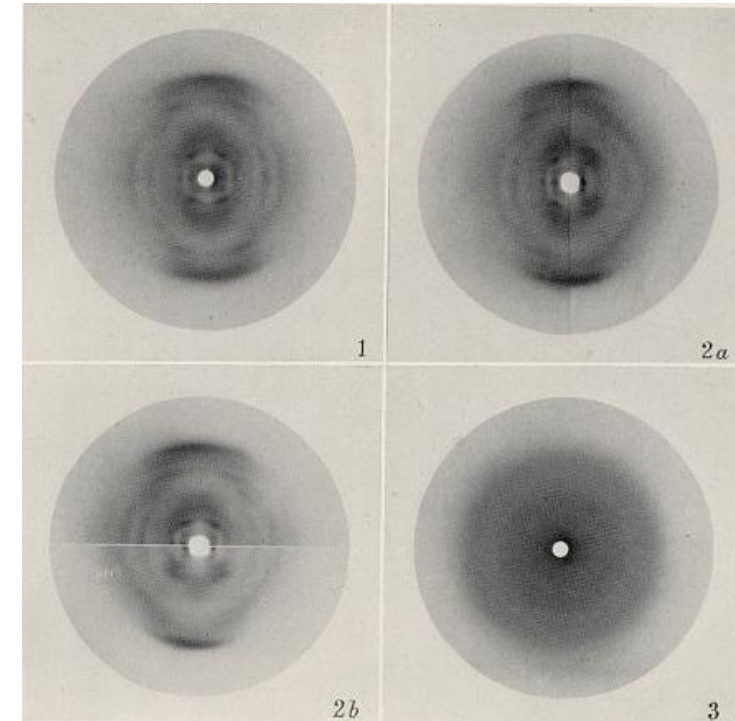
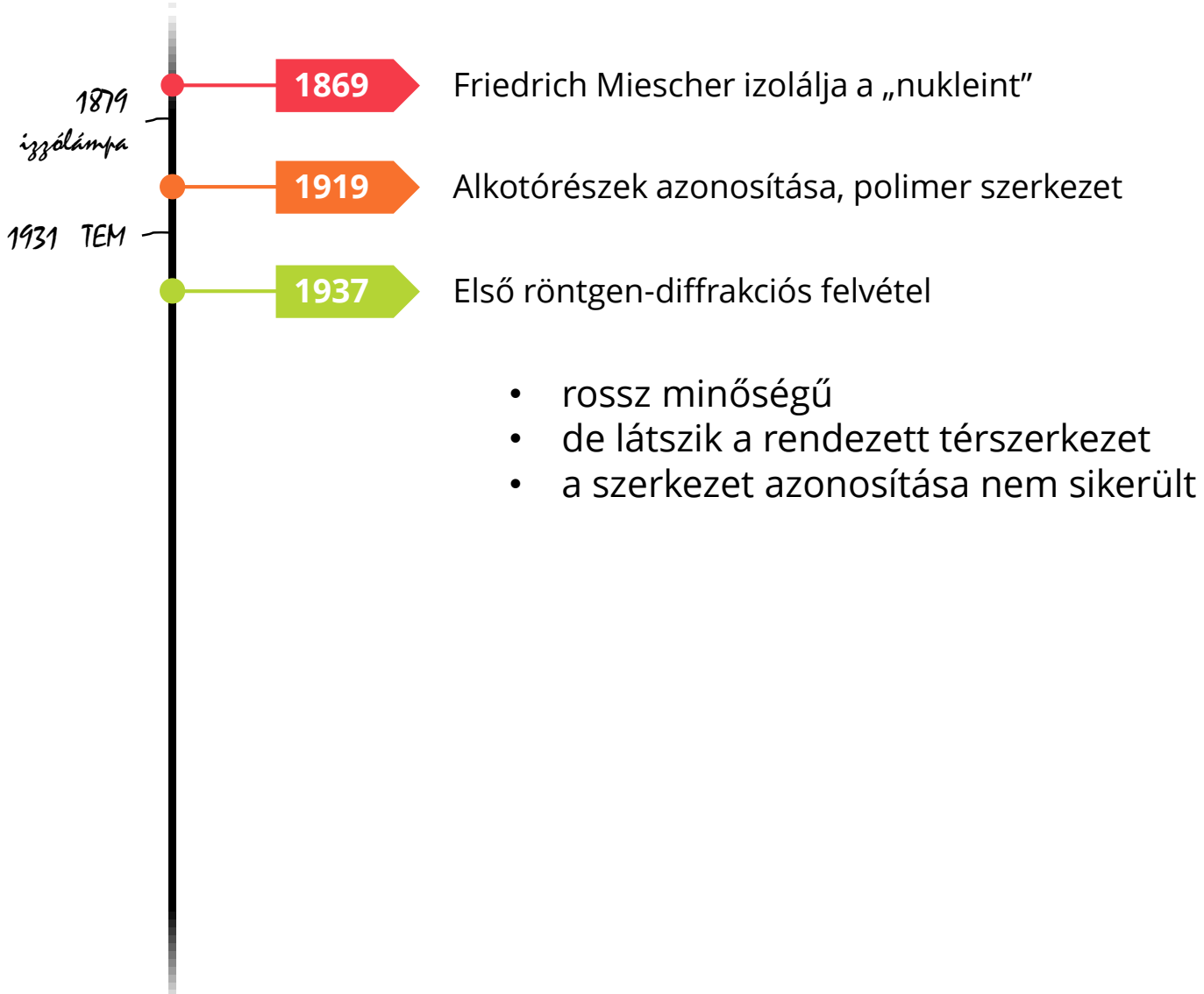
A tübingeni egyetem laboratóriuma (1879 körül)



# A DNS TÖRTÉNETE

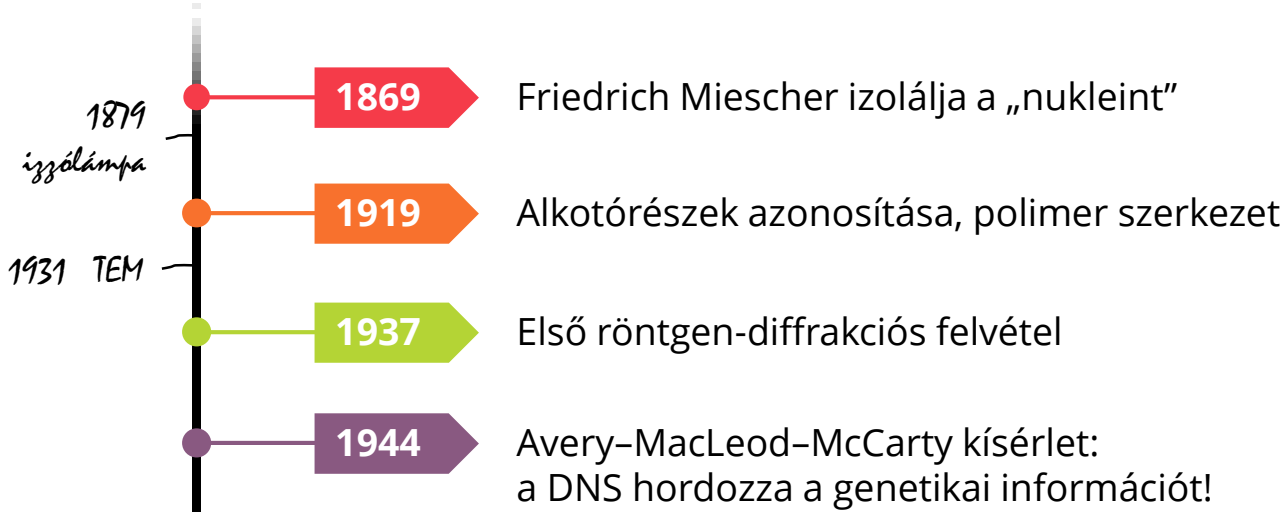


## A DNS TÖRTÉNETE



(William Astbury)

# A DNS TÖRTÉNETE

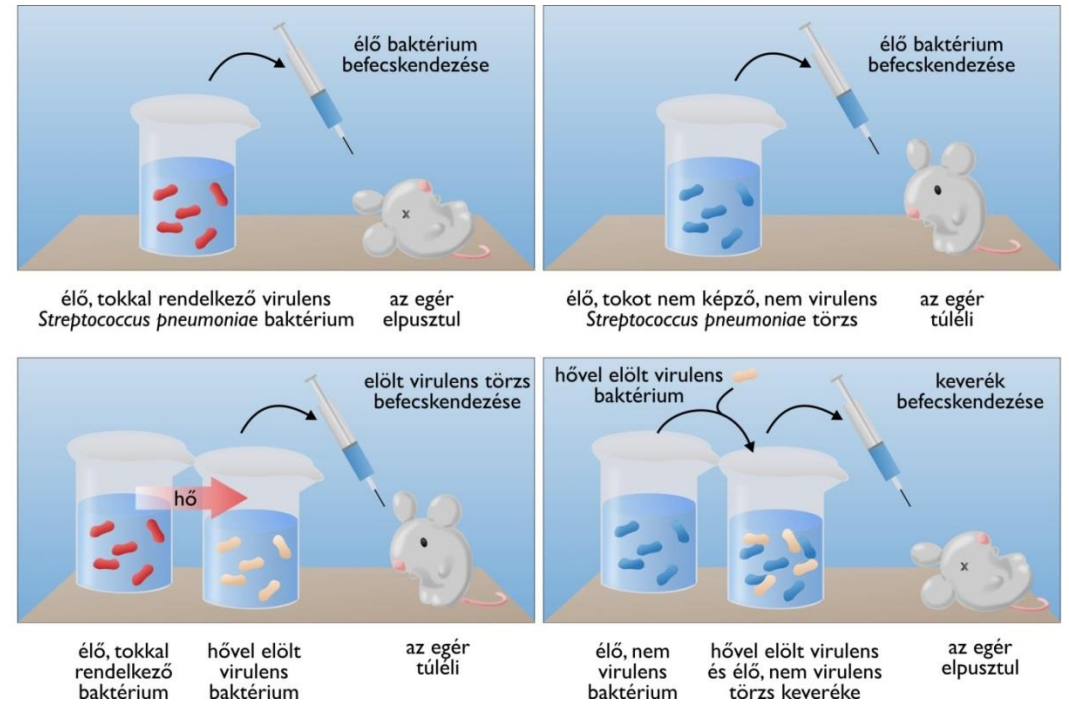


- RNS-, fehérje- és DNS-bontó enzimek

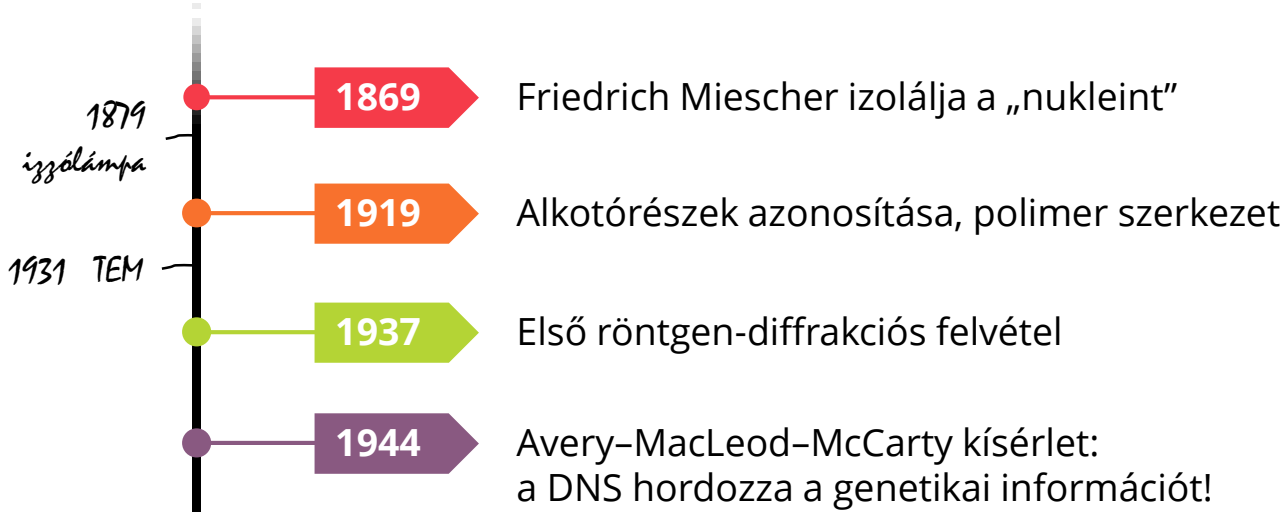


- mi okozza a transzformációt?
  - RNS?
  - fehérjék?
  - DNS?

Nyitray László, Pál Gábor: Griffith-kísérlet



# A DNS TÖRTÉNETE



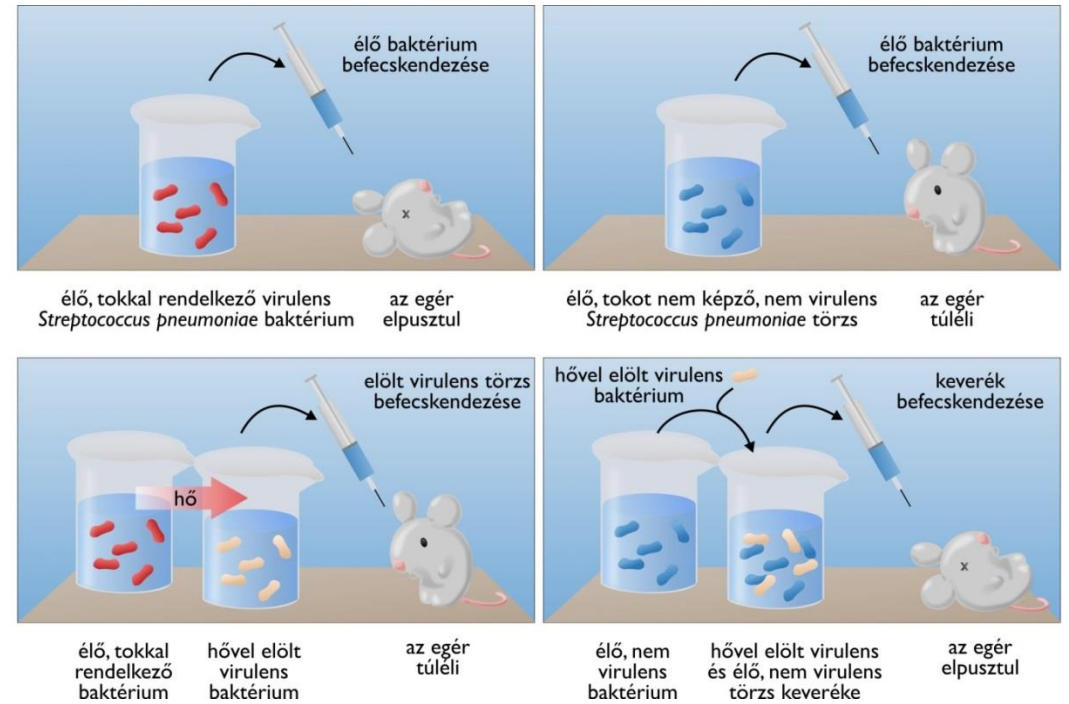
- RNS-, fehérje- és DNS-bontó enzimek



- mi okozza a transzformációt?

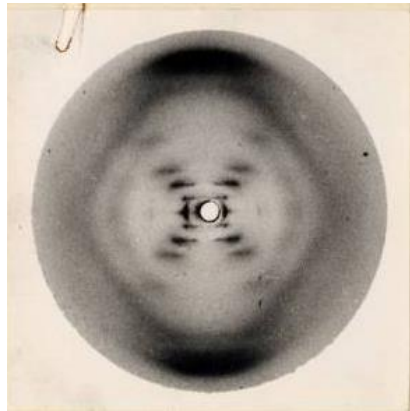
- RNS? ✗
- fehérjék? ✗
- DNS? ✓

Nyitrai László, Pál Gábor: Griffith-kísérlet



# A DNS TÖRTÉNETE

- 1879 *izzólámpa* — 1869 Friedrich Miescher izolálja a „nukleint”
- 1919 Alkotórészek azonosítása, polimer szerkezet
- 1931 TEM — 1937 Első röntgen-diffrakciós felvétel
- 1947 *transzisztor* — 1944 Avery–MacLeod–McCarty kísérlet: a DNS hordozza a genetikai információt!
- 1953 Watson-Crick kettős hélix modell (Franklin felvétele alapján)



*Franklin fizikus volt!*

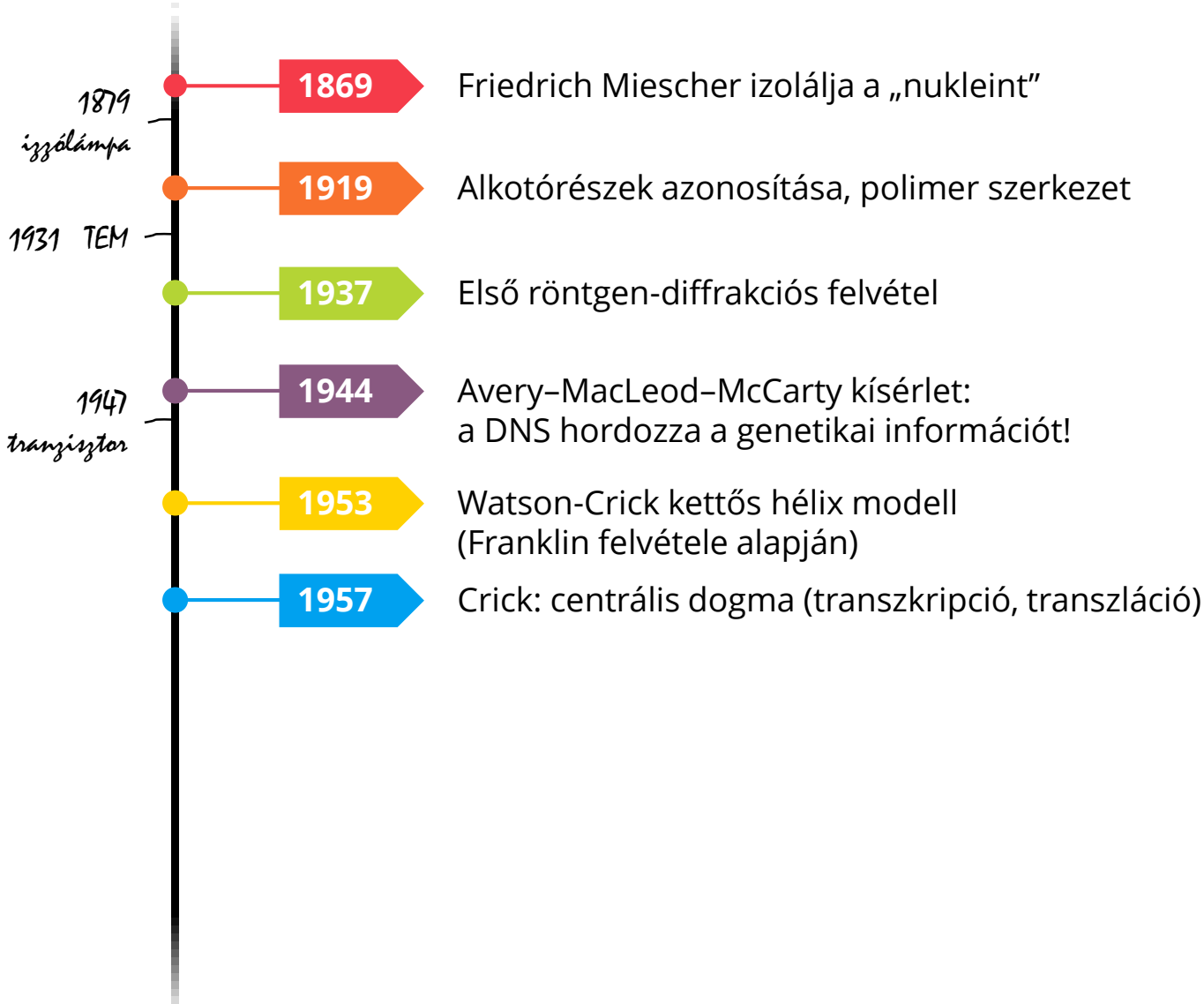
elméleti megfontolások



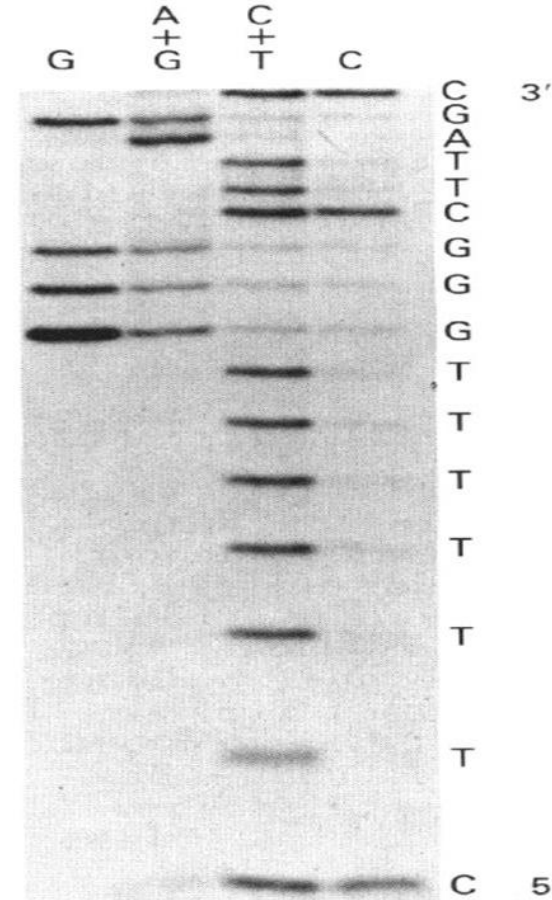
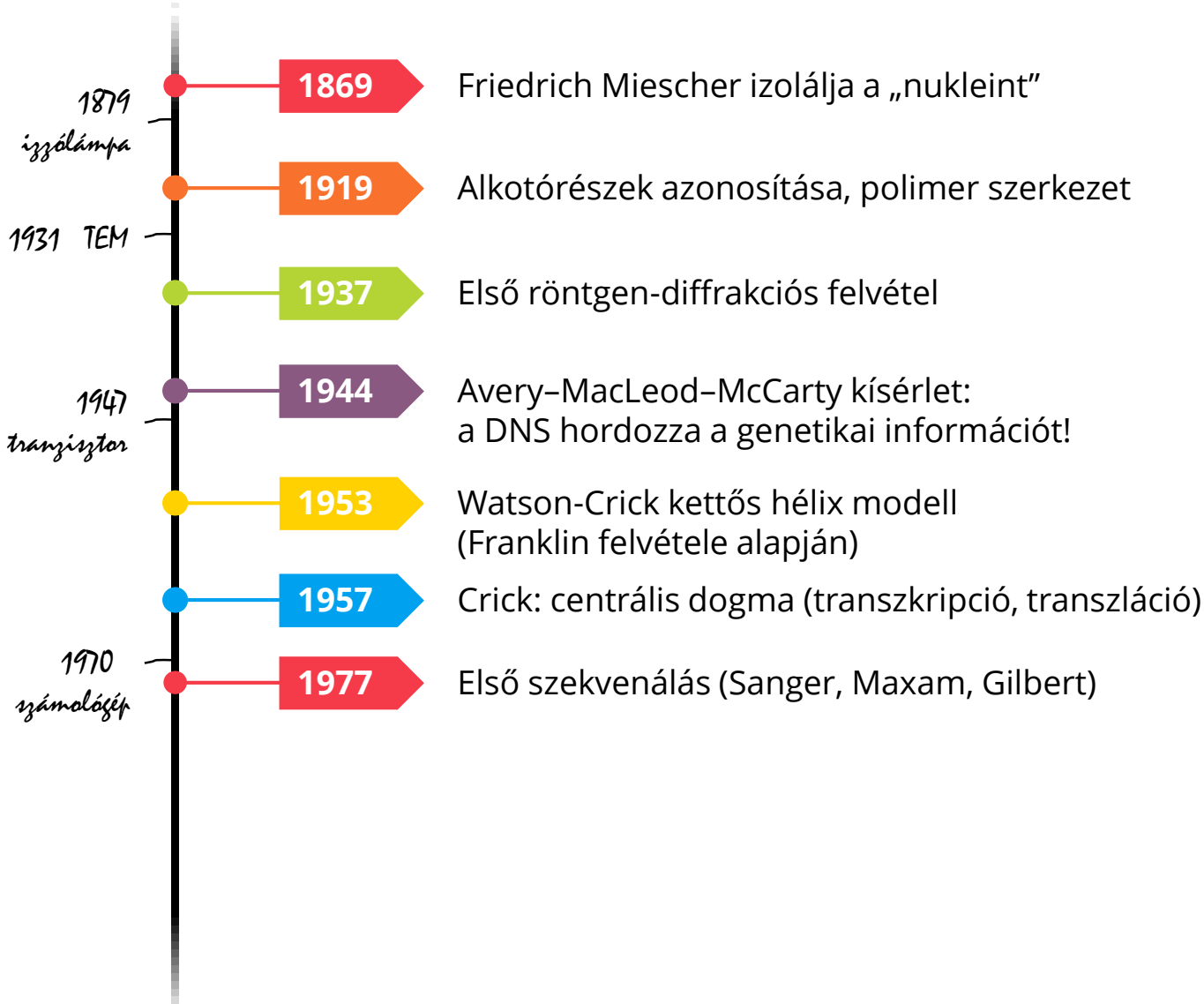
*Crick is!*



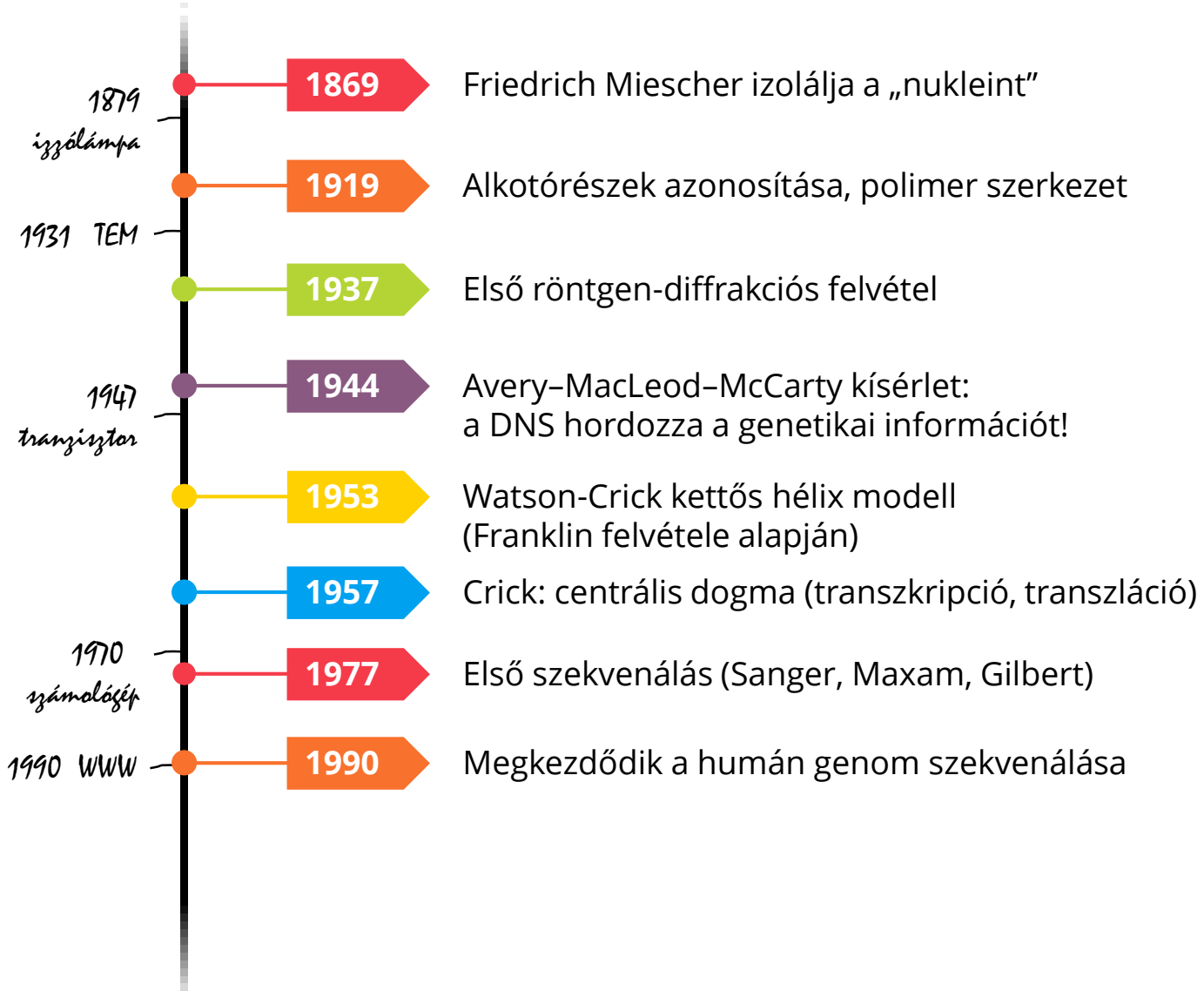
# A DNS TÖRTÉNETE



# A DNS TÖRTÉNETE

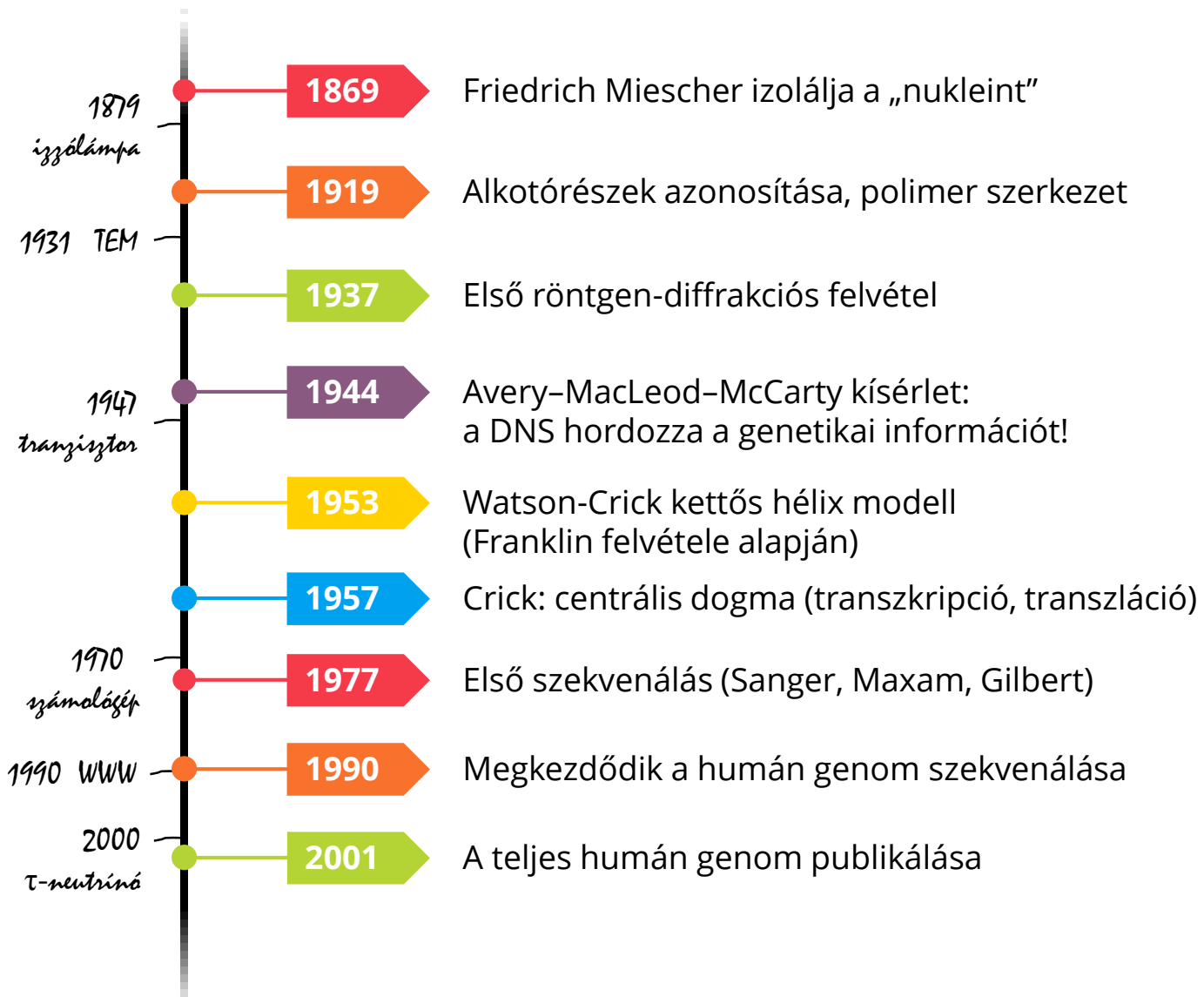


# A DNS TÖRTÉNETE





# A DNS TÖRTÉNETE



# ÚJ KUTATÁSI LEHETŐSÉGEK

## ÚJ KUTATÁSI LEHETŐSÉGEK

- különbségek DNS és DNS között: **örökletes és szerzett mutációk**
  - nagy része „ártatlan”: szemszín, hajszín stb.
- **genetikai betegségek** okozói: a DNS egyes szakaszainak hibái, **mutációi** (akár egy bázis megváltozása is!)
  - színvakság (örökletes)
  - sarlósejtes vérszegénység (örökletes)
  - laktózérzékenység (lehet szerzett)
- **naponta 500 ezer** molekuláris hiba (endogén és exogén tényezők miatt)

## ÚJ KUTATÁSI LEHETŐSÉGEK

- különbségek DNS és DNS között: **örökletes és szerzett mutációk**
  - nagy része „ártatlan”: szemszín, hajszín stb.
- **genetikai betegségek** okozói: a DNS egyes szakaszainak hibái, **mutációi** (akár egy bázis megváltozása is!)
  - színvakság (örökletes)
  - sarlósejtes vérszegénység (örökletes)
  - laktózérzékenység (lehet szerzett)
- **naponta 500 ezer** molekuláris hiba (endogén és exogén tényezők miatt)

### Miért ritkák ehhez képest a **genetikai betegségek**?

- a hibák véletlenszerűek
- **nagyon hatékony javító-mechanizmusok!**

## ÚJ KUTATÁSI LEHETŐSÉGEK

- különbségek DNS és DNS között: **örökletes és szerzett mutációk**
  - nagy része „ártatlan”: szemszín, hajszín stb.
- **genetikai betegségek** okozói: a DNS egyes szakaszainak hibái, **mutációi** (akár egy bázis megváltozása is!)
  - színvakság (örökletes)
  - sarlósejtes vérszegénység (örökletes)
  - laktózérzékenység (lehet szerzett)
- **naponta 500 ezer** molekuláris hiba (endogén és exogén tényezők miatt)

### Miért ritkák ehhez képest a **genetikai betegségek**?

- a hibák véletlenszerűek
- **nagyon hatékony javító-mechanizmusok!**

→ Nem akkor van baj, ha a DNS „elromlik”, hanem ha nem tudjuk kijavítani!

## ÚJ KUTATÁSI LEHETŐSÉGEK

Néhány lehetséges mutáció:

Kis-skálás

Nagy-skálás

- pontmutáció (T>G)



# ÚJ KUTATÁSI LEHETŐSÉGEK

Néhány lehetséges mutáció:

Kis-skálás

Nagy-skálás

- pontmutáció (T>G)
- inzerció



## ÚJ KUTATÁSI LEHETŐSÉGEK

Néhány lehetséges mutáció:

Kis-skálás

Nagy-skálás

- pontmutáció (T>G)
- inzerció
- deléció





## ÚJ KUTATÁSI LEHETŐSÉGEK

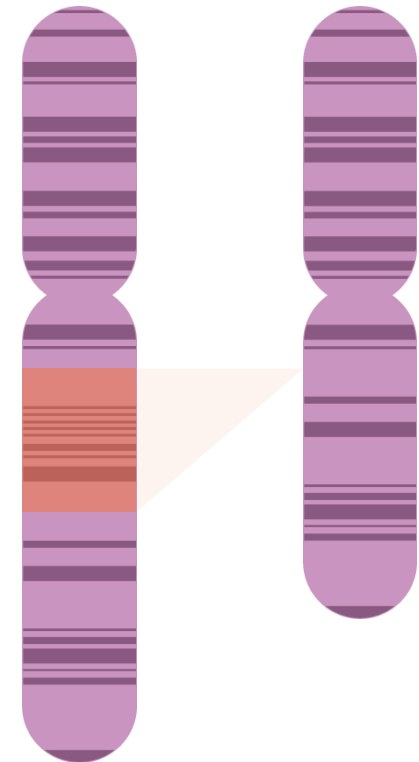
Néhány lehetséges mutáció:

### Kis-skálás

- pontmutáció (T>G)
- inzerció
- deléció

### Nagy-skálás

- kromoszómán belüli:
  - deléció



## ÚJ KUTATÁSI LEHETŐSÉGEK

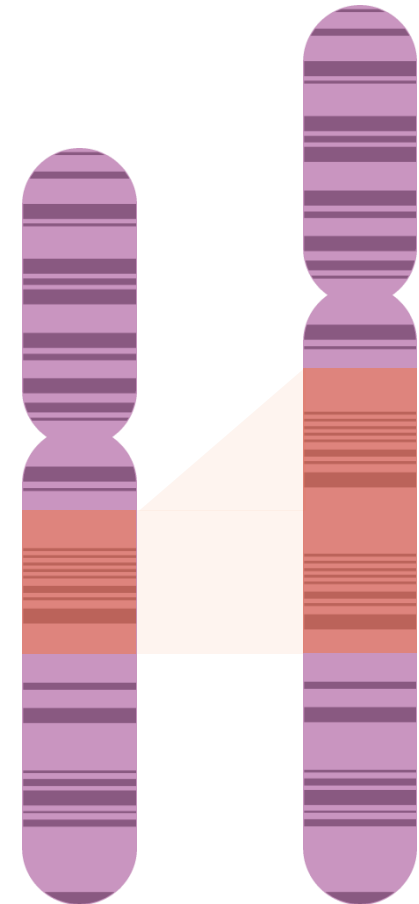
Néhány lehetséges mutáció:

### Kis-skálás

- pontmutáció (T>G)
- inzerció
- deléció

### Nagy-skálás

- kromoszómán belüli:
  - deléció
  - inzerció/duplikáció



## ÚJ KUTATÁSI LEHETŐSÉGEK

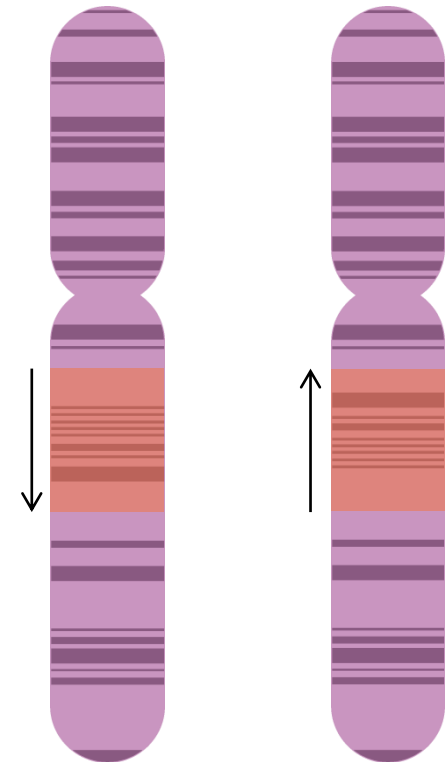
Néhány lehetséges mutáció:

### Kis-skálás

- pontmutáció (T>G)
- inzerció
- deléció

### Nagy-skálás

- kromoszómán belüli:
  - deléció
  - inzerció/duplikáció
  - inverzió



# ÚJ KUTATÁSI LEHETŐSÉGEK

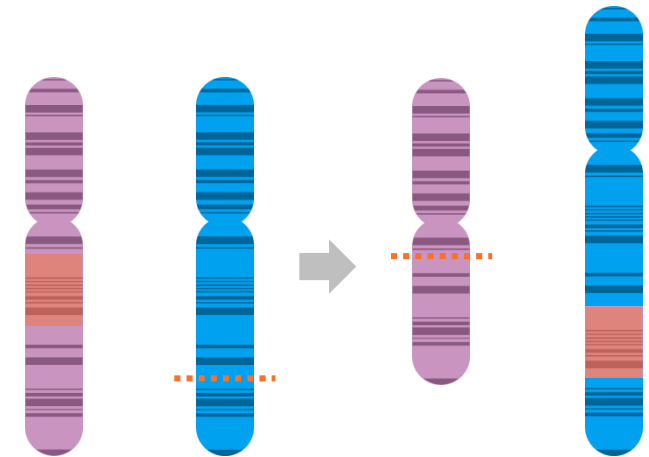
Néhány lehetséges mutáció:

## Kis-skálás

- pontmutáció (T>G)
- inzerció
- deléció

## Nagy-skálás

- kromoszómán belüli:
  - deléció
  - inzerció/duplikáció
  - inverzió
- kromoszómák közti:
  - inzerció



## ÚJ KUTATÁSI LEHETŐSÉGEK

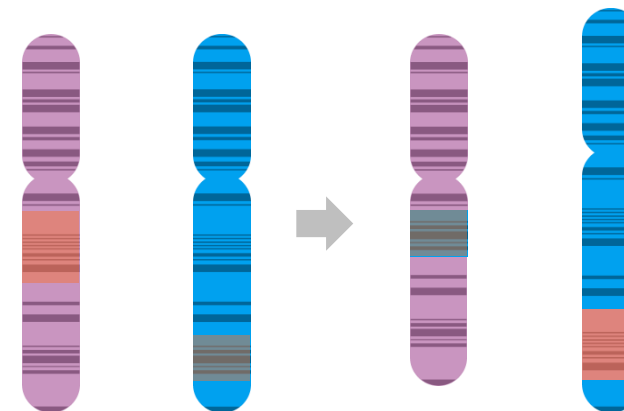
Néhány lehetséges mutáció:

### Kis-skálás

- pontmutáció (T>G)
- inzerció
- deléció

### Nagy-skálás

- kromoszómán belüli:
  - deléció
  - inzerció/duplikáció
  - inverzió
- kromoszómák közti:
  - inzerció
  - transzlokáció

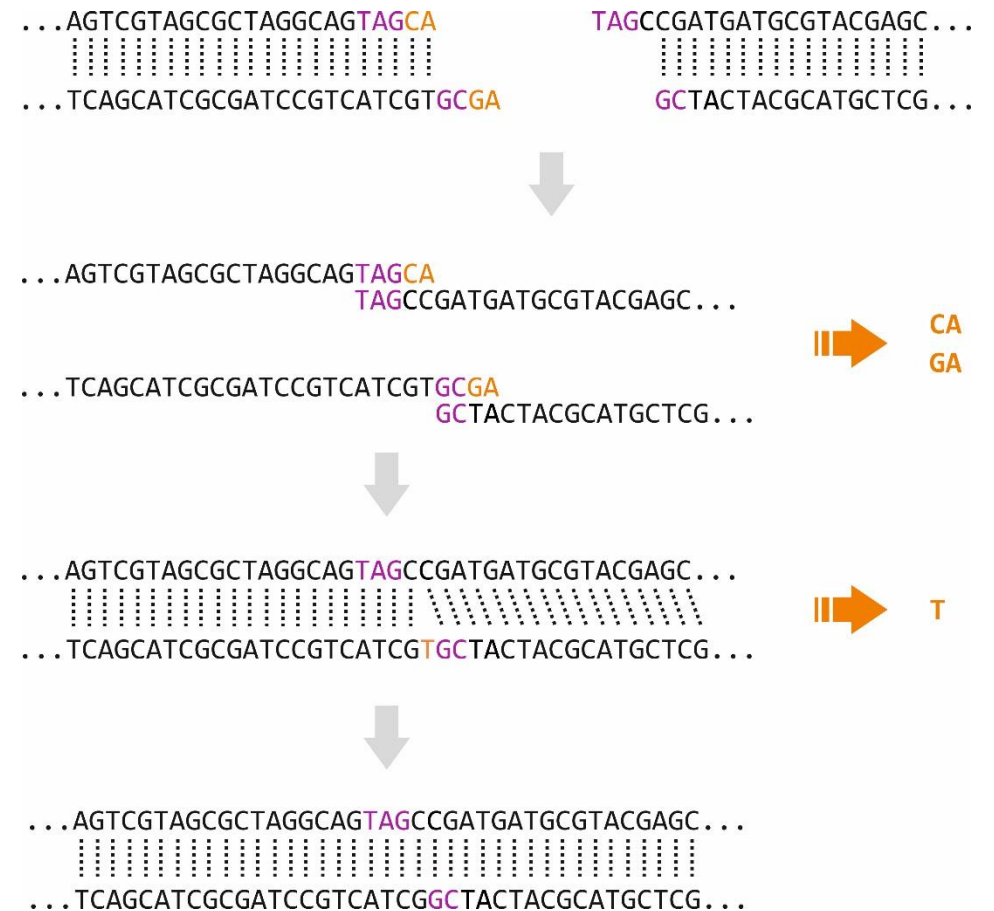


# ÚJ KUTATÁSI LEHETŐSÉGEK

Néhány DNS-javító mechanizmus:

- **egyik szálon** sérült DNS javítása:
  1. hiba felismerése, hibás szakasz eltávolítása
  2. másik szálról komplementer szakasz legyártása
  - BER: oxidálódott, alkilizálódott vagy hidrolizálódott nukleotidok cseréje
  - NER: hélixszerkezetet torzító mutációk javítása
  - MMR: nem összetartozó (tehát nem AT vagy CG) bázispárok korrekciója
- **kettős szál törés:**
  - szálak „összeragasztása”
    - NHEJ: minta nélkül, rövid átfedő szakaszt keresve (→ rövid deléciók)
    - rekombináció: minta alapján

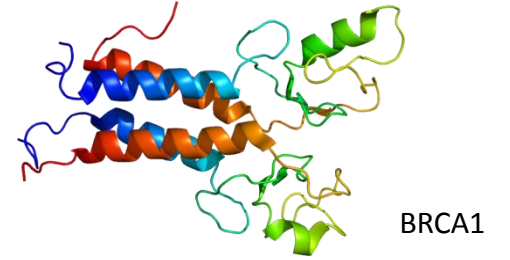
NHEJ mechanizmus



## ÚJ KUTATÁSI LEHETŐSÉGEK



## ÚJ KUTATÁSI LEHETŐSÉGEK

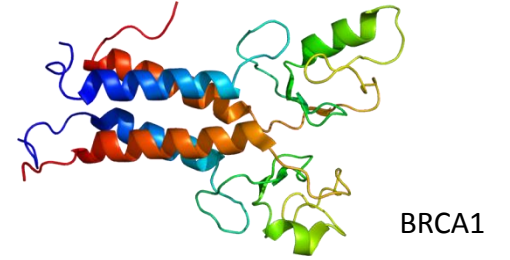


A DNS másolása és javítása nagyon komplex feladat, nagyon sok enzim és fehérje vesz részt benne!

→ egyelőre nagyon keveset tudunk a konkrét mechanizmusokról, ezek feltérképezése lenne a cél



## ÚJ KUTATÁSI LEHETŐSÉGEK

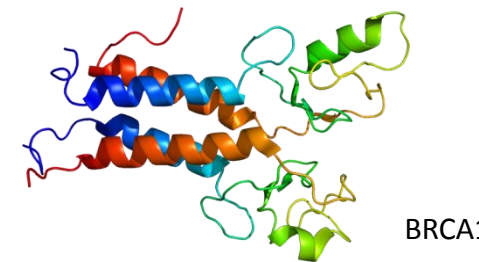


A DNS másolása és javítása nagyon komplex feladat, nagyon sok enzim és fehérje vesz részt benne!

→ egyelőre nagyon keveset tudunk a konkrét mechanizmusokról, ezek feltérképezése lenne a cél

Módszer:

## ÚJ KUTATÁSI LEHETŐSÉGEK



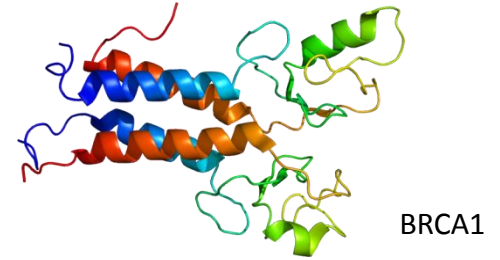
A DNS másolása és javítása nagyon komplex feladat, nagyon sok enzim és fehérje vesz részt benne!

→ egyelőre nagyon keveset tudunk a konkrét mechanizmusokról, ezek feltérképezése lenne a cél

Módszer:

1. Egy adag sejtben direkt, célzottan rontsuk el valamelyik „fontosnak gyanított” gént. (→ **mutáns sejtek**)

## ÚJ KUTATÁSI LEHETŐSÉGEK



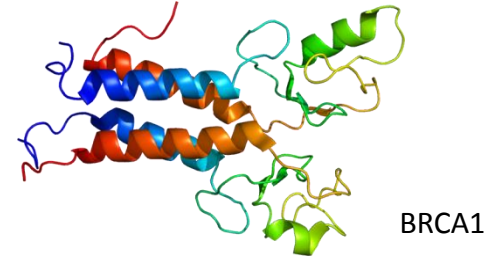
A DNS másolása és javítása nagyon komplex feladat, nagyon sok enzim és fehérje vesz részt benne!

→ egyelőre nagyon keveset tudunk a konkrét mechanizmusokról, ezek feltérképezése lenne a cél

### Módszer:

1. Egy adag sejtben direkt, célzottan rontsuk el valamelyik „fontosnak gyanított” gént. (→ **mutáns sejtek**)
2. Kezeljük az „egészséges” és a mutáns sejteket valamilyen **károsító anyaggal**, ami véletlenszerű mutációkat hoz létre.

## ÚJ KUTATÁSI LEHETŐSÉGEK



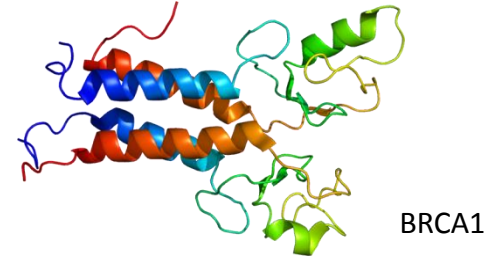
A DNS másolása és javítása nagyon komplex feladat, nagyon sok enzim és fehérje vesz részt benne!

→ egyelőre nagyon keveset tudunk a konkrét mechanizmusokról, ezek feltérképezése lenne a cél

### Módszer:

1. Egy adag sejtben direkt, célzottan rontsuk el valamelyik „fontosnak gyanított” gént. (→ **mutáns sejtek**)
2. Kezeljük az „egészséges” és a mutáns sejteket valamilyen **károsító anyaggal**, ami véletlenszerű mutációkat hoz létre.
3. Vizsgáljuk, hogy a mutáns és az egészséges sejtek **mennyire jól tudják kijavítani** a véletlenszerű mutációkat.





## ÚJ KUTATÁSI LEHETŐSÉGEK



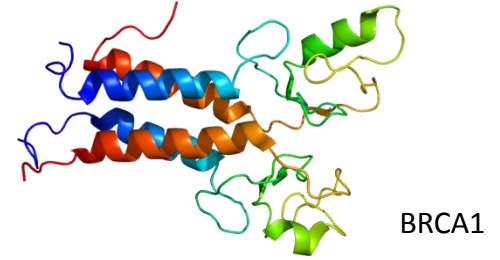
A DNS másolása és javítása nagyon komplex feladat, nagyon sok enzim és fehérje vesz részt benne!

→ egyelőre nagyon keveset tudunk a konkrét mechanizmusokról, ezek feltérképezése lenne a cél

### Módszer:

-  1. Egy adag sejtben direkt, célzottan rontsuk el valamelyik „fontosnak gyanított” gént. (→ **mutáns sejtek**)
-  2. Kezeljük az „egészséges” és a mutáns sejteket valamilyen **károsító anyaggal**, ami véletlenszerű mutációkat hoz létre.
-  3. Vizsgáljuk, hogy a mutáns és az egészséges sejtek **mennyire jól tudják kijavítani** a véletlenszerű mutációkat.
-  4. Abból, hogy a mutáns sejtekben milyen típusú végleges mutációkat találunk, következtetni lehet arra, hogy az eredetileg megrongált génnek **melyik javító-mechanizmusban milyen szerepe van**.

## ÚJ KUTATÁSI LEHETŐSÉGEK



A DNS másolása és javítása nagyon komplex feladat, nagyon sok enzim és fehérje vesz részt benne!

→ egyelőre nagyon keveset tudunk a konkrét mechanizmusokról, ezek feltérképezése lenne a cél

### Módszer:

1. Egy adag sejtben direkt, célzottan rontsuk el valamelyik „fontosnak gyanított” gént. (→ **mutáns sejtek**)
2. Kezeljük az „egészséges” és a mutáns sejteket valamilyen **károsító anyaggal**, ami véletlenszerű mutációkat hoz létre.
3. Vizsgáljuk, hogy a mutáns és az egészséges sejtek **mennyire jól tudják kijavítani** a véletlenszerű mutációkat.
4. Abból, hogy a mutáns sejtekben milyen típusú végleges mutációkat találunk, következtetni lehet arra, hogy az eredetileg megrongált génnek **melyik javító-mechanizmusban milyen szerepe van**.

### SZEKVENÁLÁS + SZEKVENÁLÁSI ADATOK ELEMZÉSE

# ÚJ KUTATÁSI LEHETŐSÉGEK

## Közös projectek az MTA Enzimológiai Intézetével:

Oncogene (2017) 36, 746–755  
www.nature.com/onc

OPEN

**ORIGINAL ARTICLE**

Loss of BRCA1 or BRCA2 markedly increases the rate of base substitution mutagenesis and has distinct effects on genomic deletions

J Zámboreszky<sup>1</sup>, B Szikriszt<sup>1</sup>, JZ Gervai<sup>1</sup>, O Pipek<sup>2</sup>, Á Póti<sup>1</sup>, M Krzystanek<sup>3</sup>, D Ribli<sup>2</sup>, JM Szalai-Gindl<sup>2</sup>, I Csabai<sup>2</sup>, Z Szallasi<sup>3,4,5,6</sup>, C Swanton<sup>7,8</sup>, AL Richardson<sup>9</sup> and D Szüts<sup>1</sup>

Szikriszt et al. *Genome Biology* (2016) 17:99  
DOI 10.1186/s13059-016-0963-7

Genome Biology

RESEARCH Open Access

A comprehensive survey of the mutagenic impact of common cancer cytotoxics

Bernadett Szikriszt<sup>1</sup>, Ádám Póti<sup>1</sup>, Orsolya Pipek<sup>2</sup>, Marcin Krzystanek<sup>3</sup>, Nnennaya Kanu<sup>4</sup>, János Molnár<sup>1</sup>, Dezső Ribli<sup>2</sup>, Zoltán Szeltner<sup>1</sup>, Gábor E. Tusnády<sup>1</sup>, István Csabai<sup>2</sup>, Zoltan Szallasi<sup>3,5,6,8\*</sup>, Charles Swanton<sup>4,7\*</sup> and Dávid Szüts<sup>1\*</sup>

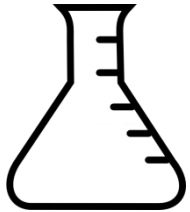


## The Genome of the Chicken DT40 Bursal Lymphoma Cell Line

János Molnár,\* Ádám Póti,\* Orsolya Pipek,† Marcin Krzystanek,‡ Nnennaya Kanu,§ Charles Swanton,§ Gábor E. Tusnády,\* Zoltan Szallasi.†,\*\* István Csabai.† and Dávid Szüts\*<sup>1</sup>

# SZEKVENÁLÁSI TECHNIKÁK

Sanger-szekvenálás





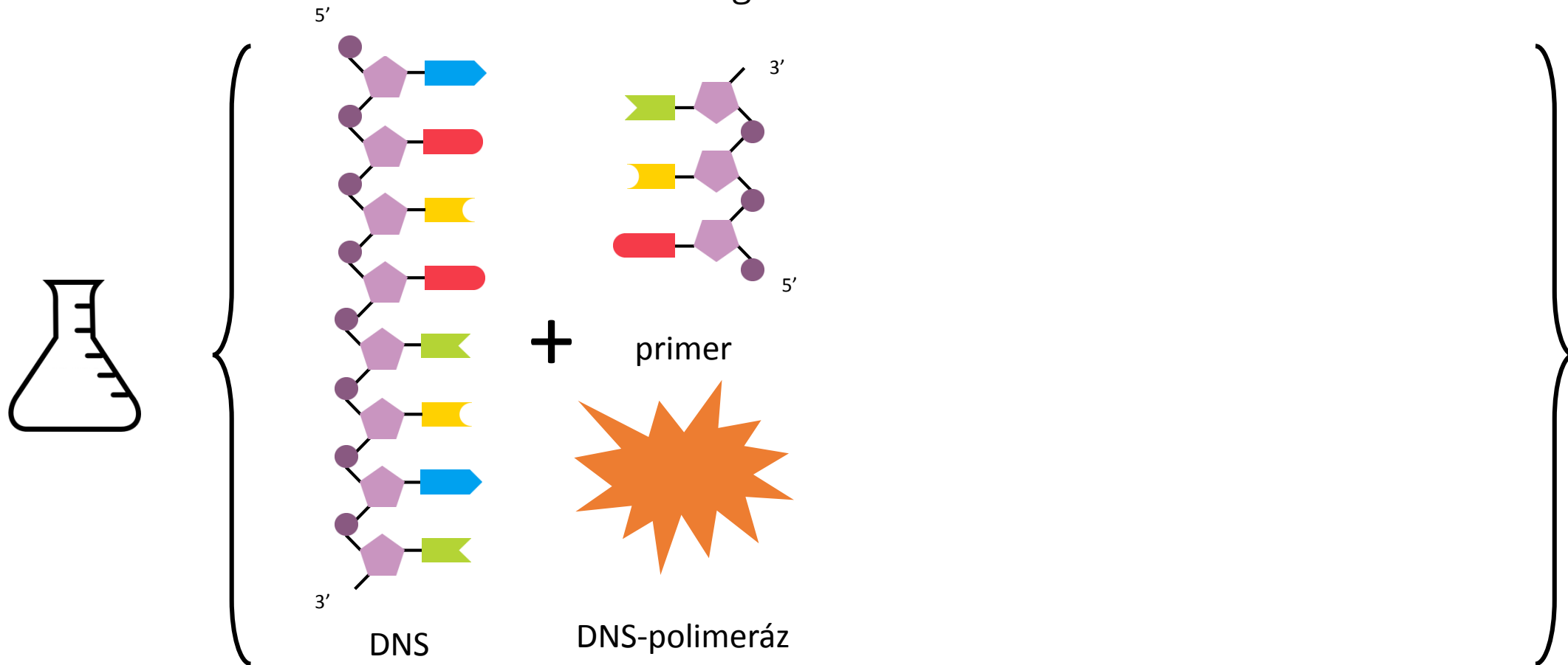
# SZEKVENÁLÁSI TECHNIKÁK

## Sanger-szekvenálás



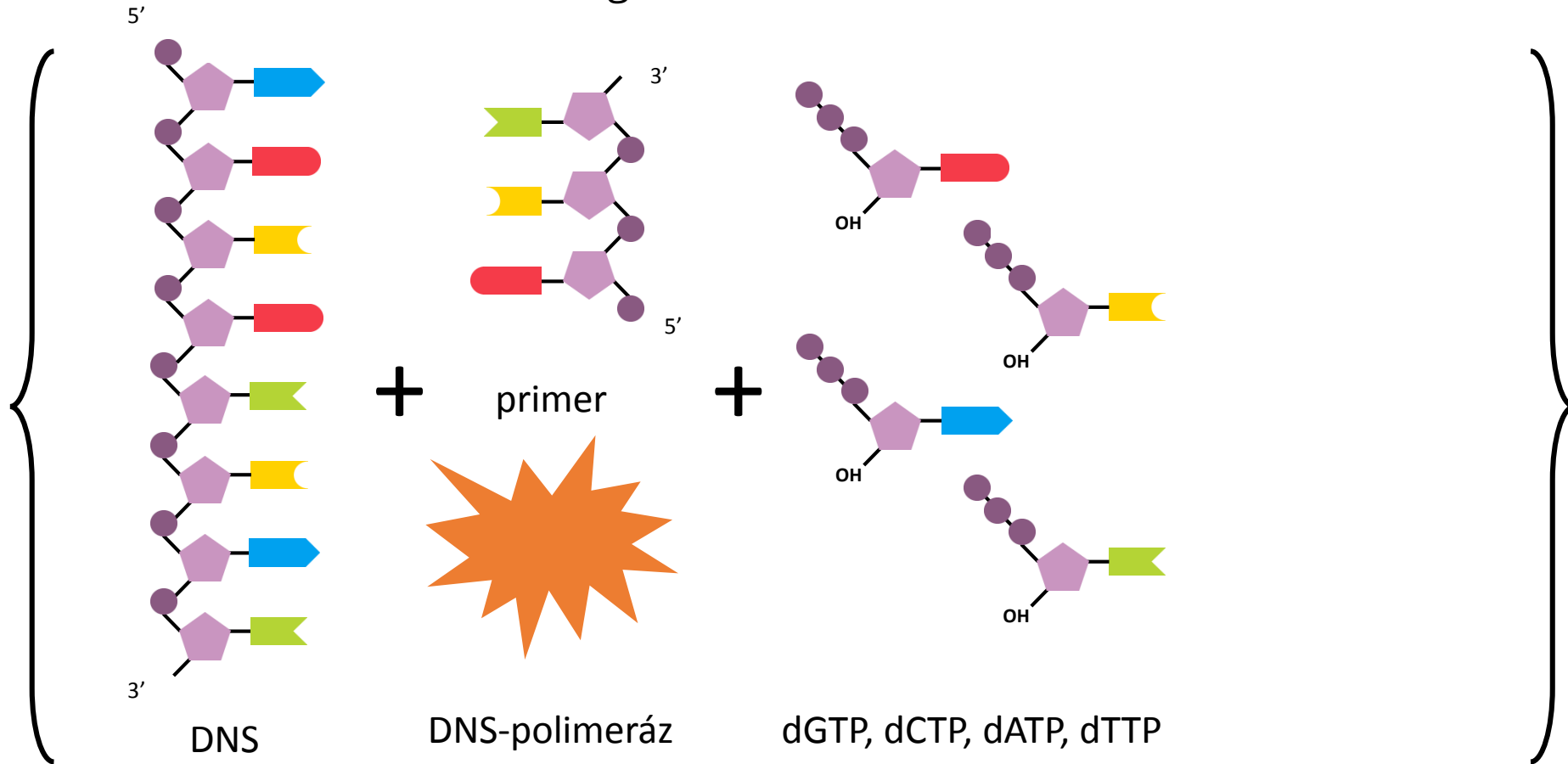
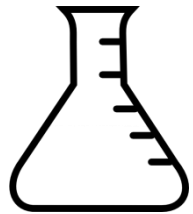
# SZEKVENÁLÁSI TECHNIKÁK

## Sanger-szekvenálás



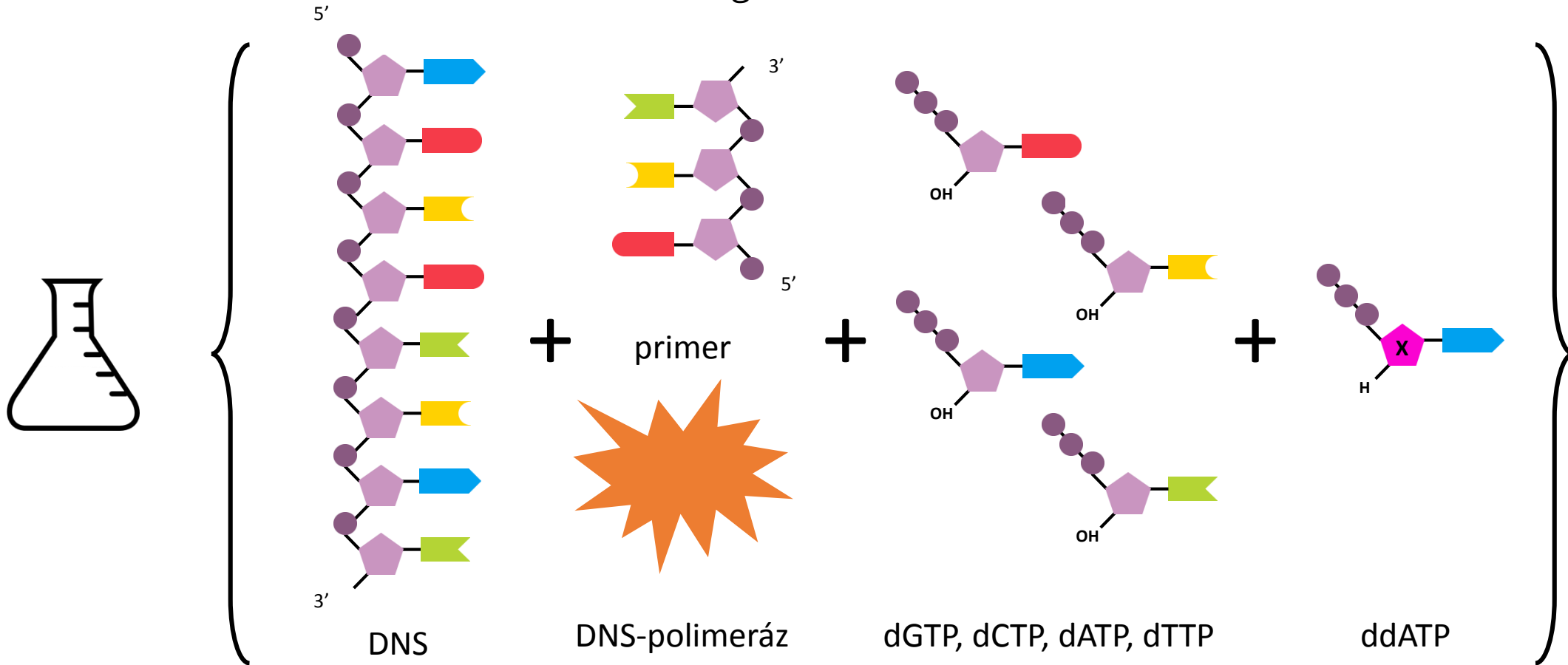
# SZEKVENÁLÁSI TECHNIKÁK

## Sanger-szekvenálás



# SZEKVENÁLÁSI TECHNIKÁK

## Sanger-szekvenálás

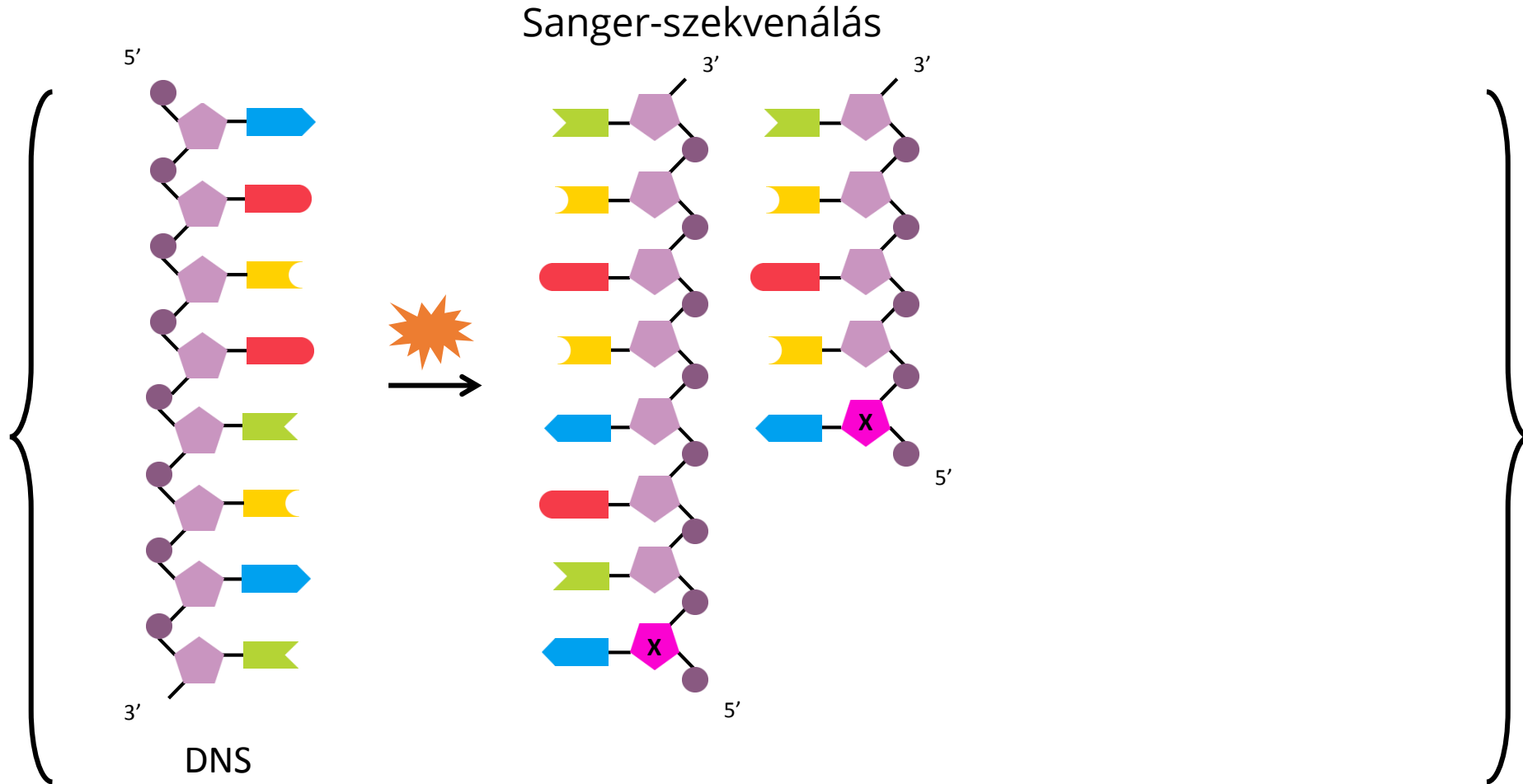
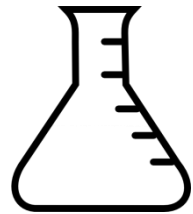


# SZEKVENÁLÁSI TECHNIKÁK

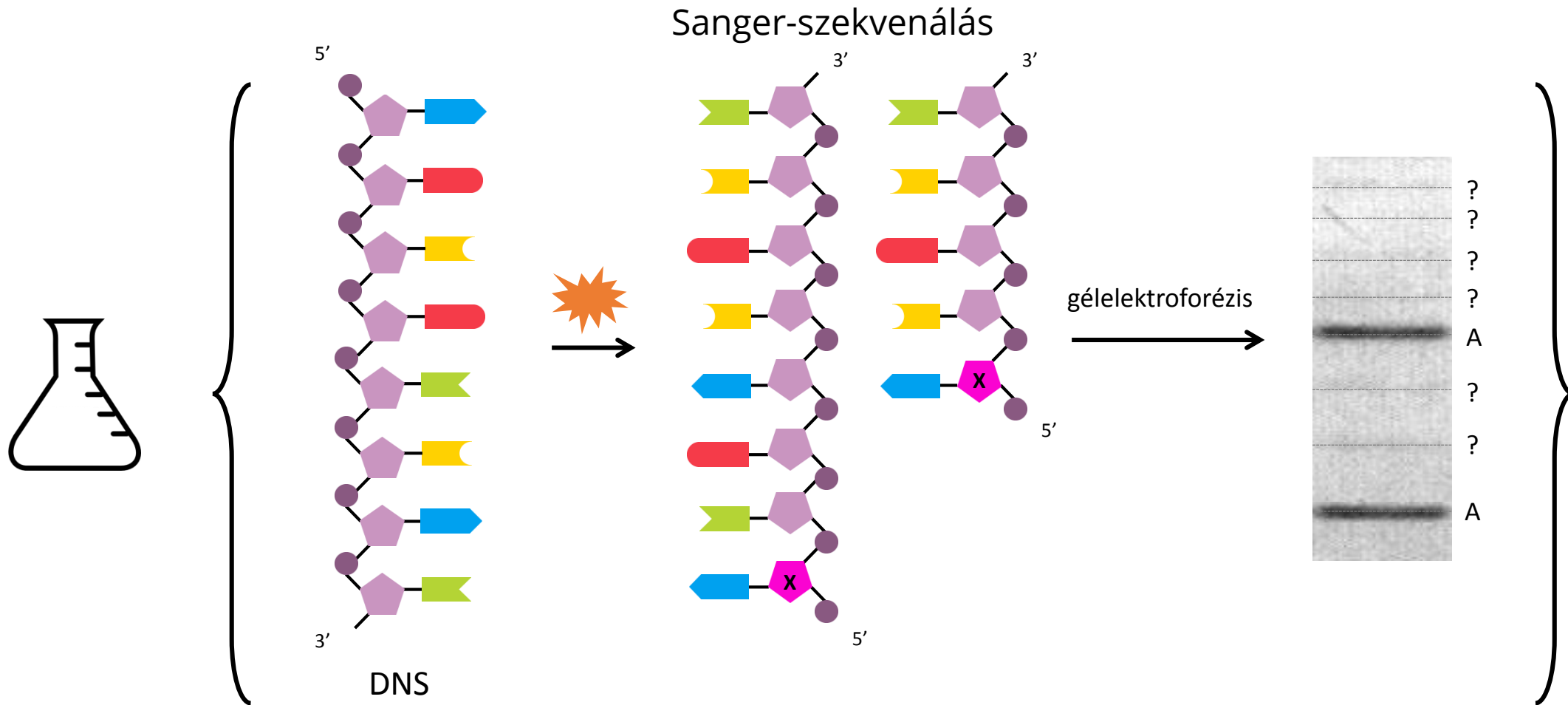
## Sanger-szekvenálás



# SZEKVENÁLÁSI TECHNIKÁK

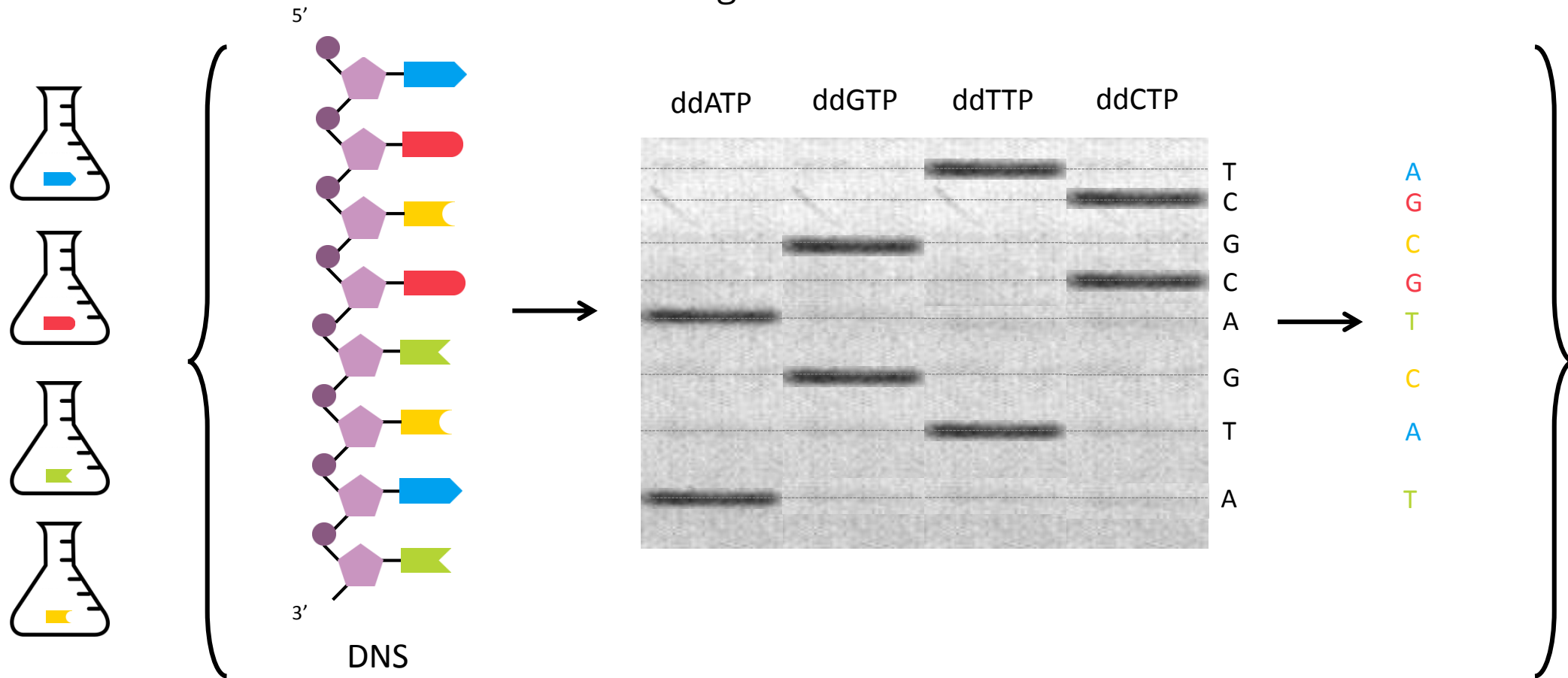


# SZEKVENÁLÁSI TECHNIKÁK



# SZEKVENÁLÁSI TECHNIKÁK

## Sanger-szekvenálás





# SZEKVENÁLÁSI TECHNIKÁK

NGS: next-generation sequencing

# SZEKVENÁLÁSI TECHNIKÁK

NGS: next-generation sequencing

1. DNS felszaporítása (PCR), darabokra („short read”) tördelése

# SZEKVENÁLÁSI TECHNIKÁK

NGS: next-generation sequencing

1. DNS felszaporítása (PCR), darabokra („short read”) tördelése
2. Egyszálú short readek rögzítése a cellán primerekkel

# SZEKVENÁLÁSI TECHNIKÁK

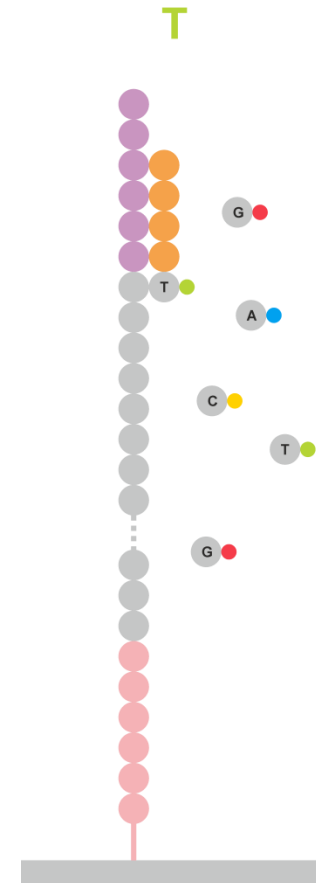
NGS: next-generation sequencing

1. DNS felszaporítása (PCR), darabokra („short read”) tördelése
2. Egyszálú short readek rögzítése a cellán primerekkel
3. Short readek felszaporítása a közvetlen környezetükben  
→ klaszterenként (1000+) teljesen azonos short readek

# SZEKVENÁLÁSI TECHNIKÁK

NGS: next-generation sequencing

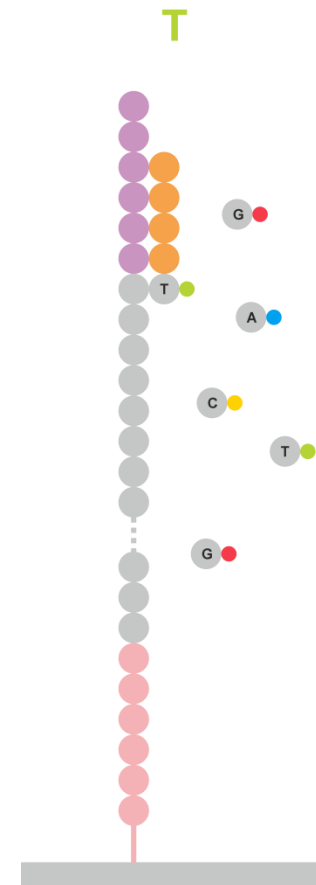
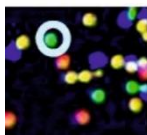
1. DNS felszaporítása (PCR), darabokra („short read”) tördelése
2. Egyszálú short readek rögzítése a cellán primerekkel
3. Short readek felszaporítása a közvetlen környezetükben  
→ klaszterenként (1000+) teljesen azonos short readek
4. Cella leöntése a négyféle festékanyaggal megjelölt ddNTP-vel és DNS-polimeráz enzimmal



# SZEKVENÁLÁSI TECHNIKÁK

NGS: next-generation sequencing

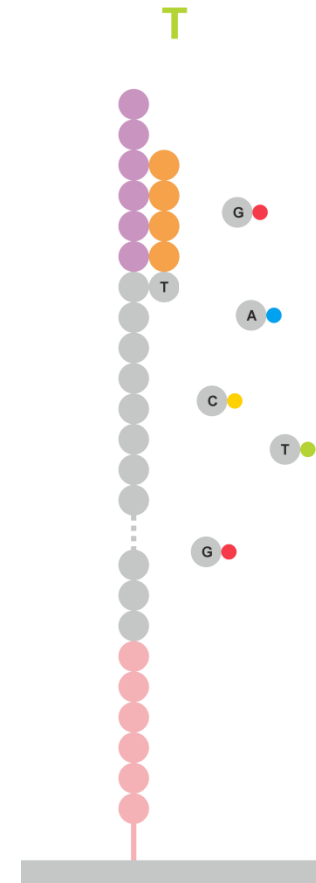
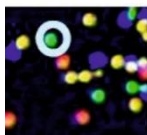
1. DNS felszaporítása (PCR), darabokra („short read”) tördelése
2. Egyszálú short readek rögzítése a cellán primerekkel
3. Short readek felszaporítása a közvetlen környezetükben  
→ klaszterenként (1000+) teljesen azonos short readek
4. Cella leöntése a négyféle festékanyaggal megjelölt ddNTP-vel és DNS-polimeráz enzimmal
5. Klaszterenkénti színes foltok rögzítése képként



# SZEKVENÁLÁSI TECHNIKÁK

NGS: next-generation sequencing

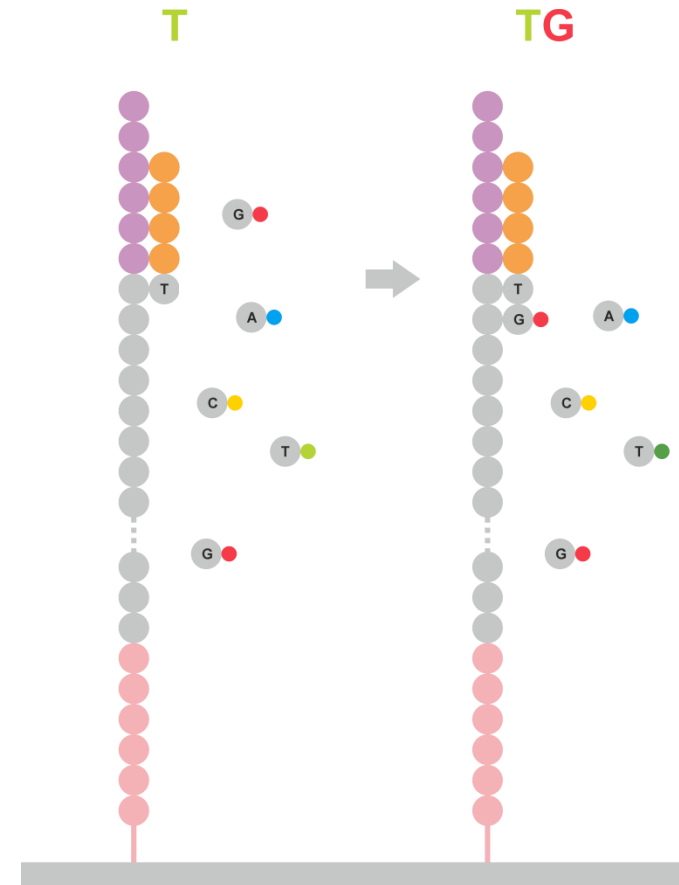
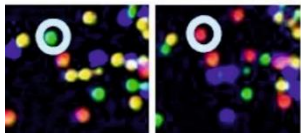
1. DNS felszaporítása (PCR), darabokra („short read”) tördelése
2. Egyszálú short readek rögzítése a cellán primerekkel
3. Short readek felszaporítása a közvetlen környezetükben  
→ klaszterenként (1000+) teljesen azonos short readek
4. Cella leöntése a négyféle festékanyaggal megjelölt ddNTP-vel és DNS-polimeráz enzimmal
5. Klaszterenkénti színes foltok rögzítése képként
6. Beépült bázisok „semlegesítése”



# SZEKVENÁLÁSI TECHNIKÁK

NGS: next-generation sequencing

1. DNS felszaporítása (PCR), darabokra („short read”) tördelése
2. Egyszálú short readek rögzítése a cellán primerekkel
3. Short readek felszaporítása a közvetlen környezetükben  
→ klaszterenként (1000+) teljesen azonos short readek
4. Cella leöntése a négyféle festékanyaggal megjelölt ddNTP-vel és DNS-polimeráz enzimmal
5. Klaszterenkénti színes foltok rögzítése képként
6. Beépült bázisok „semlegesítése”

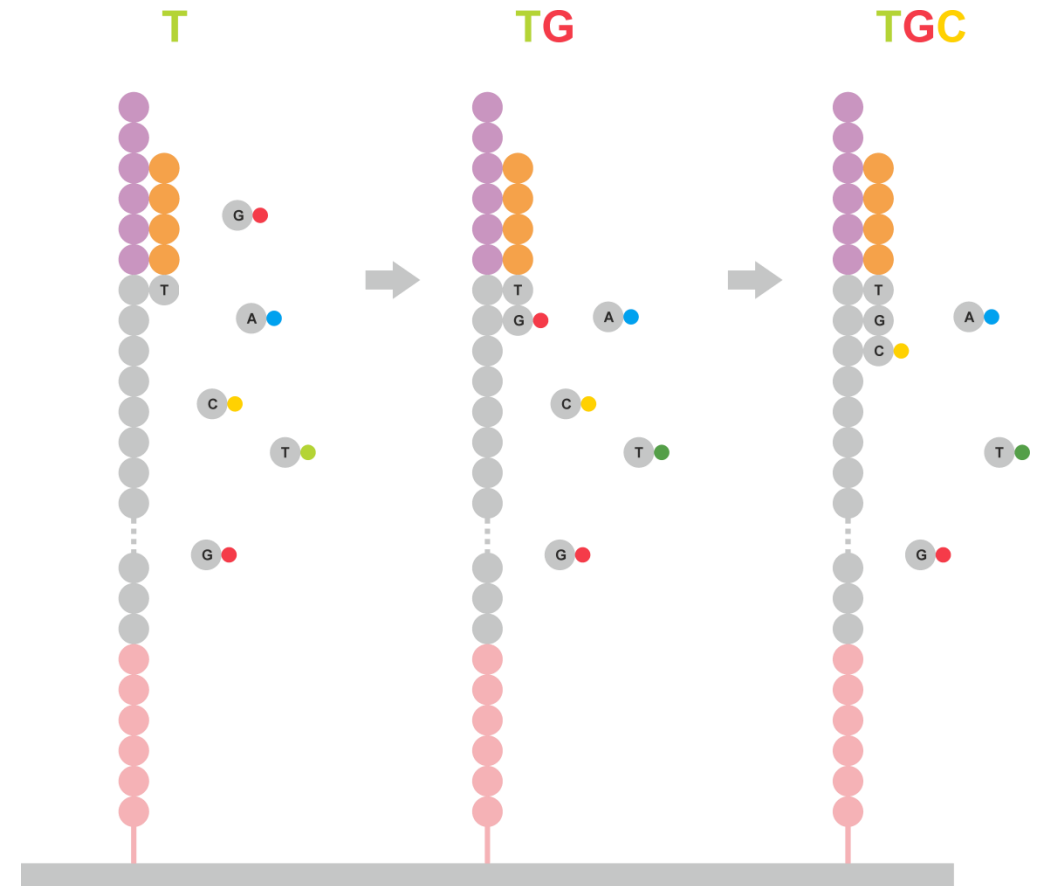
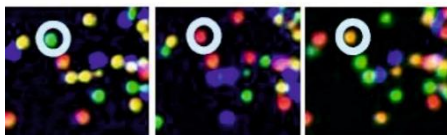




# SZEKVENÁLÁSI TECHNIKÁK

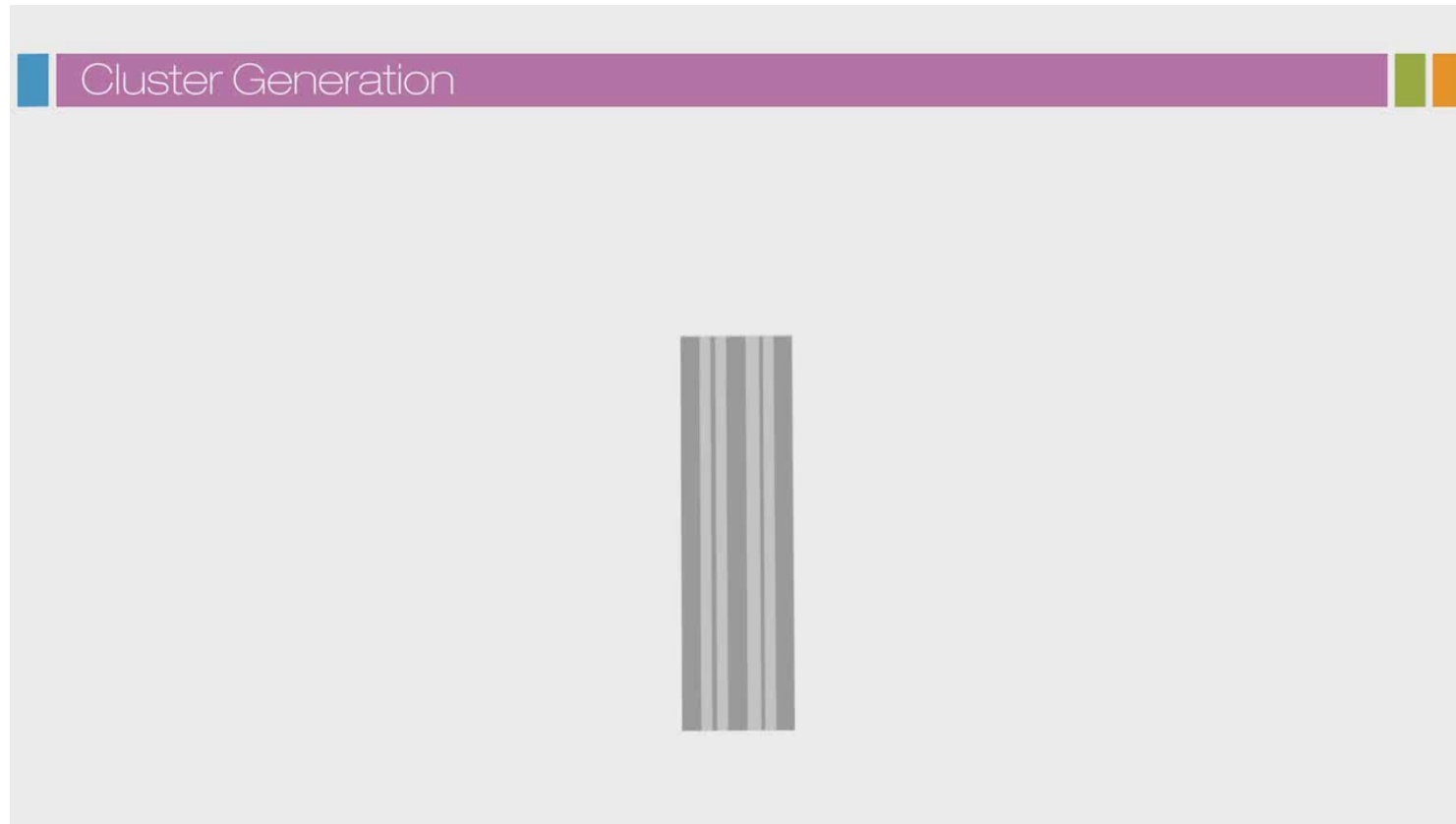
NGS: next-generation sequencing

1. DNS felszaporítása (PCR), darabokra („short read”) tördelése
2. Egyszálú short readek rögzítése a cellán primerekkel
3. Short readek felszaporítása a közvetlen környezetükben  
→ klaszterenként (1000+) teljesen azonos short readek
4. Cella leöntése a négyféle festékanyaggal megjelölt ddNTP-vel és DNS-polimeráz enzimmal
5. Klaszterenkénti színes foltok rögzítése képként
6. Beépült bázisok „semlegesítése”



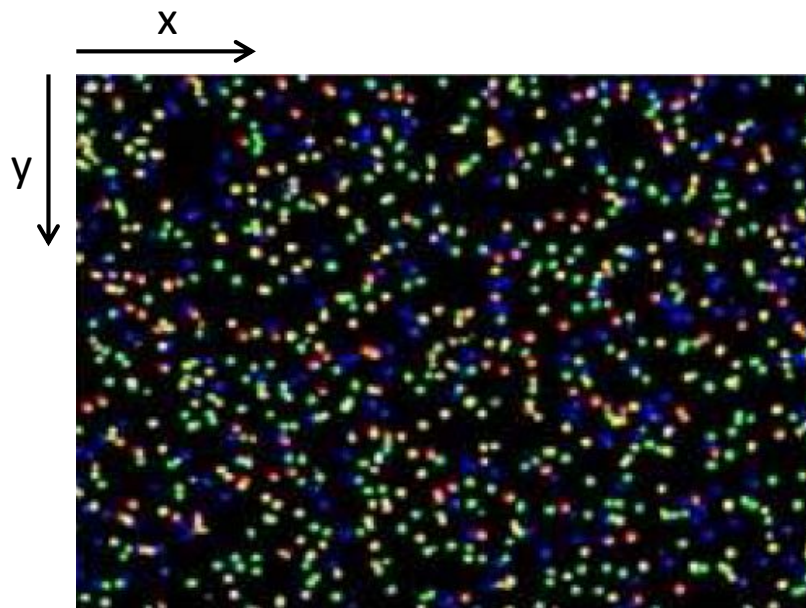
# SZEKVENÁLÁSI TECHNIKÁK

NGS: next-generation sequencing



# KIHÍVÁSOK AZ ADATFELDOLGOZÁSBAN

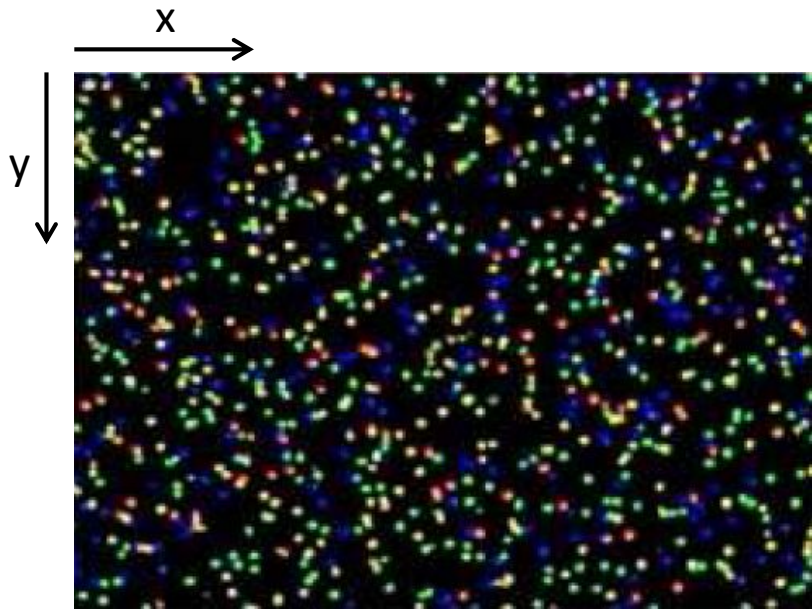
## 1. Képekből szöveges short read



koordináta		színintenzitás				
x	y	A	C	G	T	
...	...	...	...	...	...	
17	20	4	13	76	3	→ G
17	25	2	45	41	10	→ C? G?
...	...	...	...	...	...	
1001	1253	8	1	2	97	→ T
...	...	...	...	...	...	

# KIHÍVÁSOK AZ ADATFELDOLGOZÁSBAN

## 1. Képekből szöveges short read



koordináta		színintenzitás			
x	y	A	C	G	T
...	...	...	...	...	...
17	20	4	13	76	3
17	25	2	45	41	10
...	...	...	...	...	...
1001	1253	8	1	2	97
...	...	...	...	...	...

→ G  
 → C? G?  
 → T

```
@SEQ_ID
GATTTGGGGTTCAAAGCAGTATCGATCAAATAGTAAATCCATTTGTTCAACTCACAGTTT
+
!!!*(((((***+))%%%+)) (%%%) .1***-+*!') **55CCF>>>>>CCCCCCC65
```

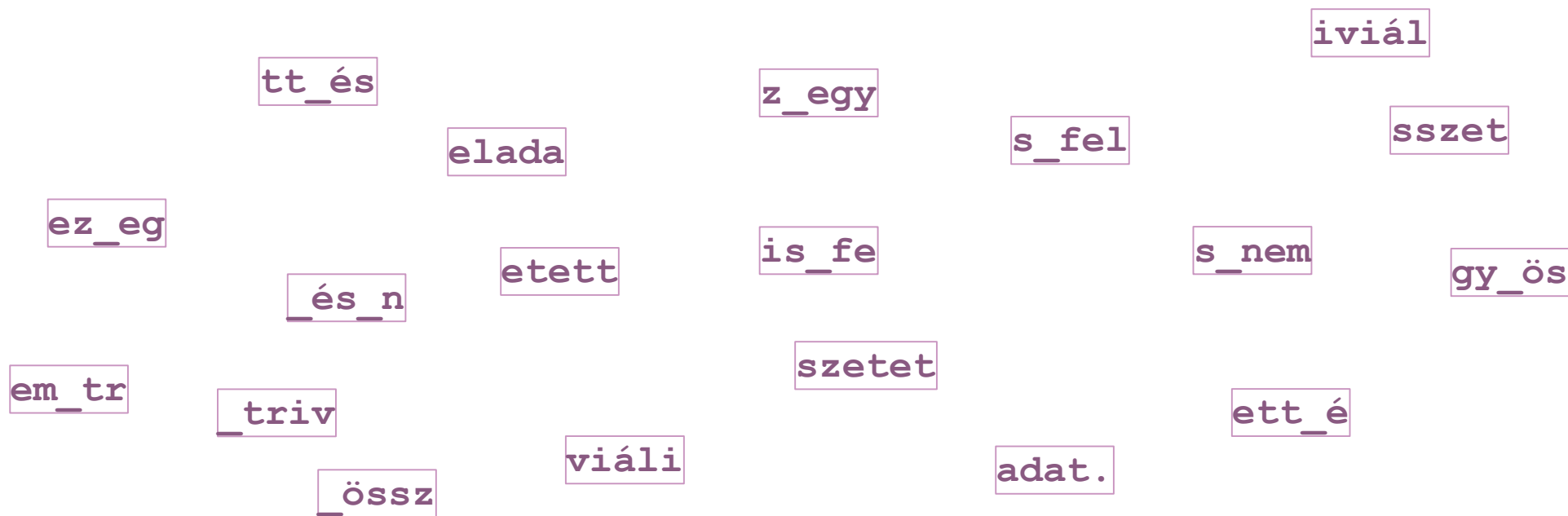
← ID  
 ← szekvencia  
 ← (ID)  
 ← „megbízhatóság” (quality)

} FASTQ formátum

# KIHÍVÁSOK AZ ADATFELDOLGOZÁSBAN

## 2. Short read-ből teljes genom: összeillesztés

A. referencia genom (minta) nélkül: **de-novo illesztés**



# KIHÍVÁSOK AZ ADATFELDOLGOZÁSBAN

## 2. Short read-ből teljes genom: összeillesztés

A. referencia genom (minta) nélkül: **de-novo illesztés**

z\_egy\_össz    tt\_és    em\_tr    viális\_feladat.  
ez\_eg    szetet    iviális\_fe  
gy\_ös    ett\_és\_nem\_triv    elada  
sszet    és\_n  
etett

# KIHÍVÁSOK AZ ADATFELDOLGOZÁSBAN

## 2. Short read-ből teljes genom: összeillesztés

A. referencia genom (minta) nélkül: **de-novo illesztés**

z\_egy\_össz    tt\_és    em\_tr    viális\_feladat.  
ez\_eg    szetet    iviális\_fe  
gy\_ös    ett\_és\_nem\_triv    elada  
sszet    és\_n  
etett



ez\_egy\_összetett\_és\_nem\_triviális\_feladat.

# KIHÍVÁSOK AZ ADATFELDOLGOZÁSBAN

## 2. Short read-ből teljes genom: összeillesztés

- A. referencia genom (minta) nélkül: **de-novo illesztés**
- B. referencia genom segítségével

ez\_így\_már\_sokkal\_könnyebb.



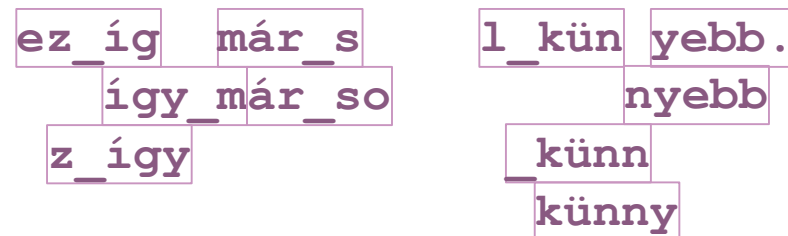


# KIHÍVÁSOK AZ ADATFELDOLGOZÁSBAN

## 2. Short read-ből teljes genom: összeillesztés

- A. referencia genom (minta) nélkül: **de-novo illesztés**
- B. referencia genom segítségével

ez\_így\_már\_sokkal\_könnyebb.



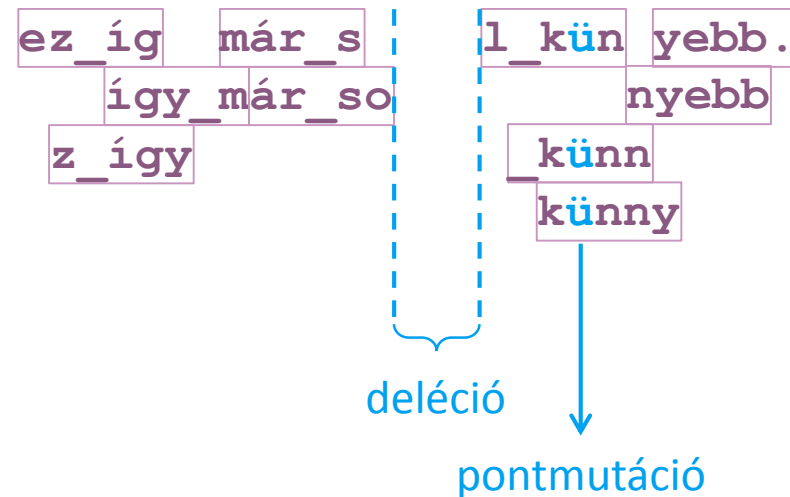
*De még ez is bonyolult!*

# KIHÍVÁSOK AZ ADATFELDOLGOZÁSBAN

## 2. Short read-ből teljes genom: összeillesztés

- A. referencia genom (minta) nélkül: **de-novo illesztés**
- B. referencia genom segítségével

ez\_így\_már\_sokkal\_könnyebb.



*De még ez is bonyolult!*

# KIHÍVÁSOK AZ ADATFELDOLGOZÁSBAN

## 3. Illesztett short read-ekből mutációk

Alapfeladat: hosszú szöveges fájl elemzése soronként (genomi pozícióként)

- eltérnek-e a readek a referencia genomtól?
- több vizsgált minta esetén (pl. kezelt/kezeletlen): eltérnek-e a readek egymástól a mintákban?  
→ ez sokkal fontosabb!

Bonyodalmak:

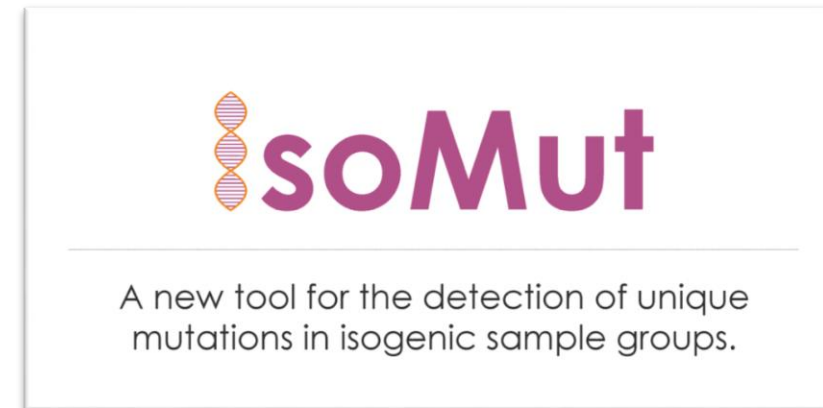
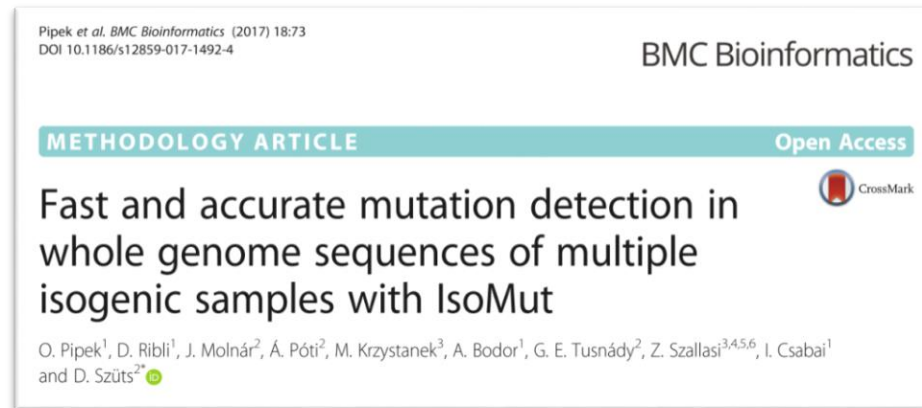
- van-e elég read az adott helyen (lefedettség)?
- mennyire megbízható a readen belül az adott bázis? („base quality”)
- mennyire megbízható a read felillesztése? („mapping quality”)
- „elégé” eltérnek-e a minták, a többi minta mennyire zajos?



rengeteg, különböző feladatra specializált mutáció detektáló algoritmus és eszköz

# KIHÍVÁSOK AZ ADATFELDOLGOZÁSBAN

## 3. Illesztett short read-ekből mutációk

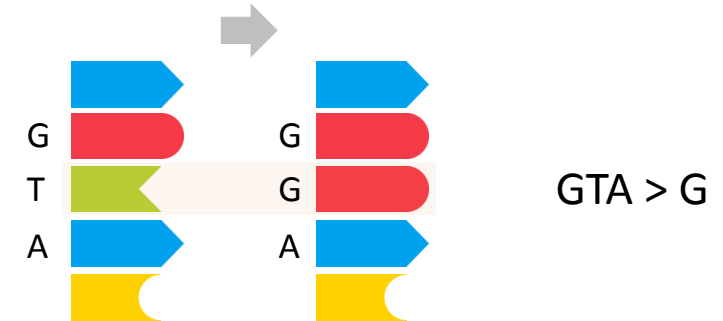


- sok izogenikus minta esetén
- referencia genomtól való eltérések, illesztési hibák korrigálódnak
- egyéni mutációk detektálására (például kezelés hatása)
- gyors
- pontos (megfelelő paraméterválasztással nagyon kevés fals pozitív eredmény)

# KIHÍVÁSOK AZ ADATFELDOLGOZÁSBAN

## 4. Mutációkból következtetések

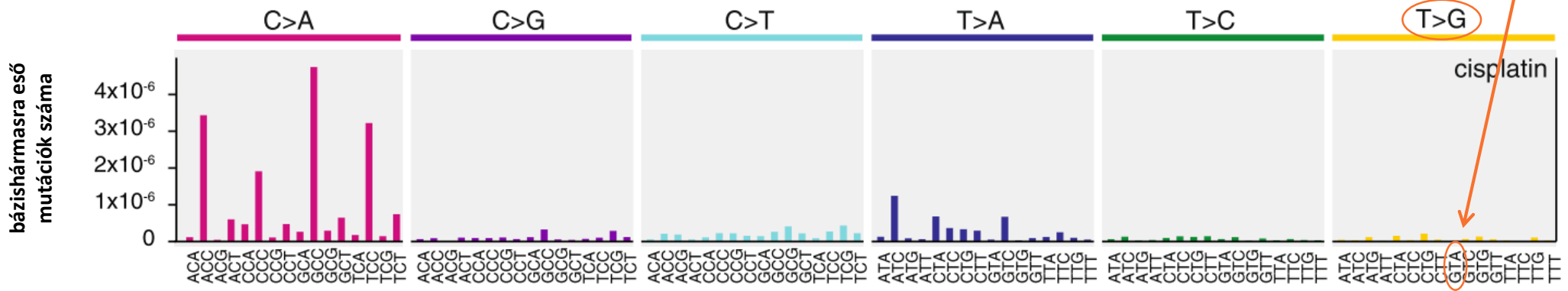
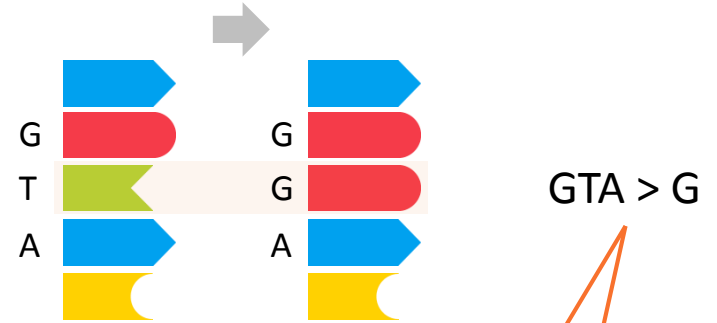
1. Mutációs spektrumok felvétele mintánként (páciensenként):



# KIHÍVÁSOK AZ ADATFELDOLGOZÁSBAN

## 4. Mutációkból következtetések

1. Mutációs spektrumok felvétele mintánként (páciensenként):



96-komponensű vektorok

# KIHÍVÁSOK AZ ADATFELDOLGOZÁSBAN

## 4. Mutációkból következtetések

1. Mutációs spektrumok felvétele mintánként (páciensenként):
2. Páciensek csoportosítása, mátrixba rendezése ráktípusonként (szervenként):

„mutációs katalógus”

$$M = \begin{pmatrix} m_1^1 & m_2^1 & m_3^1 & \dots & m_{G-1}^1 & m_G^1 \\ m_1^2 & m_2^2 & m_3^2 & \dots & m_{G-1}^2 & m_G^2 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ m_1^K & m_2^K & m_3^K & \dots & m_{G-1}^K & m_G^K \end{pmatrix} \quad K = 96$$

páciensek:                      ↓                      ↓                      ↓                      ↓                      ↓

   1.                      2.                      3.                      ...                      (G-1).                      G.

# KIHÍVÁSOK AZ ADATFELDOLGOZÁSBAN

## 4. Mutációkból következtetések

1. Mutációs spektrumok felvétele mintánként (páciensenként):
2. Páciensek csoportosítása, mátrixba rendezése ráktípusonként (szervenként):
3. Feltételezzük, hogy a konkrét páciensek spektrumai „kikeverhetők” néhány mutációs folyamatra jellemző spektrum („szignatúra”) kombinációjaként:

$$M = P \times E$$

mutációs katalógus

szignatúrák súlyfaktorai

mutációs szignatúrák

Feladat: egy ismert mátrix felbontása két másik (ismeretlen dimenziójú!) mátrix szorzatára



# KIHÍVÁSOK AZ ADATFELDOLGOZÁSBAN

## 4. Mutációkból következtetések

1. Mutációs spektrumok felvétele mintánként (páciensenként):
2. Páciensek csoportosítása, mátrixba rendezése ráktípusonként (szervenként):
3. Feltételezzük, hogy a konkrét páciensek spektrumai „kikeverhetők” néhány mutációs folyamatra jellemző spektrum („szignatúra”) kombinációjaként:
4. Kérdések:
  - Igaz-e, hogy az egyes ráktípusoknál megjelenő konkrét spektrumok visszavezethetők néhány (3-5) mutációs folyamat hatásának a kombinációjára?
  - Ha igen, találunk-e olyan folyamatokat, amik többféle ráktípusnál is aktívak?

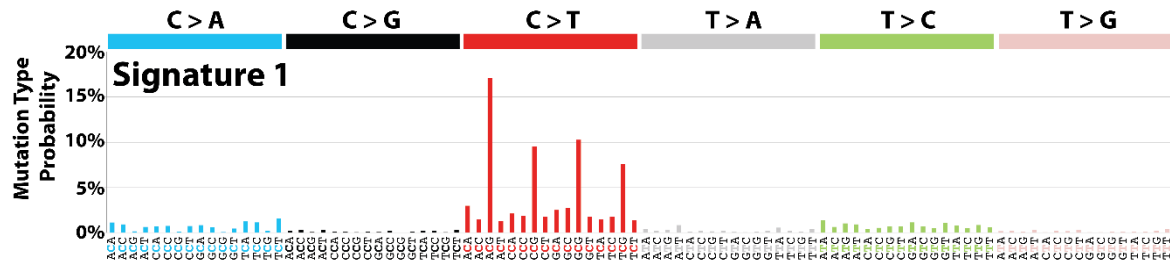
# KIHÍVÁSOK AZ ADATFELDOLGOZÁSBAN

## 4. Mutációkból következtetések

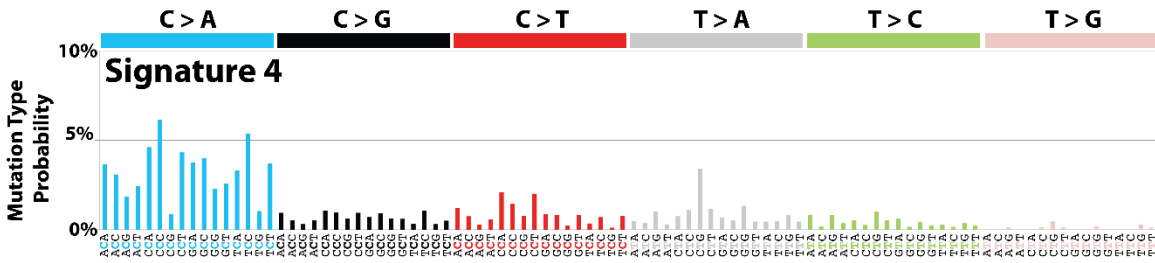
1. Mutációs spektrumok felvétele mintánként (páciensenként):
2. Páciensek csoportosítása, mátrixba rendezése ráktípusonként (szervenként):
3. Feltételezzük, hogy a konkrét páciensek spektrumai „kikeverhetők” néhány mutációs folyamatra jellemző spektrum („szignatúra”) kombinációjaként:
4. Kérdések:
  - Igaz-e, hogy az egyes ráktípusoknál megjelenő konkrét spektrumok visszavezethetők néhány (3-5) mutációs folyamat hatásának a kombinációjára? **IGEN.**
  - Ha igen, találunk-e olyan folyamatokat, amik többféle ráktípusnál is aktívak? **IGEN.**

# KIHÍVÁSOK AZ ADATFELDOLGOZÁSBAN

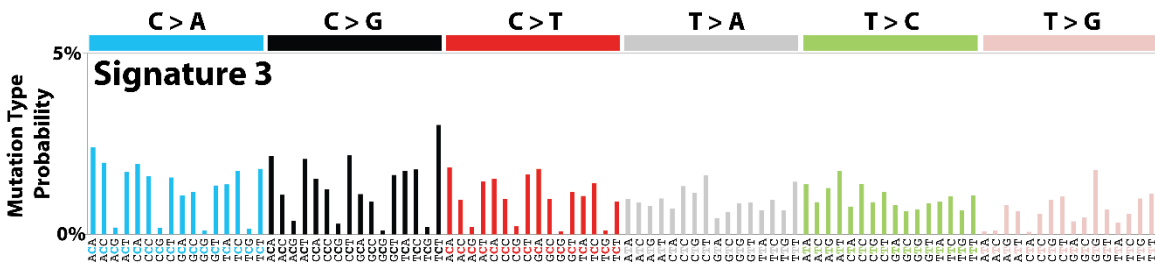
## 4. Mutációkból következtetések



öregedés



dohányzás



kettős DNS-szál törések  
javítómechanizmusának hiánya

# KIHÍVÁSOK AZ ADATFELDOLGOZÁSBAN

## 4. Mutációkból következtetések

### Mire jó ez az egész?

- Tegyük fel, hogy egy daganatban a mutációs spektrum felbontásából kiderül, hogy a kettős DNS-szál törések javítómechanizmusa hibásan működik.
- Ekkor a daganatot olyan kemoterápiás szerrel kezelve, mely gátolja az alternatív DNS-javító mechanizmusokat, a daganat megsemmisíthető. Az egészséges sejtek túlélnek, hiszen bennük az eredeti javítási útvonal jól működik.
- Viszont az olyan daganatokra, ahol ez a javítómechanizmus nem hibás, ugyanez a szer nem fog jól hatni.
  - A kezelést nem az határozza meg, hogy a daganat melyik szervben van, hanem a mutációs folyamatok összessége!

# KIHÍVÁSOK AZ ADATFELDOLGOZÁSBAN

## 4. Mutációkból következtetések



# KIHÍVÁSOK AZ ADATFELDOLGOZÁSBAN

## 4. Mutációkból következtetések

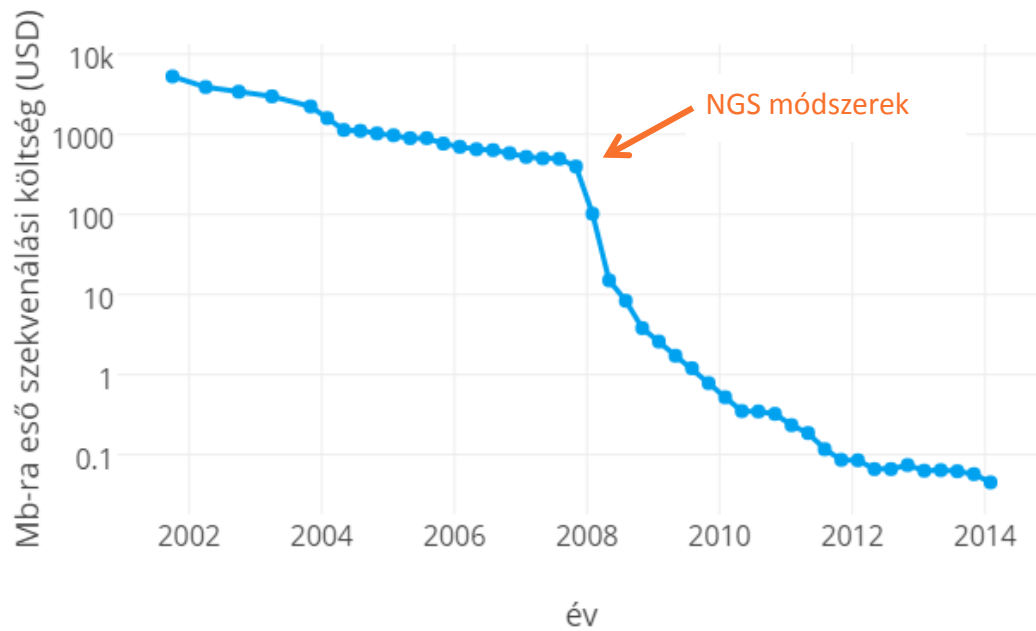
### Mit kell ehhez még tenni?

- Meg kell ismerni a különböző DNS-javító mechanizmusok hibáira utaló jeleket (szignatúrákat), hogy azonosíthatóak legyenek a mutációs spektrumokból.
- Meg kell ismerni a különböző kemoterápiás szerek hatásait és olyanokat kell fejleszteni, melyek kiaknázzák a daganatsejtek hiányosságait.

**Ez egy nagyon fontos és nemes feladat.**

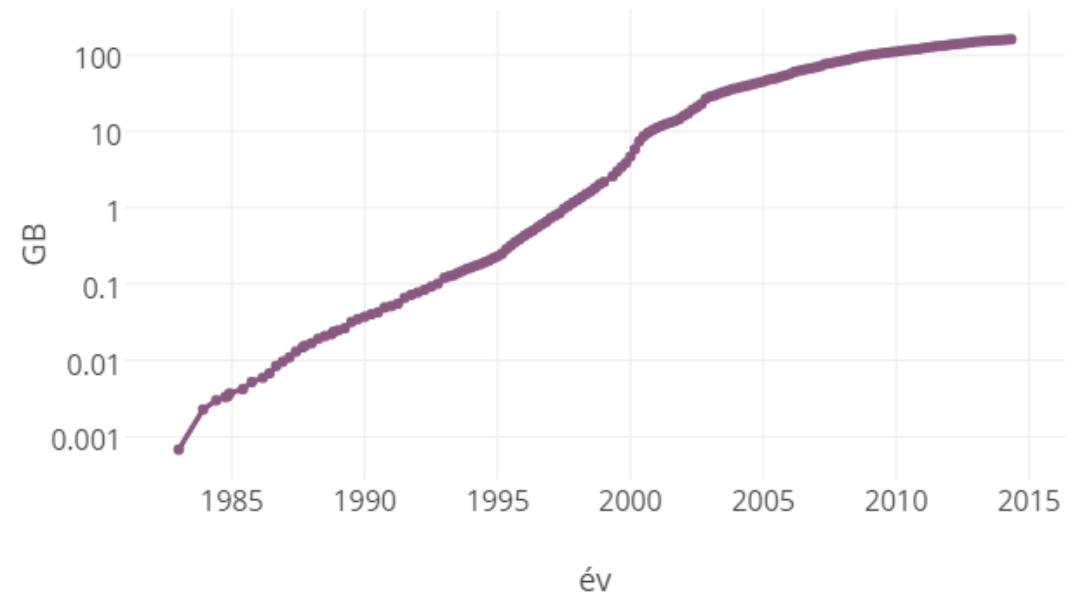
# REMÉNYTELI JÖVŐ

A szekvenálás költségének változása



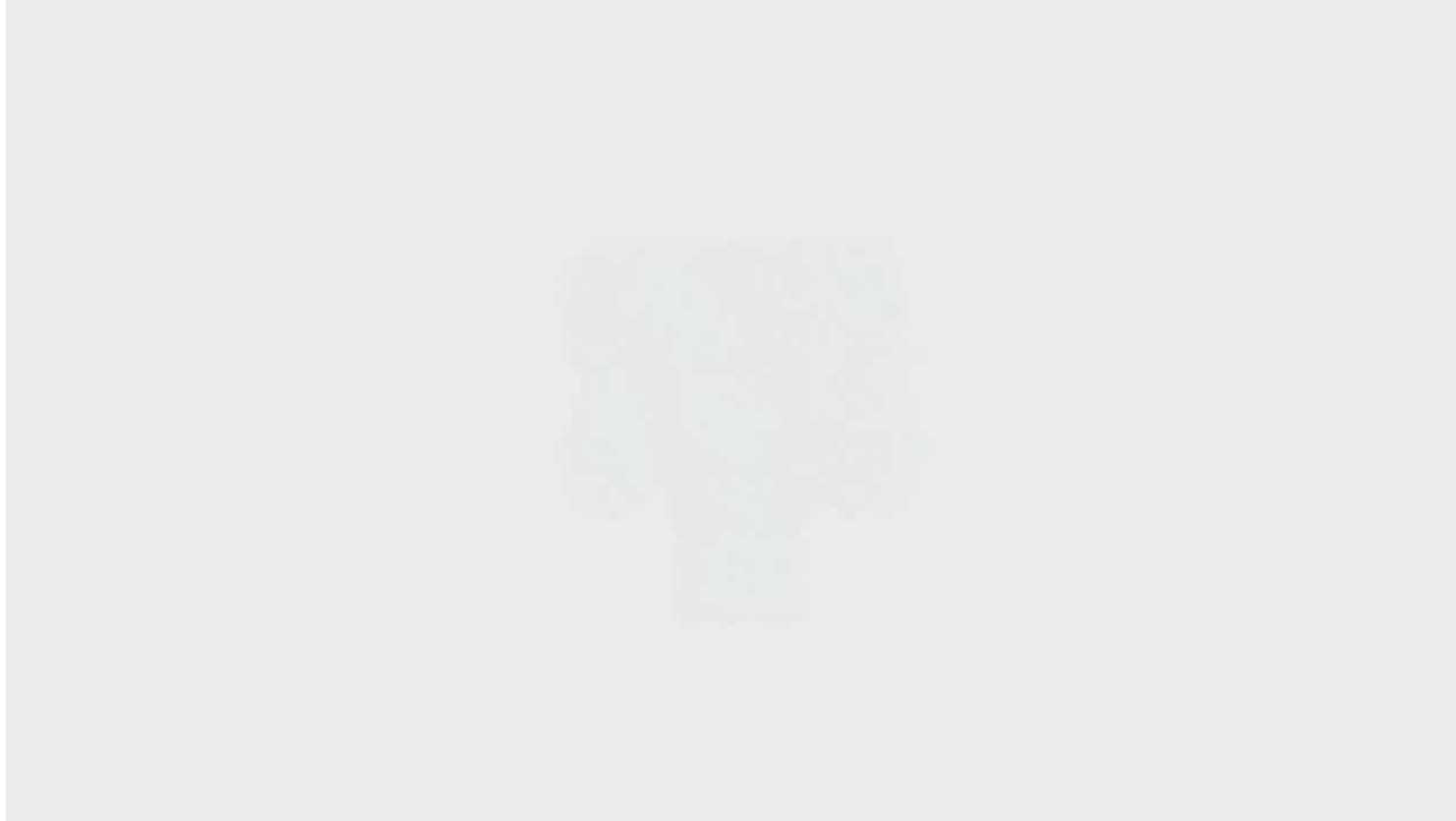
egyre olcsóbb adatot gyártani

A GenBankban elérhető adatmennyiség változása



egyre több adat publikus online

# REMÉNYTELI JÖVŐ





# REMÉNYTELI JÖVŐ



## REMÉNYTELI JÖVŐ

Nanopore MinION



- akkora mint egy pendrive
- USB csatlakozóval csatlakoztatható a számítógéphez
- „élőben” szekvenál
- a DNS megfelelő preparálása után semmilyen további laborfelszerelést nem igényel
- 1000 USD (< 300 ezer HUF)

## REMÉNYTELI JÖVŐ

- **Diagnosztika**, szűrések
  - egyes cégek 300 ezer HUF alatti összegért teljes genomot szekvenálnak, elemeznek
    - Rizikófaktorok korai felismerése
  - magzati DNS az anya véréből (180 ezer HUF)
    - főleg kromoszóma-többszörözéssel járó genetikai betegségek (pl. Down-kór, Turner-szindróma) kimutatása a terhesség korai szakaszában
- **Személyre szabott gyógyítás** („personalised medicine”)
  - a konkrét genetikai háttér nagyban befolyásolja a különböző terápiákra adott reakciót
    - ha rutinfeladattá válik a teljes genom szekvenálása a diagnosztikában, a pácienseket csak olyan jellegű kezeléseknek kell alávetni, amire várhatóan jól reagálnak
- **Kutatás**: még nagyon sok mindent nem tudunk!

## SZÜKSÉGES SZAKTUDÁS

- erős természettudományos képzettség: fizika, vegyész szak jó választás
  - egyéni problémamegoldásra tréningel
  - modellalkotás
  - statisztikai, matematikai ismeretek
- programozói készség (nem feltétlenül informatikai szaktudás!)
- angol nyelvtudás
- kreativitás (jó kérdéseket kell jól feltenni)
- hajlam a csapatmunkára (molekuláris biológiában nincs egyéni eredmény)
- érdeklődés a biológiai kérdések iránt
- adatvizualizációs készség

*Nem feltétlenül tudás!*

## ÖSSZEGZÉS

- A molekuláris biológia eszközei egyre gyorsabban fejlődnek, a kísérletek **egyre olcsóbbá** válnak.
- Az így keletkező **hatalmas adatmennyiséggel** a biológusok nem tudnak mit kezdeni.
- Rengeteg **ígéretes kutatási terület** van, amibe érdemes munkát befektetni.
- Mind a tudomány, mind az **orvoslás** szempontjában kiemelkedő jelentőségű eredmények születnek.
- Európában az adatok elemzéséhez értő bioinformatikusok kevesen vannak és viszonylag ritkán biológusok vagy informatikusok. (És soha nem orvosok.)
- Hasznos és fontos terület, melyben a **komplex rendszerek elemzéséhez értő tudósokra** van szükség.

**KÖSZÖNÖM A FIGYELMET!**

---

## ÁBRÁK, VIDEÓK

- Nap (5): <http://www.pngall.com/wp-content/uploads/2016/07/Sun-PNG-HD.png>
- Föld (5): <http://vignette1.wikia.nocookie.net/barsoom/images/1/1e/Earth.png/revision/latest?cb=20110912212924>
- DNS nagy-skálás szerkezete (6): <http://elte.prompt.hu/sites/default/files/tananyagok/ABiokemiaEsMolekularisBiologiaAlapjai/images/image476.png>
- Transzláció, transzkripció (7): [http://www.tankonyvtar.hu/hu/tartalom/tamop425/0005\\_20\\_informacioelmelet\\_scorm\\_06/20\\_k28.jpg](http://www.tankonyvtar.hu/hu/tartalom/tamop425/0005_20_informacioelmelet_scorm_06/20_k28.jpg)
- Tübingeni egyetem laboratórium (8): <http://ars.els-cdn.com/content/image/1-s2.0-S0012160604008231-gr2.jpg>
- Astbury diffrakciós felvétele (8): [http://www.codex99.com/photography/images/photo51/astbury\\_1947\\_sm.jpg](http://www.codex99.com/photography/images/photo51/astbury_1947_sm.jpg)
- Griffith-kísérlet (8): <http://elte.prompt.hu/sites/default/files/tananyagok/ABiokemiaEsMolekularisBiologiaAlapjai/images/image450.jpg>
- Watson és Crick (8): [https://cdn2.hubspot.net/hub/237126/file-660502209-jpg/watson\\_crick.jpg](https://cdn2.hubspot.net/hub/237126/file-660502209-jpg/watson_crick.jpg)
- Franklin diffrakciós felvétele (8):  
[https://askabiologist.asu.edu/sites/default/files/resources/articles/crystal\\_clear/Rosalind\\_Franklin\\_Plate\\_1\\_DNA\\_B\\_form\\_1000.jpg](https://askabiologist.asu.edu/sites/default/files/resources/articles/crystal_clear/Rosalind_Franklin_Plate_1_DNA_B_form_1000.jpg)
- Gélelektroforézis (8): [http://images.slideplayer.com/15/4823416/slides/slide\\_5.jpg](http://images.slideplayer.com/15/4823416/slides/slide_5.jpg)
- Kettős száltörés javítása, PARP inhibitor hatása (videók, 12, 30): <https://vimeo.com/128692417>
- BRCA1 komplex (13): [https://en.wikipedia.org/wiki/BRCA1#/media/File:Protein\\_BRCA1\\_PDB\\_1jm7.png](https://en.wikipedia.org/wiki/BRCA1#/media/File:Protein_BRCA1_PDB_1jm7.png)
- MTA TTK épülete (13): <http://www.maeponline.hu/files/wp/2013-11-mta-uj-epulet-01.jpg>
- Illumina szekvenálás, világító klaszterek (18): [http://www.openwetware.org/wiki/Image:BMC\\_IlluminaFlowcell.png](http://www.openwetware.org/wiki/Image:BMC_IlluminaFlowcell.png)
- Illumina szekvenálás (videó, 19): <https://www.youtube.com/watch?v=womKfikWlxM>
- Nanopore szekvenálás (videók, 33-34): <https://vimeo.com/127689865> <https://vimeo.com/127689053>
- Nanopore MinION (35): <https://nanoporetech.com/sites/default/files/s3/minion-cutout.png>