

Prokarióta és vírus genomok. Eukarióta organellek



Varga Máté
(m.varga@ucl.ac.uk)

Genom méretek

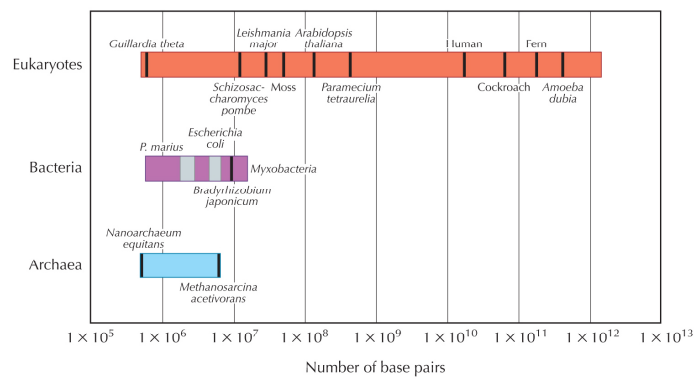
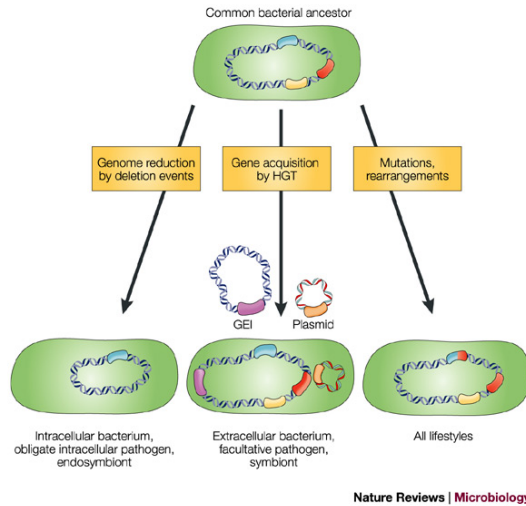


FIGURE 7.1. Genome sizes in the three domains of life. A selection of genome sizes and size ranges from specific groups of organisms is indicated.

7.1, adapted from Bentley S.D. et al., *Annu. Rev. Genet.* **38**: 771–791, © 2004 Annual Reviews, www.annualreviews.org, based on data from DOGS <http://www.cbs.dtu.dk/databases/DOGS/>

Evolution © 2007 Cold Spring Harbor Laboratory Press

A genom struktúrája a baktérium "életvitelét" tükrözi

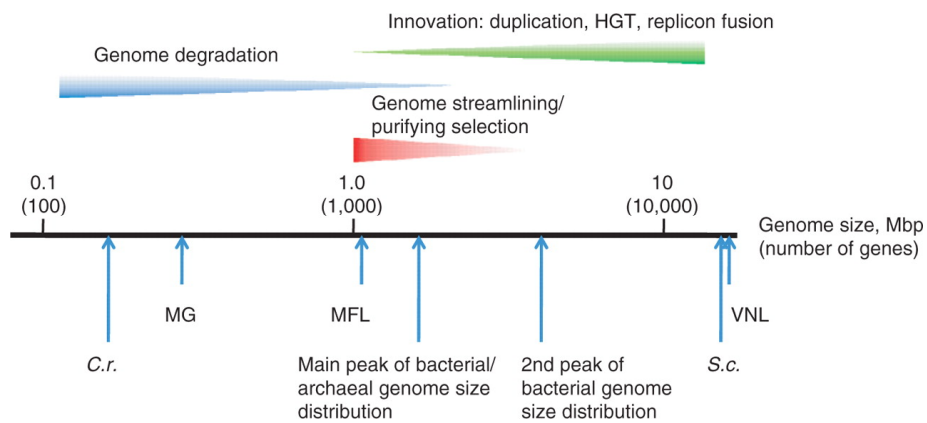


-a **stabil környezet** általában genom redukcióhoz vezet (az adott nicheben nem fontos gének elvesznek)

- a **változó életkörülmények** közt található baktériumok genomja nagyobb, mert több eshetőségre kell felkészülniük (különböző körülmények közt más-más gének biztosíthatják a túlélést.)

(Dobrindt et al. (2004) *Nat Rev Microb*)

A bakteriális genom méretét befolyásoló hatások



(Koonin and Wolf (2008) *Nuc Ac Res*)

A stabil környezet genom-redukcióhoz vezet



Table 1 General features of the genomes of *B. pertussis*, *B. parapertussis* and *B. bronchiseptica*

	<i>B. pertussis</i>	<i>B. parapertussis</i>	<i>B. bronchiseptica</i>
Size (bp)	4,086,186	4,773,551	5,338,400
G+C content (%)	67.72	68.10	68.07
Coding sequences	3,816	4,404	5,007
Pseudogenes	358 (9.4%)	220 (5.0%)	18 (0.4%)
Coding density (intact genes)	82.9%	86.6%	91.4%
Coding density (all genes)	91.6%	92.2%	92.0%
Average gene size (bp)	978	987	978
rRNA operons	3	3	3
tRNA	51	53	55
IS481	238	0	0
IS1001	0	22	0
IS1002	6	90	0
IS1663	17	0	0

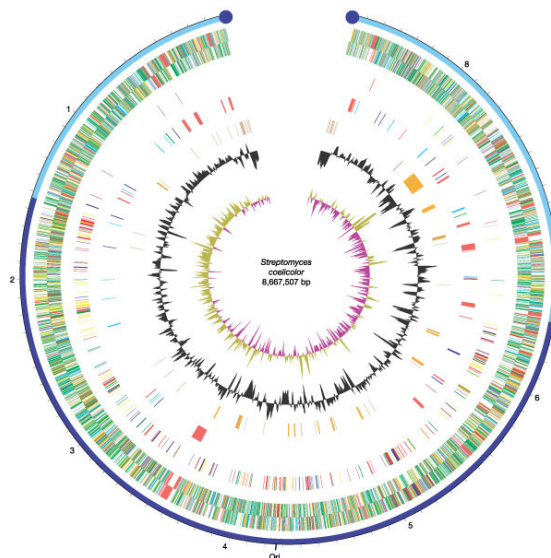
- az obligát humán patogén *B. pertussis* a kormoszómája 20%-át elvesztette (a *B. bronchiospetica* számos fajt képes megfertőzni)
- a megmaradó géneinek 10%-a pszeudogén
- az inaktiváció egyik hajtóeleme egy IS elem elterjedése

(Parkhill et al. (2003) *Nat Gen*)

A változó környezet nagy genomot igényel

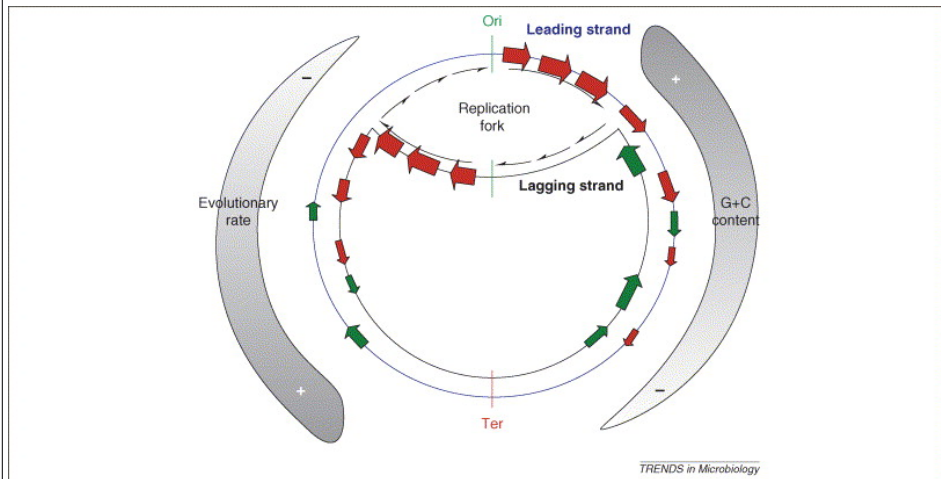


- a talajban lakó *Streptomyces coelicolor* genomja kb 8.7 Mb és ~7800 fehérjét kódol
- az esszenciális gének a lineáris kromoszóma közepén helyezkednek el
- a nem esszenciális, "készenléti" gének a karokon találhatóak



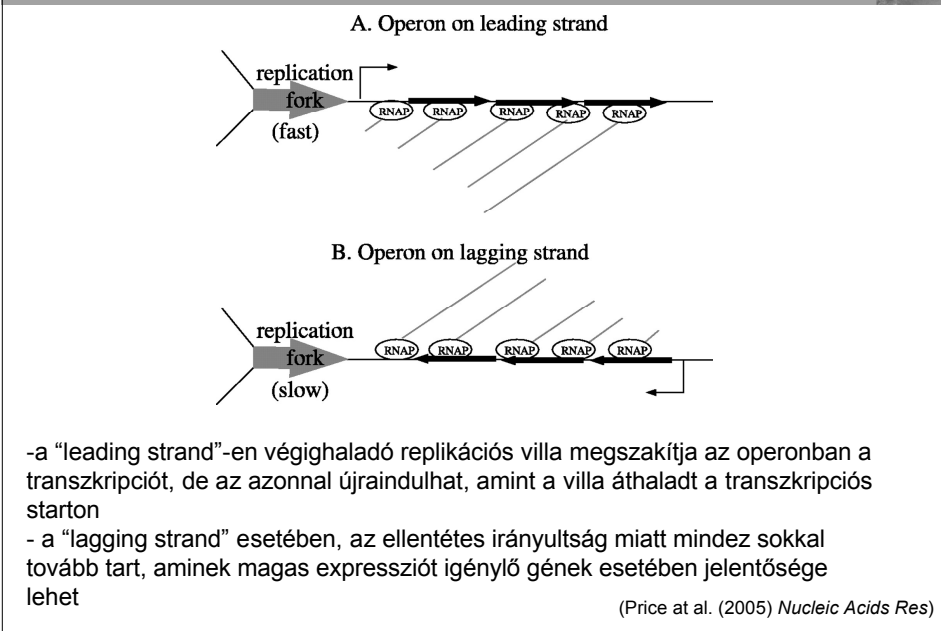
(Bentley et al. (2002) *Nature*)

A bakteriális genom evolúcióját befolyásoló belső tényezők



(Abby and Daubin (2007) *TIM*)

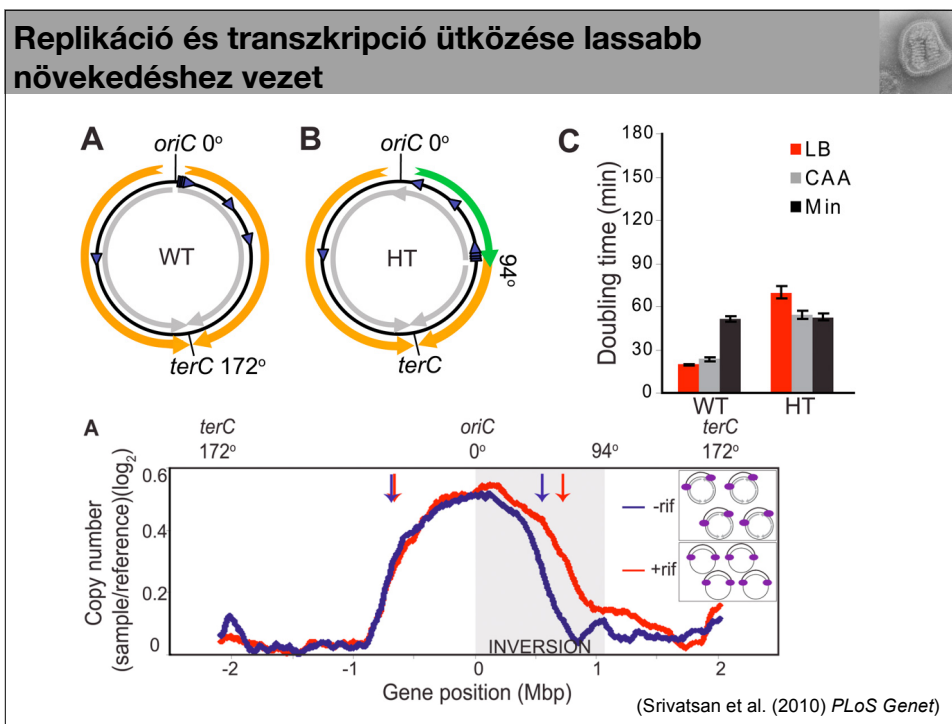
Replikáció vs. transzkripció: a szelekciós nyomás eredete



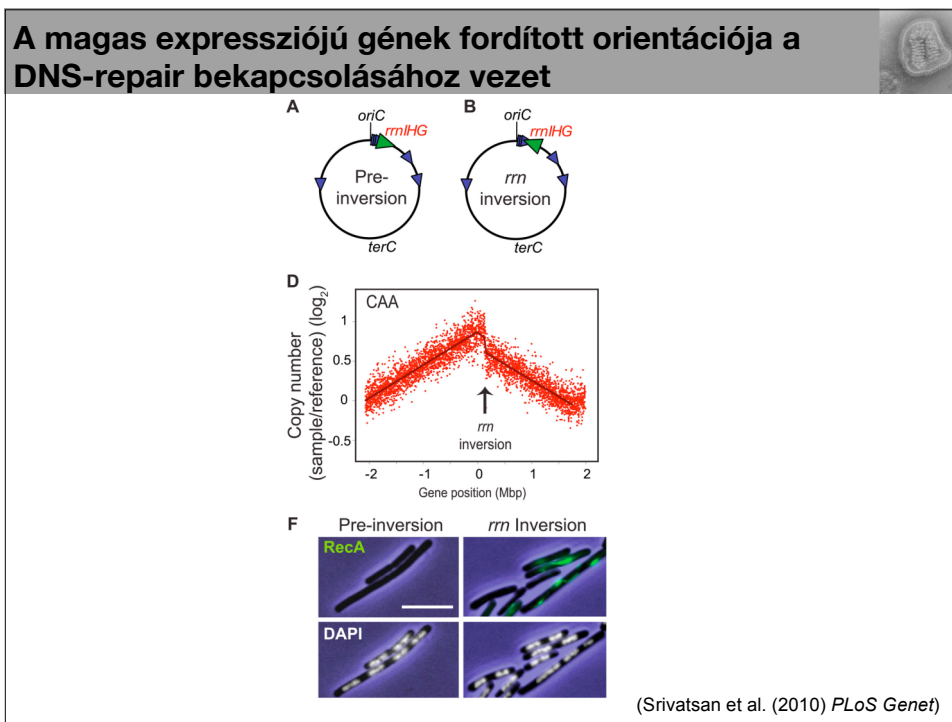
- a "leading strand"-en végighaladó replikációs villa megszakítja az operonban a transzkripciót, de az azonnal újraindulhat, amint a villa áthaladt a transzkripció starton
- a "lagging strand" esetében, az ellentétes irányultság miatt mindez sokkal tovább tart, aminek magas expressziót igénylő gének esetében jelentősége lehet

(Price et al. (2005) *Nucleic Acids Res*)

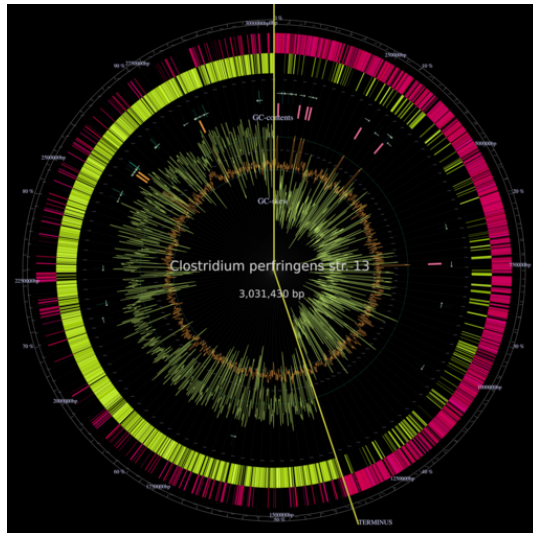
Replikáció és transzkripció ütközése lassabb növekedéshez vezet



A magas expressziójú gének fordított orientációja a DNS-repair bekapcsolásához vezet



“GC-skew” a *Clostridium perfringens* példáján át



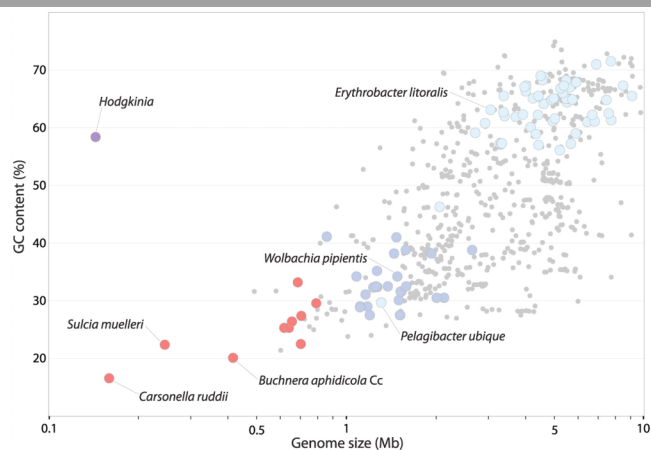
-“GC skew” a relatív C többletet írja le: $(C-G)/(C+G)$.

- mivel a bakteriális genomokban a replikáció jobban szereti a leading strandben a G-t, ezért a replikáció kezdete és vége könnyen megállapítható a GC skew érték alapján

- a *Clostridium* genom ennek az extrém példája

(Shimizu et al. (2002) *PNAS*)

GC tartalom vs. genom méret

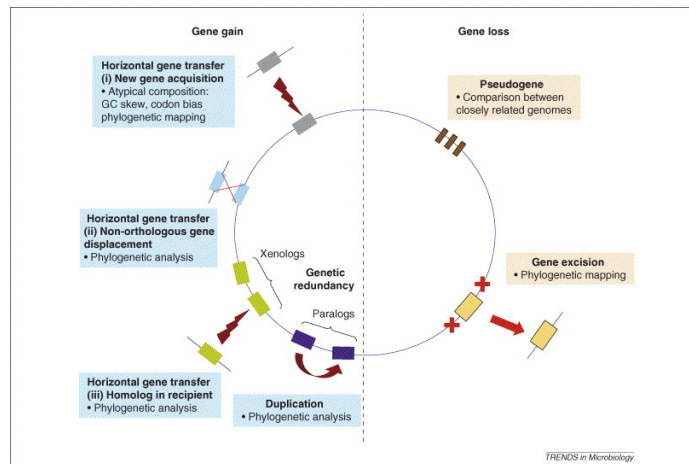


(McCutcheon et al. (2009) *PLoS Genet*)

Két lehetséges magyarázat:

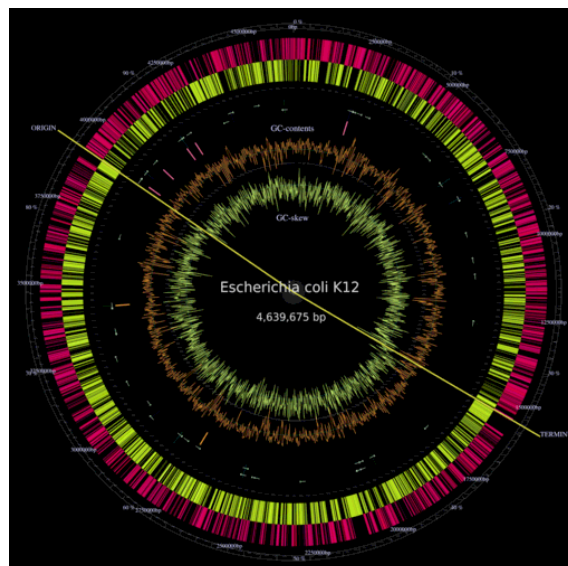
- **energetikai okok:** a GTP és CTP szintézise több energiát fogyaszt, így a kis genomú (parazita) fajok ebben is optimalizálnak
- **mutációs okok:** a kis genomú prokarióták gyakran elvesztették a hibajavító enzimeiket, a leggyakoribb mutáció típus pedig C -> T

A bakteriális genom dinamikája



(Abby and Daubin (2007) *TIM*)

Escherichia coli K12

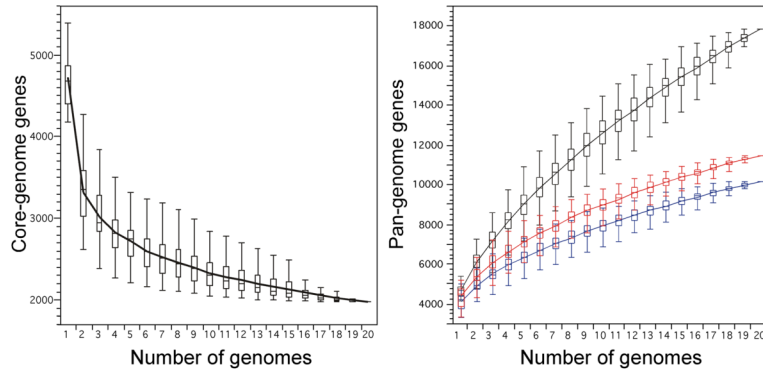


- 4.6 Mb

- 4288 fehérje kódoló gén (kb. harmadának ismeretlen a funkciója)

(Blattner et al. (1997) *Science*)

Escherichia coli - pan-genome vs core-genome

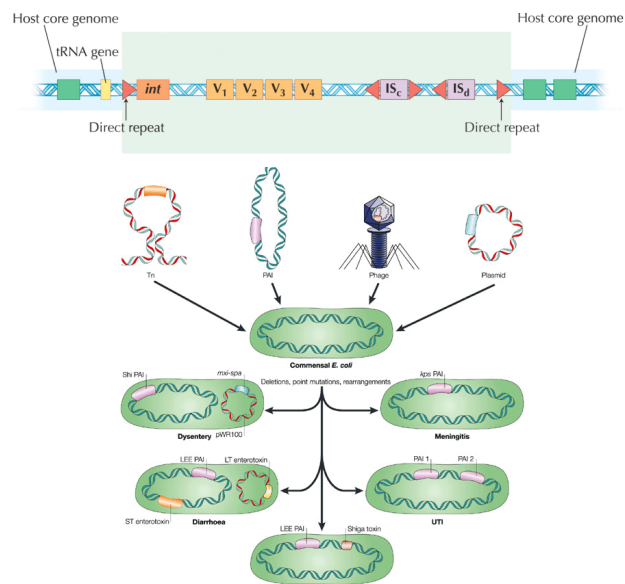


- több *E. coli* genom megszekvenálása után azonban nyilvánvaló, hogy az esszenciális gének száma csak töredéke a K12-ben találtaknak

- mai becslések szerint az *E. coli* alap (core) genom <1900 génből áll, az ún. pán-genomban található gének száma (az összes gén ami elfordul valamilyen *E. coli*-ban) pedig 17,000 felett van

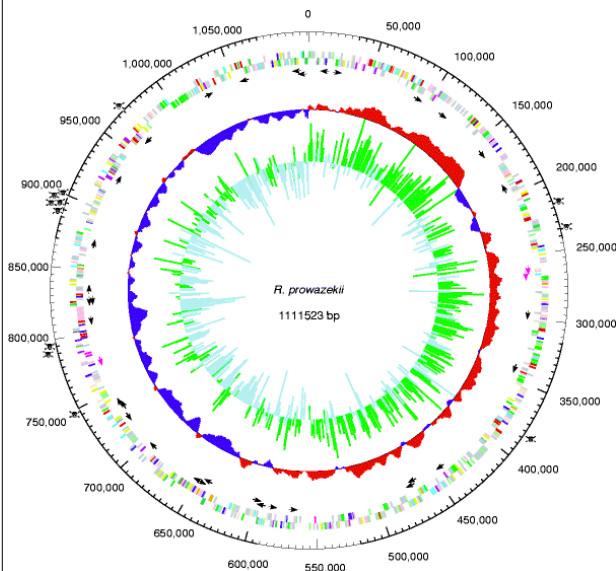
(Touchon et al. (2009) *PLoS Genet*)

Patogenitás szigetek



Nature Reviews | Microbiology

Rickettsia prowazekii



(Andersson et al. (1998) *Nature*)

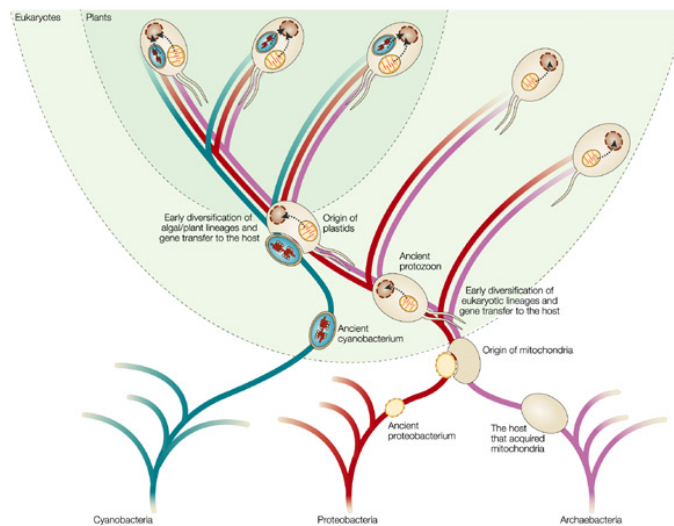
- obligát intracelluláris parazita, a tífusz okozója

- a genom szokatlanul nagy része (24%) tartalmaz nem kódoló szekvenciákat

- ezek minden jel szerint pszeudogének, amelyek fokozatosan egyre több mutációt gyűjtenek össze

- a *Rickettsia* az α -proteobaktériumokhoz tartozik, akárcsak a mitokondriumok őse, így genom degradációja informatív a mitokondriális genom evolúciója szempontjából

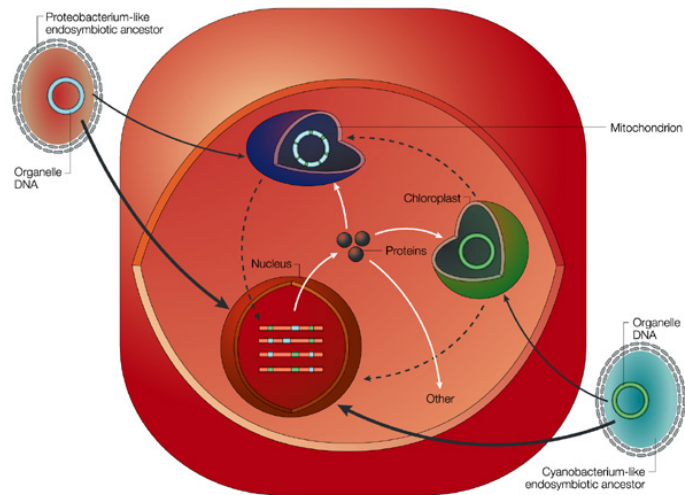
Endoszimbiotikus gén-transzfer



Nature Reviews | Genetics

(Timmis et al. (2004) *Nat Rev Gen*)

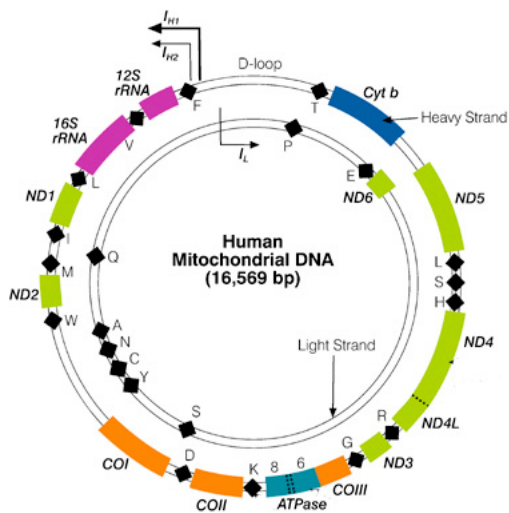
Endoszimbiotikus gén-transzfer



Nature Reviews | Genetics

(Timmis et al. (2004) *Nat Rev Gen*)

A mitokondriális genom



- dupla szálú, cirkuláris DNS molekula, ami mindkét szálán kódol

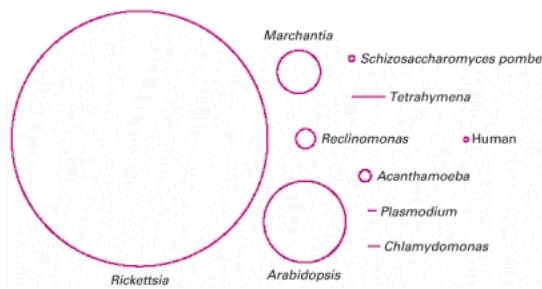
- emberek esetében 16,569 bp hosszú

- 37 gént kódol

- 13 az oxidatív foszforilációban játszik szerepet, a többi tRNS és rRNS

- minden mitokondriális mátrixban több kópia található

A mitokondriális genom mérete különböző csoportokban



- a mitokondriális genomok mérete 6000 bp (*Plasmodium falciparum*) és 300,000 bp (egyres növények) közt változik

- többségük cirkuláris, de akad lineáris is

- az állatok (Eumetazoa) mtDNS mérete eléggé stabil, kb. 16,500 bp

- a referenciaként használt *Rickettsia* genom kb 1.1 millió bp hosszú

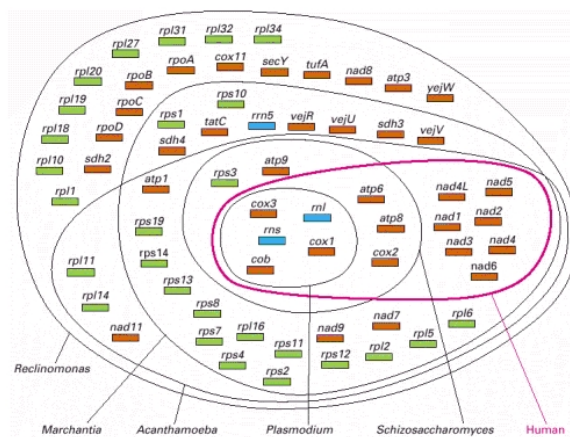
(Alberts et al.: Molecular Biology of the Cell)

A mitokondriális genom tartalma különböző csoportokban

- a kevésbé komplex mitokondriális genomok, általában az összetettebb genomokban levő gének részhalmozait kódolják

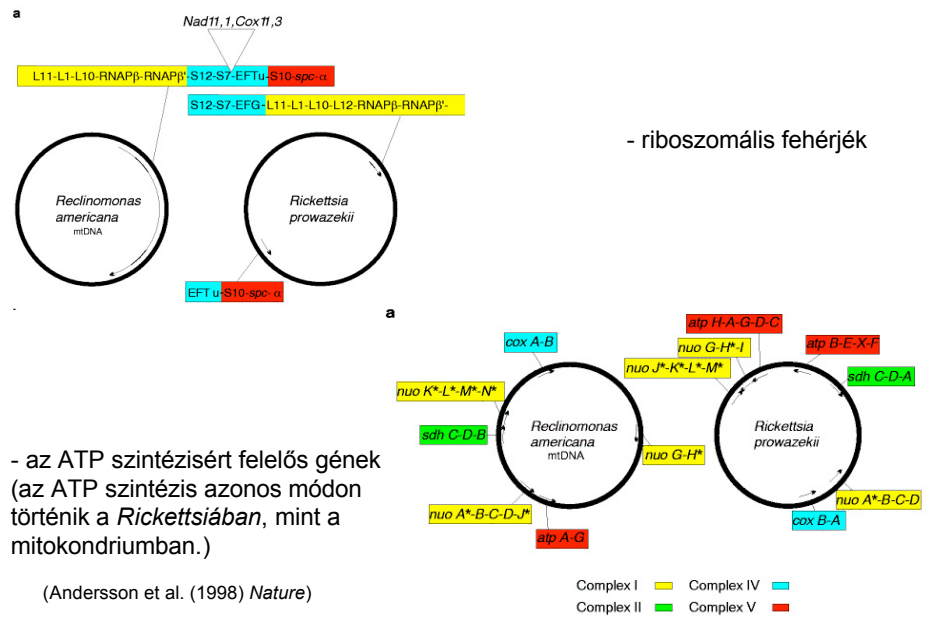
- öt olyan gén van, ami *minden* mtDNS-ben fellelhető: *cob*, *cox1*, *cox3*, *rns*, *rnl*

- hova lett a többi gén?

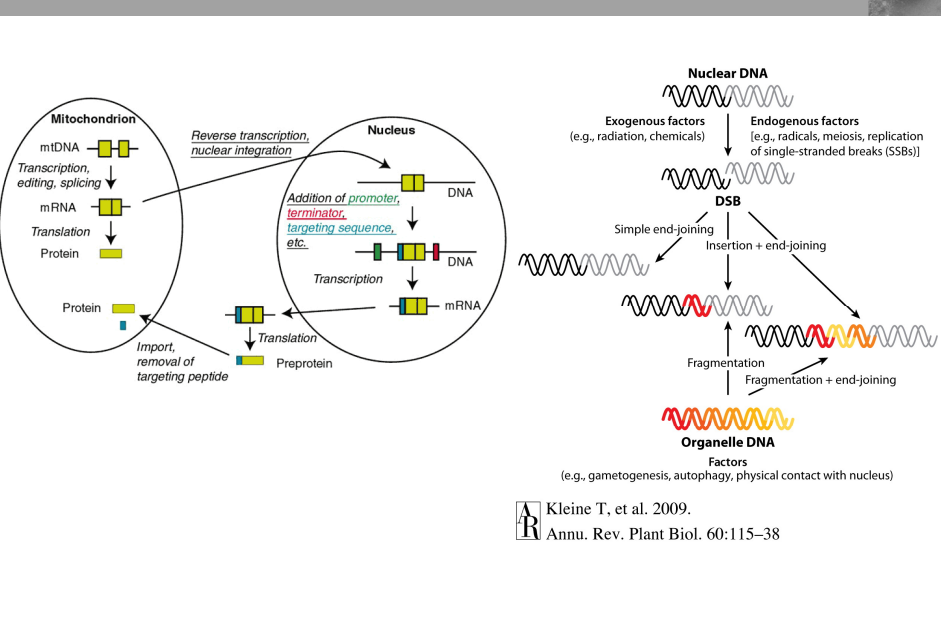


(Alberts et al.: Molecular Biology of the Cell)

A mitokondriális genom és a *Rickettsia* genom szerkezete



A gén-transzfer feltételezett mechanizmusai



mtDNS, mint intron alapanyag



Bigeloviella natans
alpha subunit guanine nucleotide binding protein gene 1

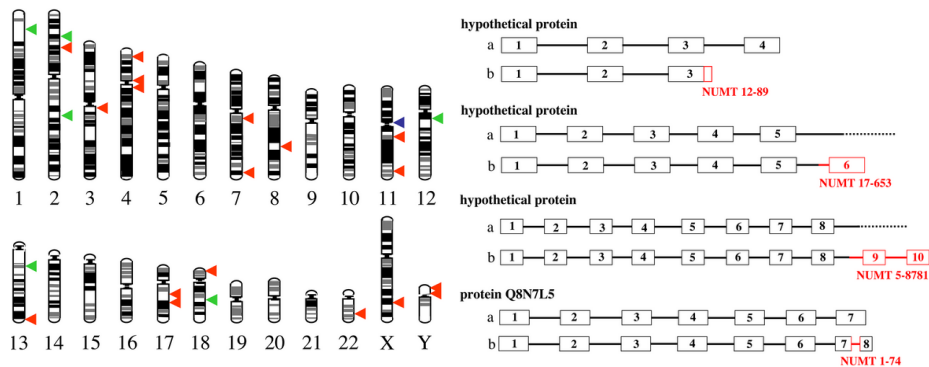


Current Biology

- egy egysejtű algafaj egyik fehérje kódoló génjének első intronja 86%-ban megegyezik a *cox1* mitokondriális gén egy szakaszával
- a splice akceptor és donor nukelotidok csak később alakultak ki, feltehetően kezdetben a szekvenca gyenge splice-site-ként működött

(Curtis and Archibald (2010) *Curr Biol*)

mtDNS, mint intron alapanyag



- csak az emberi genomban 27 darab specifikus NUMT található - ezek az utóbbi 4-6 millió évben keletkeztek és fixálódtak
- többségük feltételezett intronokba integrálódott

(Ricchetti et al. (2004) *PLoS Biol*)

A mtDNS translációs kódja bizonyos csoportokban különbözik az "univerzális kódtól"

MITOCHONDRIAL CODES					
CODON	"UNIVERSAL" CODE	MAMMALS	INVERTEBRATES	YEASTS	PLANTS
UGA	STOP	<i>Trp</i>	<i>Trp</i>	<i>Trp</i>	STOP
AUA	Ile	<i>Met</i>	<i>Met</i>	<i>Met</i>	Ile
CUA	Leu	Leu	Leu	<i>Thr</i>	Leu
AGA } AGG }	Arg	<i>STOP</i>	<i>Ser</i>	Arg	Arg

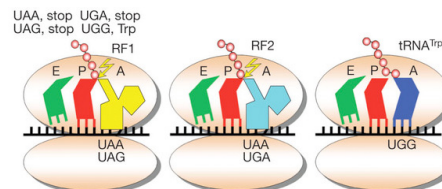
*Italics and color shading indicate that the code differs from the "Universal" code.

- növényekben és a legnagyobb mtDNS genomot hordozó protozoa *Reclinomonas* fajokban az mtDNS kódja is "univerzális"
- a STOP -> Trp változás hasonlít az egyes baktériumokban leírtakhoz
- valószínűleg az mtDNS-ben kódolt kisszámú gén jobban tolerálja egy-egy ritka kodon megváltozását

(Alberts et al.: Molecular Biology of the Cell)

Alternatív genetikai kód kialakulása egy parazita baktériumban

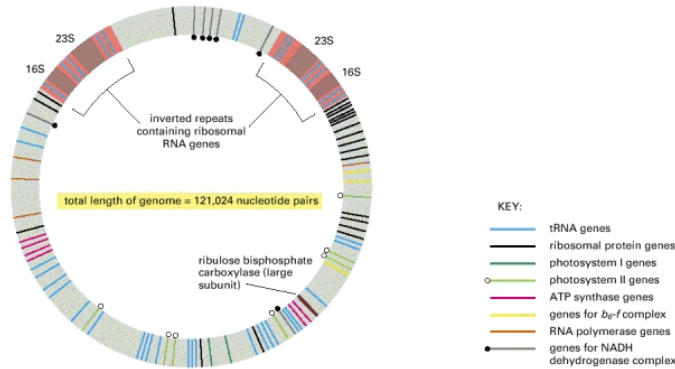
	DnaE (335)	RpoB (711)	RpoC (131)
<i>Hodgkinia</i>	SDFTLAKAHN	VAFMCNGFN	PVVHA.FHGSA
<i>Mloti</i>	ADFIKWAKAQG	VAFMPWNGYNY	PVAHIWFLKSL
<i>Cres</i>	SDFIKWKAHAG	VAFMPWNGYNY	PVAHIWFLKSL
<i>Pdeni</i>	ADFIKWAKAHN	VAFMPWNGYNY	PVAHIWFLKSL
<i>Rrubr</i>	ADFIKWAKDAD	VAFMPWNGYNY	PVAHIWFLKSL
<i>Elito</i>	ADFIKWAKDHG	VAFMPWNGYNY	PVAHIWFLKSL
<i>Pubiq</i>	SDYIKWAKNND	VAFMPWNGYNY	PVAHIWFLKSL
<i>Rrick</i>	SDFIKWAKKEG	VAFMPWNGYNY	PVAHIWFLKSL
<i>Ecoli</i>	MEFIKWAKDNG	VAFMPWNGYNY	PVAHIWFLKSL
<i>Nmeni</i>	QDFINWAKTHG	VAFMPWNGYNY	PVAHIWFLKSL
<i>Gmeta</i>	ADFIKWAKDHG	VAFMPWNGYNY	PVAHIWFLKSL



	tRNA-Trp anticodon	release factors	UGA encodes
initial state	CCA	RF1 RF2	STOP
1 mutation of tRNA-Trp gene			
some readthrough of UGA	*CCA	RF1 RF2	STOP Trp
2 loss of Release Factor 2 (RF2)			
only UAA and UAG read as stop	*CCA	RF1	Trp
3 mutation of tRNA-Trp anticodon			
UGA, UGG both read by wobble rules	UCA	RF1	Trp
4 genomic codon adaptation			
new UAA and UAG stops generated; some UGG codons changed to UGA	UCA	RF1	Trp

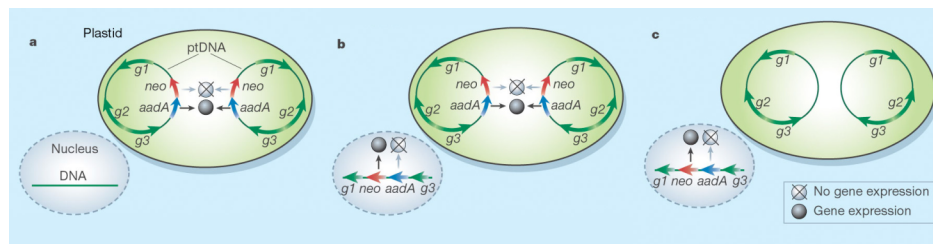
(McCutcheon et al. (2009) *PLoS Genet*)

A kloroplasztisz genom



- dupla szálú, cirkuláris DNS molekula, ami mindkét szálán kódol
- a transzkripció szabályozó elemek szinte azonosak bakteriális homológ szekvenciákkal

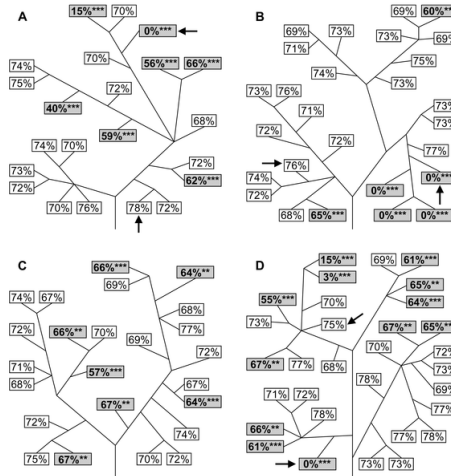
Plasztisz-sejtmag DNS transzfer



- genetikai screen plasztisz-sejtmag DNS transzfer gyakoriságának vizsgálatára:
 - a spectinomycin rezisztenciát hordozó gén (*aadA*) egy bakteriális promóter mögött van (így az a plasztiszban is aktív), míg a neomycin rezisztencia gén egy eukarióta promóter mögött van, így az csak a sejtmagban aktív
 - szomatikus sejtekben (levélben) 1 : 5 millió volt az átugrás gyakorisága, míg pollen sejtekben 1 : 16,000 (a különbség egyik oka az lehet, hogy pollenképződés közben lebomlik a plasztisz, és a felszabaduló DNS nagyobb eséllyel juthat a sejtmagba)

(Maliga (2003) *Nature*)

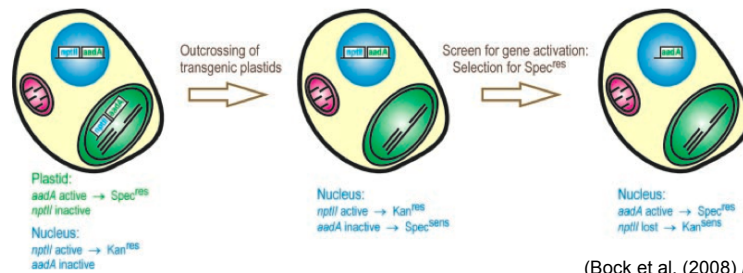
A NUPT-ok többsége instabil



- önmegporzó rendszerben vizsgálva azonban jól látható, hogy a gyakori integrációt gyakori deléció ellensúlyozza
- gyakran egyetlen generáció alatt elvesződik a rezisztencia gén - a dolog pontos mechanizmusa még nem ismert

(Sheppard and Timmins (2009) *PLoS Gen*)

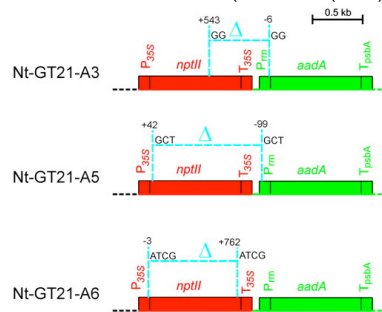
Plasztisz eredetű gének sejtmagi reaktivációja



(Bock et al. (2008) *Bioessays*)

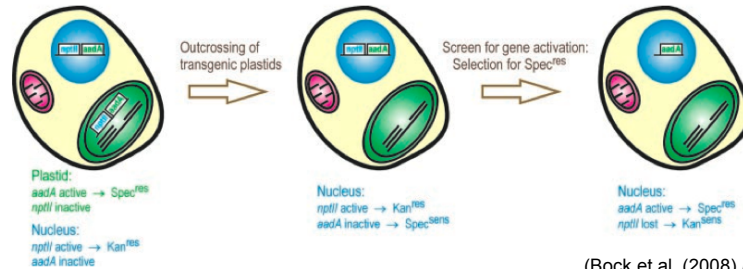
- a plasztisz szekvenicák aktiválódásának gyakorisága kb. azonos nagyságrendű a nukleáris gén-transzfer gyakoriságával

- a sejtmagba kerülő plasztisz-szekvenicák aktivációja upstream promóterek "befogásával" történik



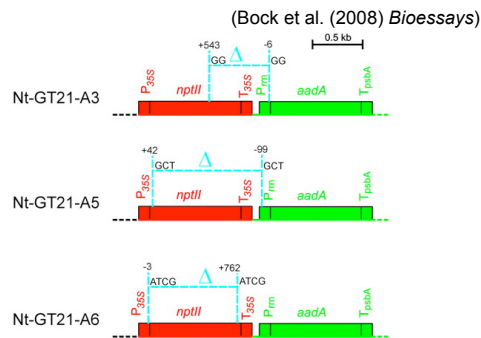
(Stegemann and Bock (2006) *Plant Cell*)

Plasztisz eredetű gének sejtmagi reaktivációja



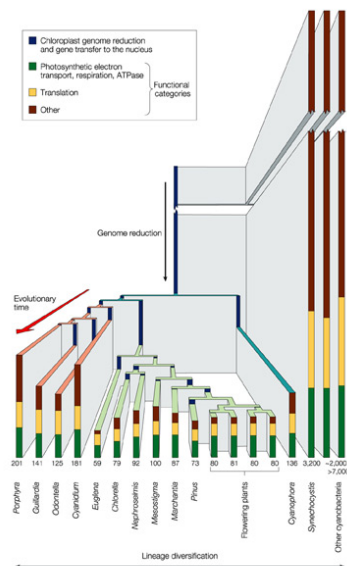
- a plasztisz szekvenciák aktiválódásának gyakorisága kb. azonos nagyságrendű a nukleáris gén-transzfer gyakoriságával

- a sejtmagba kerülő plasztisz-szekvenciák aktivációja upstream promóterek "befogásával" történik



(Stegemann and Bock (2006) *Plant Cell*)

A plasztisz-genom evolúciója



Nature Reviews | Genetics

(Timmis et al. (2004) *Nat Rev Gen*)

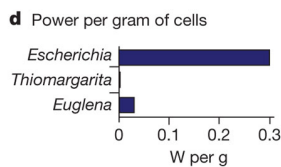
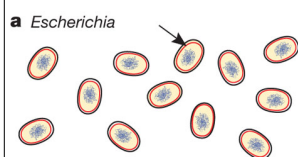
Miért nem transzlokálódott a mitokondriumok és kloroplasztiszok teljes genetikai anyaga a genomba?



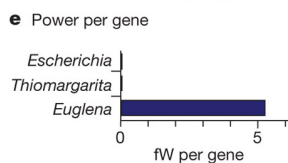
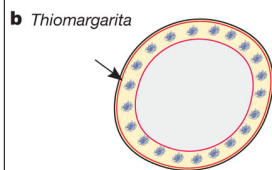
1. Az egyre kisebb genomok alternatív kodonokat hoznak létre, és az új kodontábla által kódolt fehérjék a sejtmagban nem funkcionálisak.
2. Az organelláris DNS-en kódolt gének erősen hidrofóbak, és nem lennének könnyen szállíthatóak a citoplazmán keresztül
3. CORR (COlocation of genes and gene-products for Redox Regulation of gene expression) hipotézis: bizonyos gének átíródását a bionergetikai membránok redox potenciálja szabályozza. Ezek nem kerülhetnek ki az organellumból, mert a szabályozásuk nem megoldható a sejtmagban, így az organelláris genomból való eltűnésük a mitokondrium/plasztisz működésképtelenségét okozná.

(Allen (2003) *Phil Trans R Soc Lond B*)

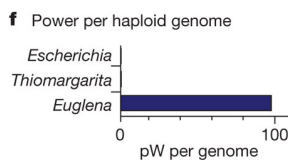
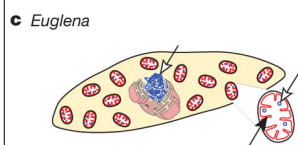
Az eukarióták komplexebb genomját a mitokondriumok teszik lehetővé



-a prokarióta sejtek energiatermelése limitált, így az egy fehérje termelésére fordítható energiamennyiség függ a fehérjék számától



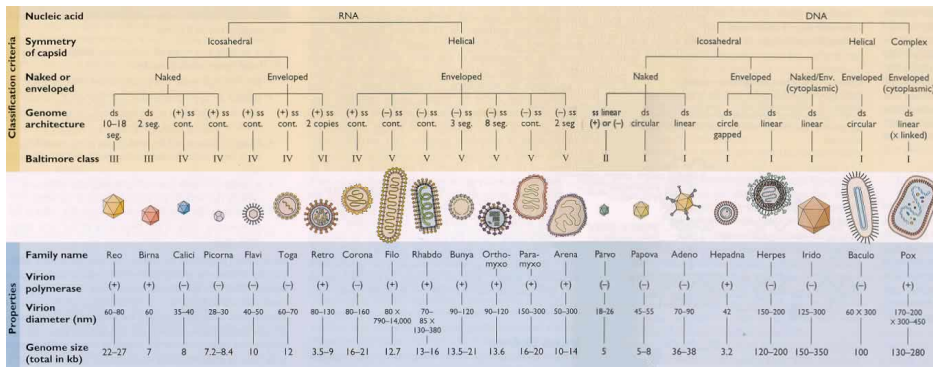
- ez felső határt szab a lehetséges kódoló szekvenciák számának



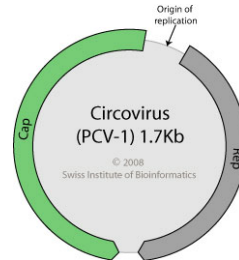
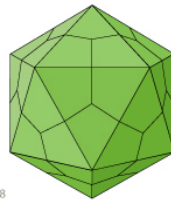
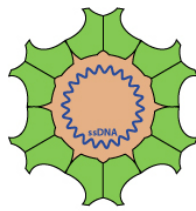
- eukariótákban a mitokondriumok révén lényegesen megnő a bioenergetikai membránok felülete és a termelt ATP mennyisége az igényekhez igazodik, ez pedig közel 400,000x nagyobb kódolókapacitáshoz vezet

(Land and Martin (2010) *Nature*)

Virális genom típusok



A legkisebb vírusok - Circoviridae



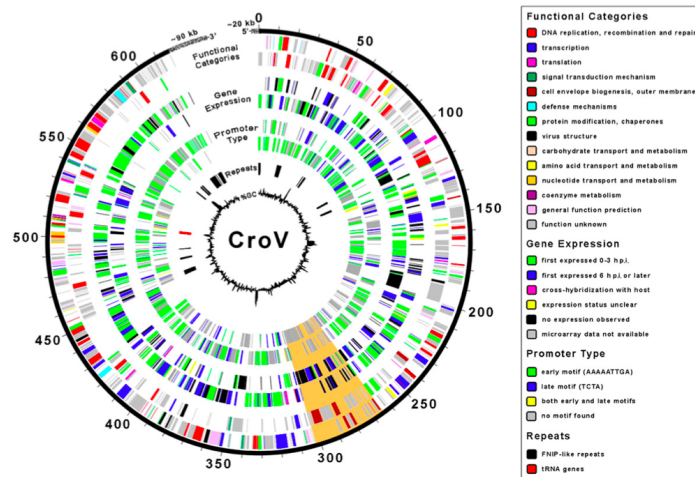
© ViralZone 2008
Swiss Institute of Bioinformatics

-<2 kb, cirkuláris ssDNS genom, mindössze két fehérjét kódol

- virális életciklus:

1. Virus penetrates into the host cell.
2. Uncoating, and release of the viral genomic ssDNA into the nucleus.
3. The ssDNA is converted into dsDNA with the participation of cellular factors.
4. viral mRNAs are transcribed and translated to produce viral proteins.
5. Replication may be mediated by a "Rep-like" protein, and would occur by rolling circle
6. These newly synthesized ssDNA can either
 - a) be converted to dsDNA and serve as a template for transcription/replication
 - b) be encapsidated by capsid protein and form virions released from the cell by budding

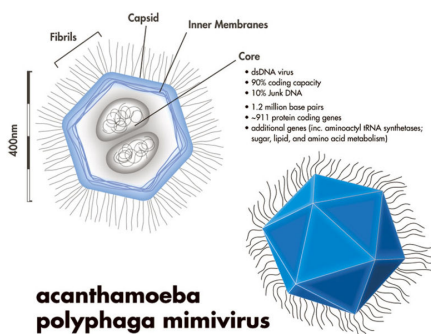
Óriás vírusok - *Cafeteria roenbergensis*



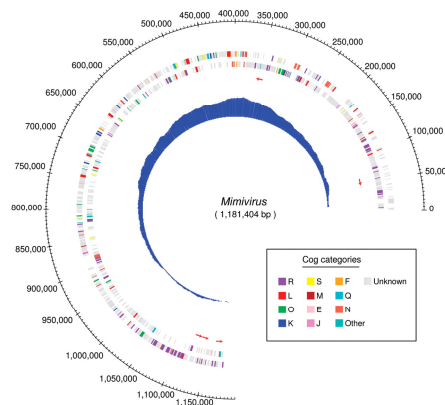
~730 kb dsDNS genom, ~550 gén köztük transzlációban szerepet játszó és DNS repair fehérjék is
 - a genom kb 5%-a repetitív DNS és a genom egy jókora része bakteriális eredetű!

(Fischer et al. (2010) *PNAS*)

Óriás vírusok - *Acanthamoeba polyphaga mimivirus*



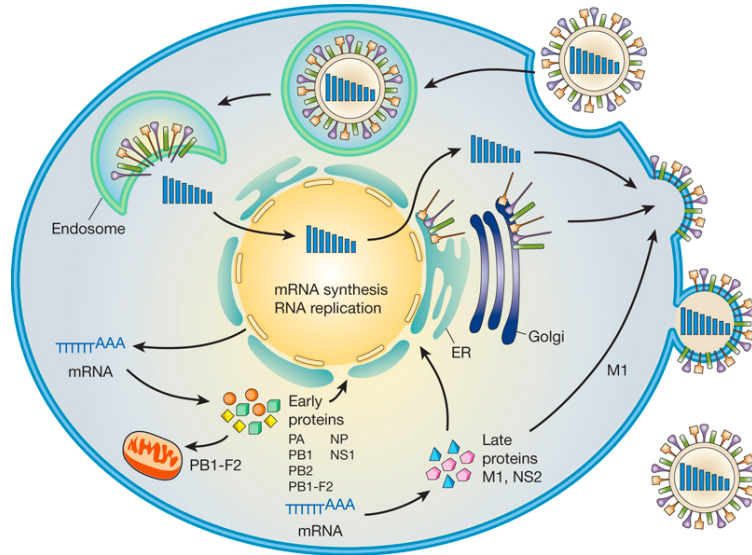
acanthamoeba polyphaga mimivirus



~1.2 Mb dsDNS genom, 981 gén, szintén sok tRNS-szintáz, transzkripció iniciációért, elongációért ill terminációért felelős fehérje, valamint DNS repair enzimek
 - a vírusok és a valódi sejtes élőlények közti határ genomikai szempontból nem létező: ezek a vírusok nagyobbak mint egyes patogén prokarióták vírusai, és komplex replikációs apparátust kódolnak

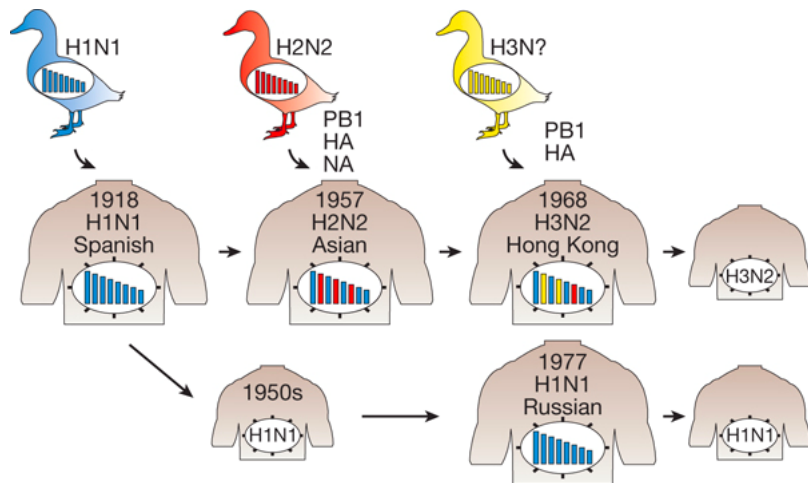
(Raoult et al. (2004) *Science*)

Influenza vírusok életrciklusa



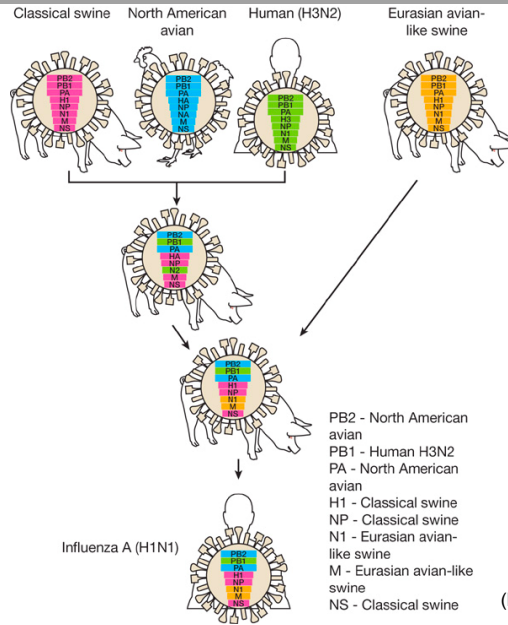
(Neumann et al. (2009) *Nature*)

Influenza vírusok evolúciója reasszortáció által



(Neumann et al. (2009) *Nature*)

Influenza - a 2009 A(H1N1) eredete



PB2 - North American avian
 PB1 - Human H3N2
 PA - North American avian
 H1 - Classical swine
 NP - Classical swine
 N1 - Eurasian avian-like swine
 M - Eurasian avian-like swine
 NS - Classical swine

(Neumann et al. (2009) *Nature*)