

Eduard Kejnovský + Roman Hobza

EVOLUČNÍ GENOMIKA

II. SVĚT RNA A POČÁTKY GENOMŮ



OSNOVA

1. Svět molekul RNA, počátky
2. Ribozymy
3. Relikty světa RNA
4. První genomy



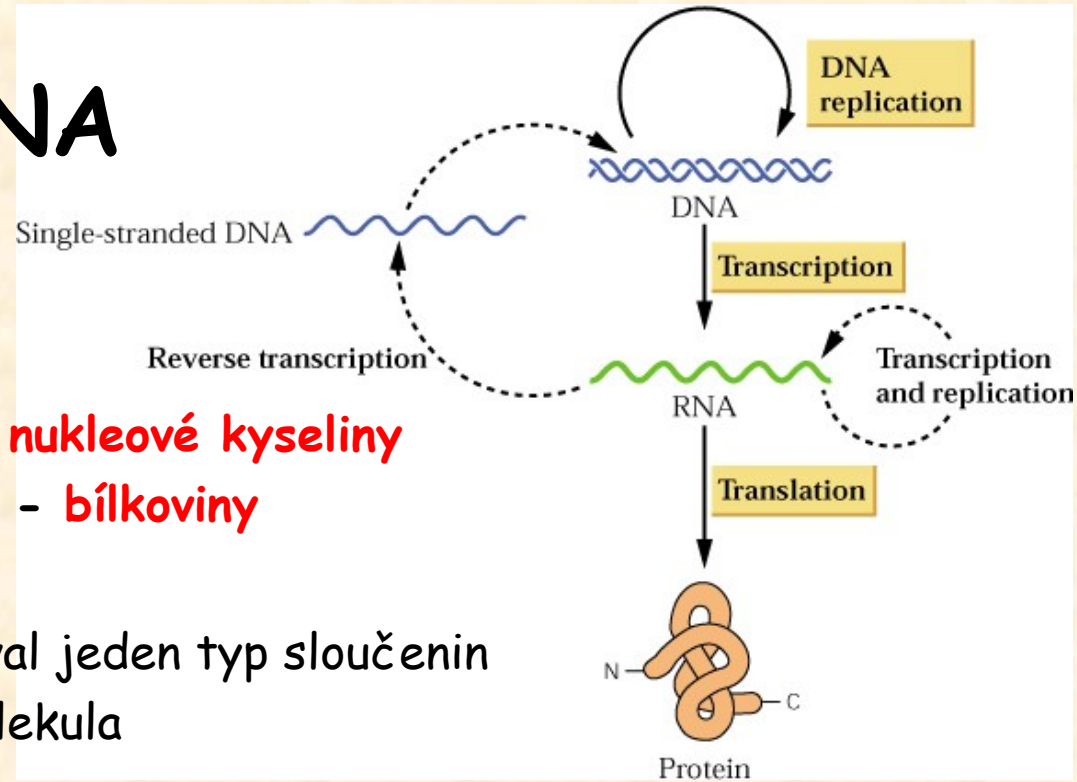
Svět RNA

Funkční specializace dnes:

- uchování genetické informace - **nukleové kyseliny**
- strukturní a katalytická funkce - **bílkoviny**

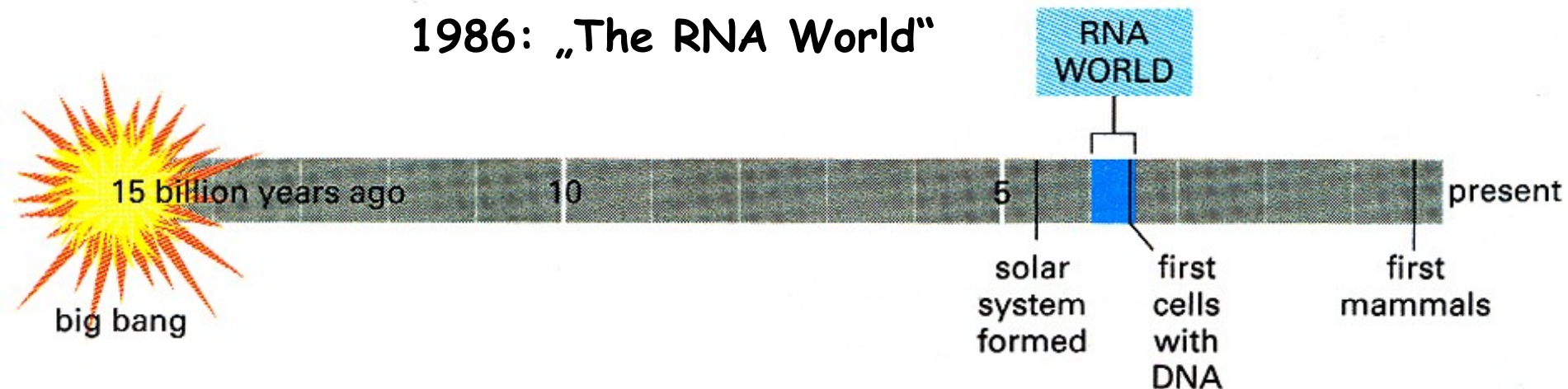
Období, kdy oba typy funkcí zastával jeden typ sloučenin
RNA - informační i katalytická molekula

1982-83: objev **ribozymů**



1986: „The RNA World“

3.5-4mld



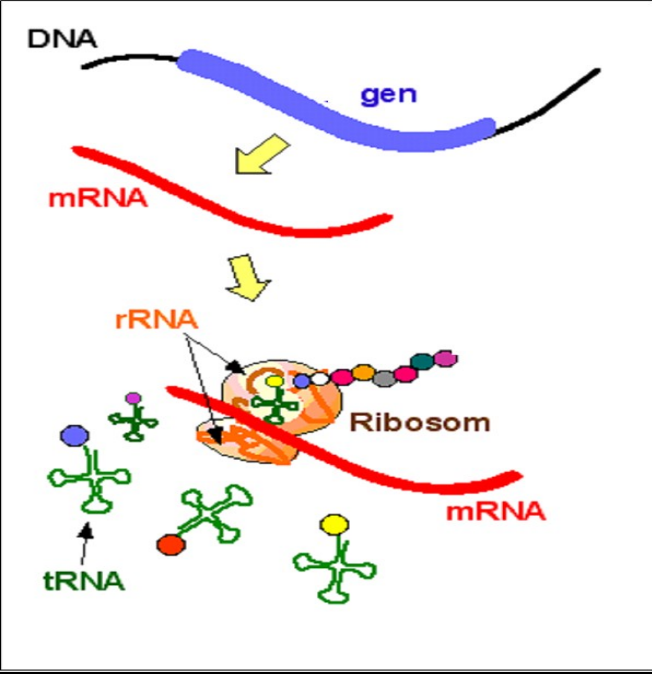
Důkazy RNA světa

1. Důležitá **role RNA** v realizaci genetické informace dnes
2. RNA viry, **retroelementy**, telomery a konzervativní mechanismus jejich replikace
3. **Ribozymy** - enzymaticky aktivní RNA

Kritéria testující zda RNA je **reliktem** světa RNA:

1. Katalytické vlastnosti
2. Všudypřítomnost
3. Centrální postavení v metabolismu

Centrální role RNA v dnešních biologických systémech



DNA

REDUPLICATION:
primer RNAs, telomerase RNA

TRANSCRIPTION

← **6S RNA** Regulation
7S K RNA, SRA RNA, Xist RNA, Air RNA

PROCESSING: sno RNAs, gRNAs, snRNAs, RNase P, self-splicing introns

mRNA

← **siRNA, miRNA** Degradation
Replication?

RIBOSOME:

16S RNA

23S RNA

tRNA

Decoding

Trans-peptidation

TRANSLATION

← **asRNAs** (*micF, CopA, OUY*) Regulation
OxyS, DsrA sRNA
← **miRNAs**

← **tmRNA** Degradation

Protein

← **T/M translocation:** SRP 4.5S RNA, 7S RNA

Pozůstatky RNA světa se dosud zachovaly a jsou skryty ve spleti buněčných procesů

Počátky světa RNA - tvorba purinů a pyrimidinů

Syntéza **adeninu**: z kyanovodíku

Syntéza **cytosinu**:
- z kvanoacetaldehydu
nebo kvanoacetylenu

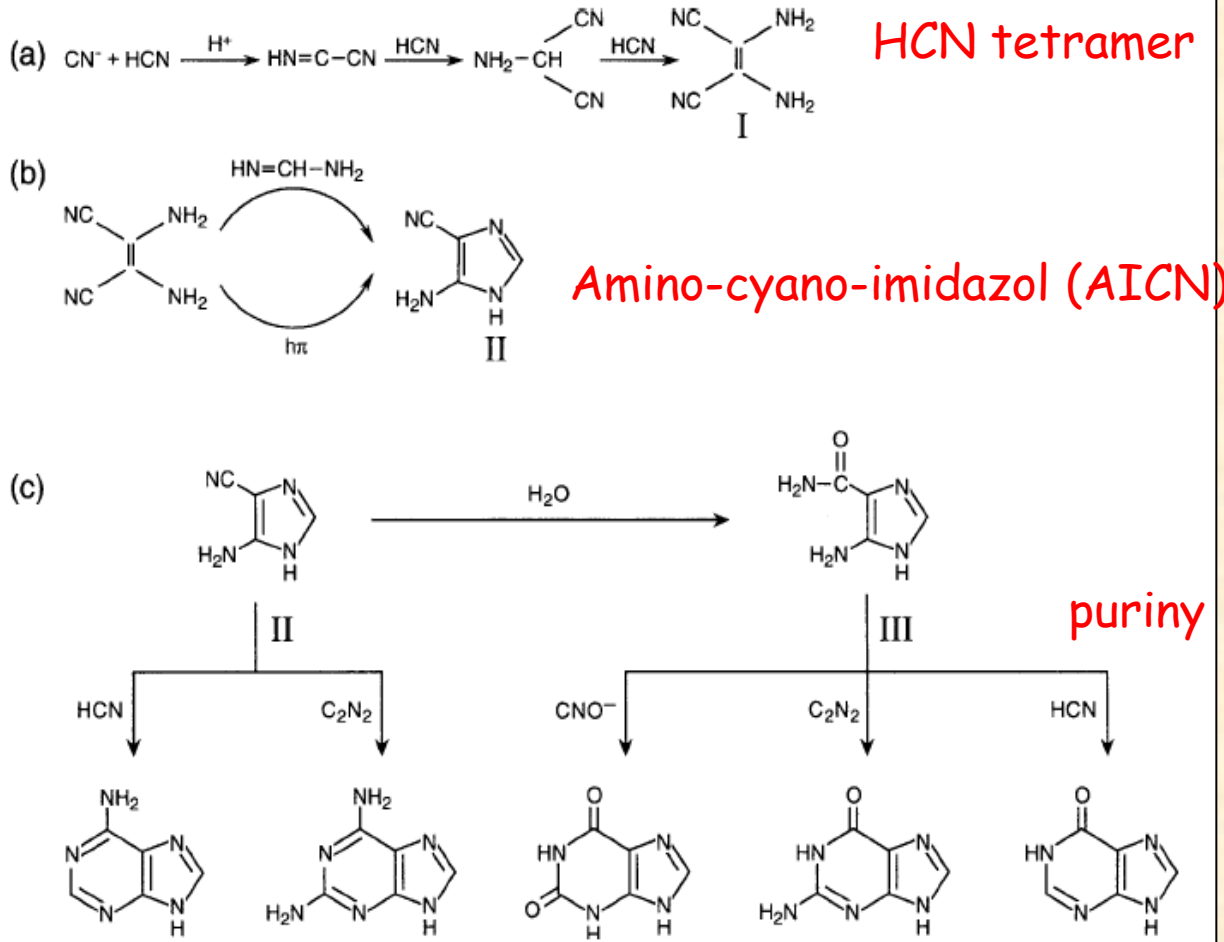


FIG. 2. Steps in possible prebiotic syntheses of adenine from HCN. (a) The formation of the HCN tetramer. (b) The conversion of HCN tetramer to AICN. (c) The formation of purines from AICN or from its hydrolysis product 4-amino-imidazole-5-carboxamide (III).

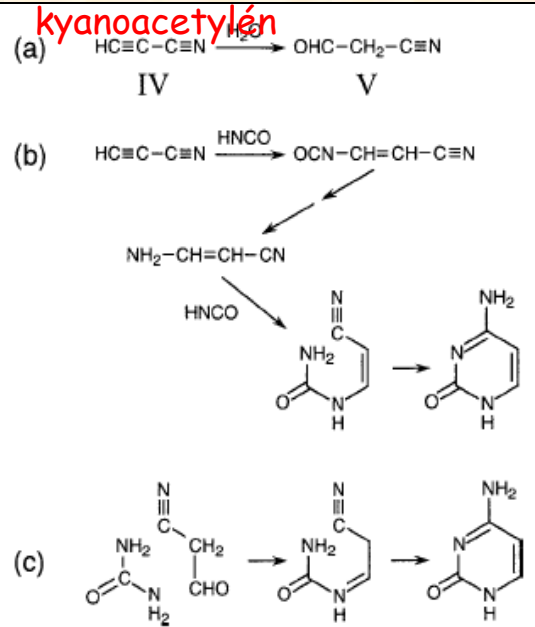
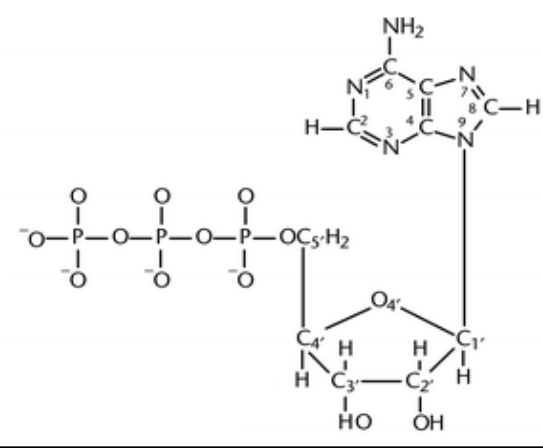


FIG. 3. Steps in proposed prebiotic syntheses of cytosine. (a) The hydrolysis of cyanoacetylene to cyanoacetaldehyde. (b) The reaction between cyanoacetylene and two molecules of cyanic acid. (c) The condensation of cyanoacetaldehyde with urea.

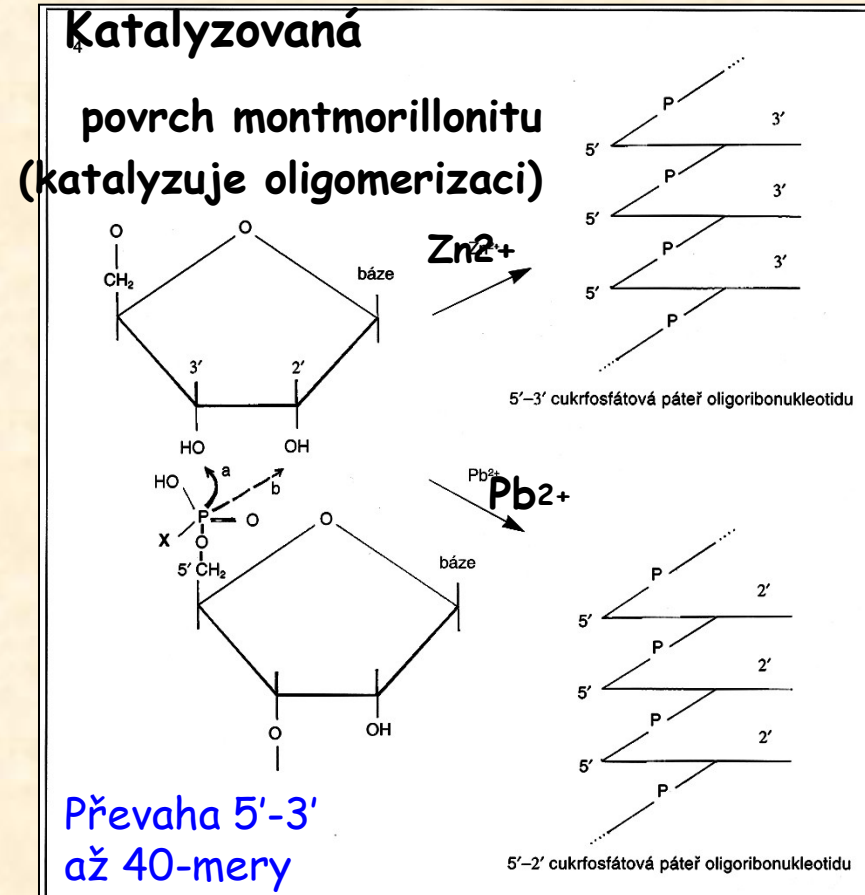
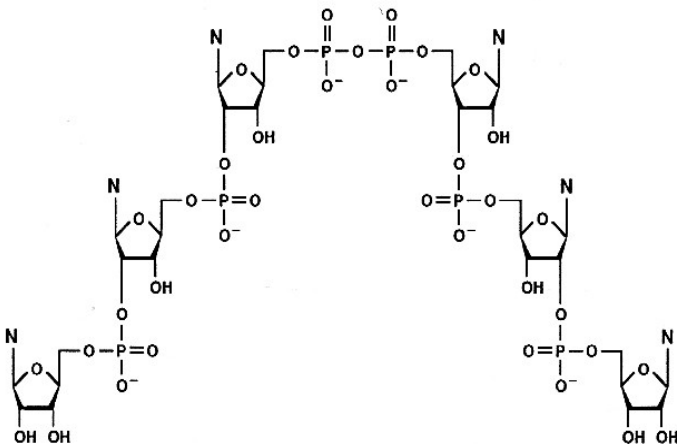
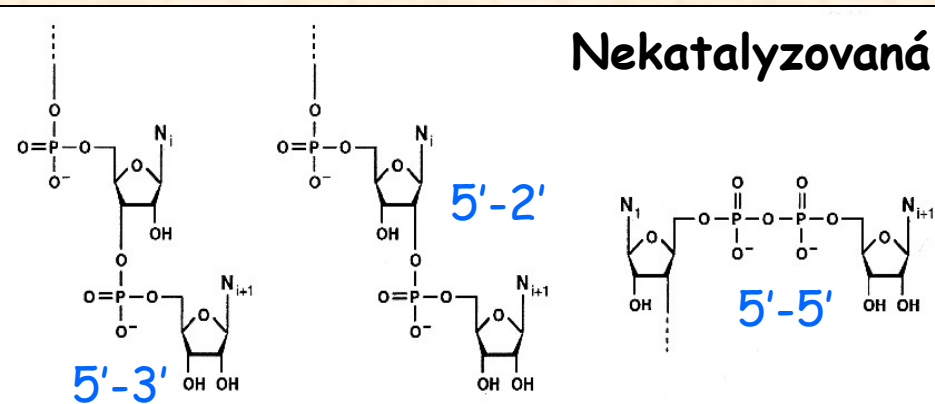
Nízká efektivita a specifita syntézy nukleotidů vedla k návržení alternativních genetických systémů

Abiotická syntéza nukleotidů a polynukleotidů

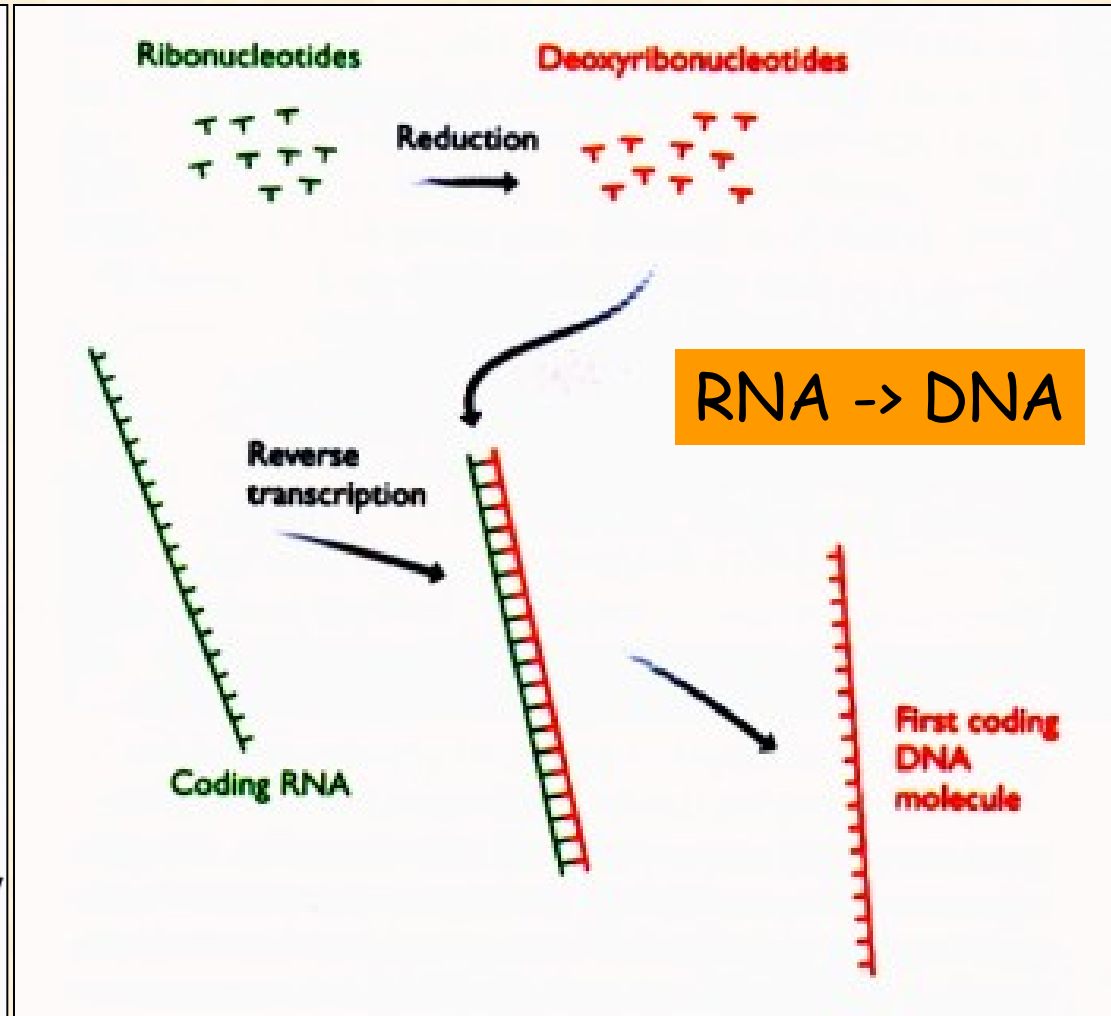
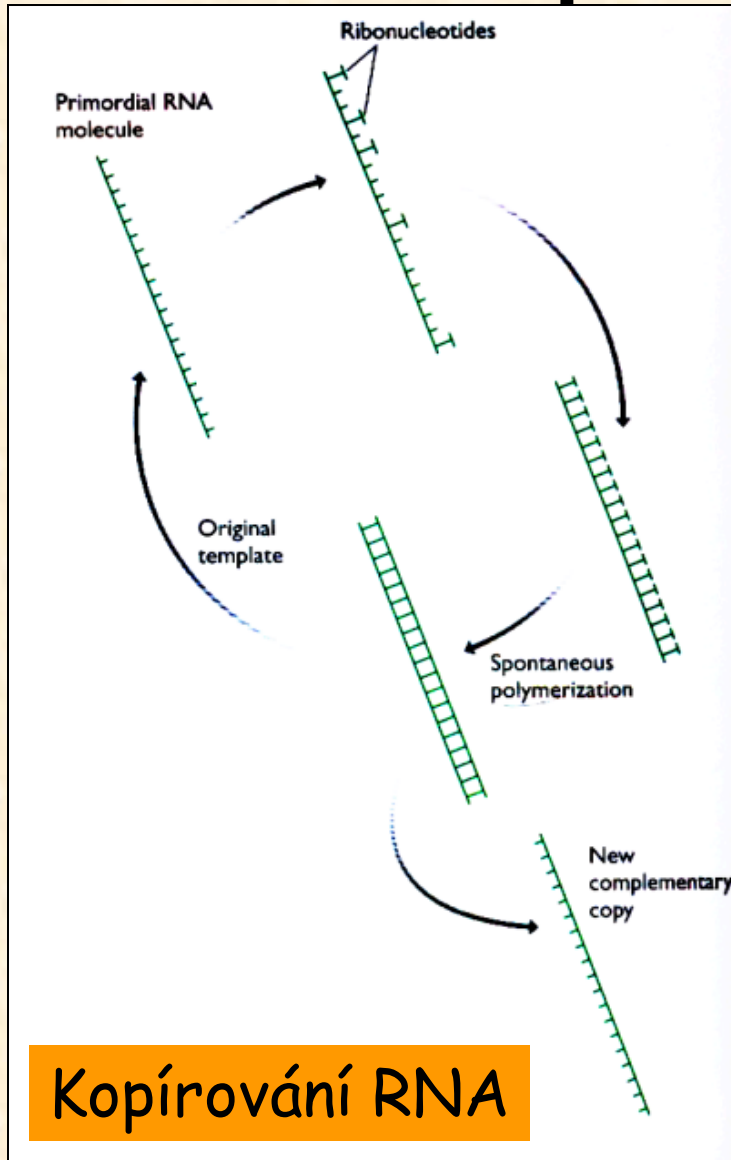
1. Syntéza **nukleozidů**: vazba bází na ribózu
2. Tvorba **nukleotidů**: fosforylace nukleozidů (racemát)
3. Tvorba **polynukleotidů** - tvorba fosfodiesterové vazby



Chemická kondenzace aktivovaných 5'-polyfosfát nukleotidů



Spontánní syntéza prvních RNA a přechod k DNA



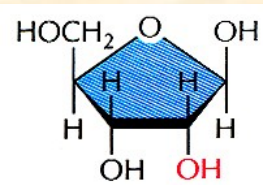
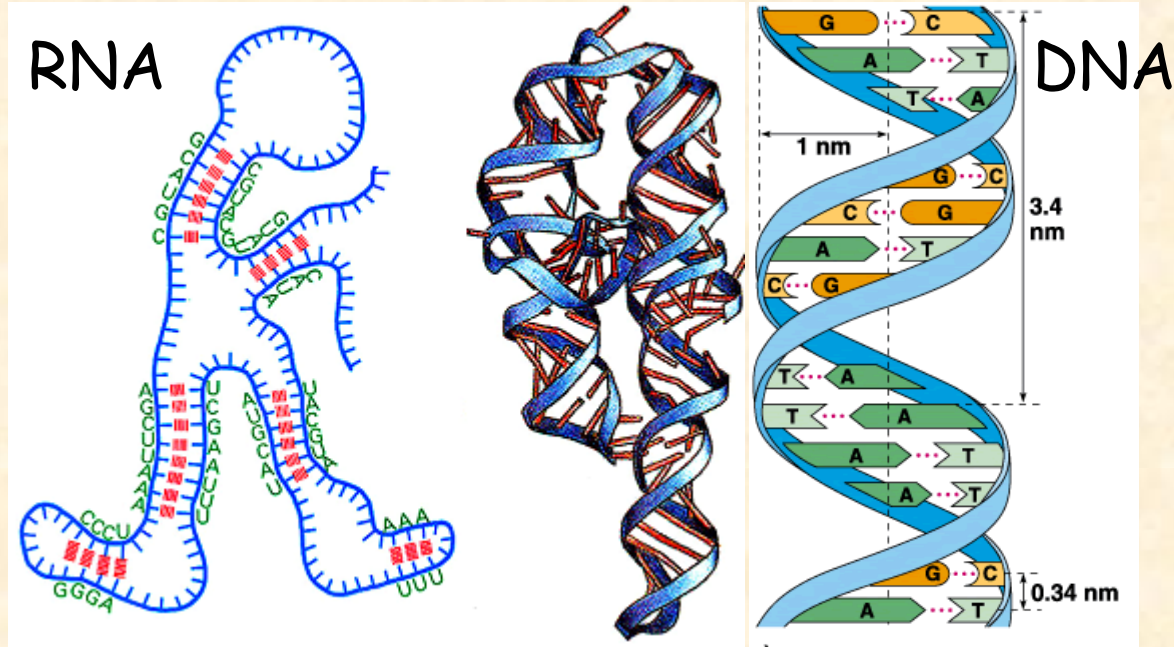
Deoxyribonukleotidy vznikají redukcí ribonukleotidů, thymin z uracilu

Proč je genetická informace uložena v DNA

Rozdíly mezi RNA a DNA:

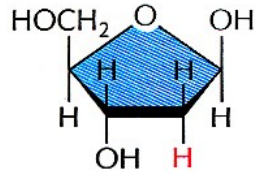
Ribosa (2'-OH skupina)

Uracil místo thyminu



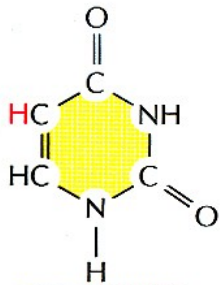
ribose

used in ribonucleic acid (RNA)



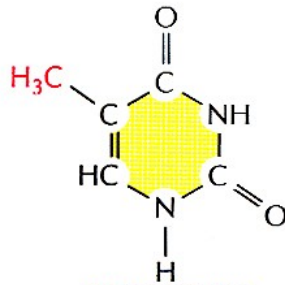
deoxyribose

used in deoxyribonucleic acid (DNA)



uracil

used in RNA



thymine

used in DNA

Důsledky:

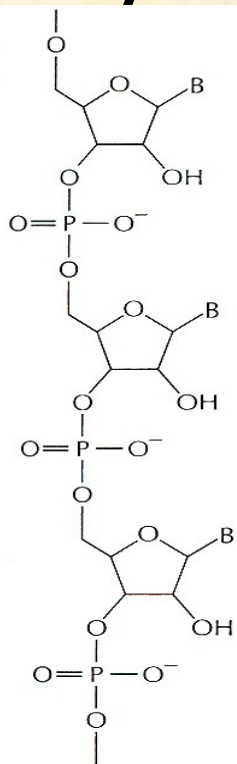
- vyšší chemická i fyzikální stabilita DNA (UV záření)
- delší molekuly (uchování komplexní informace)
- dvouřetězcová (replikace)
- méně reaktivní deoxyribóza
- konformační flexibilita - funkční relevance

Genetické systémy předcházející světu RNA

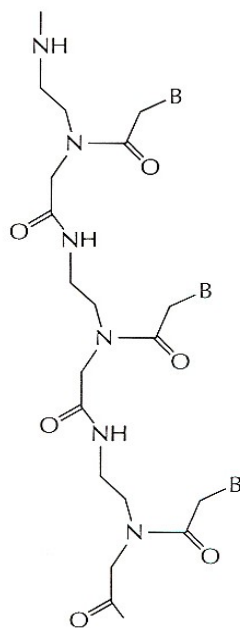
Molekuly RNA:

- chemicky **nestabilní**
- strukturně **složitá**

Peptidová nukleová kyselina (PNA)



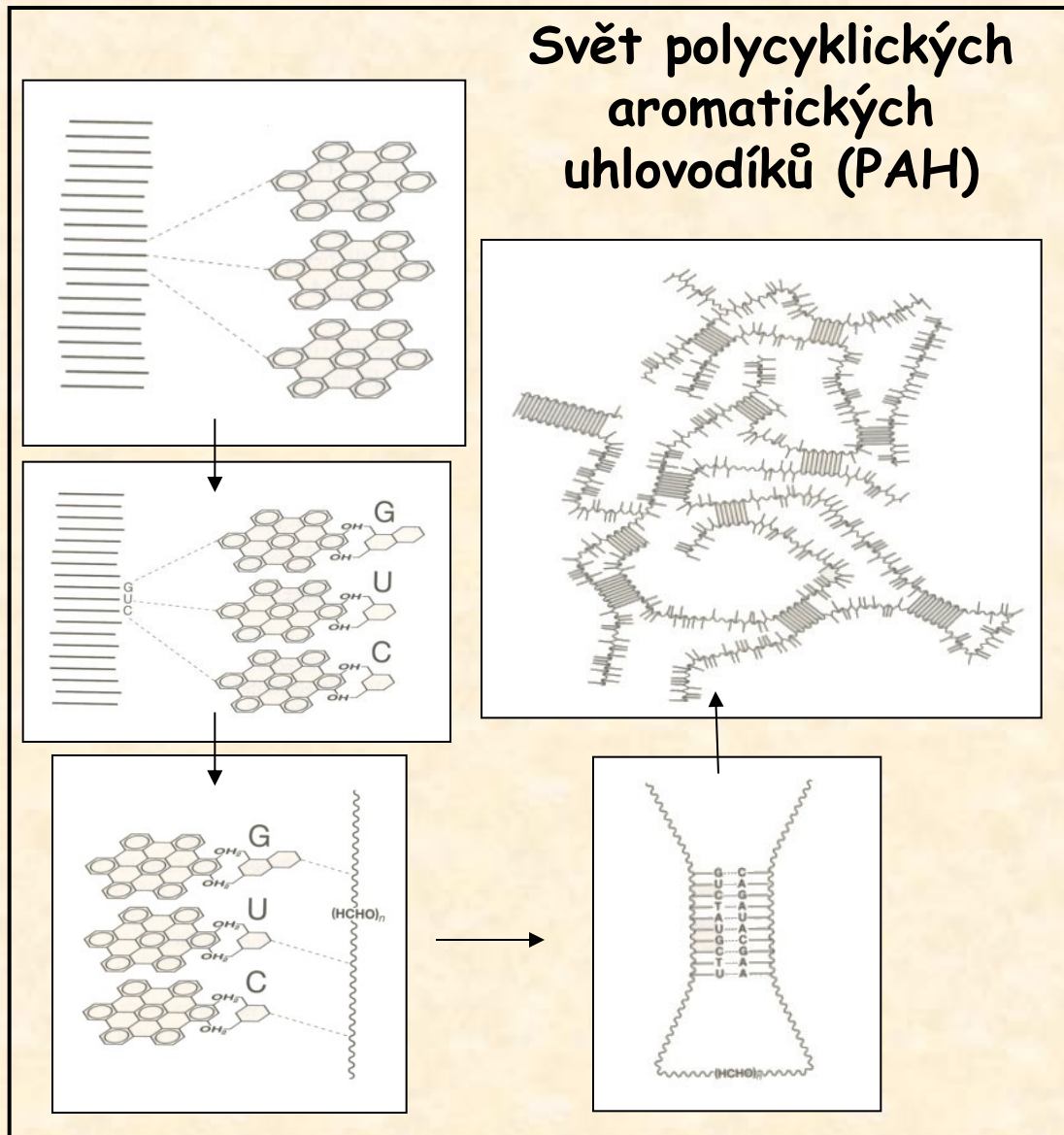
RNA



PNA

Páteř - lysin, glycin

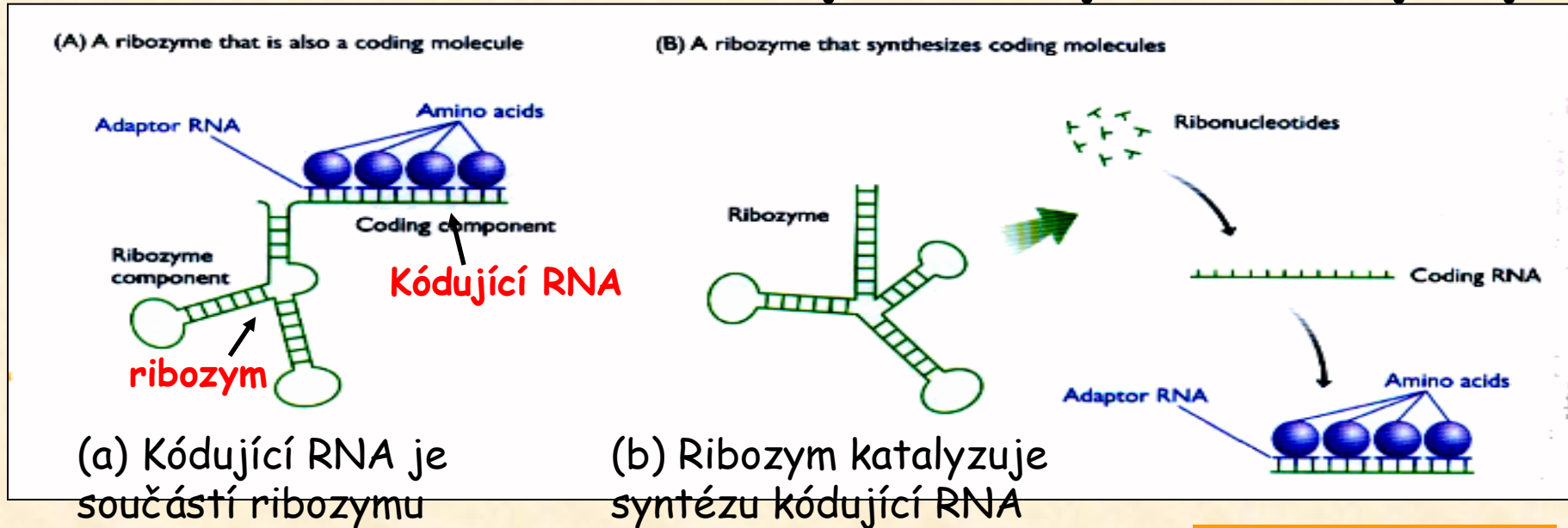
světu RNA



Svět polycyklických aromatických uhlovodíků (PAH)

RIBOZYMY

Co dokáží RNA katalyzátory (ribozomy)



Katalyzované reakce - substrátem **většinou RNA**:

1. nejčastěji **hydrolýza fosfodiesterových** vazeb (endonukleáza)
2. obrácený směr - **syntéza fosfodiesterových** vazeb (ligáza, polym.)
3. **transesterifikace** - editace, sestřih

Substrátem **není RNA** !

1. **syntéza peptidové** vazby

RNA se dokáže sama modifikovat, vystřihovat, spojovat.

Osud RNA katalyzátorů po převzetí jejich funkce proteiny

Přechod RNA → proteiny stále probíhá

1. Vymizely:

2. Převzaly nové funkce:

Ribozom: **replikace** → translace

Spliceosom: **rekombinace** → sestřih

3. Zachovaly si vysoce konzervativní funkce:

- snoRNA - úpravy **rRNA**

- RNAza P - úpravy **tRNA**

- snRNA - sestřih intronů v **mRNA**

- tyto funkce vysoce konzervativní - zachovaly se u eukaryot

- **ztráta** některých RNA reliktvů **u prokaryot** - proteiny jsou účinnější

RELIKTY SVĚTA RNA

Představitelé reliktního světa RNA

1. tRNA - od replikace k proteosyntéze
2. Ribozóm
3. Sestřih a snRNA
4. Maturace rRNA a snoRNA
5. Maturace tRNA a RNázaP
6. Signální rozpoznávací částice a srpRNA
7. Editace RNA a řídící RNA (gRNA)
8. Telomeráza a telomerická RNA
9. Vault RNA (vRNA)

1. Role tRNA a ribozómů: od replikace k proteosyntéze

Dnešní translace - složitá koordinovaná síť interakcí RNA a proteinů - vyvinula se z mnohem jednodušších systémů existujících ještě ve světě RNA (RNA-RNA interakce), důležitá schopnost replikace
Prvotní funkce ribozómů - polymerizace nukleotidů - RNA replikace

Původní role ribozómů v replikaci:

- Ribozomální protein S1 a translační elongační faktor Tu a Ts - jsou podstatnou složkou **replikázového komplexu** fága Q β
- Elongační faktory jsou složkami replikázových komplexů některých rostlinných RNA virů

Hypotéza genomových značek:

Molekuly primitivních tRNA fungovaly jako značky molekul RNA určených k replikaci RNA ribozymy, TLS (tRNA-like structures) - na 3-konci Qbeta
Aminoacylace tRNA - další značka

Pozůstatky replikační role tRNA v dnešních genomech

1. Telomeráza:

krátký fragment RNA funguje jako **templát** pro RT

2. Retroelementy:

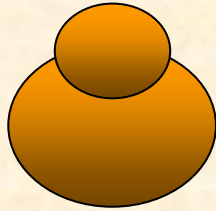
tRNA funguje jako **primer** syntézy cDNA při reverzní transkripci

3. RNA viry:

struktury podobné tRNA na 3-koncích genomové RNA TLS (tRNA like structure) jsou aminoacylovány histidinem, valinem nebo tyrosinem a fungují jako **primery** replikace RNA, jen 3-koncové CCA nebo CCCA jsou potřebné pro replikaci

2. Ribozómy: ribozymy stabilizované proteiny

Ribozómy lze považovat za relikty světa RNA



prokaryotický ribozóm

malá podjednotka

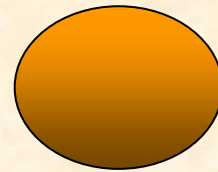
16S rRNA (1542 b)



21 proteinů (S1-S21)



S1 - role v replikázovém komplexu fágů



velká podjednotka

5S rRNA (120 b)



23S rRNA (2904 b)

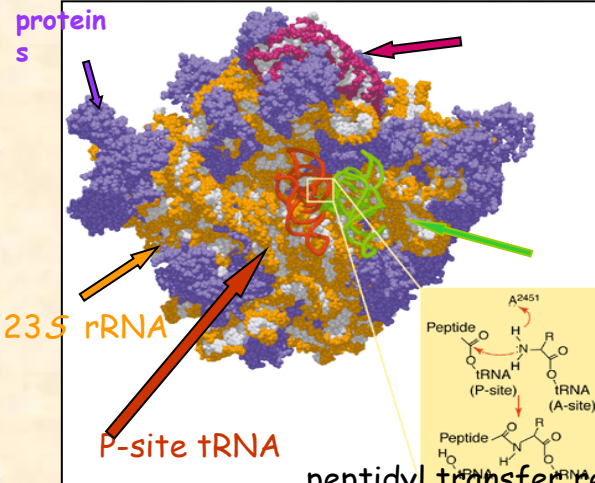


peptidyltransferázová aktivita

32 proteinů (L1-L34)



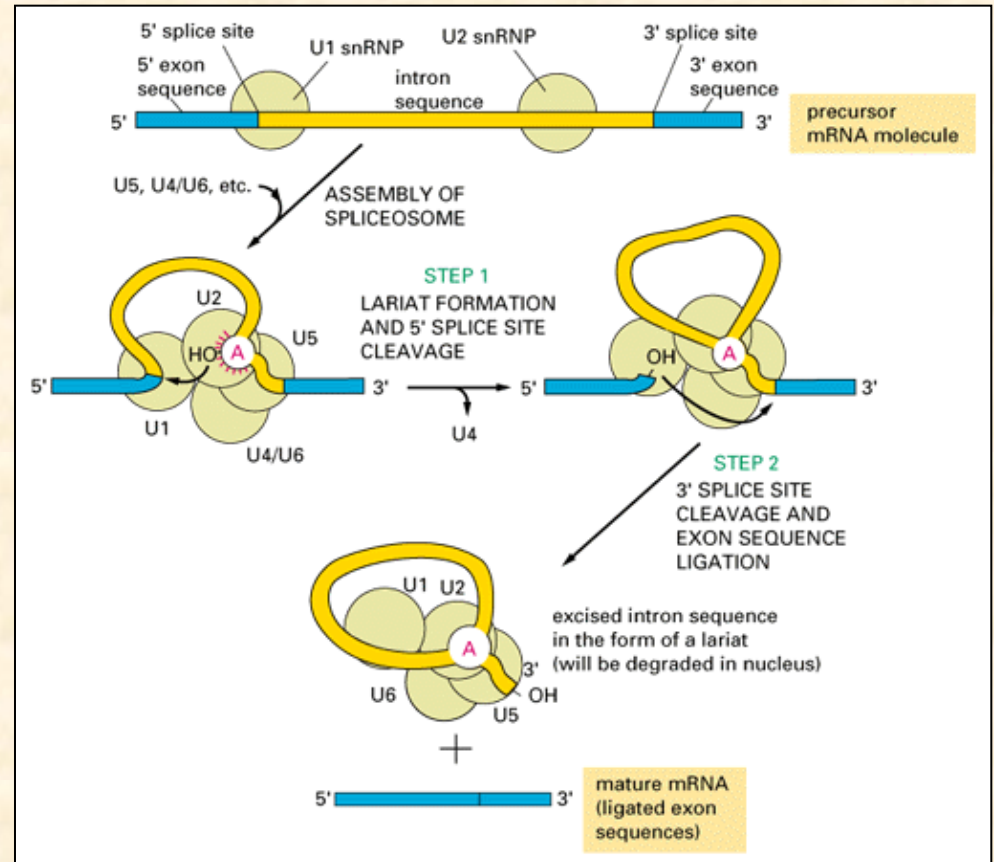
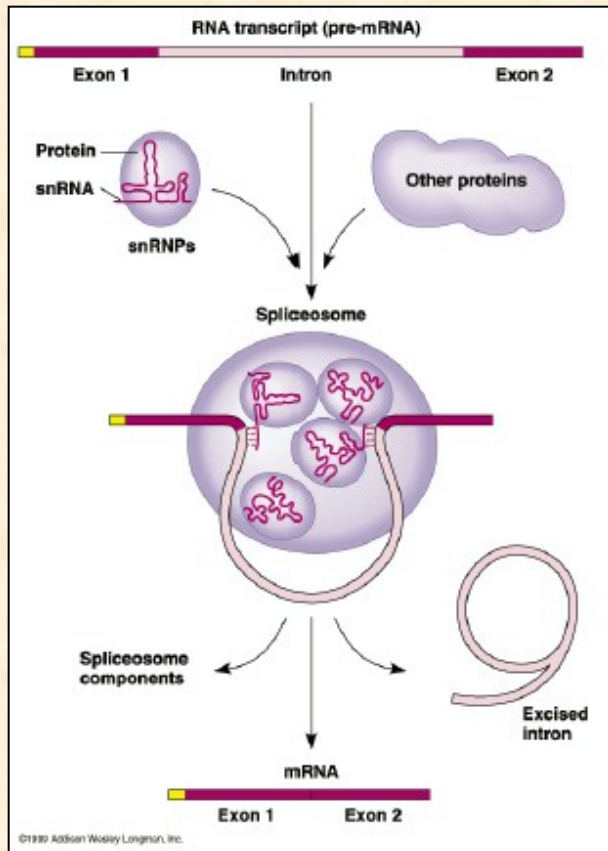
L2 - někde převzal PT roli



3. Sestřih, spliceozóm a snRNA

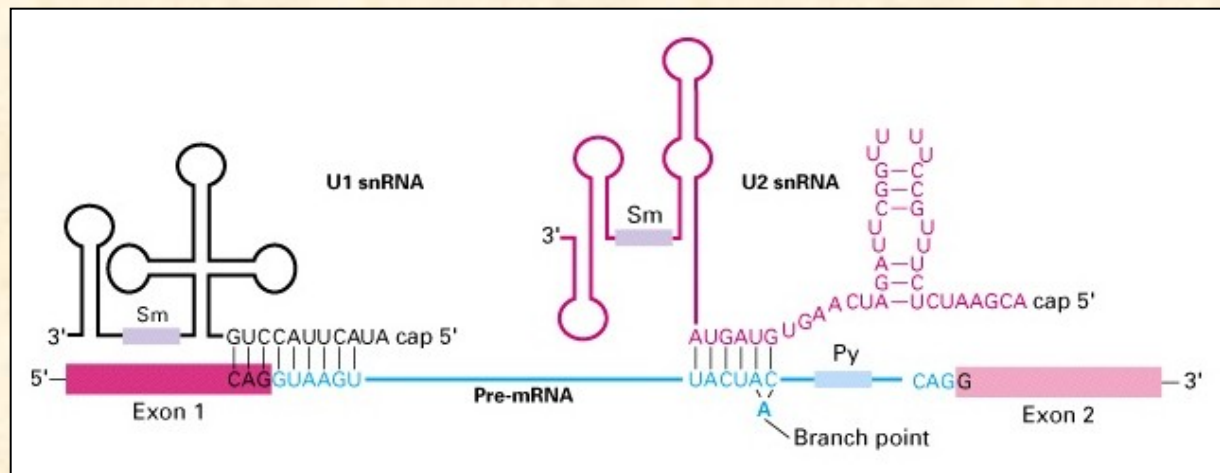
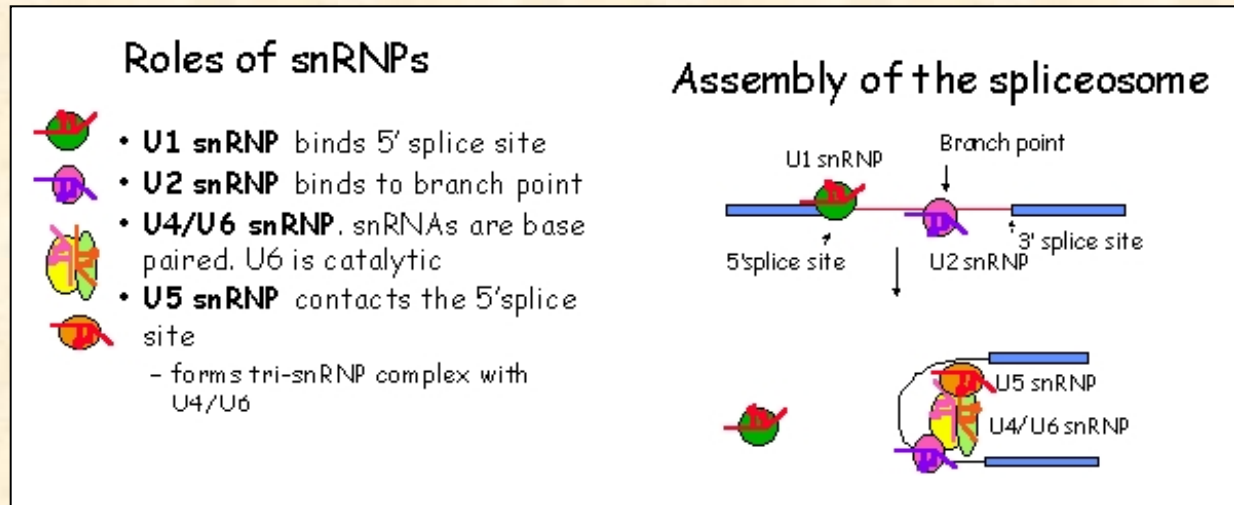
Spliceozóm:

- posttranskripční úprava pre-mRNA (hnRNA), vystřižení intronů
- účast snRNA a proteinů - snRNP
- RNA-RNA interakce
- starší než translace - původ v rekombinaci molekul RNA
- bimolekulární trans-splicing molekul ssRNA

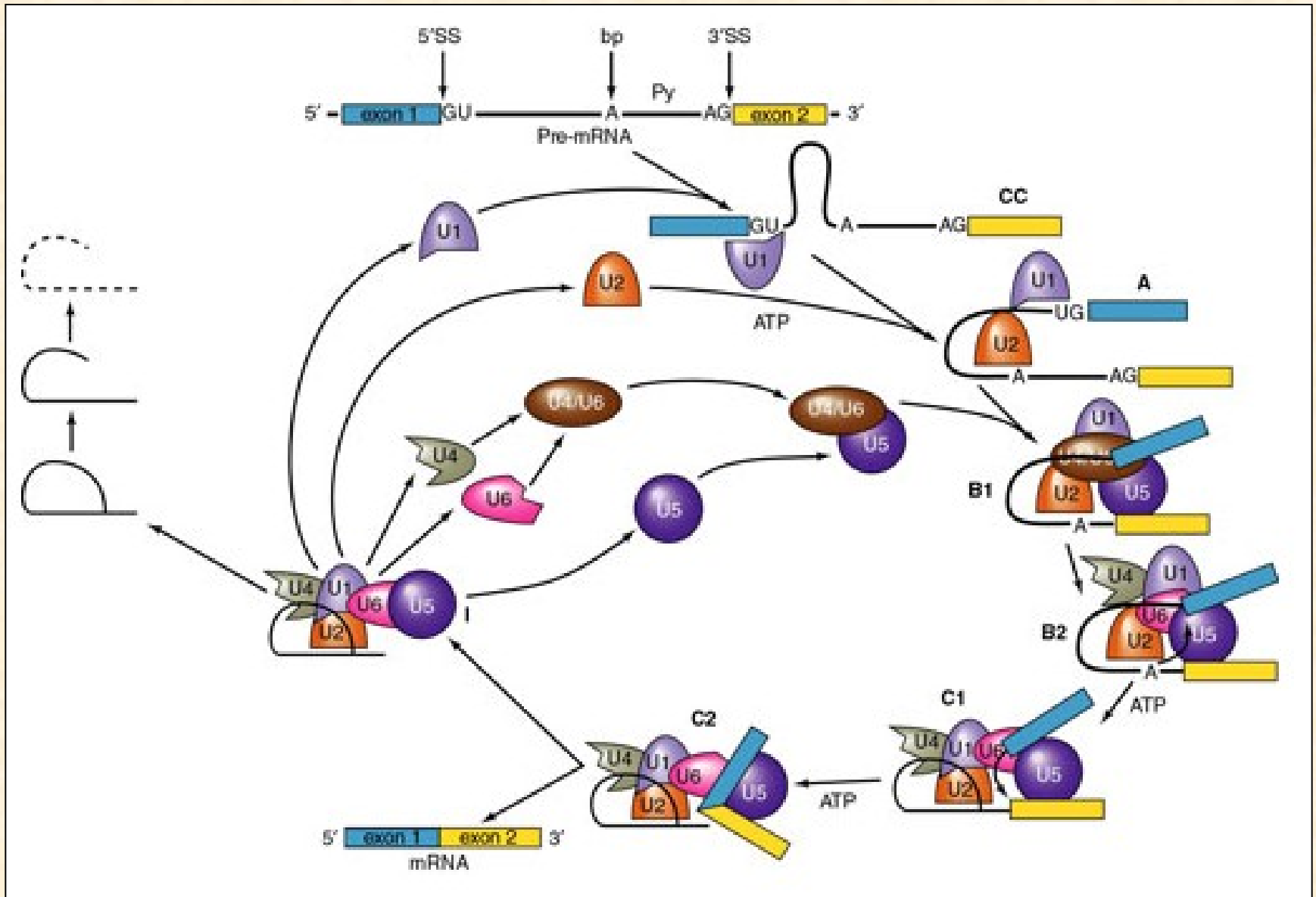


Malé jaderné RNA (snRNA)

- nacházejí se v jádře eukaryot
- účastní se sestřihu pre-mRNA a udržování telomer
- tvoří nukleoproteinové částice (snRNP = snurps), každá s více proteiny
- jsou kódovány introny
- U1, U2, U4, U5, U6
- U4+U6 se párují spolu a U6 je katalytická



Malé jaderné RNA (snRNA)



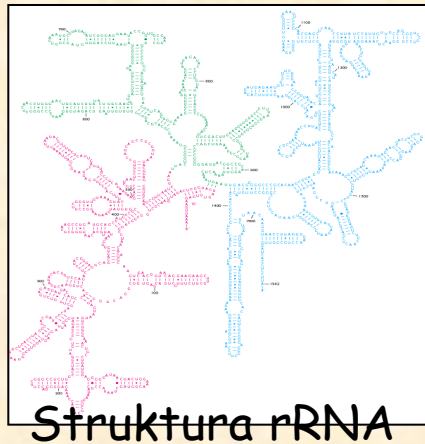
4. Maturace rRNA a snoRNA

snoRNA (malé jadéřkové RNA):

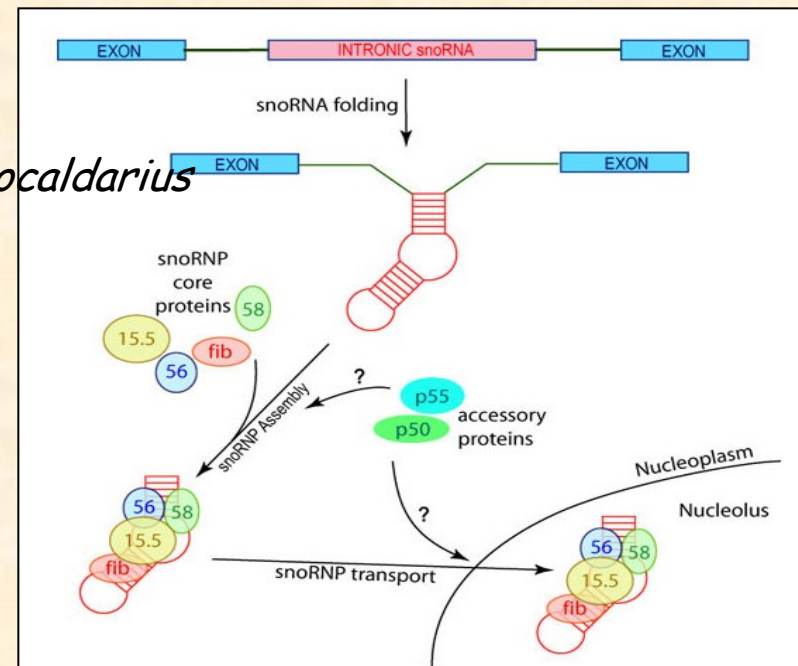
- účast při maturaci rRNA a ribozómů
- velký funkční komplex - snorpozóm
- **kódovány introny** některých genů - ribozomálních a heat shock genů
- 8 různých snoRNA kódováno 8 introny jednoho genu
- u savců 30 různých snoRNA, u kvasinky 26 snoRNA ~ 5426 b (ancestrální snorpozóm)
- homologie snoRNA s rRNA (18S a 28S), intra- i intermolekulární kontakty (kroslinkování)
- některé snoRNA potřebují spliceosom ke své maturaci

Prokaryota:

- **absence snoRNA** u prokaryot je záhadou
- maturace rRNA **jen** za účasti **proteinů**
- objev U3snoRNA u archeobakterie *Sulfolobus acidocaldarius*

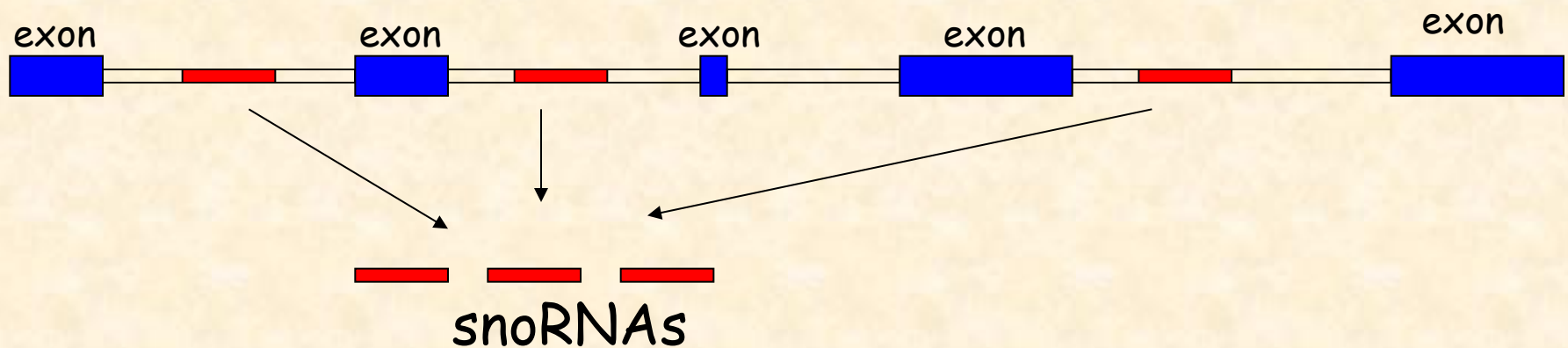


Tvorba snoRNA



Introny jsou někdy důležitější než exony (snoRNA v intronech)

- gen UHG (U22 host gene) obsahující v 8 intronech 8 různých snoRNA
- sestřihem vzniká mRNA, která je však degradována
- mRNA je málo konzervativní mezi člověkem a myší
- hlavním funkčním produktem UHG genu jsou tedy molekuly snoRNA



Tycowski et al (1996): A mammalian gene with introns instead of exons generating stable RNA products. Nature 379: 464-466.

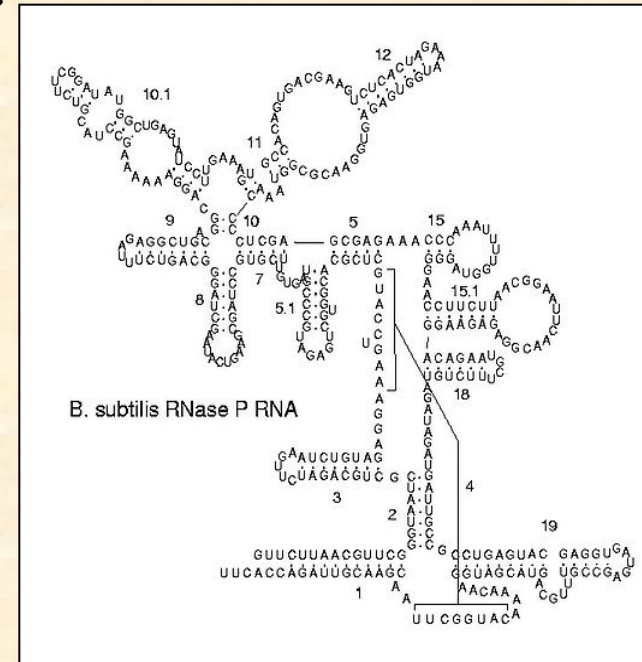
5. Maturace tRNA a RNázaP

tRNA:

- relikv světa RNA - konzervativní, všudypřítomná, centrální úloha v metabolismu
- interakce s rRNA (CCA konec tRNA interaguje s 23SrRNA)
- původní funkce v replikaci, později v proteosyntéze
- některé geny pro tRNA mají introny

RNázaP:

- úloha v maturaci tRNA
- je skutečným **enzymem**, štěpí opakovaně
- RNA katalytická podjednotka (=M1 RNA) + proteinová podjednotka (=C5)
- jediný ribozym modifikující RNA u **prokaryot**
- molekulární fosílie



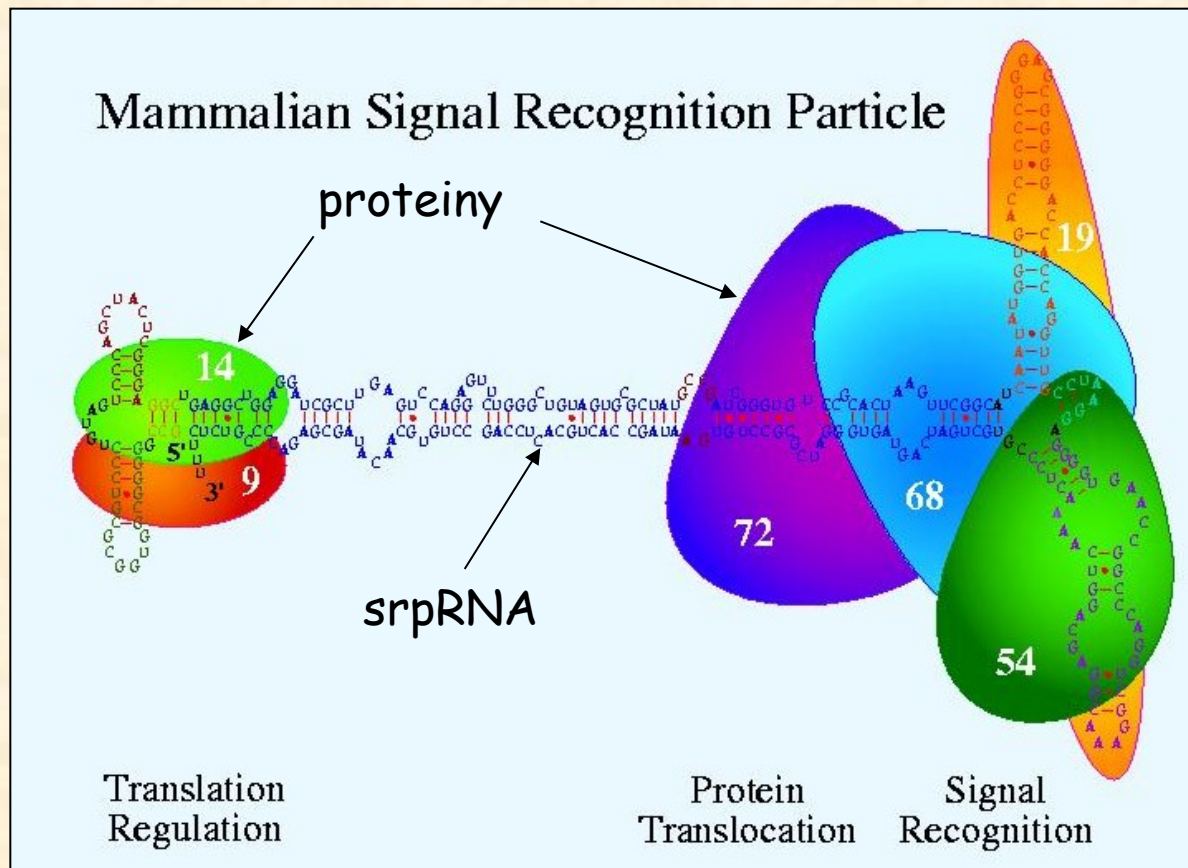
RNA složka RNAázyP
(370 bází)

RNáza MRP:

- druhá podobná molekula vzniklá duplikací a divergencí u eukaryot nebo endosymbiózou
- výskyt u *Giardia* a *Microsporidia* - nemají mitochondrie

6. Signální rozpoznávací částice a srpRNA

- RNA-proteinový komplex zajišťující vazbu ribozómu na ER a sekreci proteinů
- RNA složka 7S RNA u eukaryot a archeí, asi 300 b
- podobná struktura a funkce, homologie s Alu sekvencemi
- stimuluje hydrolýzu GTP
- původně ribozym štěpící GTP



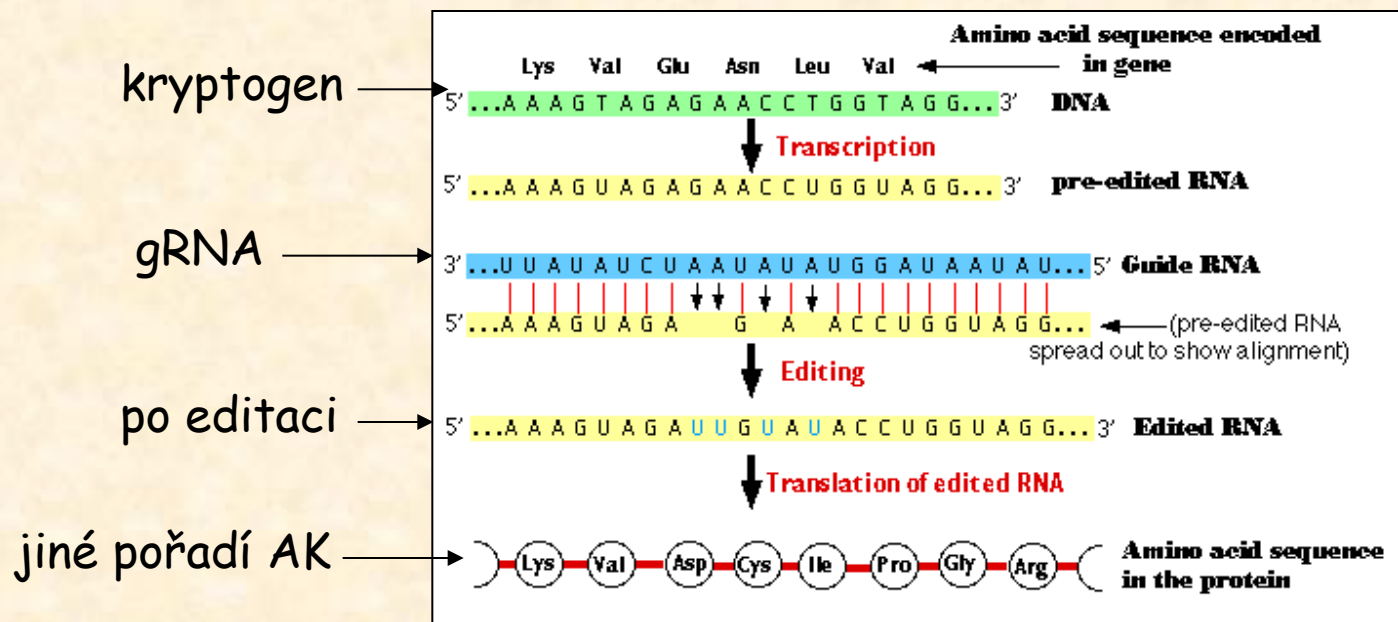
Schematic representation of the mammalian SRP depicting SRP9, SRP14, SRP19, SRP54, SRP68, SRP72 and SRP RNA. The part of SRP comprising SRP9/14 complexed with RNA forms a distinct structural domain known as the **Alu domain** due to homology of the Alu family of RNA sequences with the Alu family of repetitive DNA sequences and the small cytoplasmic Alu RNAs (scAlus). The Alu domain of SRP mediates the specific pauses(s) in the synthesis of nascent ER targeted proteins whose signal sequence has been bound by SRP54.

7. RNA editace, g-RNA, editozóm

- posttranskripční úpravy - modifikace tRNA, rRNA a pre-mRNA,
- substituce, inserce, delece, **kryptogeny**, templátem je guide RNA (g-RNA)
- eukaryota, mitochondrie trypanosom - inserce či delece polyU
- editace je podmínkou tvorby sekundárních struktur bez nichž nemůže dojít k maturaci tRNA RNázou P

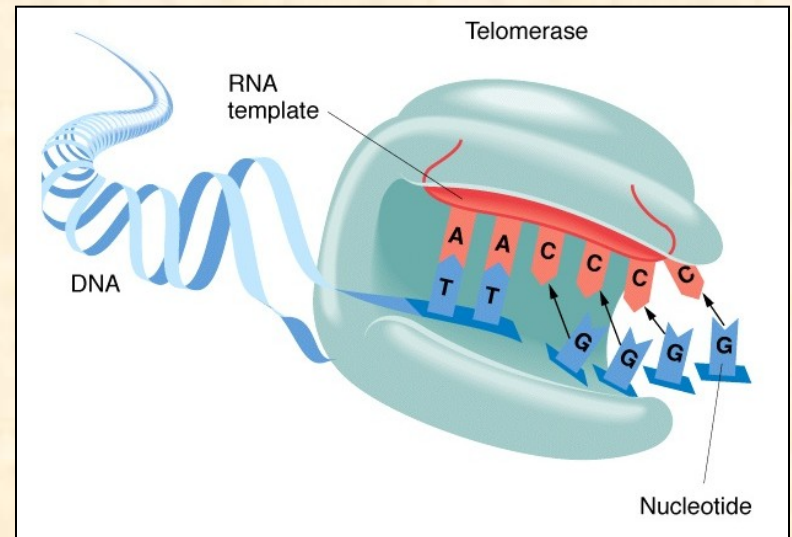
Původ editace:

- u mitochondrií - reakce na asexualitu (Mullerova rohatka), korekce
- ve světě RNA - editace jako kontrolní mechanismus exprese tRNA
- nádorová nebo neurologická onemocnění (epilepsie)



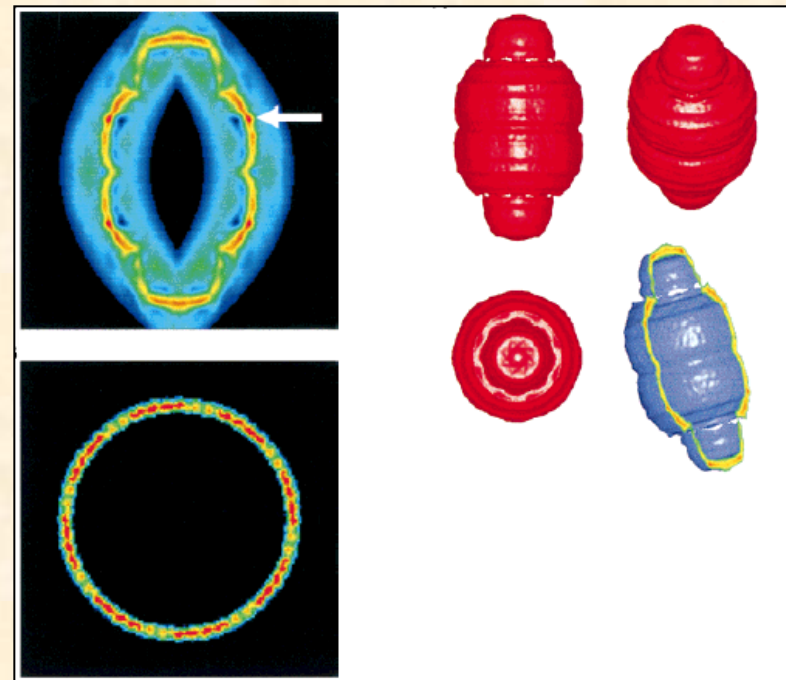
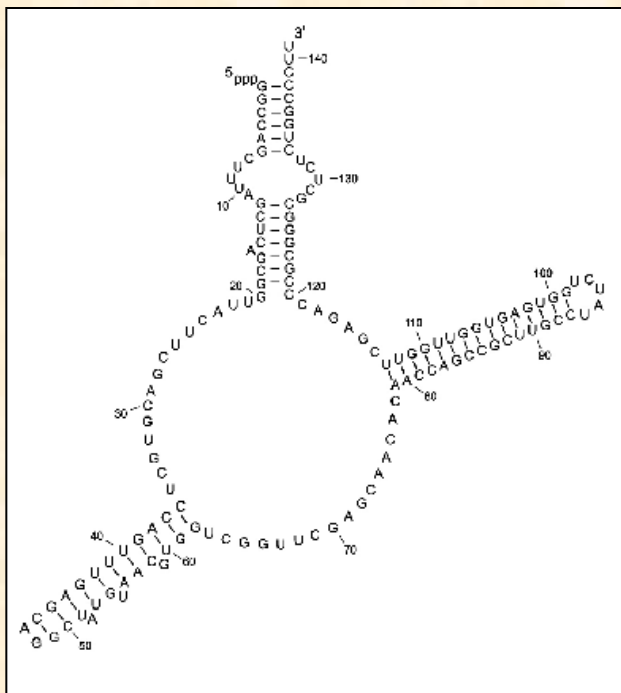
8. Telomeráza

- problém replikace konců lineární DNA u eukaryot - RNP komplexy
- RNA složka jako templát pro syntézu telomerických repetitiv
- RNA složka tvoří terciální strukturu, účast v katalýze nejasná
- nepřítomna u prokaryot, cirkulární genomy
- mutace telomerické RNA vede k prodlužování telomer
- homologie s reverzní transkriptázou
- starobylé RNA genomy byly lineární --> podpora hypotézy genomových značek



9. Vault RNA

- přilepena na povrchu jaderné membrány a asociována s komplexem jaderných pórů
- funkce neznámá, spíše funkční než strukturní (exp. odstranění RNA)
- souvisí s rezistencí rakovinných buněk k léčivům
- obsahuje RNA, sekvence konzervativní
- tvoří značku pro transport NK z jádra a do jádra
- v RNA světě existovala proto-jádro a proto-plazma, aby separovaly replikaci a transkripci, omezení šumu



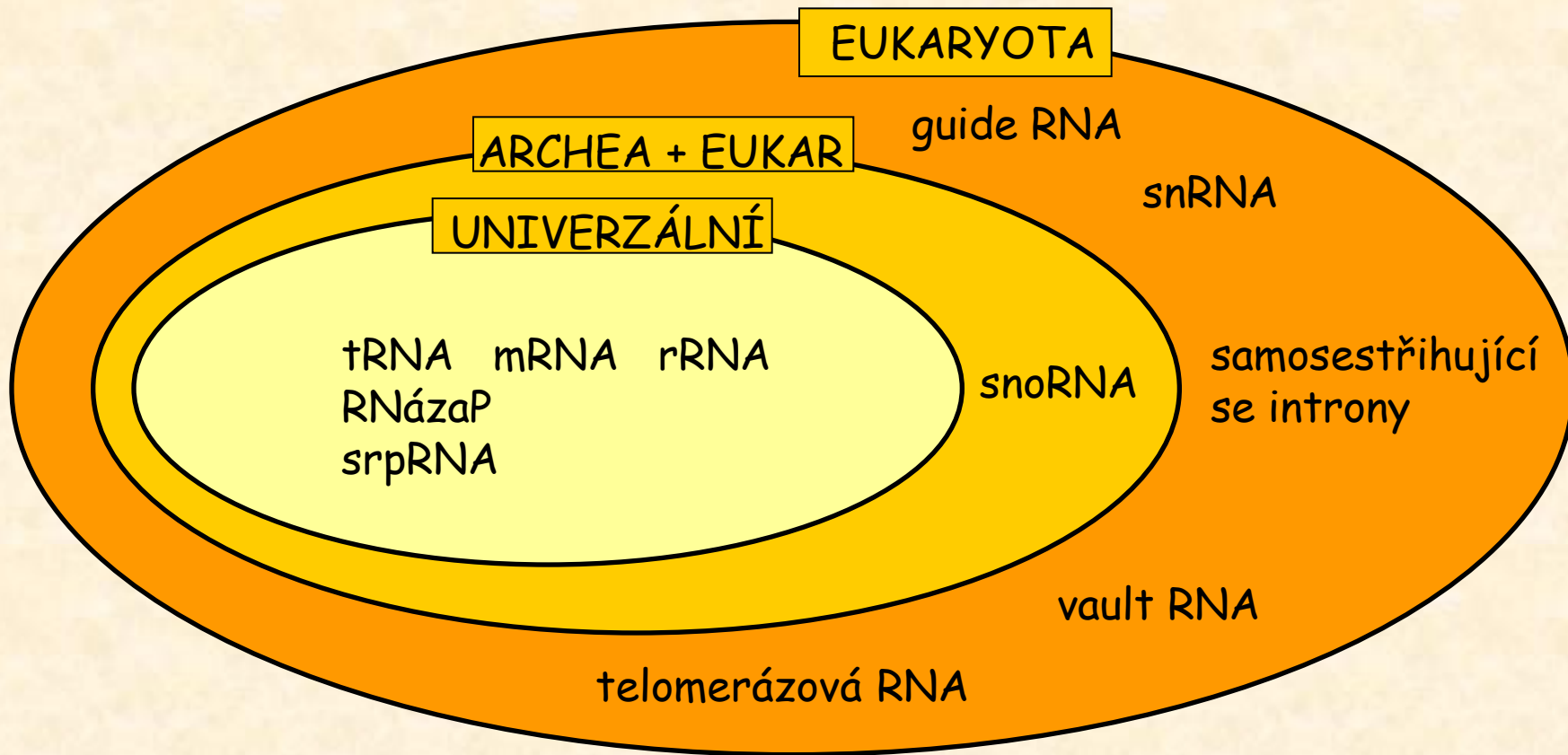
Provázanost ribozymů



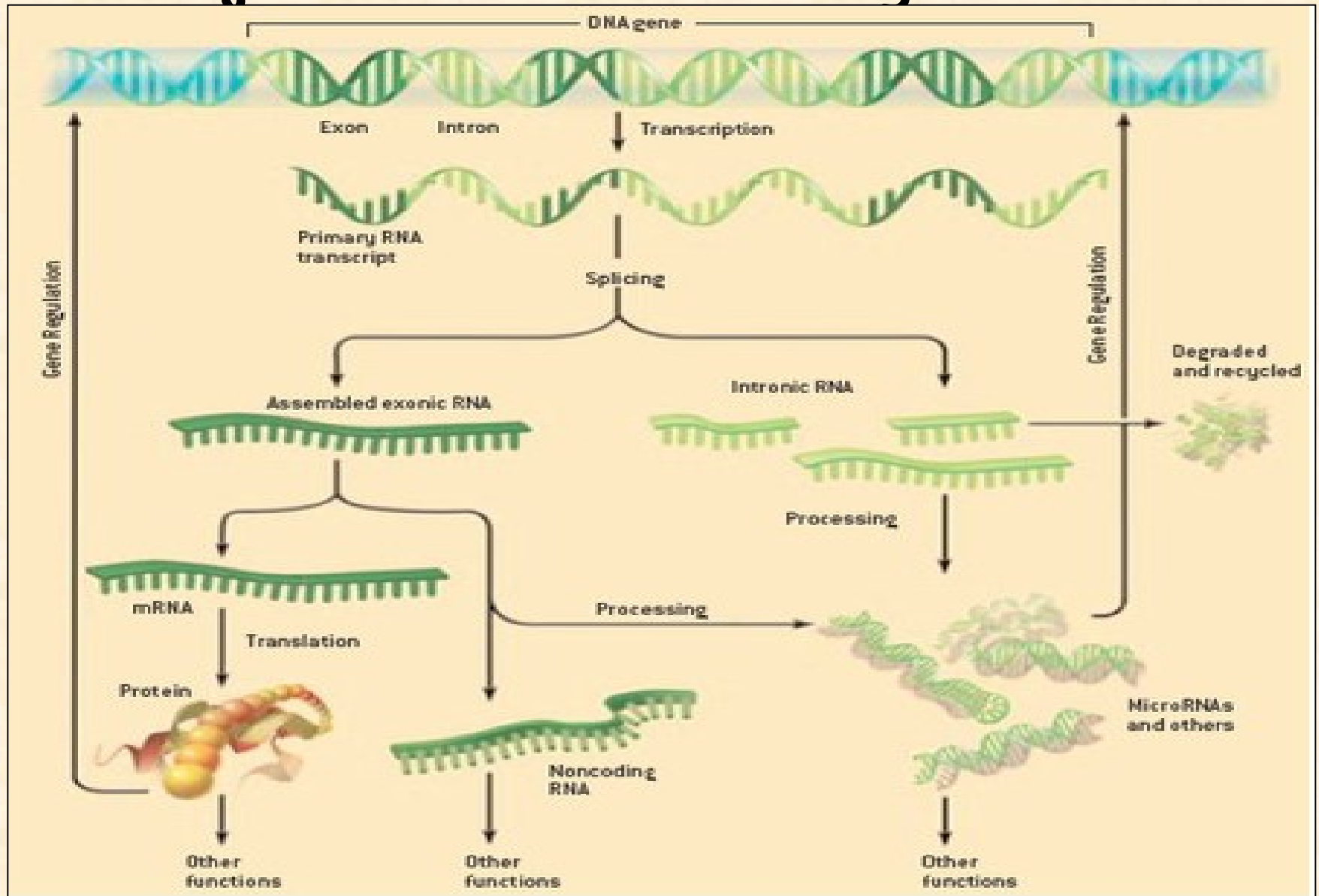
snRNA jsou potřeba pro sestřih snoRNA, které se nacházejí v intronech jiných genů

snoRNA jsou potřeba pro sestřih rRNA

Fylogenetický výskyt fosilních RNA



Nekódující RNA: důležité regulační funkce



Regulační systém na bázi RNA

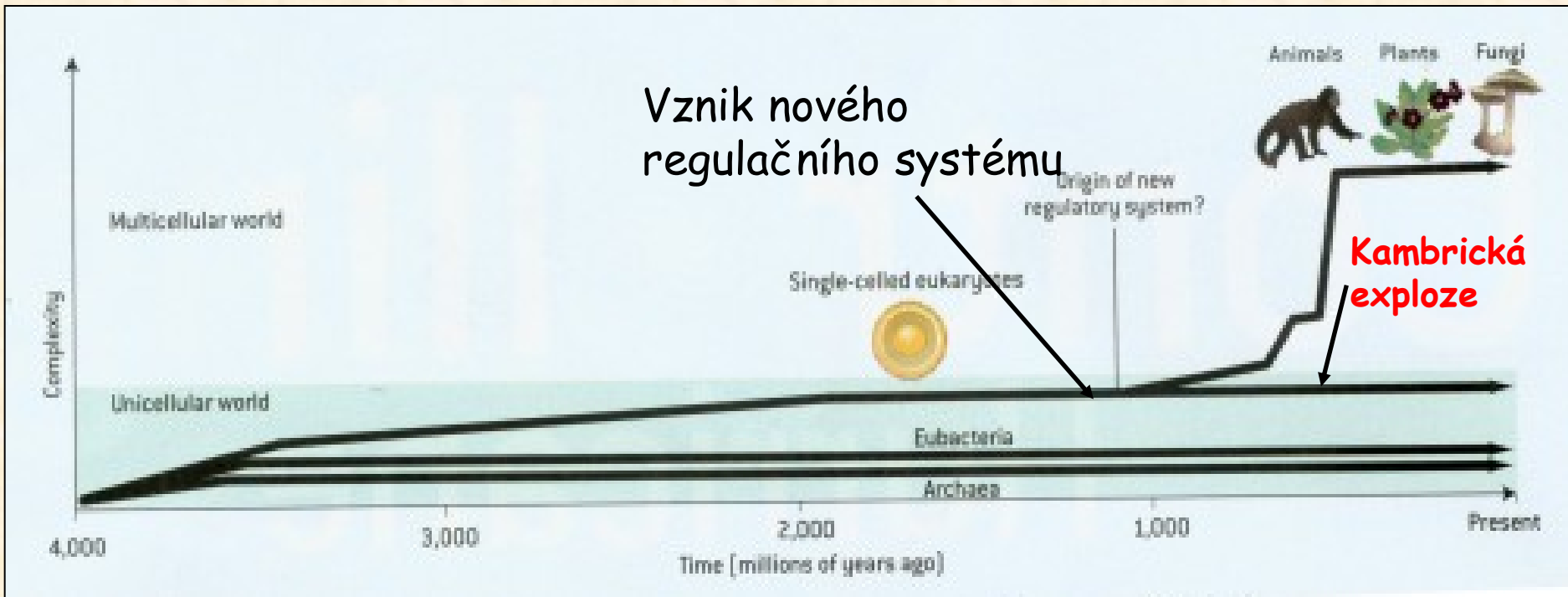


Figure 3 | **A simplified biological history of the Earth.** This graph is intended to present an overview. Some dates are still being debated and the abscissa ("complexity") has an arbitrary scale. Life appeared on Earth approximately 4,200 million years ago (mya), either arising as, or quickly streaming into, three main kingdoms — the eukarya, the bacteria and the archaea. Life remained unicellular, or at best colonial, for at least 3,000 million years. The common ancestor of the animals, plants and fungi is thought to have arisen approximately 1,200 mya, around the time at which the mitochondria entered the lineage through a rickettsial-like endosymbiont, an event that is postulated to have also brought with it type II self-splicing introns⁷⁵. Whether or not these events were coincidental, the incidence of introns (and other non-coding sequences) correlates with the complexity of the organism after that point. In the Cambrian period (~520 mya), complex animal life exploded in an event known as the metazoan radiation, in which recognizable ancestors of all modern phyla appeared only in a single strata of rock⁷⁸. What restrained the appearance of organized multicellular organisms for so long? Was it environmental or biochemical factors (such as oxygen tension and oxidative energy metabolism), or a primitive genetic operating system?

PRVNÍ GENOMY

První protein: RNA-dependentní RNA polymeráza (RNA replikáza)

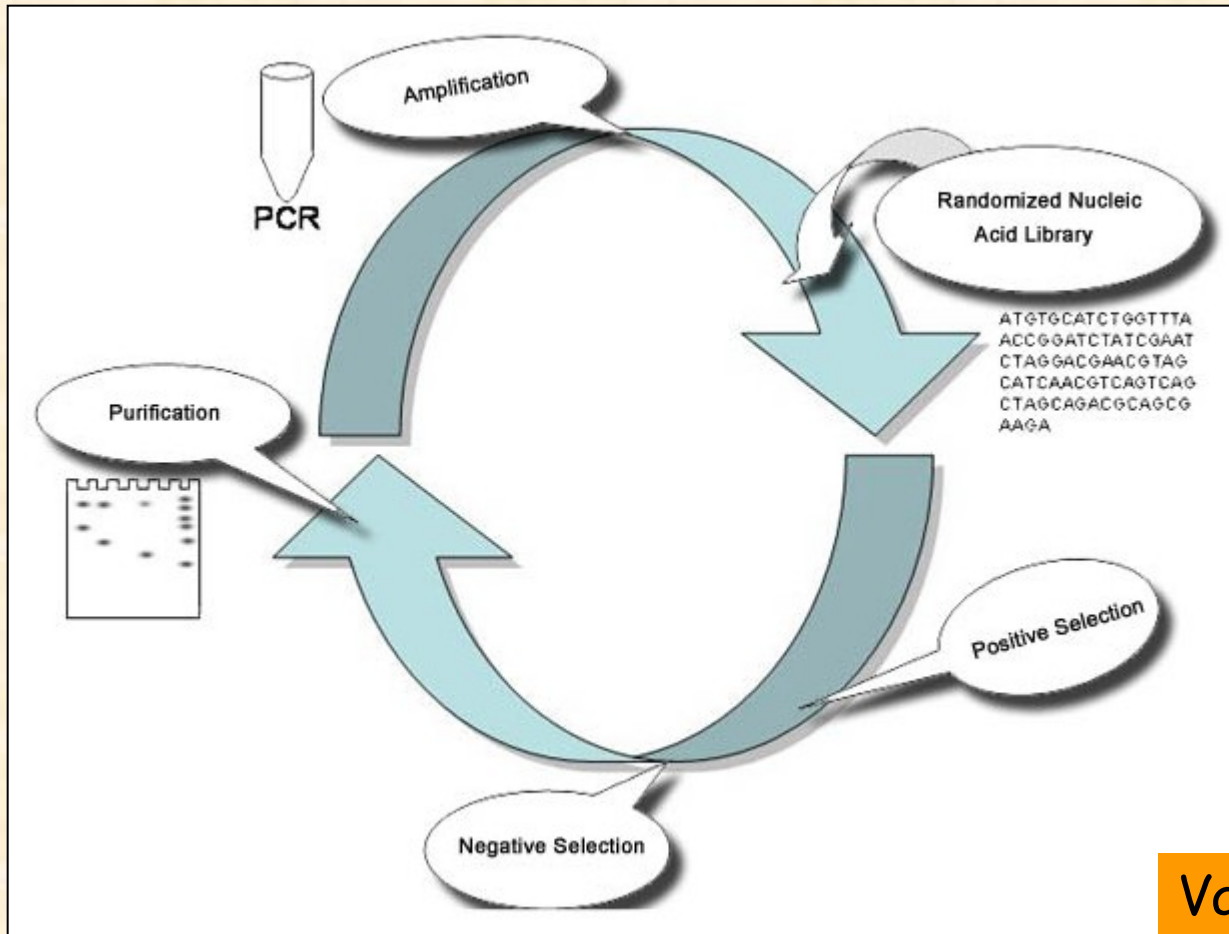
RNA → RNP → protein

Proteiny zvýšily účinnost ribozymů

- první geneticky kódovaný protein vznikl náhodou
- krátký peptid strukturně jednoduchý
- interagoval s RNA replikonem, zvyšoval jeho stabilitu či zlepšoval konformaci
- syntéza potomstva musí být **rychlejší** než degradace rodičů
- dostatečná **přesnost**, ale ne absolutní (možnost evoluce)

RNA polymeráza → Reverzní transkriptáza

Šlechtění RNA ve zkumavce



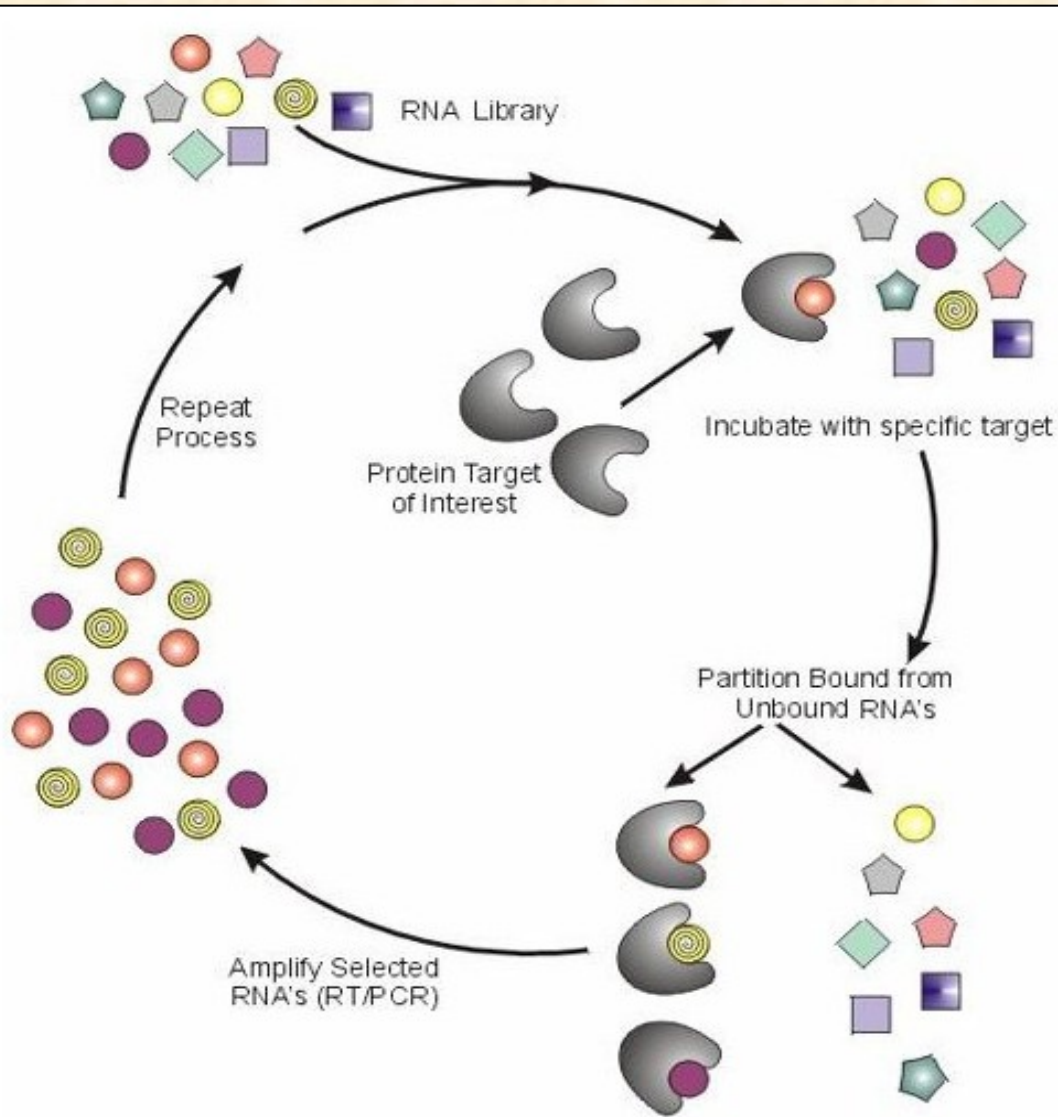
Fág Q β :

(Spiegelman 1967)
- zkracování času na replikaci --> zkracování RNA genomů (po 74 přenosech eliminace 83% genomů), i jiné selekční tlaky (jedy)

Tvorba nových proteinů nebo RNA nepřítomných v přírodě s požadovanými vlastnostmi

Variabilita/mutace
Selekce
Amplifikace

SELEX: „evoluce *in vitro*“ aneb co by teroristé neměli číst

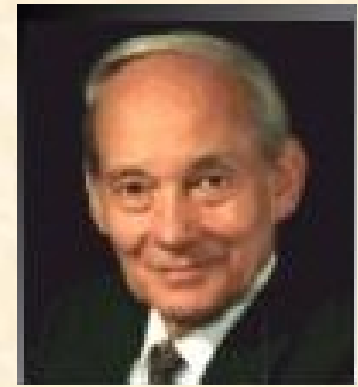


- viry pro genovou terapii
- mutace pláštěvého proteinu
- rezistence k protilátkám
- produkce supervirů

EVOLUCE:

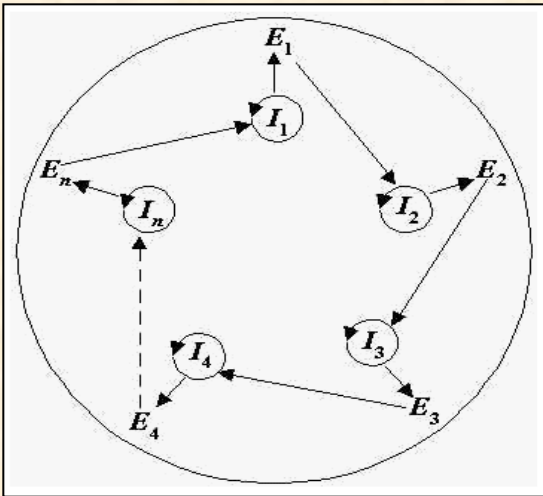
1. dědičnost
2. variabilita
3. selekce

Hypercykly aneb cesta k buňce



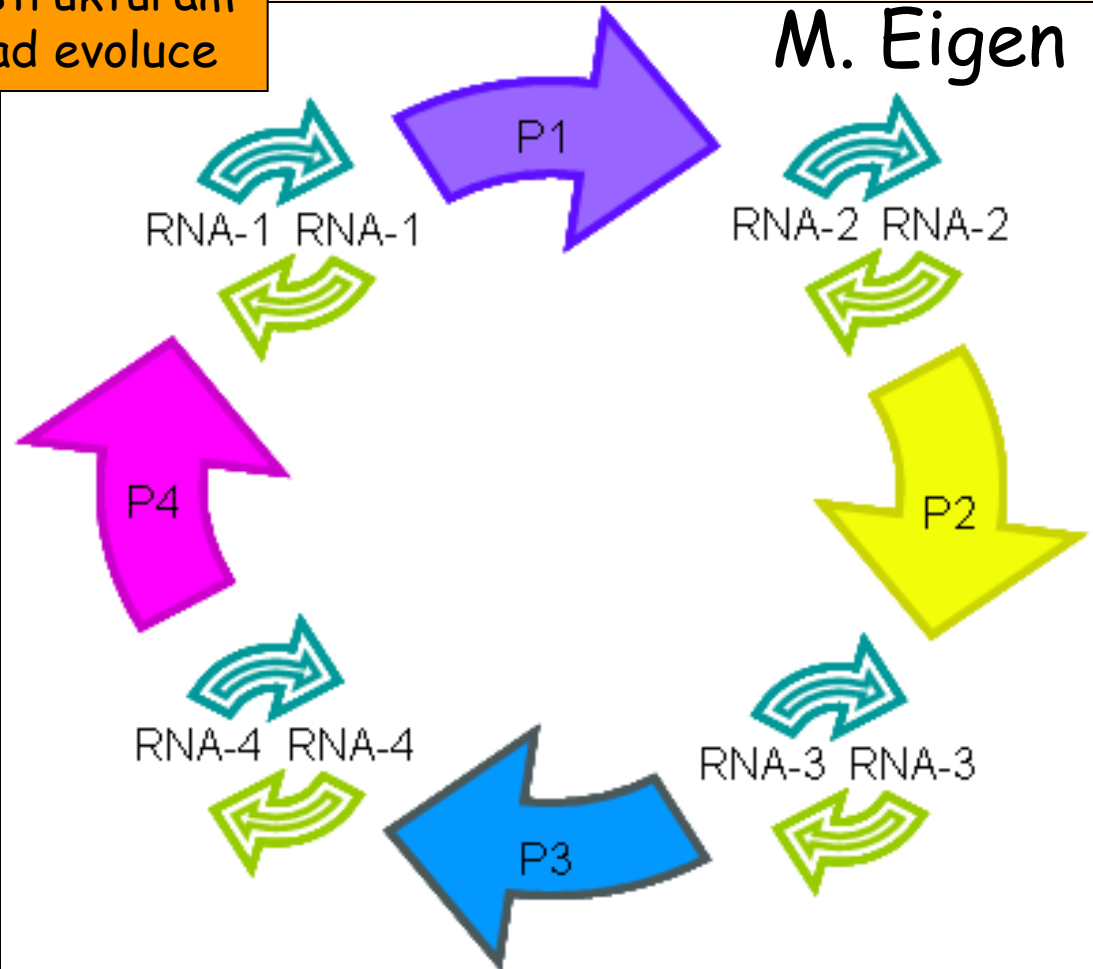
Původně jeden člen - duplikace a mutace - dva členy ...

Cesta od nukleových kyselin ke strukturám buněk = jedna z největších záhad evoluce



V určité fázi vývoje se objevily první **parazité** - zlodějské cykly. Přežily jen hypercykly schopné se bránit parazitům. Za vznikem buněk tedy možná stáli parazité (hybná síla evoluce)

M. Eigen



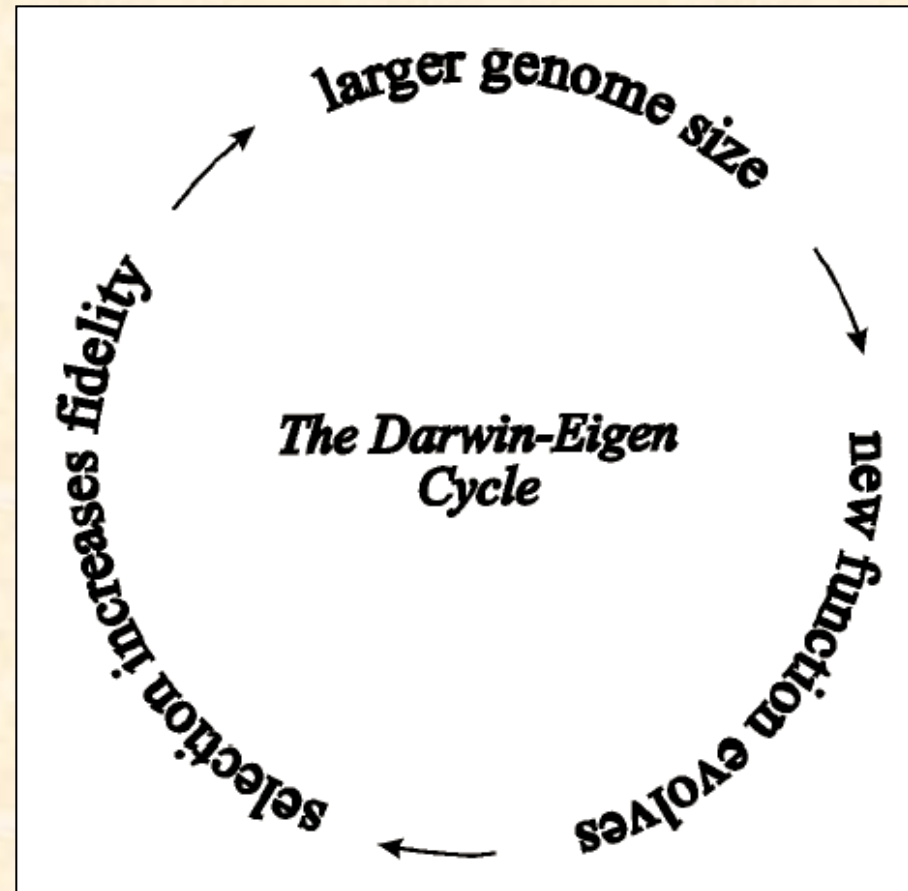
Eigenův limit: replikační přesnost je limitujícím faktorem

- **Definice:**

Čím je vyšší frekvence chyb při replikaci, tím menší genom může projít do další generace

- **Omezení katastrofických dopadů chyb replikace:**

- více kopií (ploidie)
- fragmentace genomu do chromosomů
- rekombinace



Dnešní viry: Funkční relikty časných replikonů?

Pohled na viry:

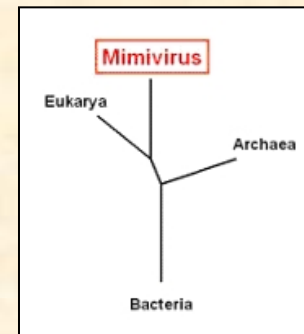
- (a) molekulární paraziti, **odvození** v důsledku způsobu života
- (b) **primitivní**, na hranici života podobně jako časné replikátory
- funkční relikty x funkční modely RNA-proteinových replikonů

RNA viry:

- minimální kódující kapacita (coronaviry 30kb)
- některé viry střídají fáze RNA a DNA - reminiscence RNA→DNA přechodu
- primerem replikace je tRNA

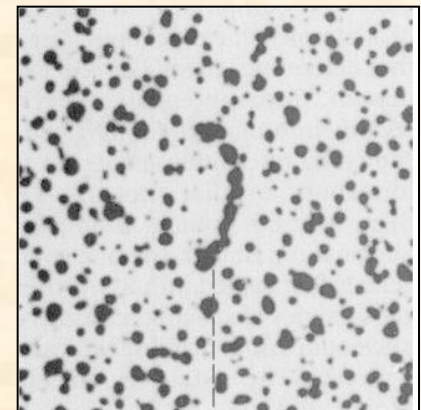
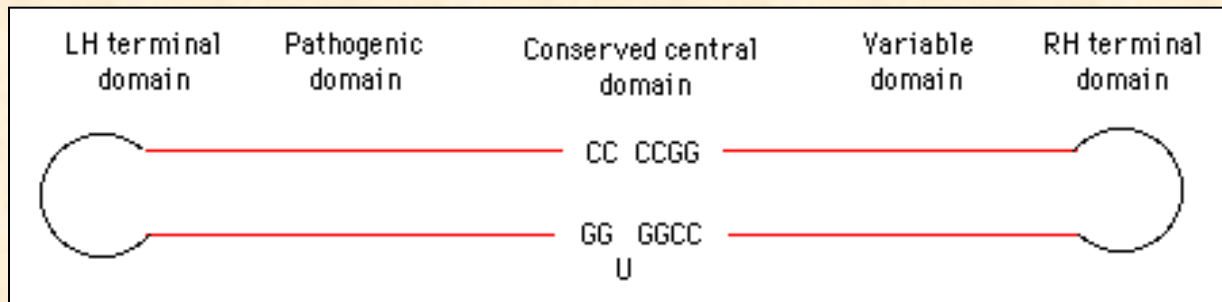
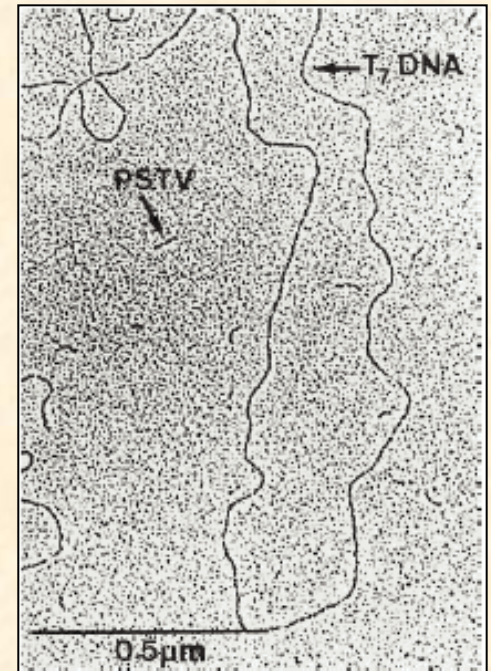
Mimiviry - hranice života:

- velikost genomu srovnatelná s prokaryoty (1.2Mb)
- metabolické geny (911 genů pro proteiny)
- 10% repetitivní DNA
- jen částečná závislost na hostiteli (proteosyntéza)



Viroidy: nejpodobnější časným replikonům

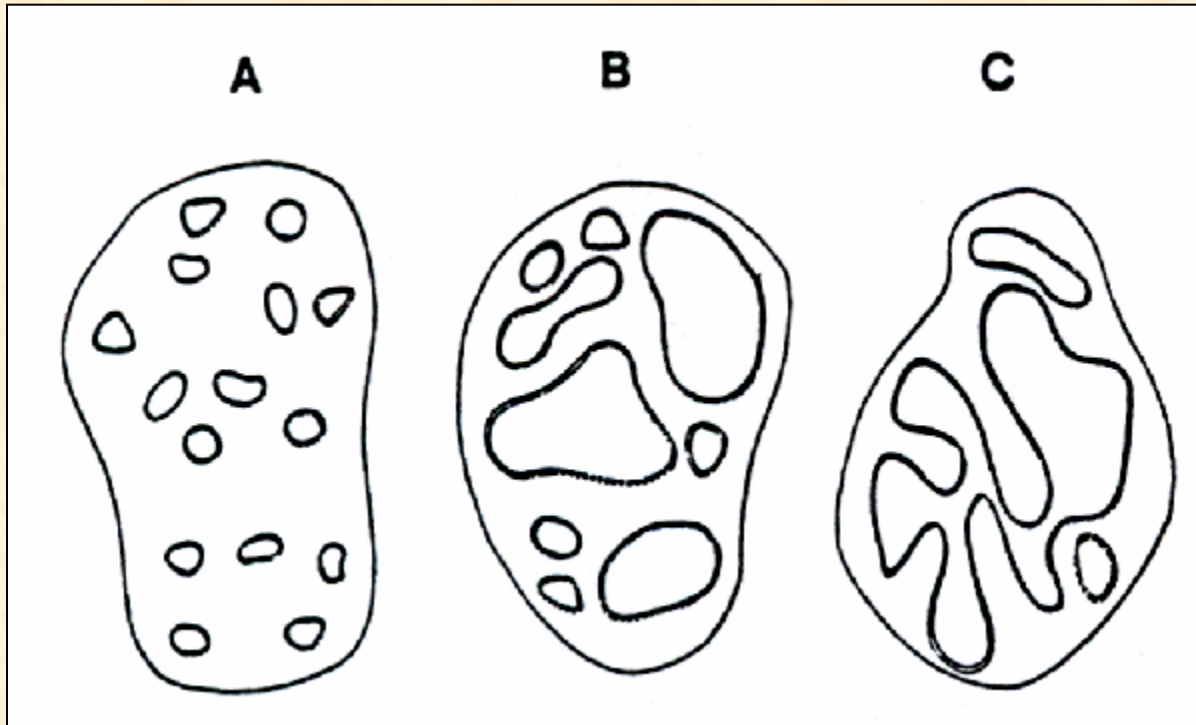
- patogeny rostlin
- 200 až 10 000 kopií na buňku
- malé RNA genomy (240-400 b), ssRNA, cirkulární,
- nekódují proteiny - jako replikony éry před proteiny
- replikovány hostitelskými RNA polymerázami
- rolling-circle mechanismus
- multimery štěpené autokatalytickými **ribozymovými** sekvencemi
- intenzivní **vnitřní párování bází** jejich genomické sekvence
- tvorba sekundárních struktur stabilizujících genomy



240 bp potato spindle-tuber viroid molecule magnified 440000 times.

První DNA genomy: vznik fúzováním malých kružnic DNA

- první genomy: lineární nebo cirkulární?
- malé kružnicové DNA genomy, disperzní genom
- fúzování, geny jako autonomní DNA
- počty kopií statisticky stejné - podobné přenosům plazmidů



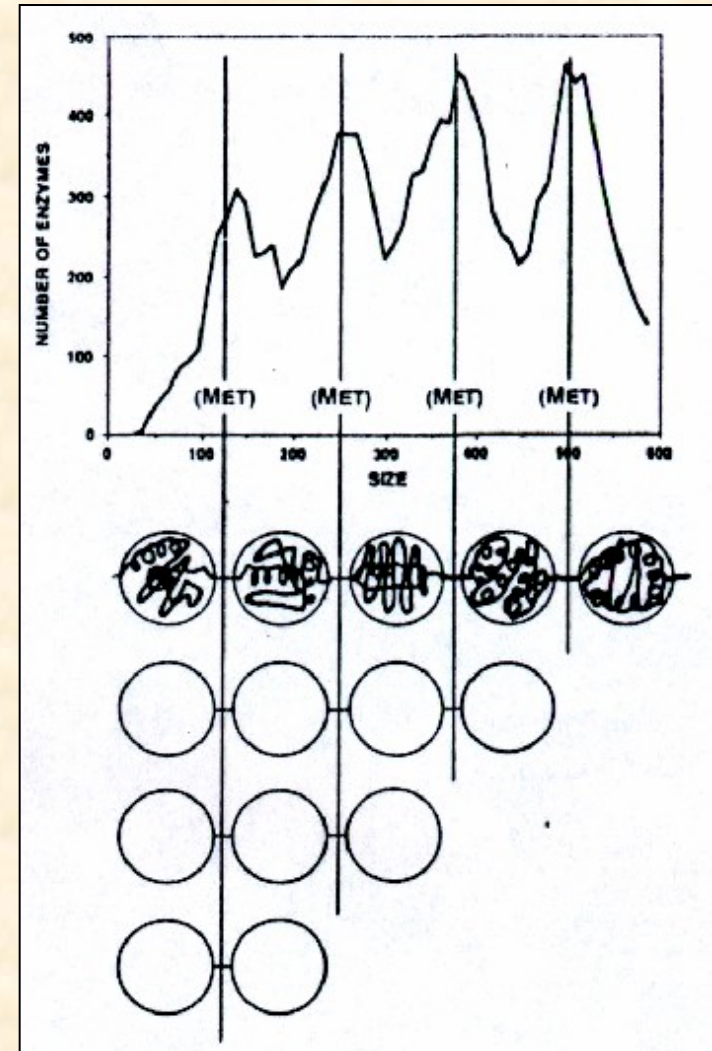
Fáze:

- A. pregenomická
- B. rekombinační
- C. genomická

První DNA genomy: vznik fúzováním malých kružnic DNA

Důkazy:

- **periodicita** délek proteinů
- nejsnazší **cirkularizace**
- periodicita výskytu **Met**
- pozůstatkem **extrachromosomální DNA**
- **mobilní** elementy, fágy, genomy organel
- replikace satelitů prostřednictvím eccDNA



Stromatolity: nejstarší stopy života

- zkamenělé útvary tvořené sinicemi (?)
- stáří až **3.8 miliardy let**
- anaerobní podmínky (archea) → činností sinic vznikl kyslík
- moderní stromatolity (hypersalinní jezera, Austrálie)

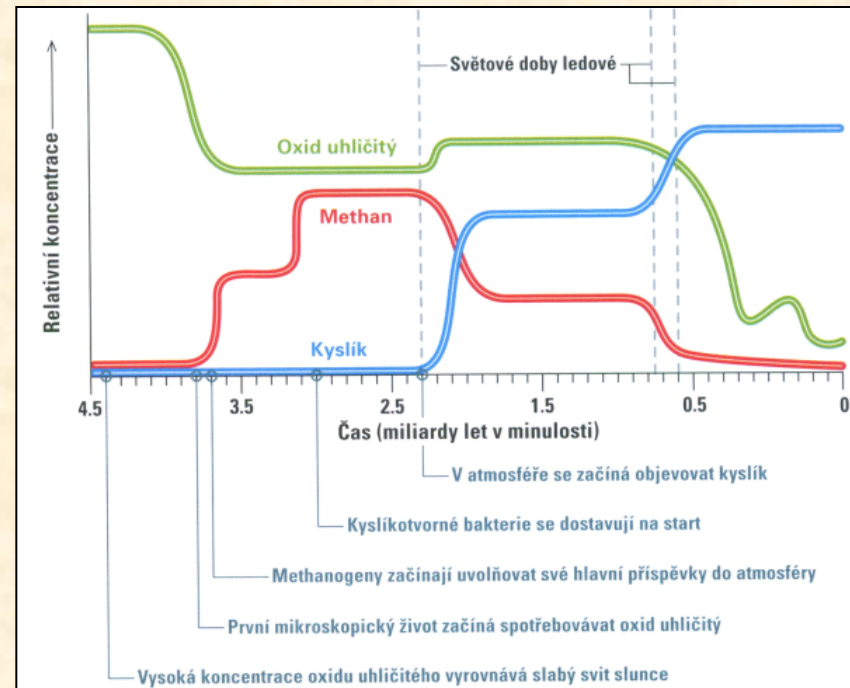
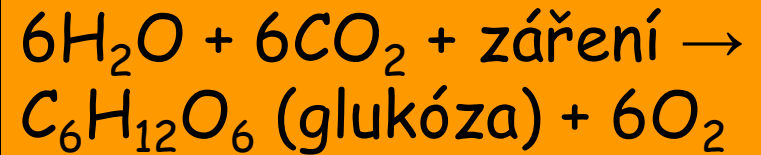


Photo P.-A. Bourque

Organizmy objevily fotosyntézu

- **fotosyntéza** = schopnost tvořit těla z vody a vzduchu za pomoci slunečního záření
- původní organizmy anaerobní, živily se organickými látkami v praprolévce
- **kyslík** jedovatý (i pro dnešní anaerobní organizmy, radikály), nahrazení CO_2 (skleníkový plyn) kyslíkem vedlo k ochlazení planety
- schopnost organismů ovlivňovat klima planety - „**Gaia** živoucí planeta“ (James Lovelock)

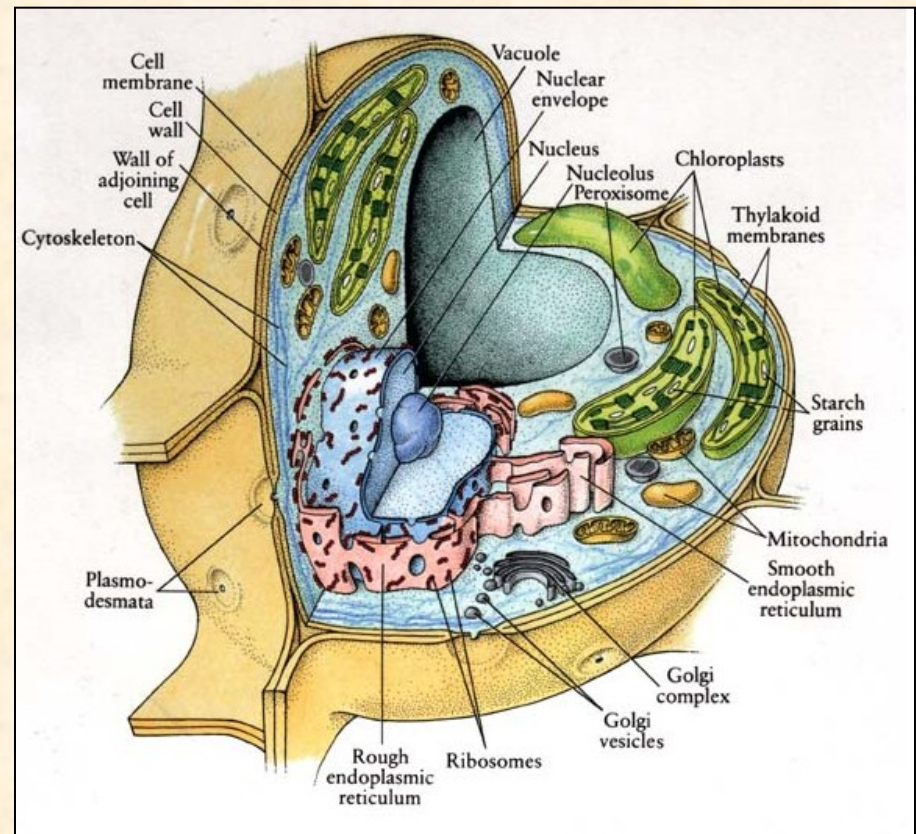
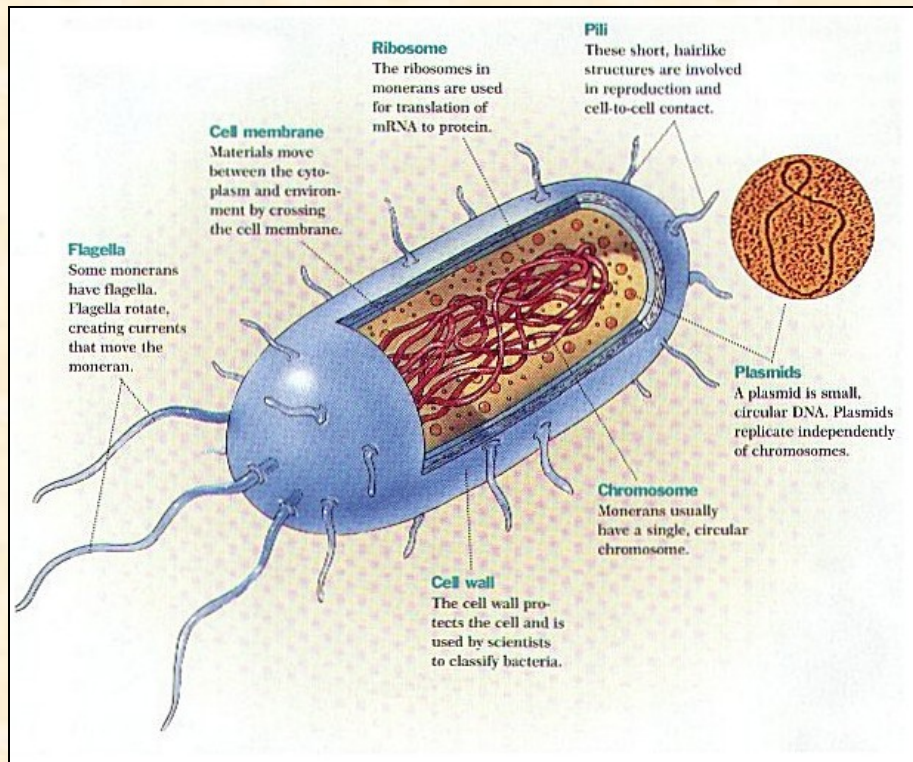
- **Symbiogeneze** (Lynn Margulis) - původně parazitismus? (J. Flegr)



POMĚRNÉ KONCENTRACE hlavních plynů ovzduší mohou vysvětlit výskyt celosvětových dob ledových v dávné minulosti Země. Methanogenní mikroorganismy na počátku vzkvétaly, ale když se před 2,3 miliardami let začal z hornin uvolňovat kyslík, zbylo pro tato stvoření jen několik málo útočišť, kde mohla přežít. Doprovodný pokles koncentrace methanu – silného skleníkového plynu – mohl ochladit celou planetu. Role oxidu uhličitého, nejvýznamnějšího skleníkového plynu v dnešní atmosféře, byla tehdy mnohem méně dramatická.

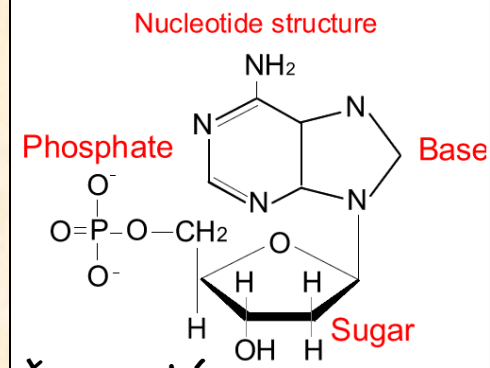
Definice genomu

- celková genetická informace organismu
- **prokaryota**: cirkulární chromosom + plazmidy
- **eukaryota**: chromosomy v jádře, mitochondrie a chloroplasty



KONEC

Počátky světa RNA



Ribóza: formozová reakce z formaldehydu (Butlerow 1861), směs pentóz a hexóz, nestabilita ribózy, Pb-katalýza, Ca-borátová stabilizace, alkalické pH

Puriny: syntéza z HCN a formamidu, adenin v chondritech

Pyrimidiny: reakcemi mezi molekulami kyanoacetylenu nebo kyanoacetaldehydu, vyšší výtěžky v přítomnosti močoviny, výtěžky lze zvýšit zmražováním nebo v podmínkách vysoké energie (hydrotermální prameny)

Nukleozid: vazba bází na cukr, beta-orientace, slabý bod hypotéz

Nukleotid: fosforylace nukleozidů v roztoku CaP (hydroxylapatitu, vysoká teplota, malá účinnost), syntéza polyfosfátů, význam trifosfátů.

Racemická směs nukleotidů: alfa a beta a L a D-izoformy

Polymerizace: pomalá a spontánní, nespecifická, racemická směs, pouze 5-3 fosfodiesterová vazba mezi beta-D-nukleotidy byla funkční a byla prodlužována při nízkých teplotách (led) katalyzováno ionty kovů, v hydrotermálních pramenech

Starobylé struktury:

Ribozóm - proteosyntéza, účast rRNA, tRNA a mnoha proteinů (původně **replikace**)

Spliceozóm - sestřih pre-mRNA, účast snRNA a mnoha proteinů (původně **rekombinace**)

Snorpozóm - sestřih pre-rRNA, účast snoRNA a mnoha proteinů