

Umělý život

Půjde vytvořit?
Jsme blízko? Nebo daleko?
Měli bychom?

Jan Pačes

`jan.paces@img.cas.cz`

Ústav molekulární genetiky AVČR
<http://www.img.cas.cz>



CZECH FOBIA

Některé důležité termíny

- Genomika: čtení, analýza a studium genetické informace
- Bioinformatika: věda zejména o tom jak se informace ukládá a šíří v živé přírodě. „Biologie v počítači“
- DNA: deoxyribonukleová kyselina (bp, base pairs)
- Proteiny (aa, aminokyseliny, Daltony)
- Umělý život <-> umělá inteligence

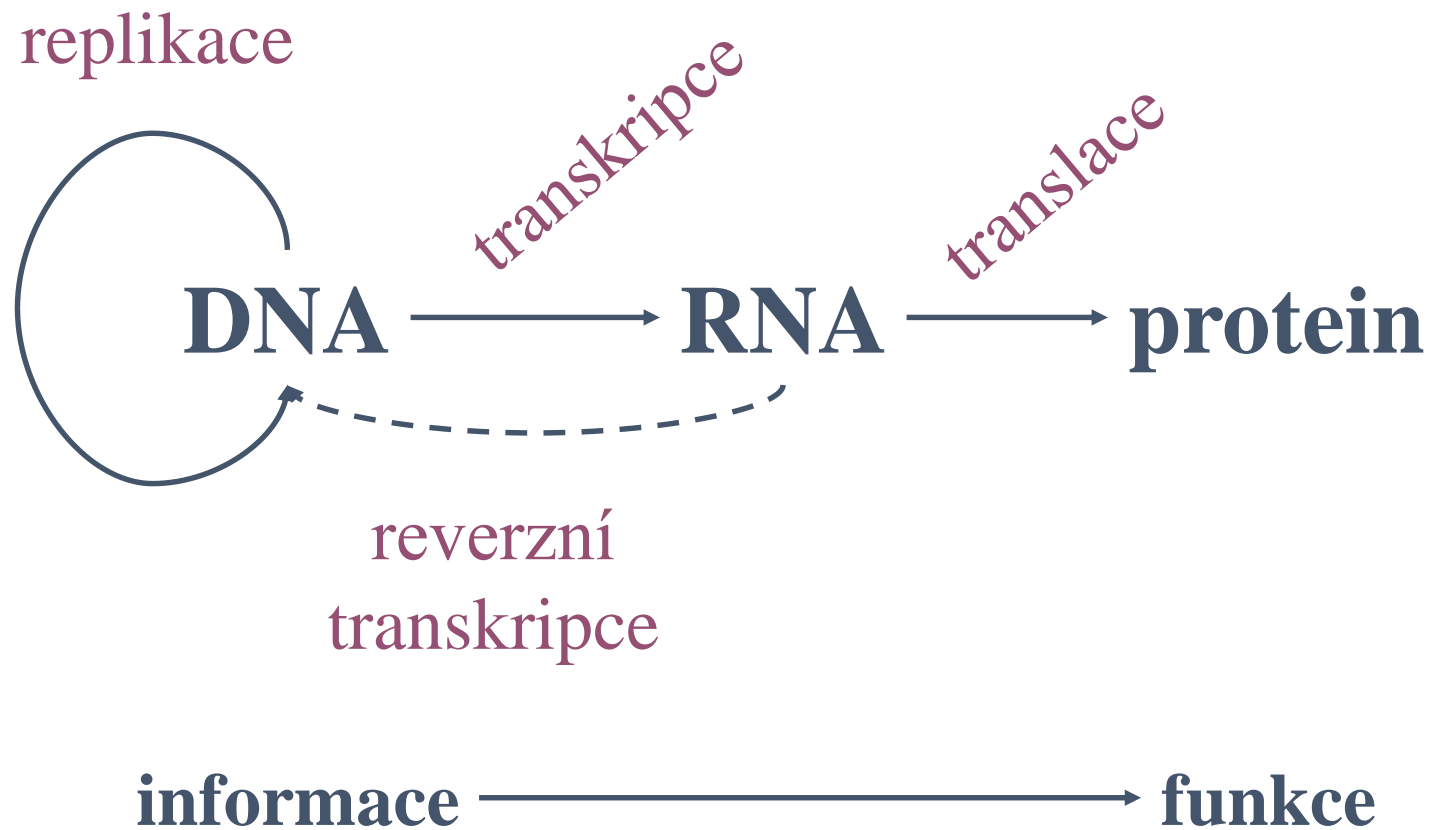
Definice života

- Musí jednoznačně rozlišit živé jedince od neživých individualit nebo systémů.
- Měla by definovat život nejenom jak ho známe ze Země v nynější podobě, ale i v časech jeho vzniku, případně i život mimozemský.
- Život není věc nebo stav, je to proces.
- NASA: Život je (soběstačný) systém schopný darwinistické evoluce

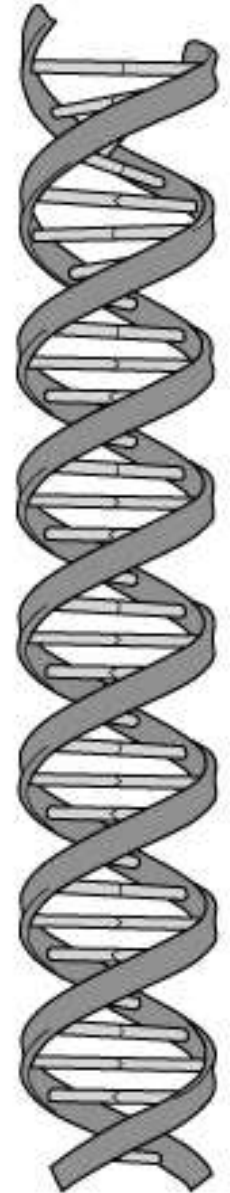
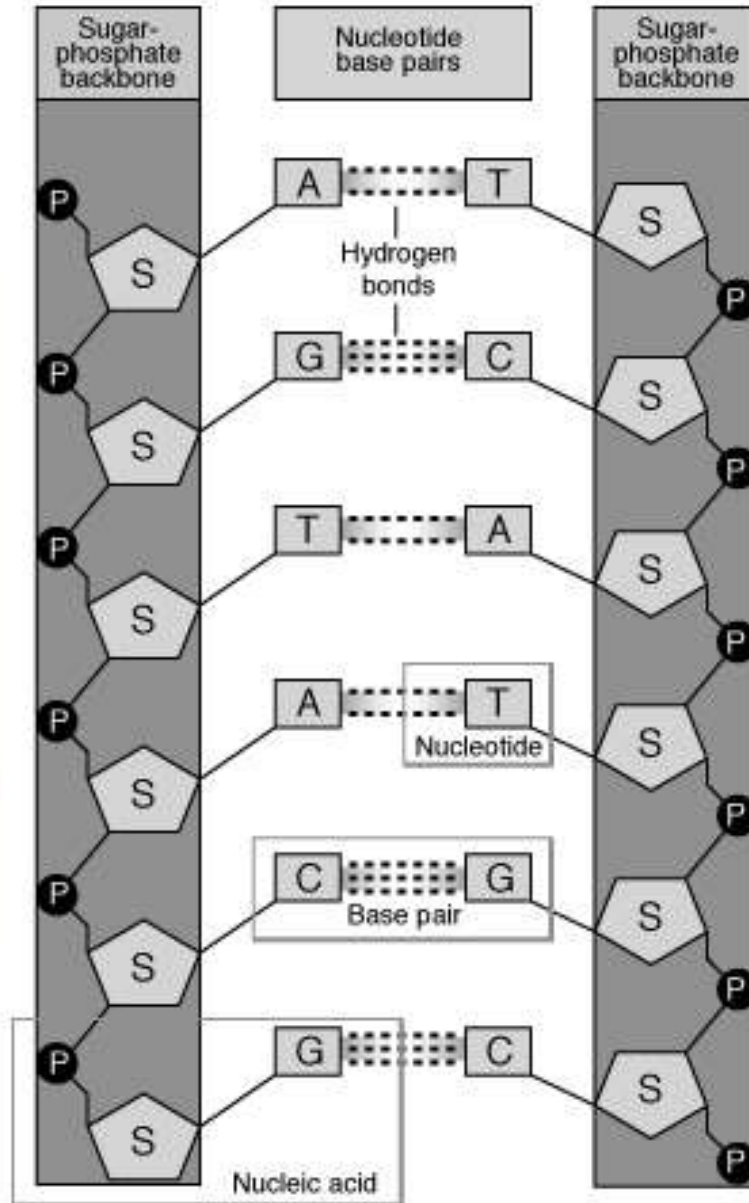
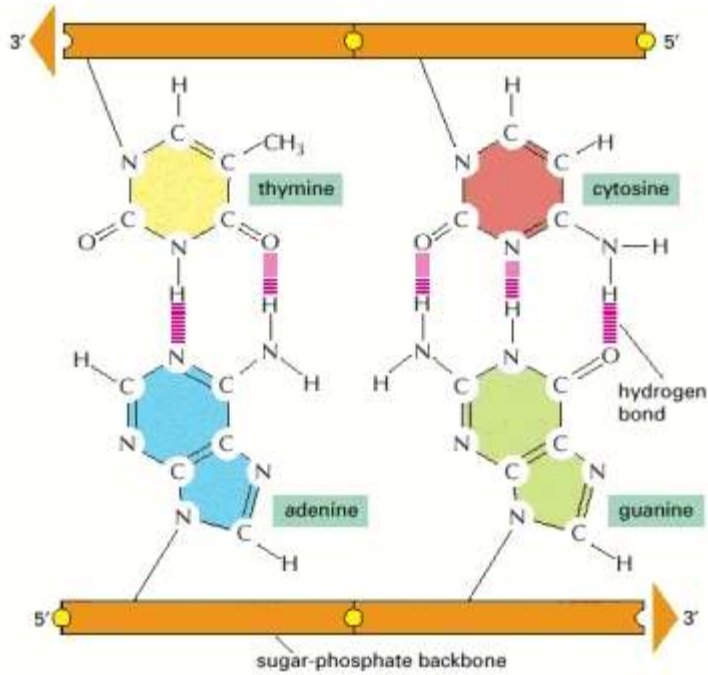
Znaky života

- Homeostáza / Metabolismus
- Růst / Rozmnožování
- Čivost
- Adaptace / schopnost evoluce
 - Ohraničenost v prostoru i čase / otevřenost
 - Přítomnost nukleových kyselin a proteinů

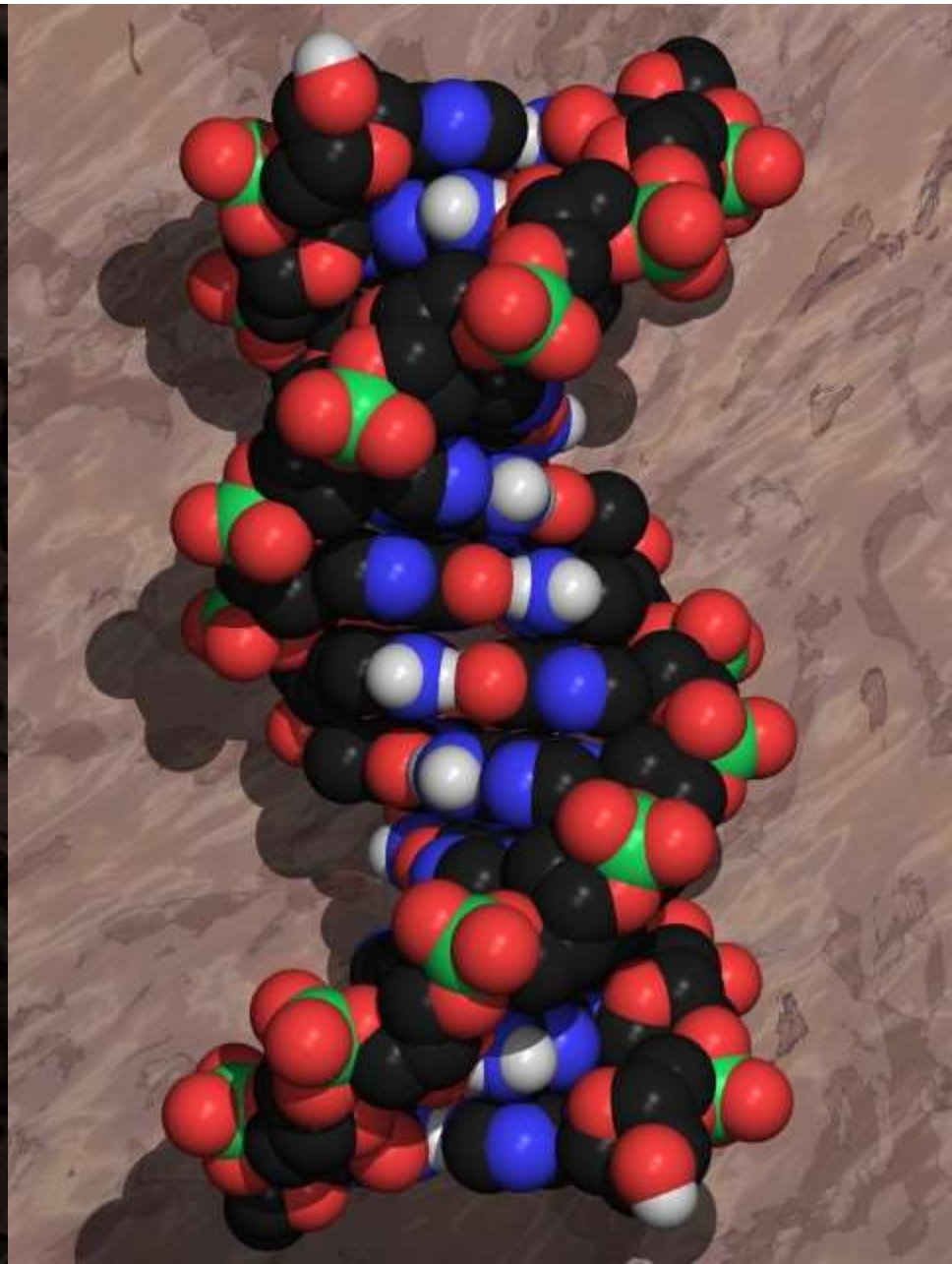
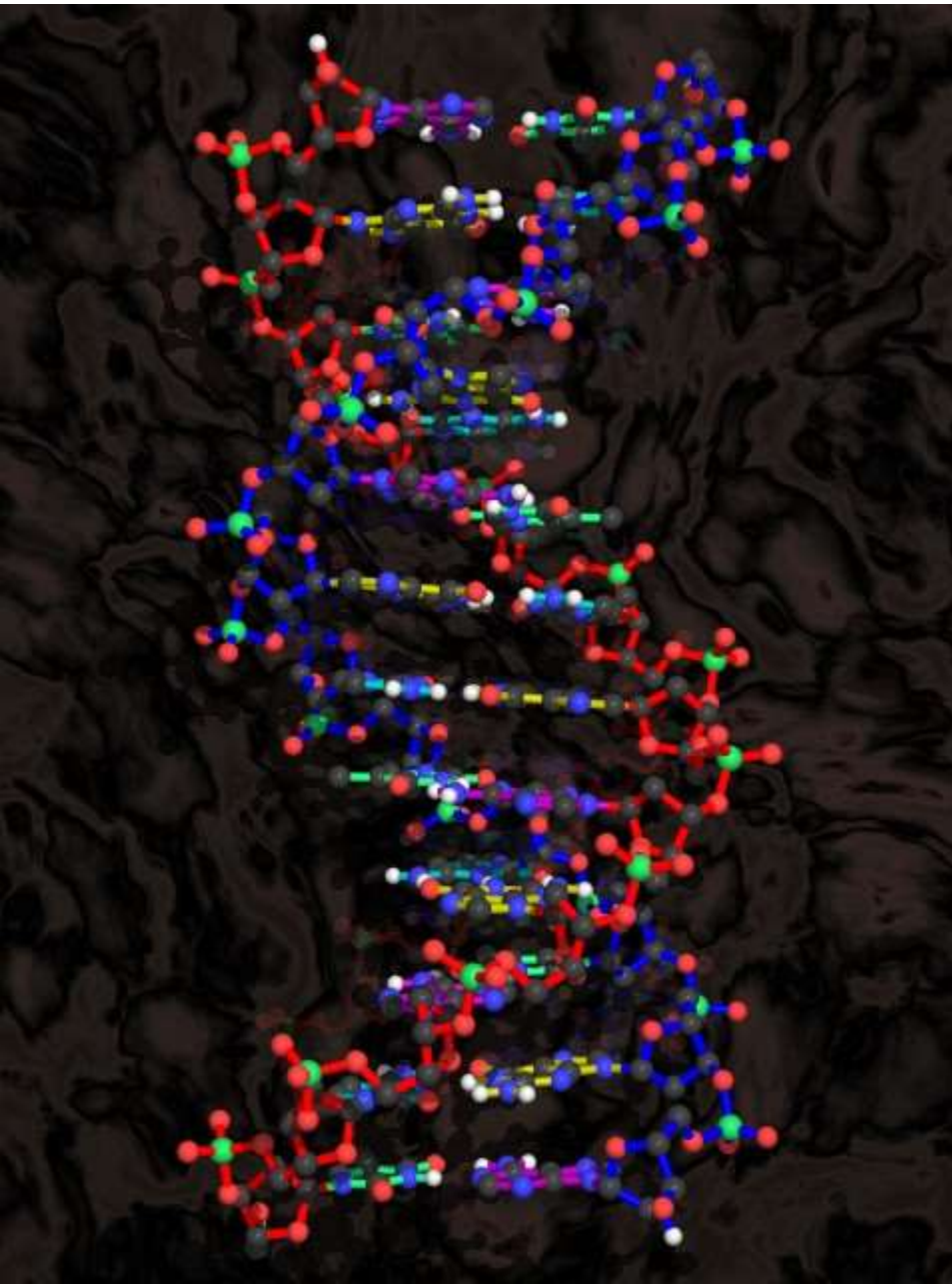
Centrální dogma molekulární genetiky



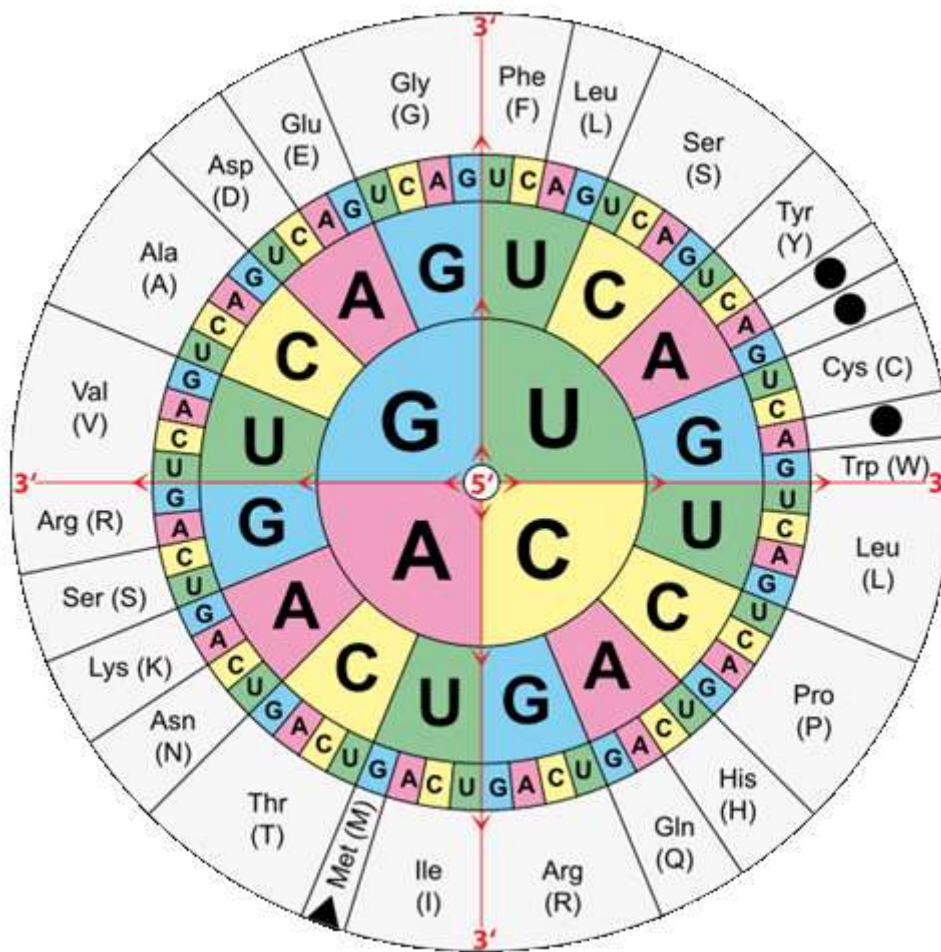
DNA



DNA



Genetický kód



Transkripce a translace

DNA 5' > ATGAAGCCGAGTCAT 3'
 3' TACTTCGGCTCAGTA <5'

transkripce

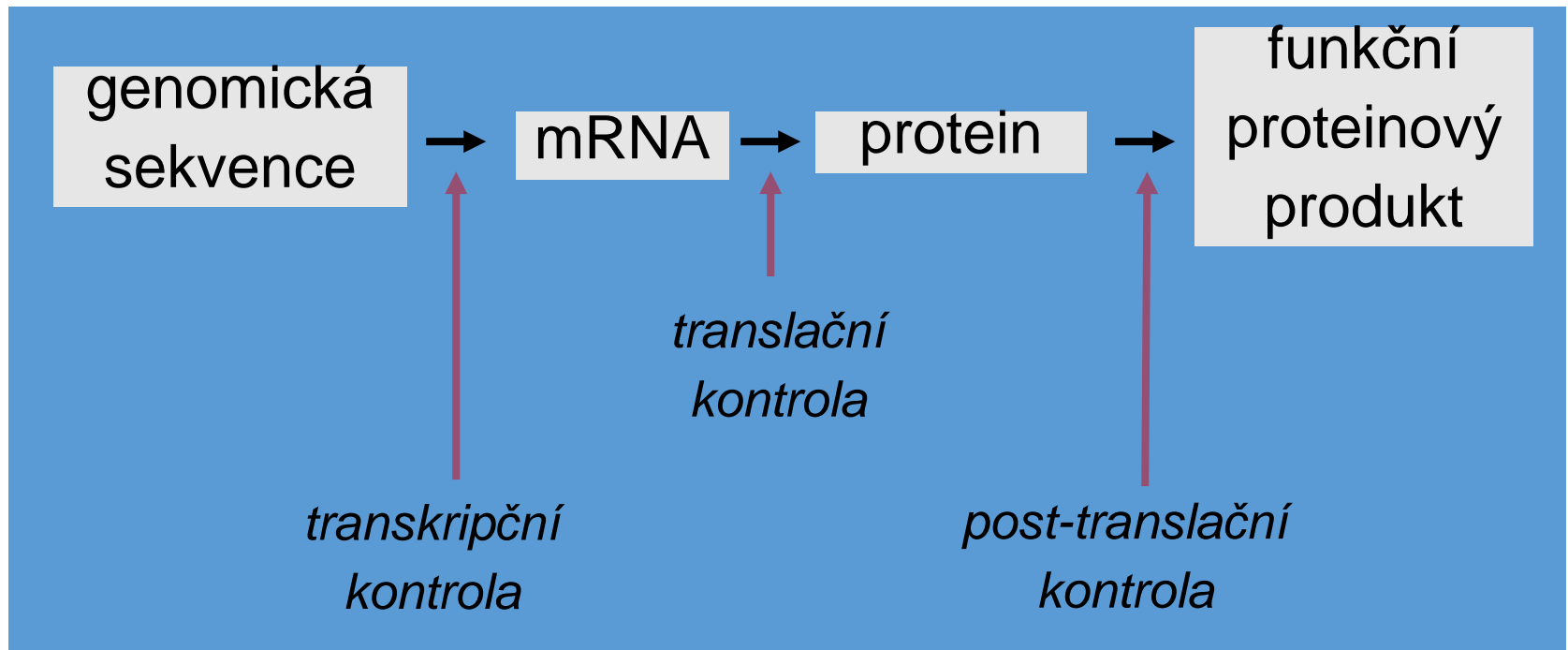
mRNA 5' > AUGAAGCCGAGUCAU 3'

translace








Protein N> MetLysProSerHis C

Úrovně kontroly

počet genů \neq úroveň mRNA \neq úroveň genové exprese \neq
množství a efektivní účinnost proteinu



Velikosti vybraných genomů

	genom	Chromozómy	velikost	geny
	<i>Mycoplasma genitalium</i>		0.58 Mbp	521
	<i>Escherichia coli</i>		4.6 Mbp (5.4 Mbp)	4 377 (5 416)
	<i>Saccharomyces cerevisiae</i>	16	12.5 Mbp	5 770
	<i>Caenorhabditis elegans</i>	6	~100 Mbp	19 427
	<i>Arabidopsis thaliana</i>	5	~115 Mbp	~28 k
	<i>Drosophila melanogaster</i>	5	~122 Mbp	13 379
	<i>Homo sapiens</i>	24	~ 3.3 Gbp	~22.5 k

Minimální genom (život)

- *Mycoplasma genitalium* G37 (nejmenší organizmus který dokážeme kultivovat) má 475 genů.
- Postupným odstraňováním genů se povedlo snížit minimální počet genů na 382.
- *Pelagibacter ubique* (nejmenší volně žijící buňka) má genom velikosti 1 308 759 bp, kódující 1 354 genů.
- Bakteriofág HTVC010P je pravděpodobně nejběžnější organizmus na zemi.
- Další nitrobuněční parazité, např *Candidatus* má jenom 169 genů.
- Simulace v počítači navrhuje, že pro funkční buňku by mohlo stačit 127 genů.

Hranice živého

- Spíše živé:
 - Chlamydie / Rickettsie (nitrobuněční parazité, mají buňku)
- Podle druhu definice:
 - Gaia, biosféra
- Spíše neživé
 - Viry, viriony
 - Transpozóny
 - Priony

„Typy“ umělého života

- **Hard:** elektromechanické, záměna biologických komponent za “technické”
- **Soft:** v počítači jako programy, které mají vlastnosti živých organismů. Simulace biologických komponent.
- **Wet:** biochemické, sestavení funkčního organismu z abiotických prvků, vnášení nových prvků do existujících živých organismů

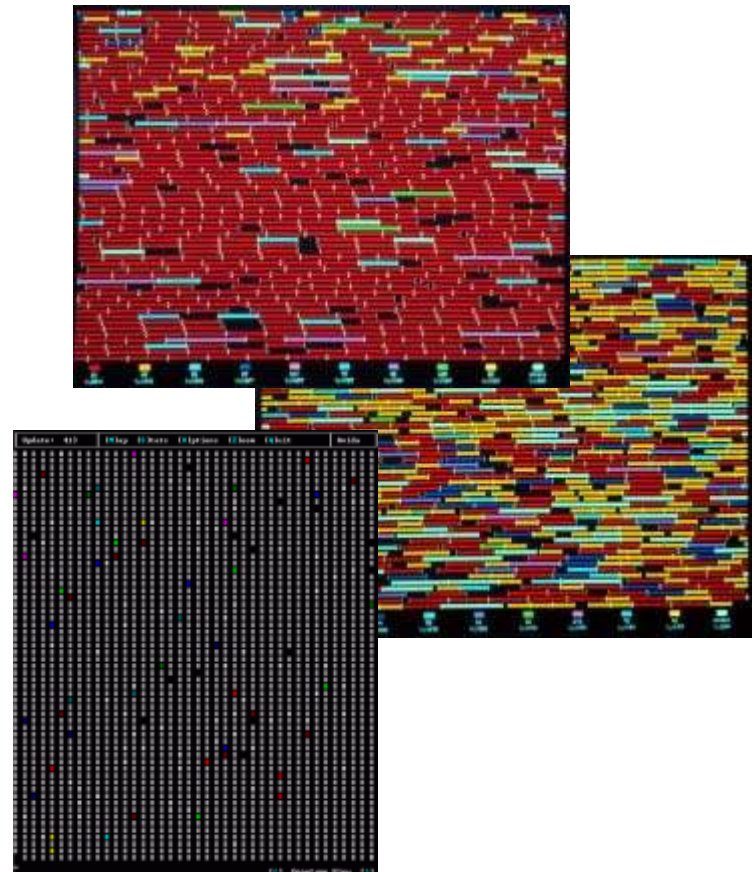
"life is a process which can be abstracted away from any particular medium" John von Neumann

Software

Digitální organizmus – samo se replikující program který mutuje a vyvíjí se

Tierra: program na simulaci evoluce v počítači. Jednotlivé programy kompetují o zdroje (paměť a CPU)

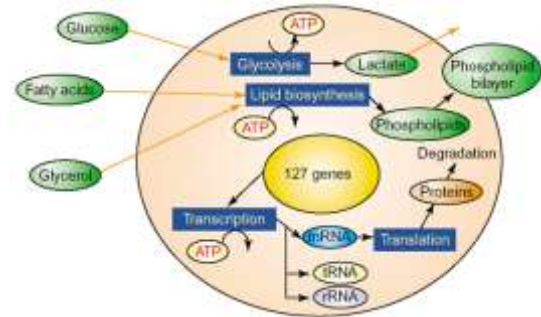
Avida a další



E-cell

<http://www.e-cell.org/>

E-cell projekt vytváří potřebné technologie a teoretický základ pro výpočetní biologii s cílem vytvořit přesnou simulaci celé buňky na molekulární úrovni.



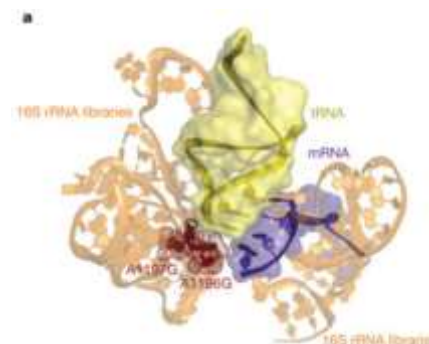
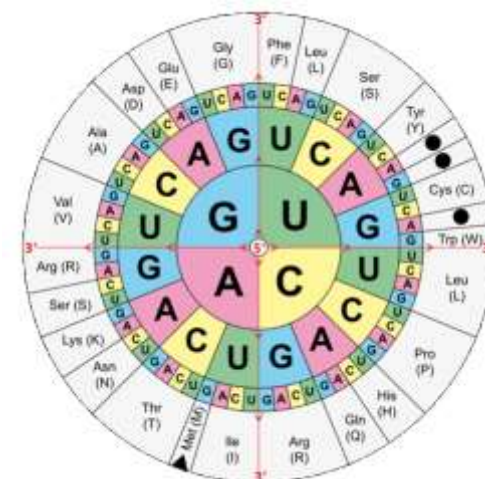
| E-CELL

Možné modifikace živého

- Modifikace stávajících organismů, proteinů
 - GMO (geneticky modifikované organizmy): přidání, odebrání nebo změna genetické informace
- Nové proteiny ze stávajících komponent (4 nukleotidy, 20 aminokyselin)
- Změna procesování genetické informace
 - Rozšíření počtu aminokyselin přeprogramováním současných kodónů
 - Principiální změna čtení DNA
 - Rozšíření informační kapacity DNA pomocí nových párů bazí

Změna genetického kódu

- tRNA s novou aminokyselinou
 - například tRNA pro vzácný AGG kodón pro aminokyselinu arginin byla modifikována, aby kódovala 6-N-allyloxycarbonyl-lysine.
- Modifikace ribozómů
 - o-ribozóm, používající jiné tRNA, jinou vazbu na mRNA
 - Ribo-Q, čte DNA po čtveřicích (256 kodónů místo 64)



Funkční příklad v myši

Stabilní integrace transgenů kódujících upravenou N^ϵ -acetyl-lysyl-tRNA syntházu a odpovídající tRNA (AcKRS)/tRNA^{PyI} do myšího genomu umožnila inkorporaci nepřírozené aminokyseliny do cílového proteinu na místě „amber“ stop kodónu.

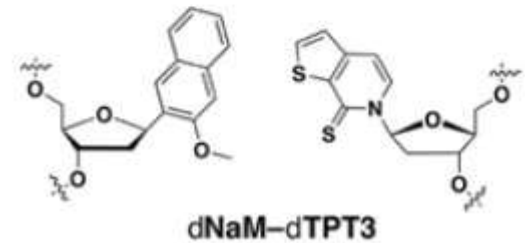
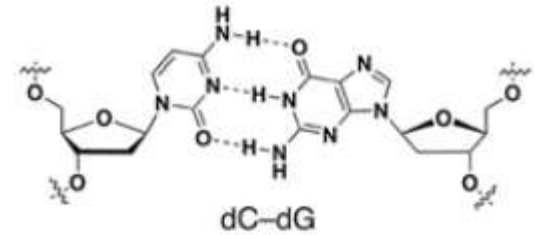
- Han, S.; Yang, A.; Lee, S.; Lee, H.W.; Park, C.B.; Park, H.S. (2017). "Expanding the genetic code of *Mus musculus*". *Nat. Commun.* 8: 14568.

Nová písmena genetické abecedy

UBP (unnatural base pairing)

XNA (xeno nucleic acid)

- **dNaM**: 3-methoxy-2-naphthyl
- **d5SICS**: 6-methylisoquinoline-1-thione-2-yl
- **dTPT3**



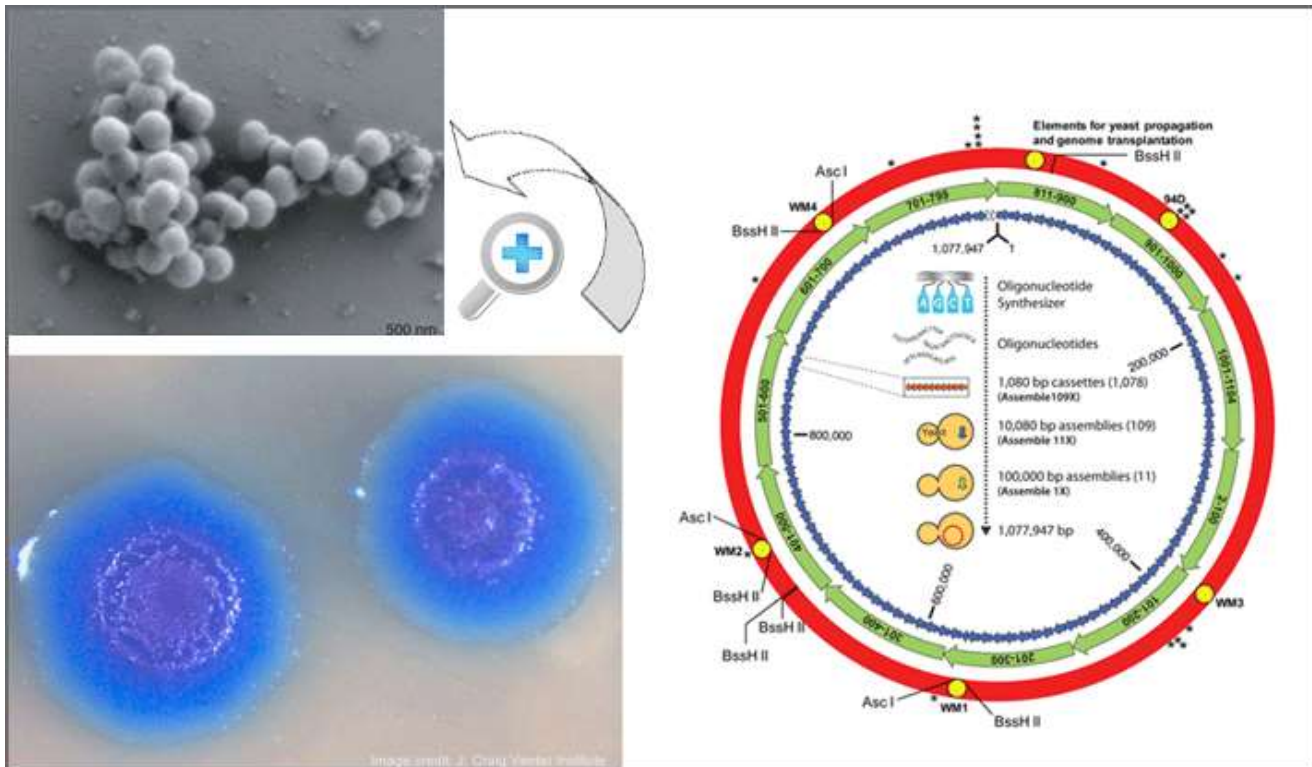
Organizmy s expandovanou DNA

- A semi-synthetic organism with an expanded genetic alphabet. Denis A. Malyshev, Kirandeep Dhama, Thomas Lavergne, Tingjian Chen, Nan Dai, Jeremy M. Foster, Ivan R. Corrêa & Floyd E. Romesberg. *Nature* vol 509, pp 385–388. 15 May 2014.
- A semi-synthetic organism that stores and retrieves increased genetic information. Yorke Zhang, Jerod L. Ptacin, Emil C. Fischer, Hans R. Aerni, Carolina E. Caffaro, Kristine San Jose, Aaron W. Feldman, Court R. Turner, and Floyd E. Romesberg. *Nature* vol 551, pp: 644–647. 30 November 2017.



Mycoplasma laboratorium

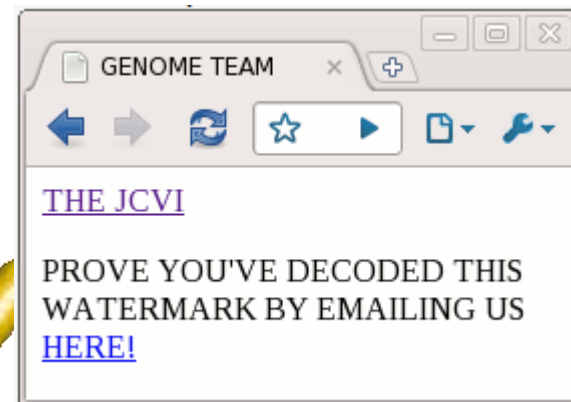
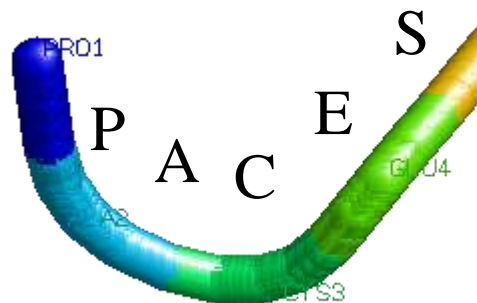
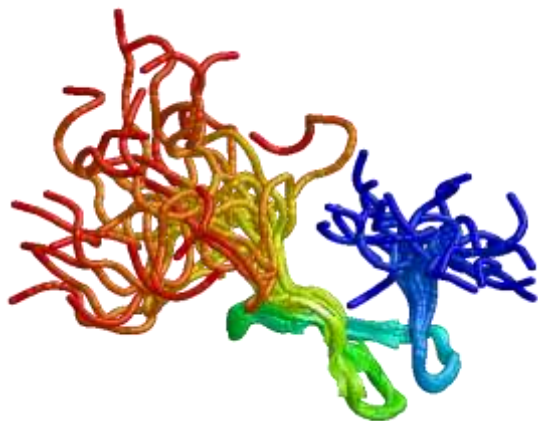
Gibson D, et al. (2008): Complete Chemical Synthesis, Assembly, and Cloning of a *Mycoplasma genitalium* Genome. Science. DOI: 10.1126/science.1151721



Ukládání informace v DNA

Watermarks:






- VENTERINSTITVTE CRAIGVENTER
HAMSMITH CINDIANDCLYDE
GLASSANDCLYDE
- Html code in *synthia*



Ukládání informace v DNA

STORAGE LIMITS

Estimates based on bacterial genetics suggest that digital DNA could one day rival or exceed today's storage technology.

	 Hard disk	 Flash memory	 Bacterial DNA	WEIGHT OF DNA NEEDED TO STORE WORLD'S DATA   ~1 kg
Read-write speed (μs per bit) >	~3,000–5,000	~100	<100	
Data retention (years) >	>10	>10	>100	
Power usage (watts per gigabyte) >	~0.04	~0.01–0.04	< 10^{-10}	
Data density (bits per cm^3) >	~ 10^{13}	~ 10^{16}	~ 10^{19}	©nature

Zakódování Shakespeareových sonetů do DNA

Thou art more lovely ...

text do ASCII

0101010101000111000101001 ...

ASCII do „trits“ (0,1,2)

20112 20200 02110 10002 ...

„trits“ do DNA

TAGAT GTGTA CAGAC TAGCG ...

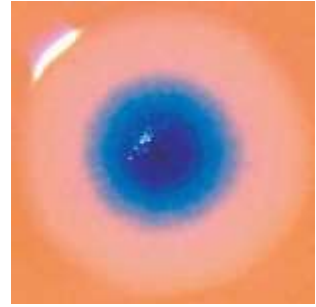
aby se každé písmenko lišilo od předcházejícího



DNA fragmenty

překrývající se a s unikátním indexem

Synthia



Synthia 1.0: 1. syntetický organizmus (901 genů)

Gibson D, et al. (2010): Creation of a bacterial cell controlled by a chemically synthesized genome. Science. DOI: [10.1126/science.1190719](https://doi.org/10.1126/science.1190719)

Synthia 3.0 - 473 genů, funkce 149 z nich je stále neznámá.

Umělá kvasinka



2014 kompletní chromozom II

- Annaluru N, Muller H, Mitchell LA, Ramalingam S, et al. 2014. Total synthesis of a functional designer eukaryotic chromosome. *Science* 344: 55–8.

2017: Sc2.0: 6 umělých chromozomů (z 12)

Metoda SCRaMbLE (*loxP*sym + *Cre* rekombináza za každým *genem*) umožňuje „míchání“ genomu

HGP-write (Human Genome Project-write)

- 2016-2026
- Chemicky syntetizovat celý lidský genom
- Ultra bezpečné buňky rezistentní k virovým infekcím (*NEW YORK, May 1, 2018 – The Center for Excellence in Engineering Biology and the leadership of Genome Project-write (GP-write) today announced its first grand-scale community-wide project, to develop “ultra-safe cells” that resist natural viruses and potentially radiation, freezing, aging and cancer.*)

Otevřené problémy

- Přejchod od neživého k živému (vznik života).
- Mohou existovat principiálně odlišné formy života?
- Etické principy pro umělý život.
- Existují nevyhnutelné události nebo body v evoluci živého? Například vznik inteligence, vědomí.

Děkuji za pozornost

Kam dál?

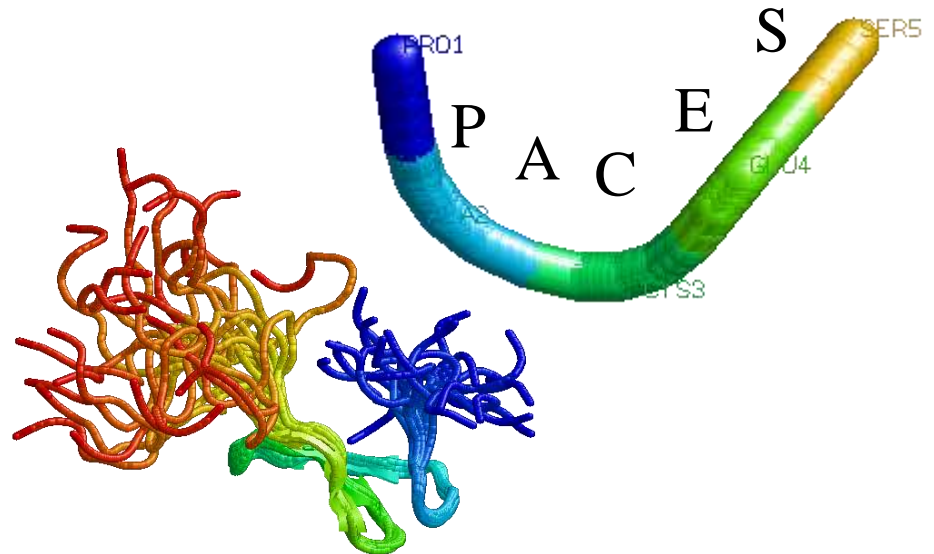
- International Society for Artificial Life
<http://www.alife.org/>
- Artificial Life (The MIT Press Journals) Impact Factor:
1.316 <http://www.mitpressjournals.org/loi/artl>

Jan Pačes

www: <http://bio.img.cas.cz>

email: hpaces@img.cas.cz

icq: #110872370

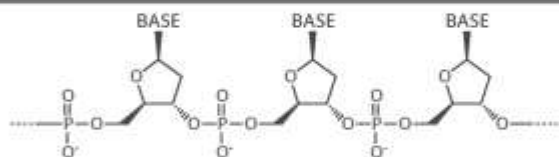




THE CHEMICAL STRUCTURE OF DNA

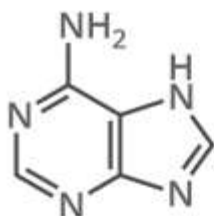
DNA (deoxyribonucleic acid) carries genetic information in all multicellular forms of life. It carries instructions for the creation of proteins, which carry out a wide range of roles in the body.

THE SUGAR PHOSPHATE 'BACKBONE'

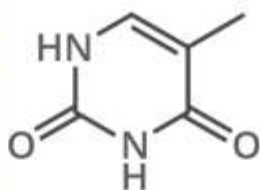


DNA is a polymer made up of units called nucleotides. The nucleotides are made of three different components: a sugar group, a phosphate group, and a base. There are four different bases: adenine, thymine, guanine & cytosine.

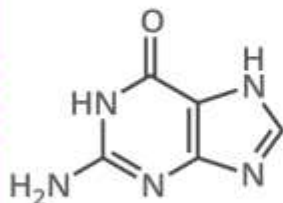
A ADENINE



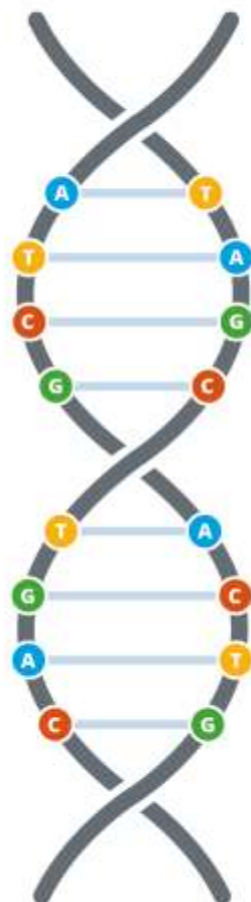
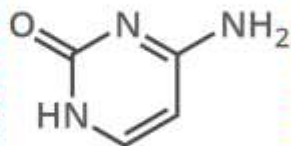
T THYMINE



G GUANINE

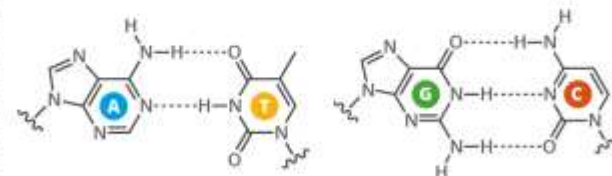


C CYTOSINE



WHAT HOLDS DNA STRANDS TOGETHER?

DNA strands are held together by hydrogen bonds between bases on adjacent strands. Adenine (A) always pairs with thymine (T), whilst guanine (G) always pairs with cytosine (C).



FROM DNA TO PROTEINS



The bases along a single strand of DNA act as a code. The letters form three letter 'words', or codons, which code for different amino acids - the building blocks of proteins.

An enzyme, RNA polymerase, transcribes DNA into mRNA (messenger ribonucleic acid). It does this by splitting apart the two strands that form the double helix, then reading a strand and copying the sequence of nucleotides. The only difference between the RNA and the original DNA is that in the place of thymine (T), another base with a similar structure is used: uracil (U).

DNA SEQUENCE	T	T	C	C	T	G	A	A	C	C	C	G	T	T	A
mRNA SEQUENCE	U	U	C	C	U	G	A	A	C	C	C	G	U	U	A
AMINO ACID	Phenylalanine		Leucine		Asparagine		Proline			Leucine					

In multicellular organisms, the mRNA carries genetic code out of the nucleus, to the cell's cytoplasm. Here, protein synthesis takes place. 'Translation' is the process of converting turning the mRNA's 'code' into proteins. Molecules called ribosomes carry out this process, building up proteins from the amino acids coded for.

Děkuji za pozornost

Jan Pačes

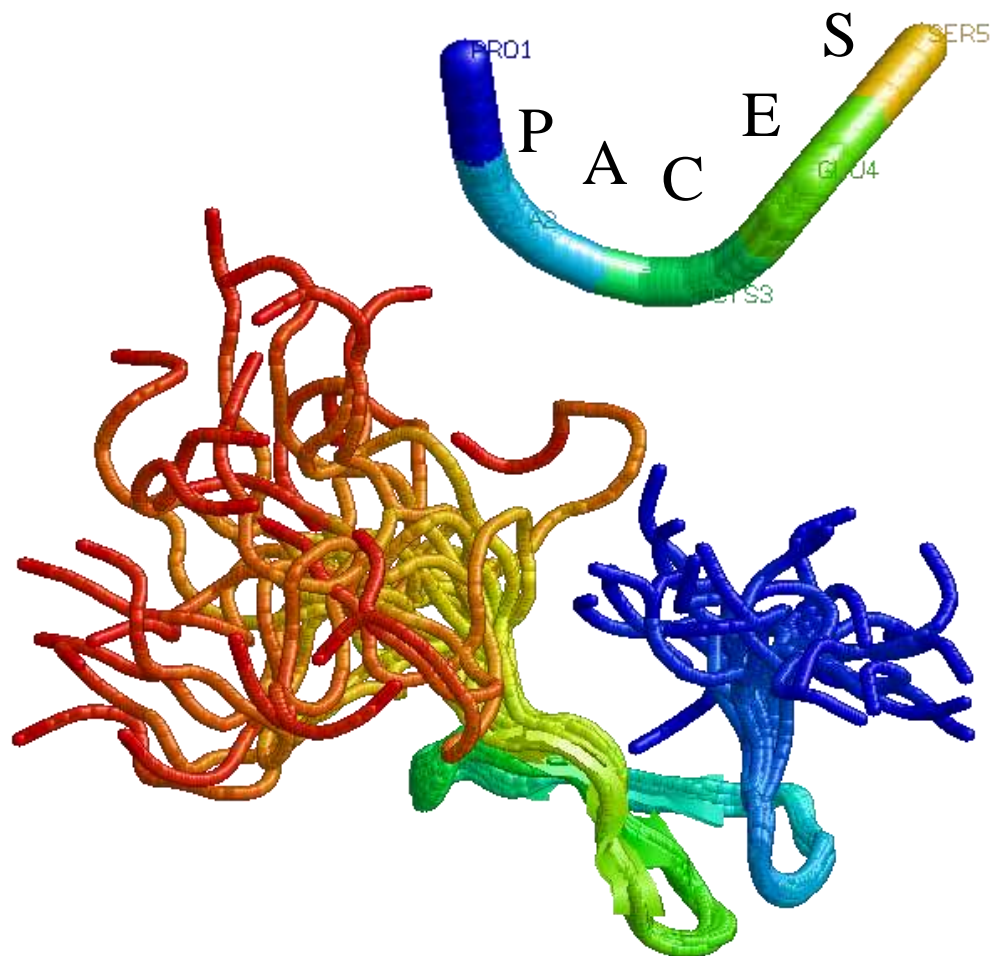
www: <http://bio.img.cas.cz>

email: hpaces@img.cas.cz

icq: #110872370

irc: efnet #hpaces

tel: +420 220183446



Synthetic Biology

- Synthetic Biology Open Language (SBOL) standard visual symbols for use with BioBricks Standard
-

Typy umělého života

- *In silico* – simulace biologických komponent
- *in silico* – programy, které mají vlastnosti živých organizmů
- *In vivo* – vnášení nových prvků do existujících živých organizmů
- *In vivo* – sestavení funkčního organismu z abiotických prvků
- *In hardware* – náhrada biologických komponent umělými

Cizorodé elementy

DNA, která se do genomu dostala jinak než vertikálním transferem, tj z předků na potomky

principy detekce:

- „cizí“ DNA je odlišná od průměrné „vlastní“ DNA
- experimentální data

možnosti

- Modifikace stávajících proteinů
- Nové proteiny ze stávajících komponent (4 DNA, 20 aminokyselin)
- Rozšíření počtu aminokyselin přeprogramováním současných kodónů
- Rozšíření informační kapacity DNA pomocí nových párů bazí

Repetitivní elementy

- Transpozóny:** transposon-derived repeats, interspersed repeats; 45% genomu
- Mikro a minisatelitey:** simple sequence repeats, opakování krátkých přímých repetitivních; 3% genomu
- Duplikace:** duplikace různě dlouhých (10-300 kb) genomových segmentů - inter i intrachromosomové; 3,3% genomu
- Jiné typy repetitivních:** centromerické a telomerické repeaty

Transpozóny

DNA transpozóny

retrotranspozóny

(RNA intermediát, reverzní transkripce)

LTR transpozóny (podobné retrovirům)

polyA (non LTR) retrotranspozóny

(kolineární s mRNA, mají polyA)

DNA transpozóny



2-3 kb

terminální reverzní repetice (50 - 100 b)

cut-and-paste mechanismus kopírování

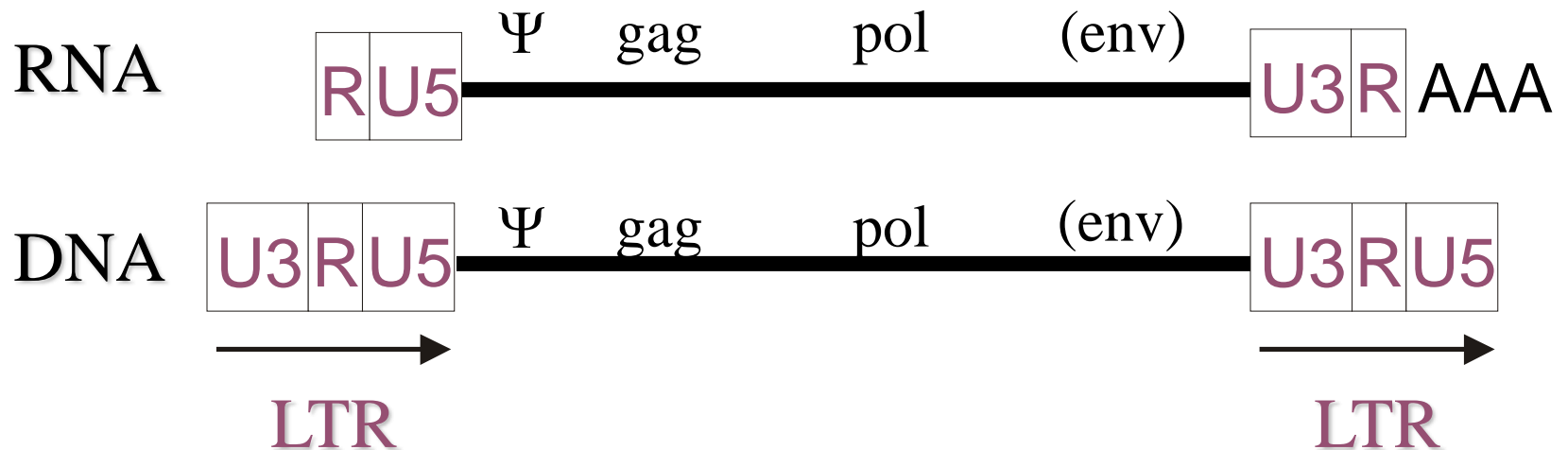
3% genomu

minimálně 7 tříd, které nejsou (blízce) příbuzné

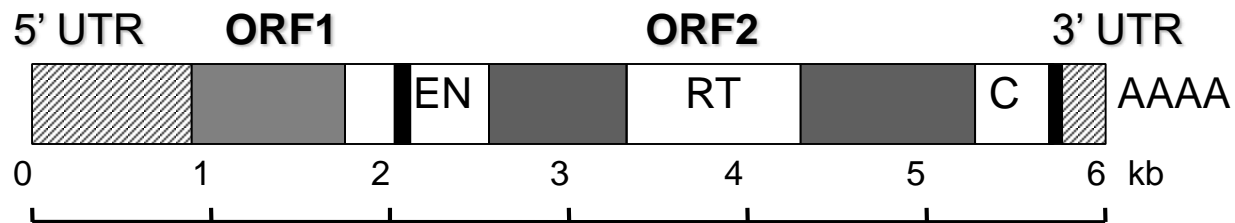
LTR retrotranspozóny

HERV:

1. 6 - 8 % lidského genomu
2. 100 000 elementů
3. desítky rodin



non-LTR (LINE1 nebo L1 elementy)



LINE – long interspersed elements

poly A (non-LTR) retrotranspozóny

RNA intermediát (interní promotor pro RNA pol. II); polyA

krátká inzerční duplikace (5-15 bp)

inzerční preference (TT|AAAA)

17 % genomu

500 000 elementů, často zkrácených na 5' konci

30-60 aktivních LINE1 elementů v genomu

Neautonomní elementy

nekódují enzymy pro svou vlastní transpozici

pro každou třídu autonomních elementů existuje neautonomní element, který používá mechanismus replikace „svého“ autonomního elementu

DNA transpozóny

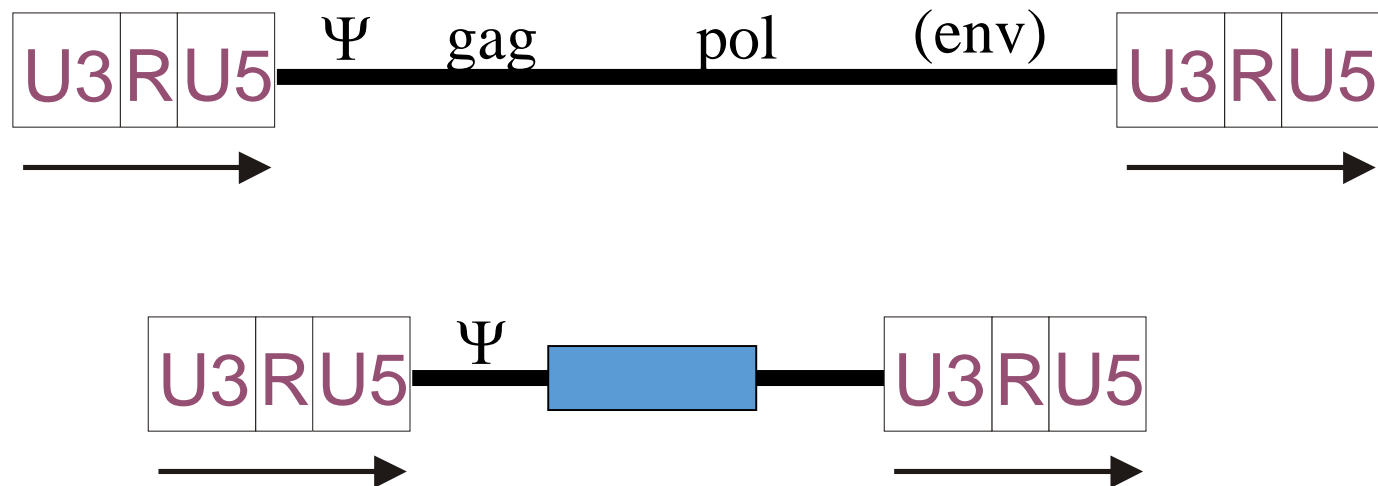


2-3 kb; terminální reverzní repetice



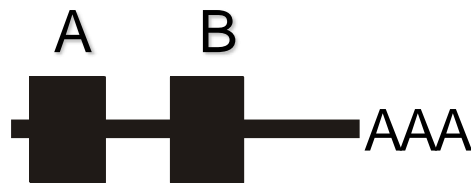
80-3000 bp; terminální reverzní repetice

Lidské endogenní retroviry (HERVs)



LTR; krátké inzerční duplikace; primer binding site

SINE (Alu) elements



SINE – short interspersed elements

poly A (non-LTR) retrotranspozóny

interní promotor pro RNA pol. III; polyA

inzerční duplikace (5-15 bp)

inzerční preference (TT|AAAA)

10 % genomu

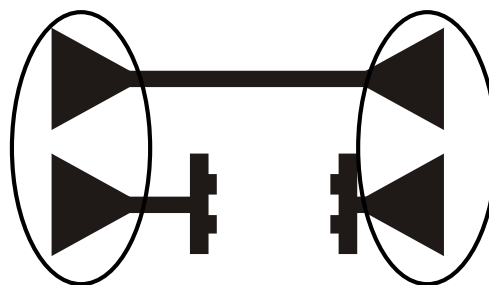
1 000 000 elementů, často zkrácených na 5' konci

Procesované pseudogeny

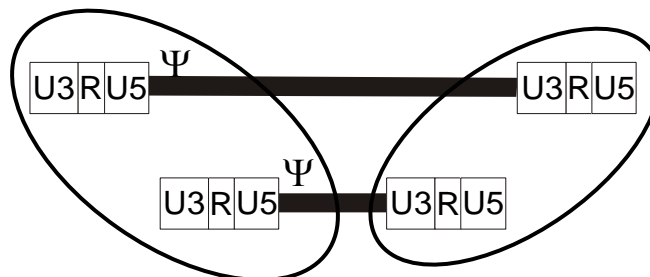
kolineární s mRNA, chybějí introny a promotory
poly A
často zkrácené na 5' konci
krátké inzerční duplikace

Koevoluce parazitů

DNA



LTR

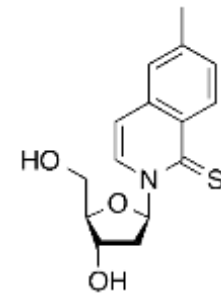
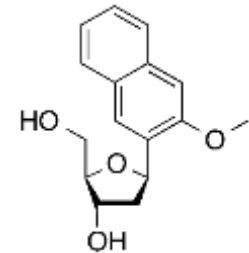


polyA



Nová písmena genetické abecedy

- UBP (unnatural base pairing)
- XNA (xeno nucleid acid)
- **dNaM** is an artificial [nucleoside](#) containing 3-methoxy-2-naphthyl group
- **d5SICS** is an artificial [nucleoside](#) containing 6-methyl[isoquinoline](#)-1-thione-2-yl



CZECH

FOBIA