

Genetický kód

Rozluštění **genetického kódu** se datuje do roku 1961, kdy bylo pokusy s translací in vitro zjištěno, že vždy **trojice nukleotidů** (triplet) kóduje určitou aminokyselinu.

- tyto triplety se nazývají **kodony**
- 4 báze DNA a RNA se mohou kombinovat jako $4^3=64$ **kodonů**, které specifikují 20 **proteínogenních aminokyselin** pro tvorbu proteinů
- **genetický kód je degenerovaný**
 - **počet kodonů** je větší než počet **aminokyselin** – všechny amk s výjimkou tryptofanu a methioninu jsou kódovány **více triplety**
 - genetický kód se tedy vyznačuje **nadbytečností** – označujeme jej jako degenerovaný
 - **kodony**, které specifikují stejnou aminokyselinu se nazývají **synonymní**
 - **variace** mezi synonymními kodony se týká hlavně **3. pozice** v tripletu, rozhodující úlohu mají obvykle první dvě báze kodonu
 - degenerace genetického kódu minimalizuje **efekt mutací**
 - díky degeneraci také nelze z proteinu určit strukturu mRNA, jež ho kódovala, naopak však **strukturu proteinů** můžeme určit zcela přesně
- ze **64 kodonů** kóduje aminokyseliny **61 kodonů** – zbývajících 3 (UAG, UGA a UAA) jsou tzv. **terminační kodony** nebo stop kodony a proteosyntézu **ukončují** (tRNA mající komplementární antikodon totiž nese místo amk obvykle vodu)
- **proteosyntéza** je naopak zahájena v místě **iniciačního kodonu** AUG, který kóduje **methionin**
- je lokalizován v některém z prvních exonů a určuje **čtecí rámec** sekvence RNA
- každá **sekvence RNA** může být čtena třemi soubory kodonů, podle toho, která báze je vybrána jako **začátek kodonu**
- soubor kodonů, který je omezen **iniciačním a terminačním kodonem** se nazývá otevřený čtecí rámec – **ORF, open reading frame**
- genetický kód je považován za **univerzální**
 - stejným způsobem ho využívají **všechny organismy**
 - existují však výjimky – zejména **mitochondrie** a některé jednobuněčné organismy
 - např. v mitochondriích UGA (obvykle terminační) kóduje **tryptofan** a naopak AGA a AGG (obvykle arginin) jsou **terminační**
- ne všechny synonymní kodony jsou užívány u různých druhů se stejnou četností – některé více, některé méně – tzv. **kodonový dialekt**

Proteosyntéza – biosyntéza bílkovin – představuje pochod, při kterém se genetická informace uchovávaná a předávaná v řeči DNA exprimuje, tj. realizuje formou bílkovin. Proteosyntéza se také nazývá translace – překlad, neboť představuje biochemický převod informace z řeči nukleových kyselin do řeči bílkovin.

Genetická informace je lineární, je uložena ve formě polynukleotidového vlákna (obvykle) DNA, což je polymer nukleotidů. Připomeňme, že používá čtyř písmen – **nukleotidy A, G, C a T** (adenin, guanin, cytozin, thymin). Naproti tomu bílkoviny jsou složeny jak z biochemického, tak i stereochemického hlediska ze zcela jiných písmen nazývaných aminokyseliny. Počet těchto aminokyselin dnes dosahuje čísla 22. Při tvorbě bílkoviny se rovněž spojují, polymerují do lineárního vlákna a dá se říci, že existuje kolinearita (souběžnost) mezi pořadím nukleotidů v DNA a sekvencí aminokyselin v bílkovině, která je touto DNA kódována. Prakticky jde tedy při translaci o vytvoření bílkovinného vlákna podle vlákna polynukleotidového. Podle konvence se polynukleotidové sekvence píšou *ve směru 5'–3'* a aminokyselinové sekvence proteinu od amino N-konce ke karboxylovému C-konci, což odpovídá směru jejich čtení.

Jak už odvodili **F. Jacob** a **J. Monod**, „přeměna“ DNA na bílkovinu neprobíhá přímo, ale přes prostředníka, kterým jsou mRNA. Každá mRNA je kopií určitého úseku jednoho z vláken DNA zhruba v délce jednoho nebo více genů. Tak jako DNA nese tedy stejnou specifickou zprávu pro výrobu jedné nebo více bílkovin. Nukleotid T je však v mRNA nahrazen jiným pyrimidinovým nukleotidem U a jako cukerná složka zde slouží ribóza a ne deoxyribóza jako v DNA. Navíc je RNA přirozeně jednovláknová^[1]. Tím vším se však RNA stává i mnohem labilnější než DNA. Dá se snadno a rychle odbourat přirozenými mechanismy buňky, když už se např. dosáhlo požadované hladiny dané bílkoviny a její přebytek by byl buňce k ničemu anebo by mohl i škodit (např. nadbytek mRNA kódující protoonkogeny). Uvádí se, že v buňce *E. coli* je průměrný poločas mRNA asi 2–3 minuty, naopak mRNA kódující v eukaryotické buňce β-globin má poločas delší než 10 hodin.

Genetický kód – kódovací pravidla pro 22 aminokyselin kódovaných geneticky. Jsou uvedeny i mezinárodní jednopísmenné zkratky aminokyselin a pro tři *stop* triplety i jejich historická pojmenování *Amber*, *Opal* a *Ochre*. Syntéza naprosté většiny bílkovin začíná methioninem, který je kódován kodonem AUG (univerzální startovací signál), výjimečně valinem (kodon GUG) případně ještě jinou aminokyselinou. Z tabulky je dobře patrná degenerace genetického kódu vyjádřená zejména proměnlivostí ve třetí pozici kodonů

Ala	A	GCU, GCC, GCA, GCG	Phe	F	UUU, UUC
Arg	R	CGU, CGC, CGA, CGG, AGA, AGG	Pro	P	CCU, CCC, CCA, CCG
Asn	N	AAU, AAC	Ser	S	UCU, UCC, UCA, UCG, AGU, AGC
Asp	D	GAU, GAC	Thr	T	ACU, ACC, ACA, ACG
Cys	C	UGU, UGC	Trp	W	UGG
Gln	Q	CAA, CAG	Tyr	Y	UAU, UAC
Glu	E	GAA, GAG	Val	V	GUU, GUC, GUA, GUG
Gly	G	GGU, GGC, GGA, GGG	Start		AUG, (GUG)
His	H	CAU, CAC	Stop		UAG (<i>Amber</i>)
Ile	I	AUU, AUC, AUA	Stop		UGA (<i>Opal</i>)
Leu	L	CUU, CUC, CUA, CUG, UUA, UUG	Stop		UAA (<i>Ochre</i>)
Lys	K	AAA, AAG	Pyl	O	UAG
Met	M	AUG	Sec	U	UGA (+ <i>SECIS</i>)

Jaký je *šifrovací klíč* neboli *kód* mezi čtyřmi nukleotidy v mRNA a 22 aminokyselinami v bílkovině? Ukázalo se, že tento tzv. **genetický kód** je třípísmenný, tj. pořadí tří nukleotidů – **triplet** kóduje vždy jedinou aminokyselinu. Ze čtyř různých nukleotidů v mRNA se tak může vytvořit $4^3 = 64$ různých tripletů neboli kódových slov, **kodonů** (viz tabulku). To je naprosto dostatečný počet ke specifikaci každé z 22 aminokyselin a navíc i k vytvoření signálů, které určí, odkud má v mRNA translace začít (obvykle od kodonu AUG) a kde má skončit (tři signály *stop* – UAA, UAG a UGA). Představují jakousi tečku za větou, tj. v našem případě za syntetizovanou bílkovinou.

Přiřazení jednotlivých aminokyselin k jednotlivým kodonům, tj. *rozšířování genetického kódu*, se experimentálně podařilo asi v polovině 60. let 20. století. Jak je vidět z tabulky, pouze dvě aminokyseliny, a to methionin a tryptofan, jsou kódovány jen jedním kodonem (AUG, resp. UGG). Pro zbytek aminokyselin platí, že jsou kódovány dvěma, třemi, čtyřmi nebo dokonce šesti kodony. Kodony pro stejnou aminokyselinu se označují jako *synonyma* a obvykle se liší ve třetím nukleotidu. Existence více kodonů pro jednu aminokyselinu znamená, že **kód je degenerovaný**. Na druhé straně je však **jednoznačný**, neboť každý kodon specifikuje vždy jen jednu jedinou aminokyselinu. Genetický kód je také skoro (viz dále) **univerzální**, neboť se používá pro translaci ve všech organismech. Tato skutečnost je jedním z nejpřesvědčivějších důkazů o společném předku všech živých organismů.

Kodony jdou v mRNA za sebou, bez mezer a bez překrývání v tzv. **čtecím rámci**. Ten se nastaví prvním – iniciačním – kodonem, což je nejčastěji AUG kódující methionin. Naprostá většina nově syntetizovaných bílkovinných řetězců tedy začíná touto aminokyselinou. Methionin se však velmi často – a to buď ještě před dokončením translace, anebo až po ní, už z hotové bílkoviny – odštěpuje. Informace v mRNA se tedy čte a překládá souvisle po úsecích dlouhých tři nukleotidy a jejich sled se nastaví prvním iniciačním kodonem^[2].

Ne všechna kódová slova pro jednu aminokyselinu jsou využívána v organismu se stejnou frekvencí. Naopak, některá ze synonym jsou v genetické informaci skoro nevyužívána (anebo jen pro speciální proteiny), na rozdíl od tzv. *běžných kodonů*. Přitom četnost využívání jednotlivých kodonů je specifická už pro každý jednobuněčný organismus, jiný typický výběr platí pro obratlovce a jiný zase např. pro vyšší rostliny. Mluvíme o tzv. **kodonovém dialektu**.

Odkazy

Související články

Reference

- Živa 2006, 5: 198
- Živa 2006, 5: 198