

Nonstop trypanozoma

ČESKÝ BIČÍKOVEC SE ZMĚNĚNÝM GENETICKÝM KÓDEM

Proč se bičíkovec, námi pojmenovaný *Blastocritidia nonstop*, tak velmi dramaticky odklonil od univerzálního genetického kódu? K čemu může být obskurnímu prvokovi taková odchylka užitečná? A jaké může mít nový poznatek důsledky pro člověka?

text **JULIUS LUKEŠ, ZDENĚK PARIS a LEOŠ SHIVAYA VALÁŠEK**

VŠICHNI SE SHODNEME, že dub se výrazně liší od bakterie a žádný z nich se ani v nejmenším nepodobá viru či žralokovi. Přesto mají tyto a další formy života, tolik odlišné svým vzhledem, něco společného. Touto dobře skrytou společnou vlastností je stejný (a proto označovaný jako univerzální) genetický kód. Lze ho charakterizovat jako soubor pravidel převádějících genetickou informaci z DNA do proteinů (neboli bílkovin), které určují vzhled a funkce daného organismu.

UNIVERZÁLNÍ GENETICKÝ KÓD DNA je známá zkratka pro deoxyribonukleovou kyselinu, kterou má každý organismus na naší planetě uloženou v jádře svých buněk a kterou si můžeme představit jako tlustou encyklopedii, obsahující místo jednotlivých hesel tzv. geny. Geny jsou jasně vymezené úseky DNA, z nichž každý nese přesnou informaci pro konkrétní protein, zapsanou jako specifické heslo podle univerzálního genetického kódu. Například lidská DNA by svojí délkou ve vytištěné formě zaplnila 400 dílů Encyclopaedia Britannica (viz **rámček**), kterých je ve skutečnosti jen 32. Geny kódované v DNA jsou přepsány do mRNA a genetická informace je přenesena z jádra do cytoplazmy, kde se setkává s továrnou na proteiny zvanou ribozom (obr. 1). To je skutečná výrobní linka buněk,

jinými slovy geniálně elegantní molekulární dekodér. Právě na něm probíhá složitě převedení čtyřpísmenového textu mRNA do dvacetipísmenového textu bílkovin (obr. 2). Prostředníkem tohoto převodu jsou molekuly transferové RNA (tRNA), které ribozomu slouží jako luštitelská příručka. Poté, co ribozom rozpozná začátek genu, překládá vždy tři po sobě následující báze (tzv. kodon nebo též triplet) mRNA do jedné aminokyseliny. Jak ribozom kodon po kodonu počítá mRNA, propojují se příslušné aminokyseliny postupně jedna po druhé do vznikajícího aminokyselinového řetězce = proteinu (obr. 1). Pro tento proces je nezbytná právě molekula tRNA. Zjednodušeně lze říci, že každý kodon má svoji „pasující“ tRNA, která ho bezpečně pozná díky svému antikodonu. Na základě tzv. komplementárního párování kodonu s antikodonem přinese tRNA na ribozom přesně tu aminokyselinu, která má být podle genetického kódu do aminokyselinového řetězce zařazena. Aby ribozom správně poznal nejen začátek genu, ale i jeho konec, tři kodony (UAA, UAG a UGA) slouží jako tzv. stop kodony (obr. 2). Ve chvíli, kdy se k nim dostane ribozom, překládá každé genu končí, syntetizovaný protein se z ribozomu uvolní a začne plnit svou specifickou buněčnou úlohu. Proč a jak se vyvinul právě tento redundantní kód a proč právě v této kombinaci čtyř a dvaceti písmen,

proč má vše živé tři stop kodony, a ne jeden či sedm, zůstává velkou záhadou. Jedno je ale jisté – výše popsaný genetický kód vznikl na Zemi pouze jedenkrát, ve společném předkovi všeho živého tvorstva, a to před několika miliardami let. Od té doby se jím řídí praneprné viry i obří velryby.

Rozlousknutí těchto dekodujících pravidel kodon-aminokyselina, která představují základ tzv. centrálního dogmatu molekulární biologie, tj. přenosu genetické informace z DNA přes RNA do proteinu, nebylo nic snadného. Důkazem budiž řada Nobelových cen, které na toto téma začaly být udělovány v roce 1962. Přímo za rozluštění genetického kódu ji získali roku 1968 Marshall Nirenberg, Gobind Khorana a Robert Holley.

UNIVERZÁLNÍ, AŽ NA VÝJIMKY I když genetický kód zůstává již po několik miliard let neměnný, naše původní představa o jeho naprosté univerzálnosti se za posledních 20 let poměrně značně změnila. Výzkum z poslední doby, včetně toho našeho, jasně ukazuje, že si už nelze dále namlouvat, že genetický kód je dokonalost sama, protože evoluce takto prostě nefunguje. Slavný biolog Stephen Jay Gould evoluci přirovnal k opilci, který při chůzi z bodu A do bodu B vrávorá, odbočuje či upadá, ale nakonec do bodu B přece jen dojde. Tak nějak to asi bylo

Prof. RNDr. JULIUS LUKEŠ, CSc., (*1963) vystudoval Přírodovědeckou fakultu UK. V Parazitologickém ústavu Biologického centra AV ČR a na Přírodovědecké fakultě JU se zabývá funkční genomikou prvoků. Působil na Amsterodamské univerzitě a na Kalifornské univerzitě v Riverside a v Los Angeles. Je členem Učené společnosti a AAAS.

RNDr. ZDENĚK PARIS, Ph.D., (*1981) vystudoval molekulární parazitologii na Přírodovědecké fakultě JU. Jako postdok působil tři roky na Ohijjské státní univerzitě. Od roku 2014 je vedoucím laboratoře RNA biologie prvoků v Parazitologickém ústavu BC AV ČR, kde se zabývá regulací translace pomocí modifikací transferových RNA (tRNA).

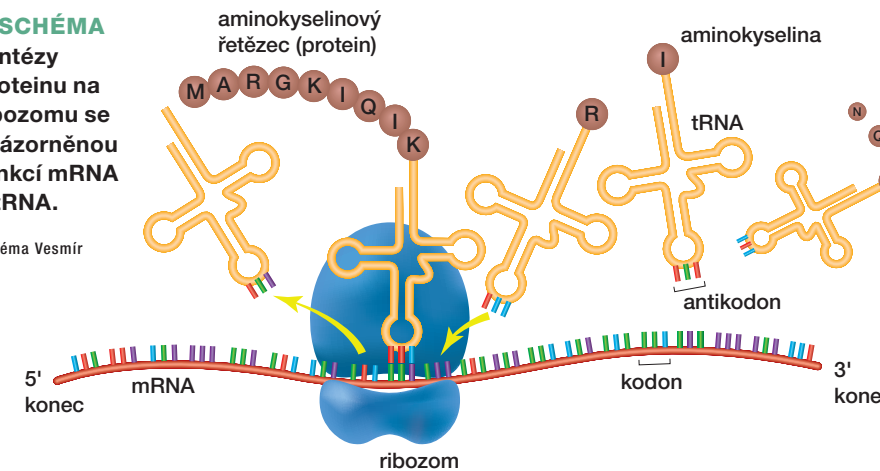
Dr. rer. nat., RNDr. LEOŠ SHIVAYA VALÁŠEK, DSc., (*1971) vystudoval molekulární biologii a genetiku na Přírodovědecké fakultě UK, doktorát získal v oboru biochemie na Viedeňské univerzitě. V Mikrobiologickém ústavu AV ČR se jeho skupina v rámci široké mezinárodní spolupráce zabývá regulací syntézy proteinů. Tým získal podporu řady prestižních zahraničních i tuzemských grantů.

Trocha přírodopisu

MY DŘÍVE NAROZENÍ jsme se dozvídali o genetickém kódu v hodinách přírodopisu, což ale pro dnešní školáky už neplatí, a proto není od věci jej ve stručné formě představit. Genetický kód DNA je čtyřpísmenový – písmena představují báze guanin = G, cytosin = C, adenin = A a thymin = T. Jsou za sebou řazena v (na první pohled) naprosto náhodném pořadí. Takový text DNA je napřed přepsán do „poslůčkové“ (anglicky messenger) RNA (mRNA), přičemž dojde jen k drobné změně, a to nahrazení thyminu (T) uridinem (U), která ale plně zachovává univerzální genetický kód. Tento překvapivě jednoduchý jazyk obou nukleových kyselin musí být poté v každé buňce převeden do výrazně složitějšího jazyka proteinů tvořeného dvaceti písmeny, z nichž každé představuje jednu aminokyselinu (obr. 1 a 2). Při čtyřech bázích existuje 64 možných kodonů (4 × 4 × 4), ale aminokyselin (písmenek) tvořících proteiny je pouze 20 (zkuste si představit, jak těžký by byl přesný překlad textu z jazyka tvořeného čtyřmi písmeny do jazyka zapsaného dvaceti písmeny). Kvůli tomuto nepoměru se genetický kód označuje jako redundantní, protože jednu aminokyselinu může kódovat třeba jen jediný kodon (např. tryptofan je určen pouze kodonem UGG), anebo hned několik kodonů, např. šest (UUA, UUG, CUU, CUC, CUA a CUG určují aminokyselinu leucin; obr. 1).

1. SCHÉMA syntézy proteinu na ribozomu se znázorněnou funkcí mRNA a tRNA.

Schéma Vesmír



i s genetickým kódem, který nevymyslel geniální programátor, nýbrž se stal produktem série nepředvídatelných událostí. Jak to vtipně nazval spoluobjevitel struktury DNA Francis Crick, jeho ustanovení lze nejlépe popsat jako „zamrzlou náhodu“.

Fascinující vlastností živého světa totiž je, že (téměř) žádné pravidlo není bez výjimek. Dříve či později se najde organismus, který si jakékoli pravidlo v průběhu evoluce tak či

onak pozmění či přímo „ohne“. Takže i když se u genetického kódu – tohoto společného jmenovatele všeho živého – dlouho předpokládalo, že jakékoliv jeho zásadní změny jsou nepřijatelné, protože každá větší odchylka musí být z definice smrtelná, ve skutečnosti tomu tak není.

Současné neuvěřitelné tempo sekvenování DNA desítek tisíc organismů ukazuje, že tu a tam si některý tvor tento genetický

kód opravdu trochu „ohnul“, v některých případech dokonce pořádně, a přesto přežil. Ačkoli stále platí, že přibližně 99,99 % organismů ho má stejný, tj. přesně takový, jaký ukazují učebnicová schémata (obr. 1 a 2), dosud bylo popsáno 32 různých typů odchylek. Tyto tzv. nestandardní genetické kódy se od univerzálního liší často jen nepatrně, např. změnou jen jednoho kodonu pro jednu aminokyselinu, ale v několika pozoruhodných případech jsou tyto odchylky poměrně dramatické. Patří mezi ně i genetický kód, v němž je stop kodon čten jako smysluplný kodon, tj. místo ukončení tvorby proteinu je mu přiřazena aminokyselina a syntéza pokračuje nerušeně dál.

PODIVUHODNÝ BIČÍKOVEC BLASTOKRITIDIE

Po tomto učebnicovém a patrně i trochu únavném úvodu je nejvyšší čas představit zatím zdaleka nejdramatičtější odchylku od univerzálního genetického kódu, k níž došlo u jednoho prvoka neboli jednobuněčného eukaryota. Jde o bičíkovec ze skupiny *Trypanosomatina*, což jsou prvoci blíže příbuzní např. trypanozomě spavičné, která je u člověka a skotu původcem smrtelného onemocnění zvaného spavá nemoc. Tohoto prvoka jsme izolovali z celkem běžně se vyskytující ploštěnce *Eysarcoris aeneus* (obr. 4A), chybně na louce vedle Podtroseckých rybníků v Libereckém kraji. Hledáním bičíkovců parazitujících na hmyzu se zabýváme téměř dvě desetiletí a během té doby jsme popsali mnoho nových druhů z různých koutů světa. Rozhodně jsme se ale nenadáli, že ten dalece nejpodivuhodnější, kterého jsme pojmenovali *Blastocritidia nonstop* (obr. 4B), se vyskytuje přímo za humny.

Dramatický odklon od univerzálního genetického kódu v jeho případě znamená, že ve svých genech nese tisíce stop kodonů UAA, UAG a UGA, které však jeho ribozom nečte jako tečky na konci věty, rozumějme genu, a tedy jako konec výroby příslušného proteinu. Kdyby tak činil, žádný z jeho proteinů by nebyl vyroben v plné délce, v nadsázce řečeno by se prvok nedožil dalšího rána. Ribozom blastokritidie těmto „tečkám“ totiž přiřazuje aminokyseliny glutamát pro UAA a UAG a tryptofan pro UGA, takže vše, co k životu potřebuje, si jeho buňka bez problému vytvoří.

Velkou záhadou však bylo, jak to dokáže. Vždyť i tento podivuhodný systém potřebuje správně určit konec každého proteinu! Podařilo se nám ukázat, že každému stop kodonu UAA je přiřazena aminokyselina glutamát pouze uvnitř genů, zatímco pokud se UAA nachází na konci genu, který má tento prvok velmi specificky definovaný, funguje jako opravdový stop kodon a ribozom zde syntézu proteinu ukončí. Příklad takového textu, který se postupně odchylí

druhé písmeno

		druhé písmeno					
		U	C	A	G		
první písmeno	U	UUU Phe UUC Ser UUA Leu UUG Leu	UCU Ser UCC Ser UCA Ser UCG Ser	UAU Tyr UAC Tyr UAA Stop UAG Stop	UGU Cys UGC Cys UGA Stop UGG Trp	U	C A G
	C	CUU Leu CUC Leu CUA Leu CUG Leu	CCU Pro CCC Pro CCA Pro CCG Pro	CAU His CAC His CAA Gin CAG Gin	CGU Arg CGC Arg CGA Arg CGG Arg	U	C A G
	A	AUU Ile AUC Ile AUA Met AUG Met	ACU Thr ACC Thr ACA Thr ACG Thr	AAU Asn AAC Asn AAA Lys AAG Lys	AGU Ser AGC Ser AGA Arg AGG Arg	U	C A G
	G	GUU Val GUC Val GUA Val GUG Val	GCU Ala GCC Ala GCA Ala GCG Ala	GAU Asp GAC Asp GAA Glu GAG Glu	GGU Gly GGC Gly GGA Gly GGG Gly	U	C A G

2. UNIVERZÁLNÍ GENETICKÝ KÓD. Počáteční metionin je vyznačen zeleně, stop kodony červeně a ostatní aminokyseliny modře.

Schéma Vesmír

STOPS MAKING SENSE
 Imagine a gene is a sentence starting with a capital letter and ending with a full stop, and a genome is a book telling an entire story. In some protozoan organisms, extra full stops have infiltrated the sentences, replacing specific letters of arbitrary words (namely E and W). As a result, readers (in this case, ribosomes) are confused as to where the sentences really end and the story becomes disjointed. In this issue of *Nature*, researchers describe a molecular mechanism that these organisms must have developed to allow the reader to navigate the sentences as if there were no extra full stops. The sentences are so specifically encrypted that readers of no other organisms but those very few can decipher the story in a proper way. The trick lies in the length of the transfer RNA molecule, and in the unique modification of a single protein that normally ensures dot recognition in cells – the precise end of protein synthesis. By the interaction of these two modified molecules, the ribosome of this protozoan knows when to correctly terminate, notwithstanding the many dots, and when to replace the dots with the two original letters (E and W), which returns meaning to the genetic information.

Engineered outcome
 Supercharged immune cells gain greater range to help combat cancer

Keeping the beat
 Wearable ultrasonic device offers real-time images of the heart

Get plugged in
 Tools and techniques that could revamp research in 2023

3

4A



4B



Snímky Miroslav Fiala (4A) a Jan Vořípka (4B)

od zavedeného standardu, a přesto jej lze přečíst, ukazujeme na obr. 5.

To ale stále neodpovídá na otázku, jak dokáže blastokritidie bez problému číst stop kodony neboli extra tečky uprostřed svých genetických vět, aniž se ztrácí jejich smysl. Vždyť jsou tisíce jasně zdokumentovaných případů, kdy nechtěný výskyt jediného stop kodonu uprostřed genu (tzv. nonsense mutace) znamená předčasné ukončení syntézy příslušného proteinu se závažnými důsledky. Zkrácený protein totiž bývá nefunkční, což vede třeba u člověka k desítkám vážných dědičných chorob; k těm nejznámějším patří například cystická fibróza.

Zjistili jsme, že stačí drobná změna ve struktuře výše zmíněné tRNA (obr. 6A), aby se tato molekula chovala jinak, než „má“, tj. aby prostřednictvím svého antikodonu četla kromě svého pasujícího (tj. komplementárního) kodonu i kodon jiný. V případě námi studovaného prvoka jde o tryptofanovou tRNA, která ve své podivně zkrácené formě umí přečíst také stop kodon UGA – prostě ho přeloží jako tryptofan (obr. 6B). Navíc takto zkrácené tryptofanové tRNA se čtením UGA ještě velmi významně pomáhá jedinečně pozměněný, tzv. uvolňovací protein (terminační faktor 1), který v buňkách za normálních okolností zajišťuje rozpoznávání všech tří teček (UAA, UAG a UGA), tj.

3. ČASOPIS Nature si tentokrát vybral článek českých autorů pro titulní stranu. 4. PLOŠTICE *Eysarcoris aeneus* (A) a z jejího střeva izolovaná *Blastocrithidia nonstop* (B – vlevo je obarvený bičíkovec ve světelném mikroskopu, vpravo bičíkovec s cystami ve skenovacím elektronovém mikroskopu).

přesné ukončení syntézy proteinu. Ve své mutované formě totiž terminační faktor 1 nedovede rozpoznat tečku UGA.

S ostatními stop kodony UAG a UAA rozptýlenými uvnitř genů si pak blastokritidie poradila tím nejjednodušším možným způsobem, tj. vytvořila si specifické tRNA, které mají svůj antikodon pozměněný tak, aby oba stop kodony vždy přečetly jako aminokyselinu glutamát. Díky součinnosti všech takto jedinečně upravených molekul pozná její ribozom, kdy má, nehledě na spoustu teček, větu správně ukončit, a kdy má tečky naopak nahradit tryptofanem či glutamátem, což genetické informaci vrací smysluplnost (viz video na vesmir.cz). Všechna tato molekulární „přenasavení“,

jichž se studovaný bičíkovec dopustil, představují naprosto bezprecedentní porušení standardních watsonovsko-crickovských pravidel pro dekódování genetického kódu. Náš objev proto zaujal i časopis Nature.¹

K ČEMU ODCHYLKA SLOUŽÍ?
 K čemu ale může být popsána obskurní úchylka obskurnímu prvokovi užitečná? V každém případě nám ukazuje, že pravidla genetického kódu lze porušit a nejen nezemřít, ale dokonce s tím i normálně žít. Proč by ale jakýkoli organismus podnikal tak riskantní změnu pravidel života? Nečekaný a zároveň elegantní důvod naznačují pokusy s *Escherichia coli*, publikované nedávno v časopise Science.² V nich vědci

této všudypřítomné bakterii pozměnili genetický kód a ukázali, že jí tím ochrání před napadením řadou virů, kteří by ji jinak zabili. Taková trvalá a nesmírně efektivní obrana před všemi viry by mohla stát za riziko dramatické odchylky od univerzálního kódu a určitě se pokusíme otestovat její platnost i u popsané blastokritidie.

Náš objev může mít ale i jiné, potenciálně dalekosáhlé důsledky. Jak jsme si řekli výše, pro ukončení translace na skutečném konci každého genu používá blastokritidie pouze jeden stop kodon (UAA), čímž uvolňuje další dva všeobecně používané stop kodony (UAG a UGA) pro jiné použití. Takové uvolnění kodonů pro syntetické, v přírodě se nevyskytující aminokyseliny je jedním z hlavních cílů dynamicky se rozvíjející syntetické biologie.

Dalším našim velice důležitým zjištěním je to, že výše popsané odchylky lze pohodlně přenést z podivné blastokritidie i do zcela standardně geneticky vybavené kvasinky *Saccharomyces cerevisiae*, která je již po několika desetiletích nejrozšířenějším modelovým organismem molekulární a buněčné biologie. Navíc se zdá, že naše práce s tímto organismem již brzy odpoví na další zajímavou otázku, tj. proč právě takto zkrácená verze tryptofanové tRNA umí přečíst tečku UGA. Ale o tom třeba příště.

Jelikož od kvasinky je už jen pomyslný krůček k člověku, domníváme se, že bude „zkrácená“ tRNA fungovat stejně i v lidských buňkách. Odtud už zase není tak daleko k představě, že by „blastokritidiovsky“ speciálně upravená tRNA mohla v lidských buňkách přepsat nesmyslné mutace (tj. stop

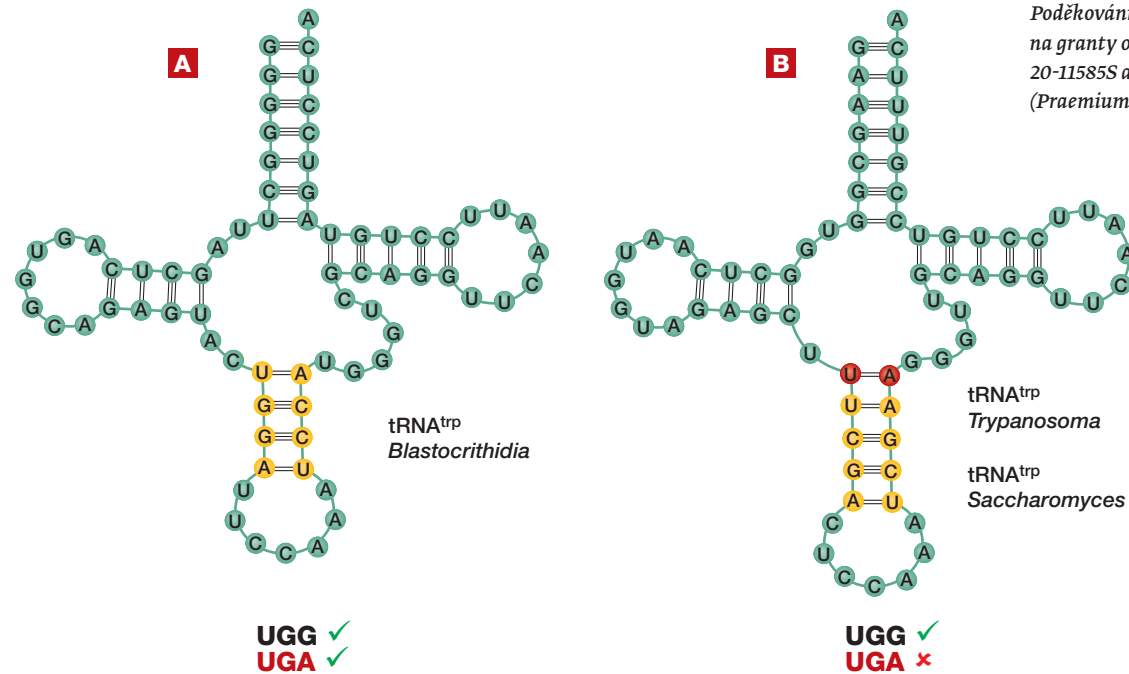
kodony) uprostřed klíčových genů aminokyselinou tryptofan, a tím pomáhat léčit řadu vážných genetických onemocnění člověka.

Až budoucnost ukáže, která z těchto představ se stane realitou. Jedno je ale jisté už nyní, a jelikož se na to poslední dobou stále více zapomíná, neodoláme to zde nezmínit. Základní výzkum prováděný třeba na zdánlivě bezvýznamných parazitech ploštic může přinést (a přináší!) objevy se značným dosahem. Tak jako evoluce dala vzniknout třeba monstrózním dinosaurům, kterým meteority a soutěživí savci nakonec odzvonili konec, dokázala evoluce vytvořit i podivnosti často schované v mikroskopických organismech, které se za poslední miliardu let moc nezměnily a skryté a v tichosti přežívají kolem nás – a nabízejí nám možnost ke zkoumání. Příroda je prostě mocná čarodějka. ●

Poděkování: Tato práce byla finančně podpořena granty od Grantové agentury ČR (22-14356S, 20-11585S a 20-00579S) a Akademie věd ČR (Praemium Academiae).

Pro jednoduchost si lze gen představit jako jednu větu začínající velkým písmenem a končící tečkou, zatímco genom je kniha vyprávějící celý příběh. Velká písmena a tečky oddělující jednotlivé věty dávají celému příběhu jasnou strukturu a smysl. Nedávno se však ukázalo, že existují organismy, do jejichž vět pronikly tečky navíc a nahradily specifická písmena (E a V) v libovolných slovech. V důsledku toho jsou četná (případě genomu jsou jimi ribozomy) zmateni, n. vědci, kde věta skutečně končí, a příběh s. stává n. srozumitelný. Č. ští. ědci obje. ili molekulární m. chanismus, kt. rý se u tohoto prvoka vy. inul, aby čt. nář. (ribozom) mohl procház. t jeho gen. tickými. ětami tak, jako by v nich žádné př. bytečné t. čky nebyly. Trik spočívá v povaze a délce molekuly transferové RNA (tRNA), která ribozomu běžně slouží jako luštitelská příručka, a v jedinečné úpravě jedné bílkoviny, která v buňkách za normálních okolností zajišťuje rozpoznávání teček. Součinností těchto dvou upravených molekul pozná ribozom prvoka – nehledě na spoustu teček –, kde má větu správně ukončit a kde má tečky naopak nahradit dvěma původními písmenky (E a V), což příběhu (genetické informaci) vrací smysluplnost.

Leoš Shivaya Valášek



6. SCHÉMA zkrácené tryptofanové tRNA (z blastokritidie) a její porovnání s normální tryptofanovou tRNA (z trypanozomy a kvasinky). Antikodonové raménko je vyznačeno žlutě a pro rozdílnou funkci klíčový pár bází červeně. Rovněž jsou vyznačeny kodony čtené příslušnou tRNA.

Schéma Vesmír

5. PŘÍKLAD TEXTU, v němž jsou vynechaná písmena podobným způsobem jako v genomu bičíkovce *Blastocrithidia nonstop*.

UGG ✓
 UGA ✓

UGG ✓
 UGA ✗