

František Katrnoška; Michal Křížek

Genetický kód a teorie monoidů aneb 50 let od objevu struktury DNA

Pokroky matematiky, fyziky a astronomie, Vol. 48 (2003), No. 3, 207--222

Persistent URL: <http://dml.cz/dmlcz/141180>

Terms of use:

© Jednota českých matematiků a fyziků, 2003

Institute of Mathematics of the Academy of Sciences of the Czech Republic provides access to digitized documents strictly for personal use. Each copy of any part of this document must contain these *Terms of use*.



This paper has been digitized, optimized for electronic delivery and stamped with digital signature within the project *DML-CZ: The Czech Digital Mathematics Library* <http://project.dml.cz>

Genetický kód a teorie monoidů

aneb 50 let od objevu struktury DNA

František Katrnoška a Michal Křížek, Praha

1. Úvod — vznik matematické genetiky

V roce 1866 augustiniánský opat Johann Gregor Mendel (1822–1884) přednesl výsledky svých mnohaletých pokusů s křížením různých druhů hrachu na půdě Brněnského přírodovědeckého spolku. V témže roce je pak shrnul ve významném díle [21], které ovšem zapadlo především proto, že vědecká veřejnost nepochopila možné souvislosti biologie a matematiky. Dnes je ovšem G. Mendel právem pokládán za zakladatele nového vědního oboru — genetiky, v němž má naše země díky němu bohatou tradici. Jeho hlavní výsledky jsou dnes známy jako tři Mendelovy zákony. Ke stejným závěrům nezávisle dospěli kolem roku 1900 Hugo de Vries, Carl Correns a Erich von Tschermak, a tak Mendelovy zákony postupně vešly ve všeobecnou známost ve světě. Počátkem 20. století vznikly další práce, v nichž se stále více objevovala snaha popsat genetické zákonitosti matematicky. Např. roku 1908 publikoval slavný matematik G. H. Hardy (1877–1947) a nezávisle též W. Weinberg výsledek (viz [15] a [32]), který je dnes známý jako zákon Hardyův-Weinbergův. Tento zákon našel mnoho aplikací v hematologii a dalších oborech. Připomeneme si jej v kapitole 2.

Prudký rozvoj matematického výzkumu v genetice nastal ve dvacátých letech minulého století. Výzkum byl zaměřen především na populační genetiku. V různých státech světa (např. Velké Británii, Rusku a USA) vznikaly týmy zabývající se aplikacemi matematiky v tomto oboru. Z předních matematiků, kteří se věnovali výzkumu v genetice, připomeňme S. Bernsteina, I. M. H. Etheringtona, L. Hogbena nebo Ju. I. Ljubiče. Zejména oni se zasloužili o vznik nového odvětví matematiky — matematické genetiky. Její význam je zdůrazněn tím, že byla zařazena jako samostatná disciplína do mezinárodní klasifikace matematických oborů označených 92D10 (Genetics), 92D20 (DNA sequences), popř. 17D92 (Genetic algebras). Těmto perspektivním oborům je také věnována celá řada specializovaných vědeckých časopisů (Genetics, Genomics, Journal of Biomathematics, American Journal of Human Genetics, IMA Journal of Mathematics Applied in Medicine and Biology a mnohé další). V následujících odstavcích se velmi krátce zmíníme pouze o dvou směrech výzkumu v matematické genetice, o genetickém kódování a jeho popisu pomocí volných monoidů a o ergodických submonoidech.

Doc. RNDr. FRANTIŠEK KATRNOŠKA, CSc. (1929), Ústav matematiky VŠCHT, Technická 5, 166 28 Praha 6.

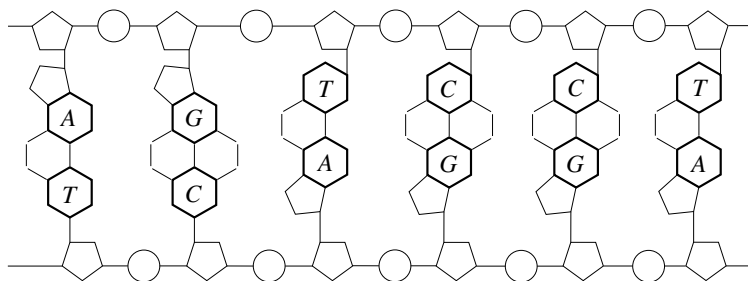
Prof. RNDr. MICHAL KRÍŽEK, DrSc. (1952), Matematický ústav Akademie věd ČR, Žitná 25, 115 67 Praha 1, e-mail: krizek@math.cas.cz

2. Struktura DNA a genotyp jedince

V roce 1944 zjistili O. T. Avery, C. M. MacLeod a M. McCarty, že nositelkou dědičné informace je deoxyribonukleová kyselina DNA, která je uložena v chromozómech¹⁾ v jádře buňky. V roce 1953, tj. právě před 50 lety, objasnili strukturu DNA James D. Watson (USA) a Francis H. C. Crick (Velká Británie), viz [30], [31] (popř. [13], [29]). Za tento významný objev získali společně s M. H. F. Wilkinsem v roce 1962 Nobelovu cenu.

Molekula DNA má tvar dvojité šroubovice tvořené dvěma polynukleotidovými řetězci (vlákny), jejichž základními stavebními kameny jsou tzv. nukleotidy. Vzdálenost řetězců je 2 nm a výška jednoho závitů činí 3,4 nm. Tento úsek obsahuje přibližně 10 párů nukleových bází. Celková délka všech molekul DNA v jádře jedné nepohlavní lidské buňky²⁾ o průměru 5 μm je přes 2 metry. Uvážíme-li, že dospělý lidský organizmus obsahuje řádově 10^{14} buněk, snadno nahlédneme, že ze všech jeho molekul DNA bychom teoreticky mohli vytvořit řetězec podstatně delší, než je vzdálenost Země od Slunce. Navíc každý metr takového informačního řetězce by obsahoval 3 miliardy nukleotidů.

Molekuly DNA jsou tvořeny třemi složkami: nukleotidy, cukrem deoxyribózou a kyselinou fosforečnou. Nukleotidy (přesněji řečeno dusíkaté báze nukleotidů) jsou *adenin A*, *thymin T*, *cytosin C* a *guanin G*. Proti sobě vždy stojí báze *A* a *T*, které jsou propojeny dvěma vodíkovými můstky, a báze *C* a *G* propojené třemi vodíkovými můstky.³⁾ Schematickým znázorněním molekuly DNA v rovině získáme útvar podobný žebříčku (viz obr. 1).

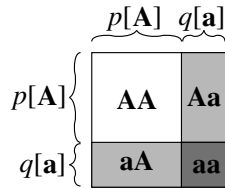


Obr. 1. Chemická struktura řetězců DNA. Nukleotidy označené symboly *A*, *T*, *C* a *G* jsou propojeny vodíkovými můstky. Molekuly cukru deoxyribózy jsou znázorněny jako pětiúhelníky a molekuly kyseliny fosforečné jako kroužky.

¹⁾ Každý chromozóm obsahuje jednu nebo dvě identické molekuly DNA, které jsou pro zpevnění a ochranu obaleny řadou podpůrných bílkovin (ve váhovém poměru přibližně 1:1).

²⁾ J. H. Tjio a A. Levan v roce 1956 prokázali, že člověk má v každé nepohlavní buňce 46 párových chromozómů, zatímco primáti jich mají 48 (polovina z nich pochází od matky a druhá polovina od otce). Lidská pohlavní buňka má jen 23 chromozómů.

³⁾ Vodíkové můstky představují jen velice slabou chemickou vazbu, která může být snadno rozštěpena. Dvojitá šroubovice ale tuto nízkenergetickou vazbu podstatně ochraňuje a zároveň zpevňuje i vlastní molekulu DNA. Pokud přesto dojde k poškození jednoho vlákna, je možné je opravit užitím informace z vlákna druhého.



Obr. 2. Grafické znázornění Hardyova-Weinbergova zákona. Na horní (resp. levé) straně jednotkového čtverce je znázorněna relativní četnost alel pro $p[\mathbf{A}] = 0,65$ a $q[\mathbf{a}] = 0,35$ samčích (resp. samicích) pohlavních buněk. Různě obarvené části čtverce odpovídají relativní četnosti alel \mathbf{A} , \mathbf{a} v nepohlavních buňkách potomků.

Jednotkou dědičné informace je gen^4), který je úsekem molekuly nukleové kyseliny. Soubor všech genů jedince se nazývá *genotyp*. Většina genů existuje alespoň ve dvou rozdílných formách, které označujeme jako *alely* (tj. alelomorfny geny). Například gen odpovídající za realizaci krevních skupin⁵) systému AB0 (viz [26]) má tři alely označované I^A , I^B a i .

Přibližme si nyní Hardyův-Weinbergův zákon. Pro jednoduchost uvažujme gen, který má jen dvě alely \mathbf{A} , resp. \mathbf{a} . Každá pohlavní buňka obsahuje buď alelu \mathbf{A} , anebo alelu \mathbf{a} . Jejich výskyt ve vyšetřované populaci je určen relativními četnostmi (pravděpodobnostmi) $p = p[\mathbf{A}] > 0$, resp. $q = q[\mathbf{a}] > 0$, pro něž platí vztah

$$p[\mathbf{A}] + q[\mathbf{a}] = 1. \quad (1)$$

Symbolsy v hranatých závorkách budou podle tradice nadále označovat, kterých alel se týká kvantitativní údaj před závorkou.

Nepohlavní buňky⁶) obsahují vždy jednu alelu od matky a jednu od otce. Nechť je vyšetřovaná populace dostatečně velká a nechť je *panmiktická*, což znamená, že je v ní zaručen naprosto náhodný výběr křížení partnerů. Dále předpokládejme, že v ní nedochází k žádným migracím ani mutacím, že průměrná plodnost všech jedinců je stejná atp. Potom pravděpodobnost vzniku jedince s dvojicí alel \mathbf{AA} je zřejmě rovna součinu $p[\mathbf{A}] \cdot p[\mathbf{A}]$. Podobně zjistíme pravděpodobnosti vzniku jedinců s dvojicemi alel \mathbf{Aa} , \mathbf{aA} , resp. \mathbf{aa} . Položme proto

$$p^2[\mathbf{AA}] = p[\mathbf{A}] \cdot p[\mathbf{A}], \quad pq[\mathbf{Aa}] = qp[\mathbf{aA}] = p[\mathbf{A}] \cdot q[\mathbf{a}], \quad q^2[\mathbf{aa}] = q[\mathbf{a}] \cdot q[\mathbf{a}]. \quad (2)$$

Umocněním rovnice (1) na druhou a pomocí (2) dostaneme rovnici

$$p^2[\mathbf{AA}] + 2pq[\mathbf{Aa}] + q^2[\mathbf{aa}] = 1.$$

⁴) Termín gen zavedl W. Johannsen již v roce 1909.

⁵) Na objevu krevních skupin se významně podílel český lékař Jan Janský (1873–1921).

⁶) Všechny nepohlavní buňky určitého jedince obsahují teoreticky naprosto totožnou genetickou informaci. Naproti tomu libovolné dvě pohlavní buňky téhož jedince obsahují obecně různou genetickou informaci, což je způsobeno zejména tzv. překřížením chromozómů při vzniku pohlavních buněk. Je pozoruhodné, že překřížení nerespektuje hranice jednotlivých genů.

Užitím vztahů (2) lze pak určit relativní četnosti alel \mathbf{A} , resp. \mathbf{a} v další generaci. Ty jsou podle (1) dány vztahy (srov. obr. 2)

$$p^2 + 2pq/2 = p(p + q) = p, \quad \text{resp.} \quad q^2 + 2pq/2 = q(q + p) = q.$$

Odtud plyne, že relativní četnosti výskytu obou alel zůstávají stejné jako původní relativní četnosti a alely pohlavních buněk nových jedinců budou opět splňovat vztah (1). Této vlastnosti se říká *Hardyův-Weinbergův zákon* (viz [16]). Např. v evropské populaci má necelých 17 % občanů negativní Rh faktor⁷⁾ a tato hodnota se v důsledku zákona Hardyova-Weinbergova v čase příliš nemění.

Existuje však genový posun (drift), což je změna relativních četností alel v následujících generacích, ke které může dojít např. vlivem mutací způsobených náhodnými jevy vnějšího prostředí. Genový posun předpokládal již Hardy ve své práci [15]. Odhad velikosti genového posunu je stanoven v knize [11].

3. Genetické kódování

Genetický kód představuje jistý recept, jak vytvářet proteiny (tj. bílkoviny) z aminokyselin. Růst živých organismů se obvykle uskutečňuje pomocí tří procesů, které vedou k syntéze potřebných proteinů. Jsou to: replikace, transkripce a translace. Tyto procesy lze vyjádřit jako zobrazení určitých typů. Pro jejich zavedení položíme

$$D = \{A, T, C, G\}, \quad R = \{A, U, C, G\},$$

kde A, T, C, G jsou již dříve uvedené nukleové báze a symbol U označuje další nukleovou bázi — *uracyl*. Množina D představuje soubor všech nukleových bází kyseliny deoxyribonukleové DNA, množina R udává soubor všech nukleových bází⁸⁾ další důležité kyseliny — kyseliny ribonukleové RNA.

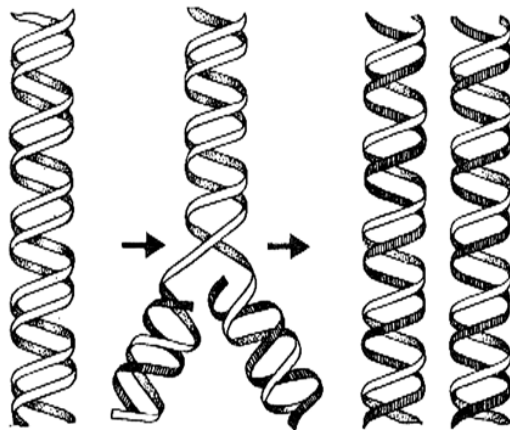
Celý proces kódování probíhá, zhruba řečeno, pomocí následujícího schématu, které modeluje složitý třístupňový proces syntézy bílkovin v živých organismech:

$$D \xrightarrow{\mathcal{R}} D \xrightarrow{\mathcal{T}} R, \quad R \times R \times R \xrightarrow{t} \mathcal{A}, \quad (3)$$

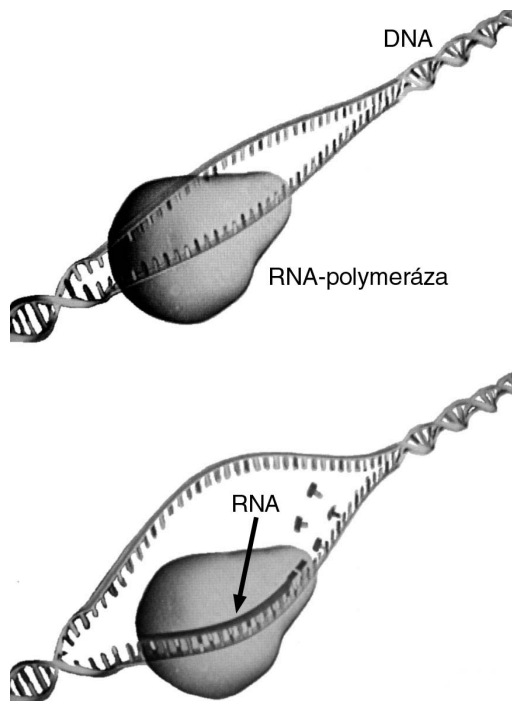
kde symboly $\mathcal{R}, \mathcal{T}, t$ označují postupně replikaci, transkripci, translaci a \mathcal{A} je množina slov obsažených v tabulce 1. Všechny tyto pojmy si nyní podrobně zavedeme.

⁷⁾ Mají-li oba rodiče pozitivní Rh faktor, může jejich dítě mít Rh faktor negativní.

⁸⁾ Nukleovými bázemi některých ribonukleových kyselin RNA mohou být i další tzv. minoritní báze, jako například pseudouridin či ribotimidin. Tyto nukleotidy ale nekódují aminokyseliny řetězce proteinu. Jejich význam spočívá pouze v tom, že se účastní procesu syntézy (včlenění) aminokyselin do proteinového řetězce.



Obr. 3. Schematické znázornění replikace DNA při dělení bakteriálních buněk (viz [10]). Dvojšroubovice se nejprve postupně rozvolňuje rušením vodíkových můstků mezi jednotlivými dvojicemi nukleových bází pomocí enzymu DNA-polymerázy. Poté se oba volné řetězce doplňují komplementárními bázemi a vznikají tak dvě totožné kopie původní dvojšroubovice.



Obr. 4. Při procesu translace nejprve enzym RNA-polymeráza „otevře“ část dvojšroubovice DNA. Poté z volně plovoucích nukleotidů postupně sestavuje řetězec RNA podle informace uložené v DNA.

Definice 1. *Replikace* je zobrazení $\mathcal{R}: D \rightarrow D$ definované předpisem $\mathcal{R}(A) = T$, $\mathcal{R}(T) = A$, $\mathcal{R}(C) = G$ a $\mathcal{R}(G) = C$.

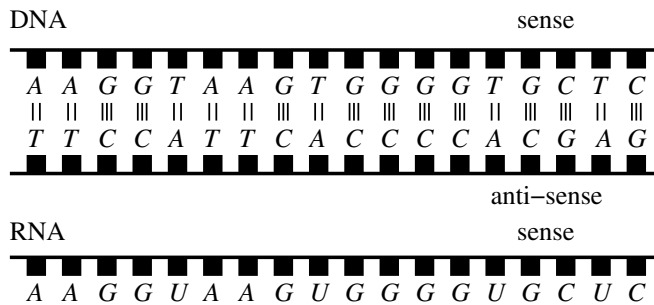
Pomocí replikace \mathcal{R} dochází k vytvoření druhého vlákna kyseliny DNA, čímž tato kyselina dostává tvar dvojšroubovice (viz obr. 3).

Nukleové báze A a T se proto nazývají *vzájemně komplementární*, stejně jako báze C a G .⁹⁾ Zřejmě $\mathcal{R} \circ \mathcal{R}$ je identické zobrazení množiny D na sebe.

Další proces — transkripce — znamená vlastně jakýsi „překlad“ informace z jazyka kyselin DNA do jazyka kyselin RNA (viz obr. 4 a 5). Pomocí něj vznikají všechny typy RNA, např. informační RNA, mitochondriální RNA, ribozomální RNA, transferová RNA (srov. obr. 6).

Definice 2. *Transkripce* je zobrazení $\mathcal{T}: D \rightarrow R$ definované předpisem $\mathcal{T}(A) = A$, $\mathcal{T}(T) = U$, $\mathcal{T}(C) = C$ a $\mathcal{T}(G) = G$.

Poznamenejme, že jen jednomu z vláken dvojšroubovice DNA odpovídá řetězec syntetizovaných aminokyselin. Toto vlákno se anglicky nazývá *sense*. Druhé komplementární vlákno se anglicky nazývá¹⁰⁾ *anti-sense* a podle něj se vlastně vytváří „sense“ vlákno RNA.¹¹⁾



Obr. 5. Schematické znázornění procesu transkripce, tj. přepisu genetické informace z molekuly DNA do molekuly RNA.

Jakmile je vlákno RNA syntetizováno, přesune se z jádra buňky do některého z ribozómů, kterých je v každé lidské buňce přibližně 50 000. Zde probíhá poslední a zároveň nejdůležitější fáze procesu kódování — translace. Během ní dochází k definitivnímu určení pořadí aminokyselin v proteinovém řetězci. Celkový počet základních (levotočivých) bílkovinných aminokyselin vyskytujících se v proteinech je 20. Jejich seznam je obsažen v tabulce 1. Označme symbolem \mathcal{A} množinu těchto dvaceti

⁹⁾ Již v roce 1949 Erwin Chargaff zjistil, že počet bází adeninu, resp. cytosinu je vždy stejný jako počet bází thyminu, resp. guaninu bez ohledu na množství a původ DNA.

¹⁰⁾ Odpovídající české názvy „smysluplné a protismysluplné vlákno“ se zatím neujaly.

¹¹⁾ Roku 1970 uveřejnili H. M. Temin a D. Baltimore zajímavé zjištění, že rakovinnotvorný virus Rousova sarkomu může informaci uloženou ve své RNA předat vyšším organismům následujícím způsobem: virová RNA \rightarrow DNA \rightarrow RNA \rightarrow protein. Zobrazení r : virová RNA \rightarrow DNA se nazývá *reverzní transkripce*, protože platí $r^{-1} = \mathcal{T}$.

aminokyselin a tří slov amber, ochre, opal z tabulky 1, jejichž význam objasníme později. Pak je proces translace definován takto:

Definice 3. *Translaci* budeme rozumět zobrazení $t: R \times R \times R \rightarrow \mathcal{A}$ definované předpisem daným tabulkou 1.

TAB. 1. Genetický kód. Nukleotidy molekuly ribonukleové kyseliny RNA jsou označeny začátečními písmeny příslušných bází U, C, A, G . Každá trojice bází kóduje určitou aminokyselinu nebo označuje počátek či ukončení syntézy proteinu. Každý protein je kódován jedním genem.

| | | Druhý nukleotid | | | | | |
|-----------------|-----|--|--|--|---|--------------------------|-----------------|
| | | U | C | A | G | | |
| První nukleotid | U | UUU fenylnalanin UUC UUA leucin UUG | UCU UCC serin UCA UCG | UAU tyrosin UAC UAA ochre UAG amber | UGU cystein UGC UGA opal UGG tryptofan | U C A G | Třetí nukleotid |
| | C | CUU CUC leucin CUA CUG | CCU CCC prolin CCA CCG | CAU histidin CAC CAA glutamin CAG | CGU CGC arginin CGA CGG | U C A G | |
| | A | AUU AUC isoleucin AUA AUG metionin | ACU ACC treonin ACA ACG | AAU asparagin AAC AAA lysin AAG | AGU serin AGC AGA arginin AGG | U C A G | |
| | G | GUU GUC valin GUA GUG | GCU GCC alanin GCA GCG | GAU asparát GAC GAA glutamát GAG | GGU GGC glycin GGA GGG | U C A G | |

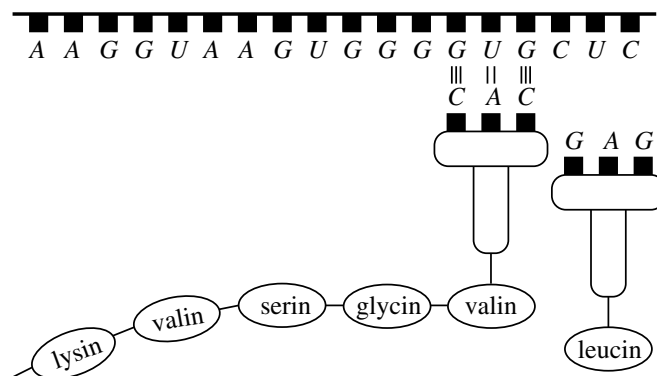
Za včlenění určité aminokyseliny do proteinového řetězce odpovídá vždy pouze trojice nukleotidů (viz obr. 6). Tuto trojici nazýváme *triplet*. Genetický kód je tedy tripletní, čemuž odpovídá definice translace. Celkem existuje $64 = 4^3$ vzájemně různých tripletů, a proto zobrazení translace t není injektivní. Většina aminokyselin je kódována více triplety. Například isoleucin je kódován třemi různými triplety: AUU nebo AUC nebo AUA (viz tab. 1). Arginin, leucin a serin jsou kódovány dokonce šesti různými triplety. Všechny kódovací možnosti pro určitou aminokyselinu ale nejsou stejně často využívány. Například u známé bakterie *Escherichia coli* je lysin kódován v 85 % tripletem AAA a pouze v 15 % AAG . U jiných organismů se tento poměr může lišit.

Dva speciální triplety AUG , resp. GUG kódují metionin, resp. valin, současně však mohou kódovat i začátek syntézy proteinu (např. u virů). Proto se nazývají *iniciační kodony*. Naproti tomu triplet UAG , UAA a UGA (označované postupně jako amber, ochre a opal — viz tab. 1) kódují ukončení syntézy a neodpovídají žádné aminokyselině.¹²⁾ Nazývají se *terminační kodony* a mají tak vlastně význam

¹²⁾ Z tohoto pravidla ale existuje drobná výjimka. Triplet UGA kóduje u jedné skupiny bakterií další jedenadvacátou aminokyselinu — selenocystein.

čárky v genetické abecedě. Jsou znamením, že syntetizovaná molekula proteinu je již kompletní. Navíc pro stanovení správného směru čtení informace z řetězce informační RNA je na jednom z jeho konců fosfátová a na druhém hydroxylová skupina.

Informace obsažená v DNA v jádře buňky se nazývá *genom*. Zde je ale nutno zdůraznit, že jen 5 až 7% této informace (tzv. exon) má skutečný význam pro tvorbu bílkovin a další činnost buňky. Zbylá část představuje jistou nadbytečnou informaci, která je důsledkem složité evoluce a kódování se neúčastní.



Obr. 6. Schematické znázornění syntézy bílkovin v ribozómech. Polypeptidový řetězec při procesu translace postupně narůstá o aminokyseliny určené pořadím tripletů v molekule informační RNA. Přenos aminokyselin zajišťují poměrně malé molekuly transferové RNA, kterých existuje více než 20 druhů.

Z matematických relací týkajících se vzájemných vztahů mezi nukleovými bázemi kyseliny RNA a aminokyselin proteinů stanovených známým fyzikem G. Gamowem v práci [12] a z práce [13] J. S. Griffitha, F. H. C. Cricka a L. E. Orgela vyplynulo, že genetický kód by měl být tripletní a měl by se načítat na RNA postupně a jen po trojicích. Genetický kód (viz tab. 1) byl pak definitivně stanoven r. 1966 na sympoziu v Cold Spring Harboru (USA) v referátu M. W. Nirenberga¹³⁾ sestavením tabulky tripletů pro všechny aminokyseliny.

Genetický kód z tabulky 1 je univerzální v tom smyslu, že platí (až na drobné nuance) v rámci celé organické říše na Zemi, tj. počínaje viry a konče člověkem. Soubor genů může být lokalizován nejen v jádře buňky, ale též v mitochondriích a v chloroplastech. Soubor genů v mitochondriích má prstencové uspořádání a u člověka obsahuje 37 genů. V genetickém kódu mitochondrií byly ale zjištěny 4 tripletety, které mají rozdílný význam proti genomu jádra (viz tab. 1). Jsou zde i rozdíly v iniciačních a terminačních kodonech.

Závěrem tohoto odstavce je třeba se ještě zmínit o významu data 12. února roku 2001 pro další rozvoj genetiky. Tento den John Craig Venter předložil tabulku lidského genomu. Určil, že lidský genom sestává přibližně z 35 000 genů a že jeho skladba

¹³⁾ Za tento vynikající výsledek získal M. W. Nirenberg Nobelovu cenu v roce 1968.

u všech lidí je v 99,8 % identická.¹⁴⁾ Definitivní verze tabulky lidského genomu byla ovšem dokončena až v dubnu 2003, tj. právě 50 let od publikování zásadního a zcela fundamentálního článku [30] v dubnovém čísle časopisu Nature v roce 1953. V současné době víme, že DNA v jádře každé lidské (nepohlavní) buňky obsahuje zhruba tři miliardy párů vzájemně komplementárních bází (viz [18]). Tedy 0,2 % lidského genomu tvoří 6 milionů párů bází, v nichž se jednotliví lidé od sebe vzájemně liší (kromě jednovaječných sourozenců, kteří mají identické genomy). Přitom je třeba zdůraznit, že změna byť jen jediného páru může způsobit i velmi těžké onemocnění.

Praktickým uplatněním poznatků molekulární genetiky se zabývá genové inženýrství. Dokonalá znalost struktury molekuly DNA má obrovské množství aplikací: při určování otcovství, při studiu dědičných chorob, při transplantaci orgánů, při výrobě geneticky modifikovaných potravin, v kriminalistice, klonování, při biosyntéze léků (např. inzulínu), při genové terapii, při klasifikaci genů v genetických bankách, při zamezování šíření epidemií, při odstraňování toxických havárií pomocí mikroorganizmů, při určování vývojových linií jednotlivých živočišných a rostlinných druhů apod. Také analogie známého překřížení (crossing-over) chromozómů či jejich mutací našly uplatnění v tzv. genetických algoritmech (viz např. [23]), které slouží k hledání globálního minima (maxima) složitých účelových funkcí.

4. Genetický kód a volné monoidy

Genetický kód je vyjádřen pořadím tripletů nukleotidů kyseliny RNA, jímž je určeno pořadí aminokyselin v proteinu. Z matematického hlediska odpovídá genetickému kódu zobrazení translace t . Čtení genetického kódu (dekódování) se řídí pořadím tripletů daného genu

$$T_1 T_2 \dots T_m, \quad m \in \{1, 2, \dots\}. \quad (4)$$

Tomu pak odpovídá konečná posloupnost

$$t(T_1)t(T_2) \dots t(T_m) \quad (5)$$

charakterizující řetězec aminokyselin proteinu. Počátečním tripletem T_1 bývá většinou triplet AUG (resp. GUG), který kóduje aminokyselinu metionin (resp. valin). Tvar (5) naznačuje, že vhodným algebraickým objektem pro studium kódování aminokyselin mohou být monoidy, přesněji řečeno volné monoidy, a jejich generátory. Uvedeme proto nyní některé definice a pojmy týkající se této oblasti (viz např. [7]).

Definice 4. *Monoidem* nazveme trojici (M, \cdot, e) , kde M je neprázdná množina s binární operací $\cdot : M \times M \rightarrow M$, která je asociativní a uzavřená v M , a prvek $e \in M$ je *neutrální (jednotkový)* prvek množiny M , tj.

$$x \cdot e = e \cdot x = x \quad \text{pro všechna } x \in M. \quad (6)$$

¹⁴⁾ Přitom lidský genom se shoduje s genomem myši domácí v 99 % a s genomem šimpanze v 99,4 %.

Monoid tvoří např. všechna přirozená čísla s operací násobení a neutrálním prvkem 1. Obecně ale monoid nemusí být komutativní. Příkladem nekomutativního monoidu je množina všech bijektivních zobrazení množiny X na sebe. Binární operací je v tomto případě skládání zobrazení a neutrální prvek je identické zobrazení. Jiným příkladem nekomutativního monoidu je množina všech binárních relací na množině X , v níž binární operací je skládání relací.

Definice 5. Necht (M, \cdot, e) a $(\overline{M}, \bullet, \overline{e})$ jsou monoidy. Zobrazení $h: M \rightarrow \overline{M}$ nazveme *homomorfizmem*, jestliže platí

$$h(x \cdot y) = h(x) \bullet h(y) \quad \text{pro všechna } x, y \in M. \quad (7)$$

Z rovnosti (6) okamžitě plyne, že neutrální prvek monoidu je určen jednoznačně. Odtud a pomocí vztahů (6) a (7) lze snadno dokázat, že $h(e) = \overline{e}$.

Je-li homomorfizmus h bijektivní zobrazení, pak h se nazývá *izomorfizmus* a ze (7) můžeme přímo odvodit, že h^{-1} je též homomorfizmus. Injektivní, resp. surjektivní homomorfizmus se nazývá *monomorfizmus*, resp. *epimorfizmus*.

Jako příklad můžeme uvést následující tvrzení (viz [17]). Pro jednoduchost budeme nadále monoid (M, \cdot, e) značit jen M .

Věta 1. Každý monoid M je izomorfní monoidu všech zobrazení $F_a: M \rightarrow M$ tvaru $F_a(x) = a \cdot x$, $a \in M$, s operací skládání.

Zavedme nyní některé další nezbytné pojmy.

Definice 6. Podmonoidem (nebo též submonoidem) monoidu M nazveme podmnožinu $M' \subset M$ splňující následující podmínky:

- (i) je-li $x, y \in M'$, pak $x \cdot y \in M'$,
- (ii) je-li e neutrální prvek M , pak $e \in M'$.

Již v roce 1926 Thomas H. Morgan¹⁵⁾ ukázal, že geny jsou uloženy v chromozómu za sebou v lineárním pořadí (tzv. první Morganův zákon). Tento fakt, vlastní struktura genu (4) a jemu odpovídající řetězec aminokyselin (5) předurčují, jak již bylo řečeno, volný monoid (viz definice 7) jako jednu z možných algebraických struktur pro vyšetřování vlastností genetického kódování. Jinou možností jsou tzv. blokové kódy s pevnou délkou kódového slova rovnou třem.

Necht X je libovolná neprázdná množina. Slovem w nad X nazveme libovolnou konečnou posloupnost $w = (x_1, x_2, \dots, x_n)$, kde $x_i \in X$, $i = 1, 2, \dots, n$. Číslo n označuje délku slova. Symbolem w_0 označme slovo odpovídající prázdné množině, jehož délka je $n = 0$. Množinu všech slov nad X označíme $F(X)$.

¹⁵⁾ T. H. Morgan získal v roce 1933 Nobelovu cenu za vysvětlení funkce chromozómů při dědičnosti. Objevil mj. také chromozómy určující pohlaví jedinců. Své pokusy prováděl na mušce *Drosophila*, která se poté stala významným pokusným objektem v genetice.

Definice 7. *Volným monoidem nad X* nazveme množinu slov $F(X)$, ve které je binární operace $\circ: F(X) \times F(X) \rightarrow F(X)$ definována vztahem:

$$w_1 \circ w_2 = (x_1, x_2, \dots, x_n) \circ (y_1, y_2, \dots, y_m) = (x_1, \dots, x_n, y_1, \dots, y_m)$$

pro $w_1 = (x_1, x_2, \dots, x_n)$, $w_2 = (y_1, y_2, \dots, y_m)$, $w_1, w_2 \in F(X)$.

Vidíme, že $F(X)$ skutečně reprezentuje monoid (viz definice 4). Slovo w_0 je totiž neutrálním prvkem $F(X)$, neboť $w \circ w_0 = w_0 \circ w = w$ pro každé $w \in F(X)$. Snadno se také ověří, že operace \circ je asociativní (ale evidentně není komutativní). Množina X se nazývá *množinou generátorů volného monoidu $F(X)$* .

Kódování je, obecně řečeno, zobrazení $K: Z \rightarrow B$, kde Z je *množina zdrojů* a B je *kódová abeceda*. Prvky B nazýváme *kódové znaky* a množinu $K(Z)$ nazýváme stručně *kódem*.

Algebraickou teorii kódování se zabývají např. práce [1], [4], [13]. Genetické kódování vychází z předpokladu, že $Z = R \times R \times R$, $B = \mathcal{A}$ a kódování $K = t$ je translace (viz definice 3). Prvky volného monoidu $F(R \times R \times R)$ jsou tedy sekvence tripletů nukleových bází z množiny R a prvky volného monoidu $F(\mathcal{A})$ jsou sekvence jim příslušejících aminokyselin a tří slov (amber, ochre, opal) odpovídajících terminačním kodonům.

Všimněme si, že zobrazení $t^*: F(R \times R \times R) \rightarrow F(\mathcal{A})$, které je přirozeným rozšířením translace t , je epimorfizmem volného monoidu $F(R \times R \times R)$ na volný monoid $F(\mathcal{A})$. Bez důkazů nyní uvedeme některá tvrzení (viz [3], [17]), která jsou důležitá pro porozumění popisu genetického kódování pomocí volných monoidů. Symbolem M_n budeme nadále označovat volný monoid s $n \in \{1, 2, \dots\}$ generátory.

Věta 2. *Volný monoid M_{20} (resp. M_{23}) lze nekonečně mnoha způsoby izomorfně zobrazit na podmonoid volného monoidu M_4 se čtyřmi generátory (srov. tab. 1).*

Věta 3. *Nechť M je monoid, který má konečný počet generátorů. Pak každý podmonoid monoidu M má rovněž konečný počet generátorů.*

Nechť M a \overline{M} jsou monoidy a necht' $h: M \rightarrow \overline{M}$ je monomorfismus. Podmonoid $h(M)$ monoidu \overline{M} se nazývá *vložením monoidu M do \overline{M}* .

Věta 4. *Libovolný volný monoid M_n lze vložit do volného monoidu M_2 vytvořeného jen dvěma generátory.*

Důkaz této věty, která je důsledkem Evansovy věty, lze nalézt např. v [9, kap. 9]. Věta 4 ukazuje, jakou důležitost mají tzv. *binární kódy* (tj. takové kódy, kde za množinu zdrojů většinou volíme $Z = \{0, 1\}$). Dává nám pozitivní odpověď na jednu ze základních otázek, totiž zda lze genetický kód redukovat na kód binární. K tomu poznamenejme, že člověk objevil dvojkovou soustavu již asi před 3000 lety (viz např. PMFA 42 (1997), s. 224), i když její praktické využití přišlo až ve 20. století. Příroda ale objevila dvojkovou nebo, chcete-li, čtyřkovou soustavu před více než 3,5 miliardami

let. Genetická informace v jádře buňky je totiž zaznamenána digitálně¹⁶), provedeme-li např. toto ztotožnění

$$A = 00, \quad C = 01, \quad G = 10, \quad T = 11.$$

Horní řetězec z obrázku 1 pak odpovídá dvojkovému zápisu ...001011010111... . Tři miliardy párů nukleových bází lidské buňky odpovídají tedy šesti miliardám bitů (= 750 MB), což je zhruba kapacita jednoho CD.

V genetice je užitečné popisovat kódování pomocí monomorfizmů. Z historického hlediska je proto zajímavé se ještě zmínit o dalším pojmu — „kódu bez čárky“, který po objevu struktury DNA vytvořil F. H. C. Crick se svými spolupracovníky [13], viz též [25]. Abychom tento kód objasnili, uvedeme nejprve jednu definici a větu:

Definice 8. Budeme říkat, že slovo w volného monoidu M je obsaženo ve slově $\bar{w} \in M$, existují-li slova $w_1, w_2 \in M$ tak, že platí $\bar{w} = w_1 \circ w \circ w_2$.

Slova w_1 , resp. w_2 mohou mít např. význam iniciačního, resp. terminačního tripletu. Důkaz následujícího tvrzení (viz [3]) je jednoduchý.

Věta 5. Nechtě M_m , resp. M_n jsou volné monoidy generované m , resp. n prvky a nechtě $h: M_m \rightarrow M_n$ je monomorfizmus M_m do M_n . Pak $h(M_m)$ je volným podmonoidem monoidu M_n a zobrazení h je izomorfizmem monoidu M_m na monoid $h(M_m)$.

Odtud okamžitě plyne, že když slovo w volného monoidu M_m je obsaženo ve slově \bar{w} téhož monoidu M_m , pak slovo $h(w)$ volného monoidu $h(M_m)$ je obsaženo ve slově $h(\bar{w})$ monoidu $h(M_m)$. Množina slov $h(w)$ s touto vlastností tvoří tzv. kód bez čárky.

Zavedení pojmu kódu bez čárky vyjadřuje shodu některých algebraických výsledků s biologickými experimenty. Jiné výsledky naopak ukazují na drobný nesoulad jednoduchých matematických modelů s biochemickými reakcemi. Například v molekule DNA bývá několik procent nukleových bází metylováno (cytosin je zde nahrazen 5-metylcytosinem apod.). Metylace bází mění jejich funkční vlastnosti. Je to jeden z mechanismů, který se uplatňuje při regulaci funkcí jednotlivých genů.

5. Ergodické submonoidy

Úvodem toho odstavce poznamenejme, že chováním fyzikálních systémů, které mají velmi mnoho stupňů volnosti, se zabývá tzv. ergodická teorie. Pro krátkou a výstižnou informaci o ní odkazujeme čtenáře na přehlednou Halmosovu knížku [14].

Nechtě M_n je volný monoid s konečným počtem generátorů g_1, g_2, \dots, g_n a nechtě w je slovo v M_n . Dále označme $\lambda(w)$ délku slova w a nechtě $\lambda_i(w)$ je počet výskytů prvku g_i objevujícího se ve slově w . Odtud plyne, že $\sum_{i=1}^n \lambda_i(w) = \lambda(w)$. Poměr

¹⁶) Naproti tomu informace přicházející do našeho mozku se uchovávají analogově.

$P_i(w) = \lambda_i(w)/\lambda(w)$ pro $w \neq w_0$ se nazývá *frekvencí prvku g_i ve slově w* . Číslo $P_i(w)$ je zřejmě racionální, $0 \leq P_i(w) \leq 1$ pro $i = 1, 2, \dots, n$ a platí, že

$$\sum_{i=1}^n P_i(w) = 1.$$

Pokud například za w zvolíme slovo příslušející částečnému řetězci pěti aminokyselin z obrázku 6, pak $\lambda(w) = 5$, $\lambda_6(w) = 2$, $\lambda_{17}(w) = 1$, kde index $i = 6$ odpovídá valinu, $i = 17$ lysinu apod. (srov. tab. 1).

Pro daná racionální čísla $P_i \in [0, 1]$, $i \in \{1, \dots, n\}$, označme symbolem E_{P_i} množinu obsahující prvek w_0 a všechna slova $w \in M_n$ s frekvencí $P_i = P_i(w)$. Pak platí následující tvrzení (viz [3]):

Věta 6. *Pro každé $i = 1, \dots, n$ je množina E_{P_i} submonoid volného monoidu M_n .*

Důkaz. (i) Předpokládejme, že $w_1, w_2 \in E_{P_i}$, a necht' $w = w_1 \circ w_2 \neq w_0$ (pro triviální submonoid volného monoidu M_n je tvrzení věty zřejmé). Pak platí $\lambda_i(w_1) = P_i\lambda(w_1)$ a $\lambda_i(w_2) = P_i\lambda(w_2)$. Odtud plyne, že

$$\frac{\lambda_i(w)}{\lambda(w)} = \frac{\lambda_i(w_1 \circ w_2)}{\lambda(w_1 \circ w_2)} = \frac{\lambda_i(w_1) + \lambda_i(w_2)}{\lambda(w_1) + \lambda(w_2)} = \frac{P_i\lambda(w_1) + P_i\lambda(w_2)}{\lambda(w_1) + \lambda(w_2)} = P_i,$$

a tedy $w_1 \circ w_2 \in E_{P_i}$. Asociativita operace \circ na množině E_{P_i} okamžitě plyne z inkluze $E_{P_i} \subset M_n$.

(ii) Prázdné slovo označené symbolem w_0 je neutrálním prvkem monoidu M_n . Proto platí rovnice $w \circ w_0 = w_0 \circ w = w$ pro každé $w \in E_{P_i}$. Odtud plyne, že $w_0 \in E_{P_i}$ a podle definice 6 je E_{P_i} submonoid volného monoidu M_n . \square

V případě genetického kódu je zdrojem biologické informace množina všech tripletů nukleových bází množiny R . Zavedeme nyní následující definici:

Definice 9. Zdroj informace Z nazveme *ergodickým*, jestliže existují racionální čísla P_i taková, že

$$P_i \geq 0, \quad i = 1, 2, \dots, n, \quad \sum_{i=1}^n P_i = 1$$

a přitom platí

$$\frac{\lambda_i(w)}{\lambda(w)} \rightarrow P_i \quad \text{pro } \lambda(w) \rightarrow \infty.$$

Bylo zjištěno, že pro genetický kód, který je dán zobrazením t ze vztahu (3) a tabulkou 1, platí

$$\frac{\lambda_i(w)}{\lambda(w)} \approx P_i$$

pro dostatečně dlouhá slova w volného monoidu vytvořeného množinou Z . Jeho zdroj má tedy ergodický charakter. Rovněž další přirozené i umělé kódy mohou mít svůj ergodický zdroj. Tak například každý jazyk může být charakterizován svou ergodickou povahou (viz [2]).

6. Závěrečné poznámky

Literatury zabývající se genetickým kódováním je velké množství. Lze doporučit například knihy [3], [8], [10], [11], [14], [19], [22], [24], [29] a [33].

Zmíňme se ještě krátce o další aplikaci matematiky v populační genetice. Počátky této vědní disciplíny sahají do dvacátých let minulého století a jsou spojeny se vznikem specializovaných matematických týmů, které se zabývaly výzkumem v genetice. Matematickým objektem použitým v populační genetice jsou genetické algebry, které jsou komutativní, nikoliv však asociativní. Základem jejich definice je pravděpodobnostní hledisko. Pojednání o nich by si ale vyžádalo další samostatný článek. Pro stručnou informaci o genetických algebrách a o příbuzných otázkách viz práce [5], [6] a [20].

Zvláštní pozornost zasluhuje také kniha [7], která obsahuje mimo jiné aplikace monoidů v teorii automatů, číslicových počítačů a umělých programovacích jazyků. Poznamenejme, že molekulární systém RNA představuje též konečný automat, jehož množinou vstupních signálů jsou 4 nukleotidy U , A , C , G a jehož množinou výstupů jsou aminokyseliny a slova označující ukončení syntézy. Množina stavů S tohoto automatu odpovídá procesu biosyntézy proteinů.

Ke studiu molekuly DNA přispívá v poslední době též topologie. Určité enzymy totiž dokáží měnit topologickou strukturu prstencové DNA. Vytvářejí na ní uzly (nebo je odstraňují), a tak mění její vlastnosti (viz [27], [28]). Občas také může dojít k poškození či přetržení molekuly DNA, např. tím, že do ní narazí nějaká vysokoenergetická elementární částice. V jádře buňky však existují složité enzymatické samoopravující procesy, které dokáží dvojšroubovici DNA opět správně napojit a zrekonstruovat poškozenou informaci, pokud vzniklé poškození není příliš velké. Opět vidíme, že příroda objevila samoopravující se kódy mnohem dříve než člověk (srov. PMFA 44 (1999), s. 18). Navíc každý gen se obvykle v řetězci DNA několikrát opakuje, aby se při poškození jedné jeho kopie mohly použít další.

Po přečtení tohoto článku si čtenář jistě položí otázku, kde se takové složité molekuly, jaké mají nukleové kyseliny, a informační procesy s nimi spojené ve vesmíru vzaly. Odpověď je nasnadě. V pozorovatelné části vesmíru je více než 10^{11} galaxií, z nichž každá obsahuje alespoň 10^{11} hvězd. Kolem hvězd obíhají menší tělesa (planety a jejich měsíce, planetky, komety) a na některých z nich panují podmínky vhodné pro život. Primitivní život na Zemi se objevil přibližně před 3,5 miliardami let¹⁷⁾, o čemž svědčí četné paleontologické nálezy. Příroda tedy měla od okamžiku velkého třesku¹⁸⁾ přibližně 10 miliard let ($\approx 4 \cdot 10^{17}$ s) času na experimentování spojené se vznikem života. Přitom stačilo, aby vznikla nějaká samoreplikující se molekula, která se mutacemi mohla dále „zdokonalovat“. Také je třeba si uvědomit, že povrch naší Země má zhruba $500\,000\,000\,000\,000\,000\text{ m}^2$, což, jak jistě uznáte, je nesmírně obrovská (bio)chemická laboratoř. Kyselina RNA je vývojově starší než DNA. Vznikla asi o jednu miliardu let dříve než DNA. Darwinova evoluční teorie o vzniku druhů

¹⁷⁾ Zatím však stále není zcela vyloučena ani možnost, že život mohl být na naši planetu zanesen odjinud z vesmíru.

¹⁸⁾ Stáří vesmíru se podle současných měření odhaduje na $13,5 \cdot 10^9$ let.

přírodním výběrem v konkurenčním prostředí pak vysvětluje, proč vůbec existuje dnešní vyspělá lidská civilizace.

Poděkování. Autoři děkují doc. RNDr. H. ILLNEROVÉ, prof. RNDr. V. PAČESOVÍ, DrSc., RNDr. P. PUDLÁKOVÍ, DrSc., a prof. RNDr. I. RAŠKOVÍ, DrSc., za cenné diskuse a dále Mgr. H. HOLOVSKÉ, Mgr. Š. KASALOVÍ a Ing. P. KRÍŽKOVÍ za technickou pomoc při zpracování tohoto článku. Práce byla podpořena grantem A 1019201 GA AV ČR.

L i t e r a t u r a

- [1] ADÁMEK, J.: *Kódování*. SNTL, Praha 1982.
- [2] BEADLE, G. W.: *The Role of the Nucleins in Heredity*. Baltimore 1957.
- [3] BEIER, W.: *Biophysik*. VEB, Leipzig 1960.
- [4] BERLEKAMP, R. E.: *Algebraic Coding Theory*. McGraw-Hill, New York 1968.
- [5] BERNSTEIN, S. N.: *Demonstration mathématique de la loi d'hérédité de Mendel*. C. R. Acad. Sci. Paris 177 (1923), 528–531.
- [6] BERNSTEIN, S. N.: *Principe de stationarite et generalisation de la loi de Mendel*. C. R. Acad. Sci. Paris 177 (1923), 581–584.
- [7] BIRKHOFF, G., BARTEE, T. C.: *Modern Applied Algebra*. McGraw-Hill, New York-London-Toronto 1970.
- [8] BRESSLER, S. J.: *Molekuljarnaja biologija*. Izd. Nauka, Leningrad 1973.
- [9] CLIFFORD, A. H., PRESTON, G.: *The Algebraic Theory of Semigroups, Vol. II*. American Mathematical Society, Providence 1967.
- [10] DOSTÁL, P., ŘEHÁČEK, Z., DUCHÁČ, V.: *Kapitoly z obecné biologie*. SPN, Praha 1994.
- [11] GALE, J. S.: *Theoretical Population Genetics*. Unwin Hyman, Boston 1990.
- [12] GAMOW, G.: *Possible mathematical relation between deoxyribonucleic acid and protein structures*. Nature 173 (1954), 318.
- [13] GRIFFITH, J. S., CRICK, F. H. C., ORGEL, L. E.: *Codes without commas*. Proc. Nat. Acad. Sci. U.S.A. 43 (1957), 416–427.
- [14] HALMOS, P. R.: *Lectures on Ergodic Theory*. The Math. Soc. of Japan, Tokyo 1956 (ruský překlad Izd. inostranoj lit., Moskva 1960).
- [15] HARDY, G. H.: *Mendelian proportions in a mixed population*. Science 28, N 706 (1908), 49–50.
- [16] HATINA, J., SYKES, B.: *Lékařská genetika: Problémy a přístupy*. Academia, Praha 2002.
- [17] LALLEMENT, G.: *Semigroups and Combinatorial Applications*. John Wiley & Sons, New York 1979.
- [18] LANDER, E. S.: *Mapping heredity: Using probabilistic models and algorithms to map genes and genomes*. Notices Amer. Math. Soc. 42 (1995), 747–753.
- [19] LEWIN, B.: *Genes VII*. Oxford Univ. Press, Oxford 2000.
- [20] LJUBIČ, J. I.: *Matěmaticeskije struktury v populjacionnoj genetike*. Naukova Dumka, Kiev 1983.
- [21] MENDEL, J. G.: *Versuche über Pflanzen-Hybriden*. Verh. naturforsch. Verein in Brünn N 4 (1866), 3–47.
- [22] NEČÁSEK, J., CETL, I. a kol.: *Obecná genetika*. SPN, Praha 1979.
- [23] POSPÍCHAL, J., KVASNIČKA, V.: *Genetické algoritmy: nic pro biology*. Pokroky mat. fyz. astronom. 40 (1995), 1–10.
- [24] RAShevsky, N.: *Mathematical Biophysics*. Chicago 1948.
- [25] ROSEN, R.: *The DNA-protein coding problem*. Bull. Math. Biophysics 21 (1959), 71–96.
- [26] VOGEL, F.: *ABO blood groups and disease*. Amer. J. Human Genetics 22 (1970), 464–475.
- [27] VOLOGODSKIJ, A. V.: *Topologija i fizičeskije svojstva kolčevyeh DNK*. Nauka, Moskva 1988.

- [28] WANG, J.: *Les enzymes qui modifient la topologie de l'ADN*. La science des nœuds, Pour la Science, Paris (2001), 130–142.
- [29] WATSON, J. D.: *Molecular biology of the gene*. W. A. Benjamin, Inc., Menlo Park 1976.
- [30] WATSON, J. D., CRICK, F. H. C.: *A structure for deoxyribose nucleic acid*. *Nature* 171 (1953), 737–738.
- [31] WATSON, J. D., CRICK, F. H. C.: *Genetic implications of the structure of deoxyribonucleic acid*. *Nature* 171 (1953), 964–969.
- [32] WEINBERG, W.: *Über den Nachweis der Vererbung beim Menschen*. Jahresber. Ver. vaterl. Naturk. in Württemberg. 64 (1908), 368–372.
- [33] WOESE, C. R.: *The Genetic Code*. Harper and Row, New York 1967.

Z historie inverzního variačního problému: Odvození podmínek silné variačnosti

Jan Kotůlek, Opava

1. Úvod

Variační počet je oblast matematiky, která se mimo jiné zabývá tzv. variačními integrály. Příkladem takového integrálu je

$$\int_{t_1}^{t_2} L(t, q(t), \dot{q}(t)) dt,$$

kde integrand L , který se obvykle nazývá Lagrangeova funkce nebo lagrangián, závisí vedle nezávislé proměnné t na diferencovatelné funkci (křivce) $q(t)$ a její derivaci $\dot{q}(t)$. Variační počet se zajímá především o extremály variačních integrálů, tedy křivky, na nichž je integrál nejmenší (resp. největší). Již J. L. Lagrange odvodil, že pro extremály platí podél křivky $q(t)$ rovnice

$$\frac{\partial L}{\partial q} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}} = 0, \quad \text{kde} \quad \frac{d}{dt} = \frac{\partial}{\partial t} + \frac{\partial}{\partial q} \dot{q} + \frac{\partial}{\partial \dot{q}} \ddot{q}.$$

Mgr. JAN KOTŮLEK (1978), Matematický ústav v Opavě, Slezská univerzita v Opavě, Bezručovo nám. 13, 746 01 Opava, e-mail: Jan.Kotulek@math.slu.cz

Tato práce vznikla za podpory výzkumného záměru MSM 192400002 „Globální analýza“, řešeného v Matematickém ústavu SU v Opavě.