

Hmotná podstata genetických informací – nukleové kyseliny

The Substance of Genetic Information – Nucleic Acids

Brdička R.

Souhrn

V historii molekulární genetiky, nezahrnujeme-li do ní samé počátky identifikace nukleových kyselin po polovině 19. století a poznatky o dědičnosti během následujících cca 70 let, můžeme za skutečně průlomovou dobu v této oblasti považovat konec druhé světové války, kdy byly za hmotnou podstatu dědičnosti uznány nukleové kyseliny. Jako téměř vždy se výzkum, zpočátku pomalý a „tápající“, postupně rozběhl takovou rychlostí a zaplavuje nás neustále takovým množstvím informací, že je již nestačíme dostatečně důkladně zpracovat a náležitě využít. Čas od času se zdálo, že o struktuře nukleových kyselin jako nositelkách genetických informací víme již vše, abychom se zakrátko přesvědčili, že víme sice leccos, ale zdaleka to není dostačující. Dalšího obrovského pokroku jsme dosáhli poznáním struktury molekul deoxyribonukleové kyseliny a také nukleotidové sekvence všech jejích molekul v našich buňkách, čímž jsme konečně přešli do tzv. postgenomické fáze. Poznáním nukleotidové sekvence výzkum lidského genomu neskončil, zvláště v oblasti mezigenových vztahů, regulačních procesů, vnějších vlivů, a stále nám zbývá dost nepoznaného i pro přicházející vědeckou generaci, která bude mít zřetelnou výhodu v technologickém pokroku a zapojení dalších, svým způsobem nových oborů, jako je např. bioinformatika.

Klíčová slova

DNA – RNA – mRNA – tRNA – miRNA – diagnostické metody

Summary

If we look at the history of our knowledge of nucleic acids, we would see in the distant past of 140 years Friedrich Miescher who had identified the acidic substance within the cell nucleus, which he called nuclein. About 70 years after his initial observation, this substance was connected with genetic information. This very substantial finding happened during the World War II. This was the impulse that research of nucleic acids received to speed up continuously growing mountain of information, which is more and more difficult to understand. Another eruption of new information about our genome was the result of ten years of intensive cooperation of many manufacturers divided into two competitive blocks which offered us knowledge of nucleotide sequence of all 46 DNA molecules. The year 2000 became the landmark marking the start of the postgenomic era. It did not mean that human genome was totally explored, but the cornerstone has been settled. Since then, we could concentrate our efforts on variability; use of the project of 1,000 genomes brought many important findings, eg. copy number variability (CNV) exceeds the single nucleotide polymorphisms (SNP). Also intergenomic relationships, studies on function and pathways began to be much more understandable by elucidation of the genome primary structure. NGS as a tool also accelerated the epigenetic research. All this improved molecular diagnostics by discovering many new markers playing their role in disease and treatment and allowed us to enter the field of multifactorial illnesses including cancer. The progress in diagnostic technologies which has happened during the last decade forced our research teams to include other professions – eg. bioinformatics.

Key words

DNA – RNA – mRNA – tRNA – mikroRNAs – diagnostic methods

Autor deklaruje, že v souvislosti s předmětem studie nemá žádné komerční zájmy.

The author declares he has no potential conflicts of interest concerning drugs, products, or services used in the study.

Redakční rada potvrzuje, že rukopis práce splnil ICMJE kritéria pro publikace zaslané do biomedicínských časopisů.

The Editorial Board declares that the manuscript met the ICMJE "uniform requirements" for biomedical papers.



prof. MUDr. Radim Brdička, DrSc.
Pšenčíkova 16
142 00 Praha 12-Kamýk
e-mail: radim.brdicka@gmail.com

Obdrženo/Submitted: 7. 2. 2012

Přijato/Accepted: 12. 3. 2012

Východiska

Při povídání o nukleových kyselinách je nezbytné smířit se s rizikem, že pro někoho půjde o zcela zbytečné sdělení, v němž neshledá nic nového, natož zajímavého. Přesto je možné spoolehnot se na šťastnou náhodu, která přinese autorovi čtenáře, který tam něco nového najde.

Připomeneme-li si tolikrát opakovanou a studentům vštěpovanou historii, pak o nukleových kyselinách víme již téměř od poloviny 19. století, tedy od doby, kdy Smetana skládal Libuši a kdy začínala stavba Národního divadla. Na Miescherův objev „kyselých“ látek v jádrech buněk z roku 1869 navázali až Avery, MacLeod a McCarty a v roce 1944 spojili nukleové kyseliny s genetickou informací.

A posléze již s vědomím, že jde o látku spojenou s dědičností, James Dewey Watson, Francis Crick a Maurice Wilkins vytvořili prostorový model struktury DNA. Nešlechtné opomenutí zásluh jejich spolupracovnice Rosalindy Franklinové kritizovala řada odborníků.

V živých živočišných buňkách, a tedy i v lidských, se nukleové kyseliny vyskytují ve dvou základních podobách, a to jako deoxyribonukleové (DNA) a jako ribonukleové (RNA) kyseliny, které se navzájem liší v mnoha směrech. Po stránce strukturní, což trochu vyplývá již z názvu, jde o odlišnost cukerného zbytku a kromě toho se liší i přítomností jedné dusíkaté báze odvozené od pyrimidinu. V řetězcích DNA je přítomen thymín, kdežto v RNA uracil. Odlišnosti jdou však ještě dále: zatímco u DNA většinou, a dá se říci, že dokonce pravidelně, sestávají její molekuly ze dvou antiparalelně uložených řetězců navzájem spojených vodíkovými vazbami mezi dusíkatými bázemi podle pravidla komplementarity (A-T, C-G), molekuly RNA jsou „jednořetězcové“. V jádře buňky se tyto velice dlouhé molekuly DNA vyskytují v lineární podobě namotané na bílkovinné cívky histonů a mnohonásobným stočením v prostoru jsou uspořádány tak, že v podobě chromozomů (chromatid) se mohou bez problému pohybovat uvnitř buněk při jejich dělení.

Mimo jádro se v buněčné cytoplazmě vyskytují dvouřetězcové molekuly DNA

v mitochondriích v cirkulární podobě. Jejich délka vyjádřená počtem základních jednotek – nukleotidů – je nepoměrně kratší než u molekuly jaderné DNA a odpovídá šestnácti a půl tisíci nukleotidů. (Na rozdíl od jaderných molekul, kdy např. odhadovaný počet nukleotidů molekuly DNA složené v prvním – a největším – lidském chromozomu je cca 220 milionů nukleotidů.)

Na rozdíl od jaderných molekul, které se vyskytují v párech, což odpovídá chromozomům, vyskytují se molekuly mtDNA v buňce ve velkých počtech dosahujících tisíců.

DNA můžeme považovat za nositelku genetické paměti, z níž jsou informace uvolňovány podle potřeby. Tato paměť je kopírována při replikaci, aby mohla být přenesena do nově vznikající buňky. Při dělení buněk, kdy z jedné vzniknou buňky dvě, se o nově sestavenou DNA obě podělí tak, že získají vlastně směs staré a nové DNA. Způsob replikace je pochopitelně odlišný, jedná-li se o jadernou DNA, nebo mtDNA.

Uvolňování informací z DNA se děje prostřednictvím jednořetězcových molekul RNA. Jejich syntéza vyvolá nezbytnost rozvolnění dvouřetězcové struktury DNA včetně uvolnění z histonového nosiče. Procesem označovaným jako transkripce vzniká jednořetězcová molekula RNA, která odpovídá jednomu z řetězců DNA. Je označována jako mRNA (anglicky messenger RNA). Pro české názvosloví byl původně vybrán termín „informační“, dnes „mediátorová“ RNA. Její další osud bývá složitý, a než se stane použitelnou pro translaci = proteosyntézu (máme-li na mysli proteinotvorné geny), jsou některé její části odstraněny – ty, které odpovídají v DNA tzv. intronům – zbylé jsou spojeny do jedné molekuly, která navíc získá na svém začátku zvláštní strukturu označovanou jako čepička a na svém konci obvykle homopolymerní úsek v podobě polyA (adenin). Tím ovšem její úpravy nekončí, stejně jako otázka jejího použití k translaci. Vedle již v rámci proteosyntézy uvedené mRNA jsou k ní zapotřebí molekuly dalších typů RNA, z nichž některé jsou součástí ribozomů (rRNA), a molekuly, jejichž úkolem je přinášet do ribozomů aminokyseliny transferové RNA (tRNA).

V jejich molekulách jsou specifická místa vytvářející vodíkové vazby, čímž v jejich molekulách vznikají úseky připomínající dvouřetězcové molekuly DNA, ale vytvořené v tomto případě jediným řetězcem, který se dostane do antiparalelní polohy a obtáčí se sám kolem sebe. Jiné části molekuly tRNA zůstávají v jednořetězcové podobě a umožňují jak navázání určité aminokyseliny, tak vytvoření vodíkových vazeb k molekule mRNA. Za takovou sekvenci je považován sled tří nukleotidů, tzv. antikodon. Ten odpovídá kodonu, který je v mRNA představován také sekvencí tří nukleotidů. Protože aminokyselin je zhruba třikrát méně, než kolik můžeme vytvořit kodonů (4³), jsou některé kodony s různou sekvencí určeny pro stejnou aminokyselinu (může jich být až 6). Je to umožněno zvláštní vlastností antikodonu vysvětlovanou tzv. kolébkovou hypotézou. Tak vznikl pojem degenerovaného kódu.

Proces realizace genetické informace je regulován již na samém uvolňování informace z genetické paměti – regulace transkripce – a vedle informace o struktuře finálního produktu je v DNA řada sekvencí, které tuto transkripci ovlivňují, mohou být umístěny v téže DNA molekule (v poloze CIS) nebo v jiné molekule (v poloze TRANS). Ty obvykle používají bílkovinné molekuly jako zprostředkovatele svého účinku – transkripční faktory. O funkční zdatnosti rozhodují i změny, které sice nemění primární strukturu DNA, ale některé z bází přeměňují na jejich deriváty např. tím, že je metylují (cytosin na metylcytosin) [1]. Tyto „epigenetické“ úpravy se mohou přenášet mezi generacemi anebo k nim dochází obdobně, jako vznikají somatické mutace – postihují jen některé buňky, respektive tkáně [1–3]. Podobně se mění i bílkovinný nosič DNA chromatin, jehož úpravy mohou být ještě pestřejší.

Na cestě k realizaci informace obsažené v DNA ke konečnému produktu se však také uplatňuje řada regulačních mechanismů, které jsou namířeny až proti mediátorové RNA – nedovolí, aby mRNA k procesu translace dospěla. Přitom se uplatňují drobné molekuly RNA, tzv. mikroRNA (miRNA), které způsobí rozpad mRNA. Těchto „nekódujících“ RNA je veliké množství a uplatňují

se v různých způsobech regulací během realizace genetické informace [4–7].

Přetvoření primárního transkriptu v molekulu RNA, která může být použita k translaci, je složitý proces, v němž zásadní roli hraje vynětí intronových sekvencí a pospojování sekvencí exonových (proces sestřihu), což je zároveň spojeno s jejich výběrem. To umožňuje, aby z jednoho „genu“ vznikalo více produktů.

Na cestě od genetické paměti k její realizaci je tak mnoho míst, která mohou být zdrojem různých poruch a odpovídat za změny ve fenotypu jak jednotlivých buněk, tak celého organismu. Vznikla tím kategorie tzv. RNA nemocí, které bývají právě projevem poruchy sestřihového mechanismu [8,9].

Kromě nukleových kyselin přítomných v buňkách nacházíme v poslední době díky moderním detekčním metodám nukleové kyseliny i mimo buňky, z hlediska lidské diagnostiky jsou nám k dispozici např. v krvi – v krevní plazmě. Domníváme se, že jde o DNA uvolněnou z rozpadajících se odumřelých buněk, DNA, která byla rozbita procesy souvisejícími s úhynem buněk, např. apoptózou vyvolanou nejrůznějšími vnitřními i vnějšími příčinami.

Povídání o nukleových kyselinách a v nich uložených genetických informacích je v našem případě nezbytné doplnit o metody, který mají medicínský význam.

Metody

Jedna z prvních metod se opírala o skutečnost, že prakticky každý z nás má svou vlastní genetickou paměť, která se do jisté míry liší od paměti ostatních lidí. Molekuly DNA, byť značně podobné, nejsou na úrovni sekvencí nukleotidů zcela identické. V rámci lidské populace existuje určitá variabilita. Vzájemné odchylky mohou být velmi vzácné, nebo poměrně časté. Vznikly během evoluce a vznikají neustále. Můžeme tedy prohlásit, že lidský genom jako soubor všech našich molekul DNA je variabilní a proměnlivý. Genetická diagnostika v době, kdy ještě neznala jednotlivé geny, byla schopna díky určení variabilních míst spojit DNA s určitým fenotypem – z hlediska dědičných chorob hledat místo v genomu, které s rozvojem

choroby souvisí [10]. Využívala k tomu tzv. genetické polymorfizmy, tedy variabilitu na úrovni DNA, kterou bylo možno odhalit použitím bakteriálních enzymů, které jsou schopny velice specificky štěpit DNA (restriktázy), čímž vznikají úseky (fragmenty), jejichž velikost závisí právě na výskytu pro daný enzym typické sekvence – obvykle o rozsahu několika málo nukleotidů. Proto tato metoda využívala tzv. polymorfismus restriktčních fragmentů (RFLP). Polymorfizmem je označována variabilita, kdy výskyt méně časté varianty přesahuje 1 %. V určitých situacích nám tento přístup sloužil donedávna [11]. Měl tu výhodu, že o konkrétní změně, která je odpovědná za daný fenotyp, jsme nemuseli vědět víc, než to, kde se přibližně v genomu nachází. Nevýhodou bylo to, že ne vždy se podařilo najít vhodné variabilní místo.

Tyto metody však měly značné nároky na množství DNA, které bylo pro analýzu nezbytné, a proto se např. odběry krve pohybovaly v desítkách mililitrů. I když jsme měli od samého začátku k dispozici způsob izolace DNA (tzv. fenolovou metodu), která byla po stránce kvantitativní i kvalitativní velice účinná a takovou zůstala do dneška, některé další postupy, které postupně mizí v dávnověku naší paměti – elektroforetické dělení fragmentů, přenos a upevnění fragmentů na membrány, hybridizace se značenými (původně radioaktivními izotopy fosforu) sondami, zachycení záření na citlivém rentgenografickém materiálu a jeho vyvolání – již zdaleka tak účinné nebyly.

Ohromným metodickým přínosem se stala řetězová polymerázová reakce (PCR), která zásadně změnila naše nároky na potřebné množství DNA, takže dnes teoreticky vystačíme s množstvím DNA obsaženým v jedné jediné buňce. Při rutinní diagnostice vycházíme většinou z množství většího, ale už nejsme omezeni na bílé krvinky jako hostitelky DNA v krvi, ale můžeme opustit invazivní přístup a jako vzorek použít třeba bukalní stěr.

Dnes disponujeme širokou paletou metod, kterými můžeme lidský genom ohledávat. Na jejich vrcholu se nepochybně nachází celogenomová sekvenace, jejíž účinnost se neustále zvyšuje –

to vede ke zkracování potřebného času i snížení nákladů. Postupně, jak se dařilo díky sekvenačním metodám odhalovat jednotlivé geny a jejich strukturu, převládly v lékařské genetické diagnostice techniky schopné provádět velké množství analýz současně – „biočipy“. Jestliže klasická PCR byla schopná množit jen jednu nebo několik málo cílových úseků DNA v jedné reakci a klasická Sangerova sekvenace s detekcí kapilární elektroforézou jen úsek 500–1 000 nukleotidů, pak současné přístroje díky využití principu, na jehož bázi se vyvinula technologie čipů, zvýšily svou kapacitu mnohonásobně – RainDance kapková technologie dovoluje momentálně proběhnout 10 000 PCR reakcí současně (<http://www.raindancetech.com/video/default.asp>) a některé sekvenační přístroje tzv. 2. generace jsou schopny produkovat během jednoho běhu i více než 600 Gb (gigabází). Jejich výkonnost neustále stoupá, a dokonce byla na trh uvedena již generace třetí, která signalizuje, že cíle v podobě snížení nákladů na 1 000 USD za jednu celogenomovou sekvenaci bylo již dosaženo.

Závěry

Souběžně s tímto technickým pokrokem se posunulo vpřed i naše nazírání na genom a jeho případnou patologii [12]. Od původního hledání odpovědných genů s vytipováním tzv. kandidátních genů a asociačních studií u jednoduše dědičných poruch jsme postupně začali hloubat nad genetickými příčinami stavů se složitou genetickou podstatou (dnes i prostřednictvím celogenomových asociačních studií – WGAS) a analyzovat podíl genetické složky tzv. civilizačních chorob [13,14]. Dnes již méně často hledáme geny, jejichž vztah k fenotypu je natolik jednoduchý, že jsme je zařadili mezi choroby s tzv. mendlovskou dědičností, ale snažíme se identifikovat „markery“ – úseky genomu, které nějakým způsobem (a po něm pátráme) ovlivňují fenotyp [15,16]. To platí i o somatických mutacích odpovědných za onkologická onemocnění, kdy indikátorem i cílem našeho terapeutického snažení bývá právě fenotyp [17].

A i když se pouštíme do analýzy chorob s komplexní dědičností, stále nám

zbývá několik tisíc patologických znaků, jejichž příčinu se nám na úrovni genomu ještě nepodařilo odhalit, přestože se zdá, že jejich dědičnost by neměla být příliš složitá. V molekulární diagnostice se díky novým technologiím dokážeme spokojit i s DNA jedné jediné buňky [18].

Literatura

1. Egger G, Wielscher M, Pulverer W et al. DNA methylation testing and marker validation using PCR: diagnostic applications. *Expert Rev Mol Diagn* 2012; 12(1): 75–92.
2. Bormann Chung CA, Boyd VL, McKernan KJ et al. Whole methylome analysis by ultra-deep sequencing using two-base encoding. *PLoS One* 2010; 5(2): e9320.
3. You YJ, Chen YP, Zheng XX et al. Aberrant methylation of the PTPRO gene in peripheral blood as a potential biomarker in esophageal squamous cell carcinoma patients. *Cancer Lett* 2012; 315(2): 138–144.
4. Cho WC. MicroRNAs in cancer – from research to therapy. *Biochim Biophys Acta* 2010; 1805(2): 209–217.
5. Faltejsková P, Slabý O, Hézová R et al. Role mikroRNA v imunitním systému. *Čas Lék Čes* 2010; 149(1): 10–15.
6. Wilmott JS, Zhang XD, Hersey P et al. The emerging important role of microRNAs in the pathogenesis, diagnosis and treatment of human cancers. *Pathology* 2011; 43(6): 657–671.
7. Whelan JT, Hollis SE, Cha DS et al. Post-transcriptional regulation of the Ras-ERK/MAPK signaling pathway. *J Cell Physiol* 2012; 227(3): 1235–1241.
8. Esteller M. Non-coding RNAs in human disease. *Nat Rev Genet* 2011; 12(12): 861–874.
9. Hall PA, Russell SH. Mammalian septins: dynamic heteromers with roles in cellular morphogenesis and compartmentalization. *J Pathol* 2012; 226(2): 287–299.
10. Cho WC. Recent progress in genetic variants associated with cancer and their implications in diagnostics development. *Expert Rev Mol Diagn* 2010; 10(6): 699–703.
11. Silva Pinto C, Fidalgo T, Salvado R et al. Molecular diagnosis of haemophilia A at Centro Hospitalar de Coimbra in Portugal: study of 103 families – 15 new mutations. *Haemophilia* 2012; 18(1): 129–138.
12. Baudhuin LM, Donato LJ, Uphoff TS. How novel molecular diagnostic technologies and biomarkers are revolutionizing genetic testing and patient care. *Expert Rev Mol Diagn* 2012; 12(1): 25–37.
13. Meuwissen T, Goddard M. Accurate prediction of genetic values for complex traits by whole-genome resequencing. *Genetics* 2010; 185(2): 623–631.
14. Wang K, Dickson SP, Stolle CA et al. Interpretation of association signals and identification of causal variants from genome-wide association studies. *Am J Hum Genet* 2010; 86(5): 730–742.
15. Coyle VM, Johnston PG. Genomic markers for decision making: what is preventing us from using markers? *Nat Rev Clin Oncol* 2010; 7(2): 90–97.
16. Katsios C, Ziogas DE, Roukos DH. New molecular oncology-changing era: prospects and challenges of cancer genome and integrative systems biology. *Expert Rev Anticancer Ther* 2011; 11(1): 5–8.
17. Shieh MP, Mitsuhashi M, Lilly M. Moving on up: Second-Line Agents as Initial Treatment for Newly-Diagnosed Patients with Chronic Phase CML. *Clin Med Insights Oncol* 2011; 5: 185–199.
18. Quinn GP, Pal T, Murphy D et al. High-risk consumers' perceptions of preimplantation genetic diagnosis for hereditary cancers: a systematic review and meta-analysis. *Genet Med* 2012; 14(2): 191–200.