

Sekvenování dlouhých nekódujících RNA v exozomech u pacientů s kolorektálním karcinomem

Sequencing of long non-coding RNAs in exosomes of colorectal cancer patients

Madrzyk M.^{1,2}, Macháčková T.¹, Trachtová K.¹, Catela Ivković T.¹, Součková K.¹, Kotouček J.³, Mašek J.³, Loja T.¹, Šachlová M.⁴, Slabý O.¹

¹CEITEC – Středoevropský technologický institut, MU Brno

²LF MU Brno

³Výzkumný ústav veterinárního lékařství, v.v.i., Brno

⁴Gastroenterologické oddělení, MOÚ Brno

Souhrn

Východiska: Prognóza pacientů s karcinomem kolorekta (colorectal cancer – CRC) závisí především na rozsahu onemocnění v době diagnózy, proto je brzký záchyt jedním z hlavních předpokladů úspěšné léčby. Současný výzkum ukazuje, že exozomální dlouhé nekódující RNA (lncRNA) jsou spojeny s rozvojem nádorových onemocnění. Jelikož jsou lncRNA často tkáňově specifické, jejich kvantifikace v exozomech se nabízí jako neinvazivní metoda pro včasnou detekci CRC. V naší práci jsme se zaměřili na optimalizaci protokolu pro analýzu exozomálních lncRNA z krevního séra pacientů s CRC jako potenciálních diagnostických biomarkerů. **Materiál a metody:** Exozomy byly izolovány pomocí gelové chromatografie ze 150 µl séra pacientů s CRC a zdravých dárců. Jejich kvalita a kvantita byla potvrzena elektronovou mikroskopií a analýzou dynamického rozptylu světla (dynamic light scattering – DLS) a proteinové markery byly detekovány metodou Western blot. Po izolaci RNA byly ze vzorků připraveny cDNA knihovny, které byly sekvenovány pomocí NextSeq 550. **Výsledky:** Úspěšně jsme izolovali exozomy a ověřili jsme jejich vlastnosti několika různými metodami. Knihovny byly připraveny ze všech vzorků i přes velmi nízký objem výchozího materiálu. Sekvenční data potvrzují přítomnost protein kódující (50 %) i nekódující RNA, kterou tvoří především lncRNA (28,2 %), pseudogeny (15,2 %) a další typy RNA (6,5 %). Výsledky dále ukázaly významně změněné hladiny některých lncRNA, na základě jejichž exprese bylo možné odlišit vzorky od pacientů s CRC od vzorků zdravých kontrol. Pomocí analýzy obohacení genové sady (gene set enrichment analysis – GSEA) jsme pozorovali významně obohacené třídy genů, které souvisejí s opravami DNA nebo regulací buněčného cyklu. **Závěr:** Naše pilotní data naznačují, že lncRNA představují významnou část RNA přítomné v exozomech a jejich rozdílné hladiny mají schopnost odlišit CRC pacienty od zdravých kontrol. Analýza obohacených genů zároveň prokázala významné zastoupení lncRNA podílejících se na regulaci buněčného cyklu a oprav DNA, což naznačuje jejich možné zapojení do procesů kancerogeneze. Výsledky je však třeba ověřit na větším souboru pacientů.

Klíčová slova

RNA dlouhá nekódující – kolorektální nádory – exozomy – nádorové biomarkery – stanovení celkové genové exprese

Tato práce byla podpořena projekty GA20-18889S, MZE-RO0518, BBMRI-CZ LM2018125 a CZ.02.1.01/0.0/0.0/15_003/0000495.

This study was supported by the following research programs GA20-18889S, MZE-RO0518, BBMRI-CZ LM2018125, and CZ.02.1.01/0.0/0.0/15_003/0000495.

Autoři deklarují, že v souvislosti s předmětem studie nemají žádné komerční zájmy.

The authors declare that they have 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 zasílané do biomedicínských časopisů.

The Editorial Board declares that the manuscript met the ICMJE recommendation for biomedical papers.



prof. RNDr. Ondřej Slabý, Ph.D.
CEITEC – Středoevropský technologický institut
Kamenice 753/5
625 00 Brno
e-mail: oslaby@med.muni.cz

Obdrženo/Submitted: 15. 7. 2022

Přijato/Accepted: 7. 9. 2022

Summary

Background: The prognosis of patients with colorectal cancer (CRC) depends mainly on the extent of the disease at the time of diagnosis; therefore, early detection is one of the main prerequisites for successful treatment. Current research shows that exosomal long non-coding RNAs (lncRNAs) are associated with cancer development. As lncRNAs are often tissue specific, their quantification in exosomes is proposed as a non-invasive method for early detection of CRC. In this study, we aimed to optimize a protocol for analyzing exosomal lncRNAs from blood serum of CRC patients as potential diagnostic biomarkers. **Material and methods:** Exosomes were isolated by gel chromatography from 150 µl of serum of CRC patients and healthy donors. Their quality and quantity were confirmed by electron microscopy and dynamic light scattering (DLS) analysis; protein markers were detected by Western blot. After RNA isolation, cDNA libraries were prepared and sequenced using NextSeq 550. **Results:** We successfully isolated exosomes and verified them by several methods. Libraries were prepared from all samples despite very low volume of starting material. The sequencing data confirmed the presence of both protein-coding (50%) and non-coding RNAs, which consisted mainly of lncRNAs (28.2%), pseudogenes (15.2%) and other RNA types (6.5%). The results also showed significantly altered levels of some lncRNAs that could distinguish samples from CRC patients and healthy controls. Using gene set enrichment analysis (GSEA), we observed significantly enriched classes of genes related to DNA repair or cell cycle regulation. **Conclusion:** Our preliminary data suggest that lncRNAs represent a significant fraction of the RNA present in exosomes and that their distinct levels can separate CRC patients from healthy controls. The analysis of enriched genes also showed a significant representation of lncRNAs involved in cell cycle regulation and DNA repair, suggesting their possible involvement in cancerogenesis. However, the results need to be verified in a larger cohort of patients.

Key words

RNA, long noncoding – colorectal neoplasms – exosomes – biomarkers, tumor – gene expression profiling

Úvod

Kolorektální karcinom (CRC) tvoří dle odhadů 10 % všech zjištěných karcinomů u mužů a žen a celosvětově představuje druhou nejčastější příčinu úmrtí na nádorové onemocnění [1]. Navzdory lepším přístupům v detekci a léčbě

se značný počet pacientů s CRC potýká s nepříznivou prognózou, která do značné míry závisí na rozsahu onemocnění v době diagnózy. Jelikož CRC se často vyvíjí prostřednictvím postupné progresy od adenomu ke karcinomu, včasná diagnostika a resekce prekancerózní tkáně by vedla ke zlepšení pacientovy prognózy. V této souvislosti mají vysokou prioritu neinvazivní biomarkery, které mohou poskytnout spolehlivou a včasnou detekci CRC.

V posledních letech se exozomy objevily jako potenciální rezervoáry klinicky užitečných biomarkerů, které se vyskytují ve všech tělních tekutinách. Tyto membránové váčky vznikají v rámci endozomálního systému, mají velikost 30–150 nm a jsou vylučovány různými typy buněk, vč. nádorových. Exozomy přenášejí informace v podobě nukleových kyselin, proteinů a lipidů a podílejí se na místní i vzdálené mezibuněčné komunikaci [2]. V rozvoji nádorů hrají významnou roli tím, že přepravují onkogeny cílovým buňkám, regulují interakce mezi nádorem a jeho mikroprostředím, podporují vznik metastáz a rozvoj angiogeneze [3]. Vzhledem k rostoucímu počtu studií, které prokázaly, že obsah exozomů odráží biologický stav buněk a že jsou produkovány nádorovými buňkami, lze těchto vlastností využít k detekci RNA biomarkerů, kterými mohou být např. exozomální dlouhé nekódující RNA (lncRNA).

Tato skupina nekódujících RNA se vyznačuje transkripty o délce nejméně 200 nukleotidů, které nejsou překládány do proteinu [4]. Doposud bylo identifikováno více než třicet onkogenních lncRNA zapojených do klíčových signálních drah souvisejících s molekulární patogenezi CRC a očekává se, že jejich počet bude narůstat [5]. Jednou z popsaných exozomálních lncRNA je např. CRNDE-h, jejíž zvýšené hladiny u pacientů s CRC významně korelovaly s nepříznivou prognózou, metastázováním do lymfatických uzlin i přítomností vzdálených metastáz [6]. Jelikož jsou lncRNA často tkáňově specifické a mohou sloužit jako signální molekuly v mezibuněčné komunikaci, jejich kvantifikace v exozomech se nabízí jako neinvazivní metoda pro včasnou detekci CRC [4,7]. V naší práci jsme se zaměřili na analýzu exozomálních lncRNA jako potenciálních biomarkerů pro časnou diagnostiku onemocnění.

Materiál a metody

Exozomy byly izolovány ze 150 µl krevního séra pacientů s CRC (n = 31) a zdravých kontrol (n = 15) metodou vylučovací chromatografie (qEVsingle, IZON). Tab. 1 charakterizuje kohortu pacientů s CRC. Vzorky od pacientů před chirurgickým zákrokem byly odebrány podle standardního odběrového protokolu Masarykova onkologického ústavu (MOÚ), stejně jako vzorky od zdravých

Tab. 1. Charakterizace pacientů s kolorektálním karcinodem.

Muži, n	19
Ženy, n	12
Medián věku, (rozmezí)	67 (38–86)
Diagnóza, n	
C18	14
C19	4
C20	13
Klinické stadium, n	
I	9
II	10
III	8
IV	4
Histopatologický stupeň, n	
1	2
2	21
3	8

Tab. 2. Přehled nejvýznamnějších lncRNA rozdílně exprimovaných v exozomech pacientů s kolorektálním karcinodem a zdravých kontrol.

lncRNA	p-hodnota	adjustovaná p-hodnota	log ² násobná změna	míra exprese (počet kopií)
DUX4L19	1,49 × 10 ⁻⁴	0,017	0,59	132,7
AC016590.1	5,13 × 10 ⁻³	0,048	0,63	95,2
AC027601.5	1,32 × 10 ⁻⁴	0,017	0,60	57,6
AC087664.2	2,22 × 10 ⁻⁴	0,020	-0,59	43,2
CYP2D8P	3,99 × 10 ⁻⁵	0,014	0,66	37,3
DUX4L17	1,33 × 10 ⁻⁴	0,017	0,66	36,8
AC127496.1	1,41 × 10 ⁻⁴	0,017	0,63	36,6
NALT1	6,07 × 10 ⁻⁴	0,026	0,75	34,3
PCDH8P1	1,77 × 10 ⁻⁴	0,018	0,67	31,8
ECEL1P1	1,59 × 10 ⁻⁴	0,018	0,76	30,2

kontrol odebráné v rámci preventivní onkologické prohlídky. Všechny vzorky byly zpracovány stejným způsobem a byly uloženy v Bance biologického materiálu MOÚ. Všichni účastníci studie podepsali informovaný souhlas schválený etickou komisí Masarykovy univerzity. Velikost a koncentrace exozomů byly potvrzeny elektronovou mikroskopií a analýzou dynamického rozptylu světla (dynamic light scattering – DLS). Metodou Western blot byly detekovány charakteristické proteinové markery, které jsou asociovány s extracelulární membránou nebo jsou cytosolického původu. Před extrakcí RNA (Monarch Total RNA Miniprep Kit) byly vzorky exozomů ošetřeny proteinázou K, RNázou a následně inhibitory RNázy. Po izolaci RNA byly ze vzorků připraveny knihovny cDNA pomocí komerčního kitu (NEB-Next Ultra II Directional RNA Library Prep Kit, NEB), které byly následně sekvenovány za použití chemie NextSeq 500/550 High Output Kit v2.5 (75 cyklů) v módu single-read na sekvenátoru NextSeq 550 (obojí Illumina). Bioinformatická analýza získaných dat zahrnovala vyhodnocení rozdílné genové exprese, které bylo provedeno v programu R s použitím balíčku DESeq2, a zároveň analýzu funkčního obohacení. Genová ontologie (GO) významně deregulovaných genů byla zkoumána analýzou nadměrného zastoupení po-

mocí R balíčku ClusterProfiler a analýza obohacení genové sady (gene set enrichment analysis – GSEA) byla realizována v R balíčku fgsea s použitím referenčních sad genů Hallmark (MsigDB).

Výsledky

Úspěšně jsme optimalizovali a standardizovali protokol pro izolaci exozomálních RNA z velmi malého množství vstupního materiálu s následnou bioinformatickou analýzou sekvenačních dat. Získaná data potvrzují přítomnost protein kódujících a nekódujících RNA v exozomech krevního séra. Nejpočetnější skupinu identifikovaných genů tvořily protein kódující RNA (50 %), druhou nejvíce zastoupenou třídou byly lncRNA (28,2 %), dále pak pseudogeny (15,2 %) a jiné typy RNA (6,5 %). Analýza odhalila 85 významně deregulovaných lncRNA (adjustovaná p-hodnota < 0,05; násobná změna > 1,5), z nichž 21 vykazovalo sníženou a 64 zvýšenou expresi ve vzorcích od pacientů s CRC v porovnání se zdravými kontrolami. V tab. 2 je uvedeno 10 nejvýznamnějších lncRNA s mírou exprese větší než 30 kopií. Pro získání nejvýznamnějších tříd funkčně příbuzných genů byly provedeny dva typy analýz obohacení. GO analýza identifikovala geny kódující proteiny, které jsou nadměrně zastoupeny v biologických procesech, a GSEA ukázala nejvýznamněji obohacené či ochuzené mRNA sady. Pomocí GSEA jsme

pozorovali významně obohacené třídy genů, které souvisejí s opravami DNA nebo s molekulárními cíli signální dráhy MYC (adjustovaná p-hodnota < 0,05). Nadměrně zastoupené biologické procesy souvisejí se zvýšenou frekvencí transkripce či vývojem endokrinního systému (p-hodnota < 0,001).

Diskuze

lncRNA jsou stále častěji považovány za kritické regulátory mnoha buněčných funkcí. Ve střevní tkáni modulují několik signálních drah, které jsou klíčové pro udržení její homeostázy. Naopak jejich deregulace u nádorových onemocnění může tyto signální kaskády změnit a umožnit maligním buňkám proliferaci a šíření [4]. Expresním profilováním lncRNA lze identifikovat potenciální cíle, které mohou sloužit pro včasnou detekci onemocnění. Sekvenování exozomální RNA navíc nabízí možnost vývoje biomarkerů získaných neinvazivním způsobem z krevního séra nebo plazmy.

V naší studii nám optimalizace protokolu umožnila úspěšnou izolaci exozomů a přípravu cDNA knihoven ze všech vzorků, a to i přes velmi nízký objem výchozího materiálu. S využitím vysokokapacitního expresního profilování RNA byly odhaleny významně změněné hladiny lncRNA ve vzorcích pacientů s CRC v porovnání se zdravými kontrolami a dále byly identifikovány

nádorově specifické lncRNA, které ještě nebyly v souvislosti s CRC popsány. Je zajímavé, že zvýšená exprese detekovaná lncRNA NALT1 byla v minulosti asociována s rozvojem karcinomu žaludku a vznikem metastáz [8]. Provedené analýzy GSEA a GO navíc podpořily významnost našich zjištění v kontextu CRC.

Závěr

Naše předběžná data naznačují, že exozomální lncRNA mají schopnost odlišit pacienty s CRC od zdravých kontrol, a mohly by tak v budoucnu sloužit jako slibné neinvazivní biomarkery pro včasný záchyt tohoto onemocnění. Výsledky je

však třeba ověřit pomocí nezávislé metody na větším souboru pacientů a kontrol. Biologická role vybraných lncRNA v patogenezi CRC bude dále ověřena funkčními *in vitro* a *in vivo* experimenty.

Literatura

1. Sung H, Ferlay J, Siegel RL et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* 2021; 71(3): 209–249. doi: 10.3322/caac.21660.
2. Yáñez-Mó M, Siljander PRM, Andreu Z et al. Biological properties of extracellular vesicles and their physiological functions. *J Extracell Vesicles* 2015; 4(2015): 1–60. doi: 10.3402/jev.v4.27066.
3. Whiteside TL. Tumor-derived exosomes and their role in cancer progression. *Adv Clin Chem* 2016; 74(3): 103–141. doi: 10.1016/bs.acc.2015.12.005.
4. Fang Y, Fullwood MJ. Roles, functions, and mechanisms of long non-coding RNAs in cancer. *Genomics, Proteomics Bioinforma* 2016; 14(1): 42–54. doi: 10.1016/j.gpb.2015.09.006.
5. Yang Y, Yan X, Li X et al. Long non-coding RNAs in colorectal cancer: Novel oncogenic mechanisms and promising clinical applications. *Cancer Lett* 2021; 504(2): 67–80. doi: 10.1016/j.canlet.2021.01.009.
6. Liu T, Zhang X, Gao S et al. Exosomal long noncoding RNA CRNDE-h as a novel serum-based biomarker for diagnosis and prognosis of colorectal cancer. *Oncotarget* 2016; 7(51): 85551–85563. doi: 10.18632/oncotarget.13465.
7. Derrien T, Johnson R, Bussotti G et al. The GENCODE v7 catalog of human long noncoding RNAs: analysis of their gene structure, evolution, and expression. *Genome Res* 2012; 22(9): 1775–1789. doi: 10.1101/gr.132159.111.
8. Piao H-Y, Guo S, Wang Y et al. Long noncoding RNA NALT1-induced gastric cancer invasion and metastasis via NOTCH signaling pathway. *World J Gastroenterol* 2019; 25(44): 6508–6526. doi: 10.3748/wjg.v25.i44.6508.