Jan Hamalčík

Esej
Vznik spliceosomálních intronů u eukaryot

Introny u eukaryot představují majoritní část protein kódujících genů. Jejich vystřižení spliceosomem a napojování exonů na sebe zabráňuje vzniku nežádouce fungujících bílkovin a umožňuje provedení alternativního splicingu, který zvětšuje proteom bez nutnosti zvětšení genomu. Absence spliceosomálních intronů u bakterií a archaeí nicméně nabádá k otázce, jak se tyto u eukaryot vůbec objevily? Odkud pochází snRNA a proteiny spliceosomu? Tato esej se pokouší vnést do těchto nejasností trochu světla na základě dat z článku „Domestication of self-splicing introns during eukaryogenesis: the rise of the complex spliceosomal machinery“ od Juliana Vosseberga a Berenda Snela.

Na základě homologie mezi eukaryotickými spliceosomálními introny a self-splicing introny skupiny II vyskytujících se u prokaryot či eukaryotických organelách lze usuzovat, že prvně jmenované se vyvinuly právě z těchto parazitických elementů schopných se samy vystřihnout z nukleové kyseliny. Tyto se do eukaryotických buněk vzniklých endosymbiózou archeální a bakteriální buňky pravděpodobně dostaly horizontálním genovým transferem před či po endosymbióze samotné, nebo vnitřním genovým transferem z mitochondrie. Přítomnost spliceosomu u téměř všech eukaryotických linií dále napovídá, že se spliceosomální introny vyvinuly již u společného předka všech eukaryot, tzv. LECA. Z toho vyplývá, že self-splicing introny stihly evolvovat do intronů spliceosomálních již před rozdělením eukaryot. Bez spliceosomálních intronů by totiž spliceosom neměl žádnou funkci a bez spliceosomu by vznikaly chybné proteiny.

Přesné pořadí jednotlivých událostí při evoluci spliceosomálních intronů u LECA je spekulativní, v naprosté většině scénářů se nicméně objevuje nábor proteinů, genové duplikace, ztráta samovystřihovací aktivity a další proliferace nově vznikajících intronů spliceosomálních. Pozastavme se nyní právě nad samovystřihovacími introny skupiny II. Jejich celková struktura a domény jsou místy nápadně podobné jednotlivým molekulám snRNA u spliceosomů. Například U5 snRNA a doména I mají funkčně vyměnitelné místa vazby exonů, U6 snRNA a doména V obě obsahují katalytickou triádu, váží stejné kovové ionty, atp. Existují také homologie mezi skupinou II introny kódovaným proteinem s reverzní transkriptázovou aktivitou a Prp8 proteinem přítomným ve spliceosomálním katalytickém jádře, taktéž s doménou připomínající RE aktivitu. Přítomnost introny kódovaného proteinu s RE aktivitou mimo jiné podporuje představu, že introny jsou ze své podstaty parazitické elementy.

Jelikož spliceosomy jsou ribonukleoproteinové komplexy, je nutno objasnit původ i dalších proteinů, bez kterých by spliceosomální introny bez schopnosti samovystřižení nemohly být odstraněny z pre-mRNA. Velmi důležité jsou Sm a Lsm heteroheptamerické proteiny, které fungují jako lešení pro U6 snRNA (Lsm) a ostatní snRNA (Sm), a umožňují interakce mezi snRNA a ostatními ribonukleoproteiny snRNP. Jejich homology jsou přítomné jak u bakterií, tak u archeí, a co víc, každý Sm gen je párovaný s Lsm geny. Oba tyto proteiny pravděpodobně vznikly z homoheptamerických Lsm kruhů, přičemž se do dnešní podoby dostaly genovými duplikacemi.

Řada proteinů byla také získána z jiných, již existujících struktur. Například GTPasy Snu114 jsou homology EF-2 translokáz v ribosomech. U5 snRNP specifický Brr2 protein, ATPasová helikáza ve spliceosomu, pochází z rodiny eukaryotických translačních iniciačních faktorů. RNA rozpoznávací motiv u SR proteinů, bílkovin bohatých na serin a arginin regulující sestřih, pochází původně z mnohých prokaryotických linií. Jsou to tedy převážně proteiny účastnící se translace či degradace RNA, které byly do spliceosomálních komplexů rekrutovány.

Po eukaryogenezi se introny v eukaryotických genomech a spliceosomy začaly vyvíjet nezávisle na sobě. Mnohé linie, např. živočichové a rostliny, prodělaly proliferaci intronů po

genomu a s tím související rozšíření množiny proteinů účastnící se splicingu, tím pádem narůstající komplexitu celého spliceosomálního komplexu. Na druhou stranu, v jiných eukaryotech, převážně v těch jednobuněčných, se spliceosomy zjednodušovaly, v některých dokonce společně s introny vymizely kompletně. V redukovaných spliceosomech často nejsou přítomny pro spliceosom zdánlivě esenciální komplexy, jako například Sm proteiny, či U1 snRNP. U1 snRNA mimo jiné nemá žádného analoga v samovystřihovacích intronech skupiny II, což implikuje, že U1 snRNP byl přidán až v pozdní eukaryogenezi, přičemž splicing intronů spliceosomem probíhal již před tímto „vylepšením“.

Podívejme se nyní na evoluční modely, které by tyto jevy mohly popsat. Brzcí eukaryoté, pravděpodobně i LECA, měli velmi komplexní spliceosom, což ovšem nepřinášelo žádné zřejmé výhody. Původ tohoto komplexu by tedy mohla osvětlit neutrální teorie molekulární evoluce. Vznik trans působících snRNA z fragmentů intronů skupiny II a jejich maturas, tedy proteinů účastnící se splicingu, udělal samovystřihovací schopnost intronů skupiny II nepotřebnou, což způsobilo její ztrátu a závislost brzkých eukaryot na těchto trans působících RNP komplexech, u kterých následně vznikl komplexní spliceosom neutrálním přidáváním proteinů s funkcí vazby RNA. Evoluce spliceosomů po eukaryogenezi se dále zdá být adaptivní, což mimo jiné dokládá i existence alternativního splicingu, který, jak již bylo řečeno, kombinatoricky rozšiřuje variabilitu proteomu bez nutnosti zvětšit genom organismu. Potřeba tohoto zvyšuje komplexitu spliceosomu u mnohobuněčných eukaryot, na druhou stranu, u protist nepotřebujících tak obsáhlý transkriptom je alternativní sestřih energeticky nevýhodný a adaptivně dochází k redukci intronů i komplexity spliceosomu.

Mnohé nicméně naznačuje, že se komplexita celé spliceosomální mašinerie nevyvíjela kontinuálně, nýbrž v krátkodobých boomech následovaných dlouhými periodami redukce, což lze nazvat bifázovým evolučním modelem. Za první takový boom lze považovat již popsaný iniciální vznik spliceosomu při eukaryogenezi. Ten byl následován rozsáhlou redukční periodou, která u některých eukaryot trvá dodnes, u některých již dokonce dosáhla konce kompletním vyřazením spliceosomu. Jiné eukaryotické linie na druhou stranu za svou historii prodělaly další skoky ve vývoji spliceosomu, čemuž například živočichové a rostliny vděčí za pokročilý alternativní splicing.

Introny tvoří u většiny eukaryot majoritní část genů a spolehlivě najít hraniční místa intronů a exonů v sekvenci RNA je i pro dnešní pokročilou bioinformatiku prakticky nemožné. O to víc fascinující je schopnost spliceosomů tyto místa spolehlivě rozeznat, nukleofilním atakem připojit 5’ konec intronu na 2’ OH skupinu určitého adenosinu někde uprostřed intronu fungujícího jako branch point, tento lariát vypustit a spojit exony tak, aby vznikl fungující protein. Rozluštění evoluce spliceosomu, spliceosomálních intronů a self-splicing intronů skupiny II, nám jistě některé z těchto tak řečeno záhad pomůže objasnit. Kompletní odhalení vzniku spliceosomálních intronů u eukaryot je ovšem stále předmětem výzkumu.

Zkratky použité v textu:
RE: reverzní transkriptáza
LECA: last eukaryotic common ancestor