

Oponentský posudek magisterské práce Filipa Husníka "Evolutionary origins of intracellular symbionts in arthropods"

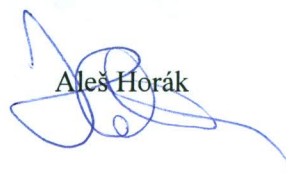
Hned na okraj se musím přiznat, že s tak kvalitní magisterskou prací jsem se za svou, pravda nepříliš dlouhou, oponentskou kariéru ještě nesetkal. Doufám, že není poslední. Bohužel je to také zatím nejméně vděčná práce na oponování. Sestává se totiž z rozsáhlého literárního přehledu, krátkého shrnutí, plánů do budoucna a manuskriptu, který již prošel peer-review procesem a byl přijat do poměrně prestižního periodika BMC Biology. Když ale odhlédnu od sobecké potřeby oponenta, hledat nedostatky, jeví se mi tento model velmi sympatický. Teď ještě přinutit studenty magisterského stupně k publikační činnosti...

Přiznám se bez mučení, že v otázce bakteriálních symbiontů členovců jsme se až do přečtení této práce neorientoval vůbec. Nemohu tedy posoudit faktickou stránku a musím se spolehnout, že úctyhodných 178 literárních pramenů existuje, že si je autor nevymyslel a že v nich je skutečně to, co tvrdí. Jinak je ale úvod psaný velmi čtivě a pro daný stupeň studia téměř nevidanou angličtinou. Kéž by většina našich doktorandů byla alespoň stejně tak zdatná, obávám se, že tomu tak není. Samozřejmě, že roditelý mluvčí by asi celý text napsal jinak (rozuměj lépe), po prvním přečtení jsem však byl příjemně šokován minimem do-očí-bijících kostřbatých spojení či překlepů a chyb. Většinu úvodu tvoří spíše faktografický přehled současných vědomostí o problematice, místy však textem prostupuje autorův pohled na věc. Pokud mám úvodu něco vytknout, tak snad to, že takových míst by tam mohlo být víc. V literárním přehledu se nachází dva obrázky bez uvedení zdroje. Jsou skutečně dílem autora předkládané práce? Dále se zde vyskytují dvě tabulky, jejichž formát se poněkud liší a měl by být sjednocen. A vůbec, grafická stránka by asi snesla více snahy. Celá problematika je pak přehledně shrnuta formátem otázek-odpovědí. Zde nemám připomínek

Místo, které je u běžným diplomových prací vyhrazeno výsledkům a diskuzi, „zabírá“ publikace, u které, vzhledem k důkladnému peer-review řízení, které předchází přijetí manuskriptu do periodika této úrovně, lze jen těžko čekat nějaké (vážnější) nedostatky. Z textu je vidět autorův vhlad do fylogenetické metodiky. Oceňuji množství postupů, kterými se rozhodli testovat pracovní hypotézu i obezřetnost při interpretaci. Přesto mám několik otázek. Jak autor vybíral geny pro konkatenovanou analýzu? Je 69 ortologních genů maximem, které šlo z datasetu 50 genomů vytěžit, nebo použil další kritéria pro selekci těch nejvhodnějších? Zkoušel autor i alternativní metody pro detekci rychle mutujících pozic (compatibility analysis, detekce rychle mutujících pozic na základě jejich rozdělení do jednotlivých gamma kategorií)? Proč v analýzách (alespoň u redukováného či překódovaného datasetu) chybí některá z variant CAT mixture modelu v kombinaci s některým tzv. covarion modelem, které nabízí specializovaná verze programu phylobayes CM?

Doufám, že z výše uvedeného je patrné, že předkládanou práci shledávám naprosto obhájení-schopnou. Hodnotím ji stupněm **Výborně** a Filipovi přeji hodně štěstí v další vědecké kariéře.

V Českých Budějovicích 28. 5. 2012


Aleš Horák

Posudek k diplomové práci Filipa Husníka

Evolutionary origins of intracellular symbionts in arthropods

Magisterská práce Filipa Husníka má formát publikace opatřené 22 stránkovým review. Publikace vyšla v impaktovaném časopise BMC biology (IF 5,2), takže o její kvalitě není třeba pochybovat. Toto uspořádání mě staví do nepříjemné pozice hledat chyby na výsledcích, které prošly kritickým čtením tří spoluautorů, několika recenzentů a editora. Úvodní review, bohužel pro mě, je také řekl bych profesionálně napsané. Není zbytečně rozsáhlé, ale přesto poměrně vyčerpávající co do obsahu. Cituje více než 170 prací a nenašel jsem tam téměř žádné formální nedostatky. Chtěl bych vyzdvihnout zejména dvě tabulky shrnující jednak známé obligátní endosymbiózy u hmyzu (tabulka 1) a publikované kompletní genomy symbiotických bakterií členovců (tabulka 2). Nevím, jaká se situace na poli hmyzích endosymbiontů, ale domnívám se, že souhrnný článek tohoto typu je vždy dobrým počinem a zasloužilo by publikovat, samozřejmě po doplnění výsledků z publikace Husník a kol. 2011.

V souvislosti s úvodním review mě napadají tři otázky:

1. Napadá Vás nějaké vysvětlení, proč mezi endosymbionty hmyzu nenajdeme žádná Archea. U prvoků jsou takové endosymbiózy známé mezi nálevníky s hydrogenozomy a methanogeny.
2. Je známo něco o tom, jak staré je endosymbiotické soužití mezi hmyzem a bakteriemi? Jsou známy fosílie hmyzu s endosymbiotickými bakteriemi? Pokusil se stáří těchto endosymbióz někdo datovat molekulárně a je nebo bylo by takové datování podle Vás věrohodné.
3. Není podivné, že při takovém množství endosymbiontů nedošlo zatím ke vzniku jediné „pravé“ organely – tj. kompartmentu, do kterého bude zaveden transport proteinů a který může být posléze více zapojen do metabolismu buňky. Je možné, že u hmyzu a živočichů obecně existuje nějaké evoluční „constrain“ bránící takové organelogenezi, pokud ano jaké?

Jak jsem již uvedl, přiložený rukopis byl publikován v kvalitním časopise, a proto v něm těžko najdeme větší nedostatky. Mě osobně tento článek zaujal především z metodického hlediska. Zjednodušeně řečeno se v něm autoři věnují boji proti fylogenetickým artefaktům – přitahování dlouhých větví a problému s velice rozdílným zastoupením GC párů. Sám jsem se s těmito problémy ve své práci setkal. Existuje mnoho způsobů, jak s nimi bojovat, ale ani jeden nepřináší zaručené výsledky. U této práce oceňuji především vynalézavost, která vedla k navržení vlastní metody filtrování dat – odstraňování sloupců obsahujících záměny AT z GC nad určitou mez. Tento přístup má jasnou logiku a dokážu si představit zaslouženou radost tvůrců, když to fungovalo. Vzhledem k tomu, že u prezentovaných analýz různé více či méně sofistikované, klasické nebo nově navržené postupy pro boj s fylogenetickými artefakty konvergovaly k podobnému výsledku – rozpadu endosymbiontů do 4 nepříbuzných skupin – jsou výsledky podle mého názoru velmi věrohodné.

Zde bych se rád zeptal, jestli je možné u těchto čtyř skupin endosymbiotických bakterií nalézt nějaký jednotící znak, jiný než molekulárně fylogenetický – např. ultrastrukturální podobnost buňky, podobnost struktury genomu, fúzované geny atp. – nebo jestli je pátrání po takových znacích v tomto případě podle Vás naprostá utopie.

Dále mám jednu drobnou výtku k textu článku. Na straně 4 v pravém sloupci uprostřed píšete, cituji: "...analyses with two different starting trees resulted in two different topologies. When compared by the approximately unbiased test, the topology with four independent origins of endosymbiotic bacteria prevailed over the topology with monophyly of P-symbionts, which **therefore** correspond to a local minimum due to a tree search failure."

Je to velmi nepodstatné, ale nedá mi to poznamenat, že slovo *therefore* zde není na místě. To, že test ukázal, že první topologie je signifikantně lepší, neudělalo z té druhé lokální minimum. Ta druhá totiž lokálním minimem již byla, protože měla nižší log likelihood. Test ukázal, což je důležité, že toto minimum je signifikantně horší než globální minimum, takže je možné jej zavrhnout.

V diskusi na straně 10 vlevo dole mluvíte o metodě odstraňování *heteropecillous sites*. Mohl byste prosím nastínit, o jaké pozice se jedná? Ještě jsem se s touto metodou nesetkal.

Jako nevýhodu metody slow-fast uvádíte celkem správně potřebu zvolit si předem ve fylogenezi dobře definované skupiny. Chtěl bych na tomto místě upozornit na metodu, kterou někteří nazývají *rate fast* a která tímto neduhem netrpí. Pozice se roztřídí do rychlostních kategorií pomocí analýzy maximum likelihood.

Přestože autoři vyzkoušeli mnoho metodických postupů, některé základní obvykle používané postupy v článku postrádám. Je však možné, že je autoři do článku prostě jen nezařadili, protože nepřinesly nic zajímavého. Jde zejména o separátní analýzu, tedy analýzu konkatenátu mnoha genů při zachování nezávislých parametrů evolučního modelu, někdy i délek větví, pro jednotlivé genové úseky, ze kterých se konkatenát skládá. Takový přístup se velmi doporučuje, protože evoluce každého genu má svá specifika, která jsou při „průměrování“ parametrů nedostatečně zohledněna. Zajímalo by mě, jestli jste si této možnosti vědom, zda víte, který software toto umožňuje a proč jste tímto směrem nešel?

Další velmi jednoduchou a rychlou metodou, která by měla pomoci především s artefaktem rozdílného nukleotidového/aminokyselinového složení je distance LogDet. Zkoušeli jste použít tuto distanci.

Konečně, při zjišťování pozice dlouhých větví ve stromu se často používá metoda long-branch exclusion, tedy provedení analýz, ve kterých jsou vždy vyloučeny všechny dlouhé větve až na jednu, aby se nemohly přitahovat. Pozice osamocené dlouhé větve na stromu by měla odpovídat její skutečné poloze, a pokud se dvě dlouhé větve ve „svých vlastních“ analýzách objeví na stejném místě stromu, je možné, že si jsou skutečně příbuzné. Pokoušeli jste se o něco podobného?

Závěrem bych chtěl konstatovat, že předloženou magisterskou práci považuji za práci velmi kvalitní, která bezpochyby splňuje nároky kladené na magisterské práce.


Vladimír Hampel