

Oponentský posudek na magisterskou diplomovou práci

Bc. Magdalena Hrdinová: Zhodnocení cytotypové a morfologické variability spolu s mírou hybridizace u rodu stolístek (Myriophyllum L.) na území ČR a sousedních států

Předložená diplomová práce se zabývá variabilitou rodu stolístek. Literární přehled na začátku práce dobře uvádí do problematiky rodu a logicky z něj vyplývá, proč byly studovány vybrané okruhy témat a proč právě zvolenými metodami. Vyzdvihnout musím také velké množství citované a zjevně skutečně prostudované literatury. Vlastní práce zahrnuje čtyři hlavní metodické přístupy a typy dat, které jsou vhodně kombinovány: stanovení ploidního stupně / velikosti genomu pomocí průtokové cytometrie, morfometrické analýzy, odhad stanovištních nároků a klíčící pokusy s cílem odhadnout intenzitu generativního rozmnožování. Sběr v terénu byl prováděn širším okruhem spolupracovníků, ale vlastní cytometrické a morfologické analýzy, klíčící pokusy a analýzu všech dat prováděla uchazečka; množství dat je na diplomovou práci dostatečné, zejména v cytometrické části nadprůměrné. Závěrečná diskuse vhodně shrnuje výsledky; oceňuji, že se nejedná o prostý souhrn nebo opakování úvodu, ale výsledky jsou porovnávány s existující literaturou a někdy jsou navrhována alternativní vysvětlení, což do diskuse patří.

K použitým metodám a vyhodnocení dat nemám zásadní výhrady (drobnosti viz na konci posudku), byly použity postupy v oboru obvyklé. Výsledky jsou prezentovány velmi detailně mnoha tabulkami a grafy, což by například v článku v časopise nebylo možné, ale v diplomové práci to nevadí. Jediným mírně problematickým bodem je časté používání neparametrických metod zdůvodňované striktně tím, že data nemají normální rozdělení. Biologická data tuto podmínku málokdy splňují a je nutné posoudit a dobře zdůvodnit, zda je narušení tak vážné, aby běžné parametrické testy nešly použít a zda by nebylo možné použít některou z běžných transformací (např. logaritmování). Automatické použití neparametrických testů není nejlepší cesta, i tyto mají svoje předpoklady, které ne vždy musely být splněny (těžko říct bez znalosti primárních dat).

Pokud jde o formální stránku, práce má standardní strukturu i grafickou úpravu. Bohužel je vidět (obvyklý) nedostatek času při dokončování práce – větší množství překlepů a nejasných formulací, volba barev v některých grafech je vysloveně špatná (zejména diskriminační analýzy – překryv skupin není vidět), atd. Zvláště nepřehledné jsou výsledky klíčících pokusů – podle metodiky měly být mimo jiné testovány i vliv způsobu uchovávání plodů (na suchu vs. ve vodě) a vliv uvolnění plodů z plodenství, ve výsledcích ale toto rozdělení nelze dohledat a zdá se, že tyto faktory nakonec analyzovány nebyly a výsledky byly seskupeny podle teplotních režimů bez ohledu na způsob uchovávání. Jazyk by také zasloužil korekci; v této souvislosti si neodpustím poznámku k přejímání anglických výrazů do češtiny – například „distribution“ není vhodné překládat jako „distribuce“, toto slovo má v češtině přece jen jiný význam a zároveň existuje zavedený výraz „rozšíření“.

Při obhajobě prosím o odpověď na následující otázky a připomínky:

1) Během cytometrických analýz nebyly potvrzeny některé ploidní stupně udávané v literatuře, např. di- a tetraploidi u *M. spicatum*. Nemůže v těchto případech jít o chybné určení – jaký je názor autorky? U některých druhů byla zjištěna variabilita ve velikosti genomu, někdy dokonce větší než jsou mezidruhové rozdíly. Jsou diskutovány různé možné příčiny, ale zároveň (str. 118) je tato variabilita zmateně dávana do souvislosti s vysokými koeficienty variance píků u jednotlivých méně kvalitních analýz, což je něco zcela jiného. Domnívá se tedy autorka, že zjištěná vnitrodruhová variabilita skutečně existuje nebo jde pouze o artefakt způsobený sekundárními metabolity apod., které zhoršují přesnost analýz?

2) Výběr znaků pro morfometrické analýzy. Použitá sada znaků obsahuje opakované měření podobných struktur – např. délka 1., 3., 5. a 7. úkrojku listu, všechny vzájemné poměry, rozestupy mezi těmito úkrojky a zároveň průměrný rozstup (délka celého listu / počet párů úkrojků); podobně to platí pro délky internodií na lodyze, atd. Výsledkem je sada vzájemně silně korelovaných znaků, což podle mého názoru způsobuje problémy v některých analýzách (PCA, *knn* - viz níže). Jaký byl důvod takového výběru znaků?

3) Finálním krokem morfometrických analýz byla kanonická diskriminační analýza s cílem odlišit vybrané skupiny a najít znaky, které je nejlépe definují. Kromě celkových testů byly provedeny i dílčí analýzy – rozdíly mezi cytotypy v rámci druhu, mezi americkými a evropskými rostlinami, apod., to velmi oceňuji. Dále

byla provedena klasifikace pomocí metody *k-nearest neighbours*. Důvodem pro použití této metody namísto parametrické (probability-based) klasifikační diskriminační analýzy bylo, že znaky vesměs nemají normální rozdělení – ale viz výše. Výsledky *knn* klasifikace jsou extrémně špatné (nízká úspěšnost), a to i v případech, kdy kanonická diskriminační analýza skupiny jednoznačně odděluje (např. obr. 53 a tab. 35 nebo obr. 61 vs. tab. 46). Na těchto výsledcích jsou potom založeny poněkud pesimistické závěry, že zejména druhy *M. spicatum* a *M. sibiricum* nelze dobře morfologicky poznat – přestože například obr. 59 ukazuje několik znaků, v kterých se tyto taxony liší. Nejde jen o důsledek mimořádně smolně vybrané sady znaků? Metoda *knn* totiž počítá Euklidovské distance, takže přítomnost většího počtu vzájemně korelovaných znaků, které ale skupiny neodlišují a jsou variabilní uvnitř každé ze skupin, může „přebít“ rozdíly mezi skupinami. Bylo by u obhajoby možné – alespoň pro výše uvedené 2 extrémní případy – uvést, jak by vyšla „klasická“ diskriminační analýza a výsledky srovnat? A jak tyto výsledky odpovídají zkušenosti uchazečky se studovanými rostlinami? Skutečně není schopná uvedené taxony rozlišit – tak, jak zkušený botanik rostliny určují „dle celkového vzhledu“?

4) V diskusi (str. 121–122) nalezneme znepokojující informaci, že jedinci byli do druhů *M. spicatum* a *M. sibiricum* klasifikováni podle jediného znaku (počet úkrojků listu), ale že tento znak je v rámci obou taxonů variabilní a rozmezí variability se může překrývat. Z toho podle mne vyplývá, že část rostlin mohla být určena chybně, což může velmi ovlivnit výsledky všech porovnání. Je to tak? Jaký byl podíl měřených jedinců, pro které bylo k dispozici nezávislé určení pomocí molekulárních metod? (Podotýkám, že molekulární analýzy nebyly součástí této diplomové práce a jejich absence není chybou studentky, jde o výsledky širšího kolektivu v rámci grantového projektu.) Jaké další znaky podle literatury odlišují oba druhy (v textu se zmiňuje rozdíl ve tvaru listenů v květenství, ale konkrétní hodnoty chybí)?

5) Diplomová práce má v názvu i v cílech (str. 11) mezidruhovou hybridizaci. Informace o křížencích jsou ovšem i v literárním přehledu kusé, ve vlastních analýzách se kříženci nezmiňují vůbec a v diskusi je jim věnováno půl stránky. Postrádám zejména informace o morfologické variabilitě kříženců – je vůbec něco známo, jsou kříženci intermediární nebo se mohou blížit jednomu z rodičů nebo jsou variabilní a mohou mít unikátní morfologii? Jak se podle názoru autorky stalo, že kříženci nebyli zaznamenáni ani v americkém materiálu, přestože z tohoto území jsou opakovaně doloženi?

Shrnutí. Předložená práce jednoznačně splňuje nároky kladené na magisterské diplomové práce a přináší zajímavé výsledky, které jistě budou v blízké době publikovány. **Práci doporučuji k obhajobě**, vzhledem ke zde uvedeným připomínkám a většímu množství formálních nedostatků navrhuji klasifikaci **velmi dobře**.

V Českých Budějovicích, 8.9.2015

Mgr. Petr Koutecký, Ph.D.

Další připomínky technického rázu, drobnosti, návrhy do budoucna:

Obecné:

- Číslování obrázků a tabulek je trochu chaotické, často jsou za sebou jdoucí obrázky citovány v textu v jiném pořadí, než jsou číslovány a vytištěny v práci.
- Krabicové diagramy (*box & whisker plots*) – není uvedeno, jaké hodnoty jsou vynášeny, předpokládám, že jde o medián, kvartily a rozsah (bez odlehých hodnot)
- U některých výsledků (např. průměrné velikosti genomu pro jednotlivé druhy / cytotypy) by bylo vhodnější uvádět střední chybu průměru (tj. přesnost odhadu) spíše než směrodatnou odchylku (tj. charakteristiku variability).

Literární přehled:

- Pořadí kapitol v literárním přehledu je trochu nelogické, zejména kapitola 2.4 (karyologie) logicky patří hned za morfologickou charakteristiku (2.2).
- Mezi literaturou, která byla použita při výběru znaků pro morfometrickou analýzu, postrádám některé moderní a velmi kvalitní určovací klíče, například pro Rakousko a Německo.
- Tabulka 1 má podle textu obsahovat publikované chromosomové údaje pro 13 taxonů, ale ve skutečnosti obsahuje i nepublikované údaje, takže taxonů tam je 18. Některé údajné ploidní úrovně nebyly cytometrií potvrzeny, ty by stálo za to nějak označit – což by mělo být někde v práci výslovně diskutováno (chyby v určení?)
- U map rozšíření v ČR z databáze Florabase.cz (obr. 5, 8, 11) postrádám informaci, nakolik jsou uvedené údaje ověřené, resp. jaký může být podíl chybných určení. Zejména u druhu *M. verticillatum* bude skutečné rozšíření zřejmě mnohem omezenější než to, co ukazuje mapa.

Cytometrie:

- Při cytometrických analýzách ploidního stupně byly použity směsné vzorky až 10 rostlin, přičemž bylo analyzováno 3000 částic. Do budoucna bych u směsných vzorků doporučoval tuto hodnotu navýšit – vzhledem k tomu, že v analýzách je někdy zřetelné pozadí a poměrně velký podíl jader (odhaduji, že třeba i 50%) připadá na standard, zbývá na každého jedince ve směsném vzorku 100-150 jader, což je na hranici viditelnosti (pokud by byla jedna rostlina z deseti jiná než zbytek).
- Porovnání velikosti genomu Evropa vs. Sev. Amerika (obr. 37 a tab. 11) je zbytečně komplikované; to nejzajímavější (rozdíly mezi kontinenty v rámci jednotlivých druhů) je ve výsledcích poněkud skryto – místo celkového Kruskal-Wallisova testu s následnými mnohonásobnými porovnáními by byly lepší 3 samostatné testy Evropa vs. Amerika pro každý druh zvlášť.

Morfometrika:

- Pro morfometrické analýzy byly chybějící hodnoty některých znaků nahrazovány populačními průměry. To je diskutabilní přístup (vnáší do dat reálně neexistující hodnoty), akceptovatelný jen při malém podílu chybějících – informace, jaký podíl hodnot scházel, ale v práci chybí.
- Tabulka 18 a jí podobné. Skutečně jde o hodnoty kanonických korelačních koeficientů? Spíše jde o koeficienty diskriminační funkce nebo *biplot scores*.
- Je škoda, že výsledky morfometrie nejsou shrnuty do podoby určovacího klíče. Nalezení rozdílů mezi taxony bylo jedním z cílů práce a klíč obsahující přesné hodnoty znaků založené na docela velkém souboru dat by jistě v blízké budoucnosti našel využití.

Klíčení:

- Ve výsledcích klíčících pokusů jsou nápadné mezipopulační rozdíly – některé populace prakticky ve všech teplotních režimech klíčí lépe než jiné. Tento fakt není v práci diskutován, ale pro interpretaci výsledků i podobné pokusy v budoucnu je zásadní. Pokud je velká variabilita mezi populacemi, výběr populací může výsledky zcela náhodně ovlivnit a může způsobit až zcela protichůdné závěry mezi různými studiemi nebo opakováními (ať už je variabilita daná čímkoliv – nějakými „vnitřními“ příčinami jako genetická variabilita nebo „vnějšími“ jako jsou příhodnější podmínky prostředí pro úplné dozrání semen nebo úplně náhodou, například mírně různou dobou sběru semen)