



## Oponentský posudek na diplomovou práci

Bc. Gabriela Pravcová: Vztah pylové viability, ploidie a velikosti genomu u střeoevropských zástupců rodu *Ornithogalum* (snědek)

Tato diplomová práce se zabývá jednou konkrétní fází reprodukčního cyklu rodu *Ornithogalum* a dává do souvislosti úspěšnost tvorby pylu a ploidii (resp. velikost genomu, která ale ploidii odráží). Kromě popisu stavu u konkrétních studovaných druhů testuje některé obecné předpoklady (např. očekávaná nižší životaschopnost u kříženců a lichých ploidních stupňů).

Práce má zcela standardní strukturu, přiměřenou délku, slušnou jazykovou úroveň, a po formální stránce zcela vyhovuje požadavkům na magisterské diplomové práce.

Úvod práce obsahuje rešerši k hlavním předmětům práce – velikost genomu, polyploidie, viability pylu a základní informace o studovaném rodu. Zejména první dvě témata jsou na rešerši vlastně velmi těžká, protože k nim existuje obrovské množství literatury a různých pohledů. Studentka shrnuje ty nejčastější, občas trochu zkratkovitě a nepřesně, ale to je pochopitelné, vlastně si zvolila nejtěžší možnou cestu. Obecně napsaná rešerše nutně musela zůstat na povrchu a obsahuje i informace, které nejsou dále využity (např. genetické nebo fyziologické aspekty). Podle mě lepší a jednodušší by bylo vybrat cíleně jen témata, která s prací přímo souvisí, a těm se věnovat víc do detailu – např. jev označovaný jako *genome downsizing*. Určitá nedotaženost a nepropojenost úvodu a zbytku práce je vidět i v dalších částech, uvedu jen několik příkladů. Například v části o viability pylu (str. 7 a 8) je uvedeno, že termín viability má spoustu možných významů a že barvitelnost pylu je pouze odhadem životaschopnosti a tyto termíny by se neměly zaměňovat – a přesto se v celé práci termín viability v rozporu s tímto tvrzením používá. V popisu rodu jsou citována různá taxonomická pojetí, ale vlastně mi z celé této části není jasné, které z nich tedy bylo použito a proč. V popisu skupiny snědku chocholičnatého (*O. umbellatum* agg.) jsou popisovány dosud známé ploidie, ale bohužel zde chybí informace o existenci heptaploidů – přestože takové rostliny byly studovány a objevují se ve výsledcích. Shrnutí - celkově se mi úvod práce líbí, obsahuje spoustu informací a hlavní témata jsou dobře zvolená, ale k dokonalosti mu ještě něco málo chybí.

Metodika práce je přiměřeně podrobná, kromě statistického vyhodnocení, které by mělo být popsáno podrobněji (test X byl použit na data Y, atd.). Chybí také popis, jak byla počítána "procentuální variabilita", která je pak vynášena do grafů a analyzována a interpretována – není mi jasné, co je 100%, při zkoumání výpočtu z přiložených tabulek to přesně nevychází ani na minimum, ani na maximum, ani na průměr.

Výsledky práce jsou dobře a poctivě dokumentovány tabulky a grafy, s jedinou výjimkou. Nikde v práci není přesně uvedeno, na základě čeho je stanovena ploidie jednotlivých taxonů, což je zcela nezbytný údaj použitý pro výpočet monoploidní velikosti genomu. Dozvíme se pouze, že "u části vzorků byla ploidie kalibrována počtem chromozomů z předchozího výzkumu (M. Hroneš, osobní sdělení) a u části vzorků byly použity údaje z literatury." Toto mělo být přehledně shrnuto v nějaké tabulce – pro které taxony který typ dat, odkud údaje pochází (ze stejných populací, jako byly analyzovány zde?). Nejde sice o vlastní výsledky studentky, ale s řádnou citací v práci mohly (měly) být.

Dále by šlo vylepšit některé drobnosti v prezentaci dat. V grafech, kde je na horizontální ose ploidie, se body pro diploidy překrývají s vertikální osou a nejsou vidět. U krabicových grafů pro taxony by bylo dobré uvádět ploidie ke všem taxonům nebo je třeba odlišit barevně (čtenář, který nemá studovaný rod "v malíčku", si to jinak musí stále dohledávat).

Samostatnou kapitolou je statistické vyhodnocení dat. Zde je bohužel napáchána řada chyb, které značně ovlivňují interpretaci výsledků. Pokusím se shrnout ty hlavní:

a) Pro porovnání hodnot více než dvou skupin byla použita analýza variance. To je správně, ale bohužel v práci vesměs chybí výsledky hlavního testu a jsou prezentována až mnohonásobná porovnání (ta se ale



dělají, pouze pokud je hlavní test průkazný, čili bez jeho výsledku nemají smysl). Pro mnohonásobná porovnání byl použit někdy Tukeyho test a někde Dunnův test (proč dva různé testy, souvisí to nějak s charakterem dat?). Dunnův test je velmi slabý, zejména při nízkém počtu pozorování, takže jeho výsledky jsou v tomto případě dost nespolehlivé – nejlepším příkladem je obr. 7 s nepřekrývajícími se skupinami, ale zároveň mnoho neprůkazných porovnání (tabulka 2), což jednoznačně víc odráží vlastnosti použitého testu než biologickou realitu.

b) Použití t-testů a analýzy variance a stejně tak lineární korelace pro data v procentech je problematické. Tato data dost výrazně narušují předpoklady testů (nehomogenita variance a nesymetrické rozdělení, zejména v blízkosti 0% a 100%) a hladiny významnosti jsou velmi nespolehlivé. Oba efekty by byly jistě dobře vidět, pokud by byla provedena příslušná regresní diagnostika. Zde by určitě byla na místě transformace závislé proměnné (arcsin) a ještě lépe použití zobecněných lineárních modelů.

c) Podobně to platí i pro data zahrnující "procentuální variabilitu" – zde je rozdělení nápadně nesymetrické. Kromě problémů v analýze variance zkresluje odlehle body sklon modelovaných lineárních závislostí (parametr označovaný jako *leverage* – např. na obr. 18b, 20b, 20c je sklon přímky do značné míry dán jediným bodem, kterým je triploidní kříženec *O. ×wildtii*; podobně hodnoty *O. pyrenaicum* silně ovlivňují sklon přímky na obr. 2). A v neposlední řadě je problém, pokud se porovnávají absolutní hodnoty této "procentuální variability": pro taxon s rozmezím 10–20 vyjde "procentuální variabilita" 100%, pro rozmezí 90–100 vyjde 11%, i když jsou rozpětí vlastně stejná (o 10).

d) Použití lineární závislosti (korelace a regrese) není vždy automatické – viz otázky na konci posudku.

Diskuse práce je poměrně rozsáhlá a konfrontuje jednotlivá zjištění s literaturou. To velmi chválím, přesně to je podstata diskuse. Jen je škoda, že vlastně zůstalo u jednotlivostí, a chybí mi nějaké celkové shrnutí – tak se na něj alespoň zeptám u obhajoby.

Otázky k obhajobě:

- 1) Pokud pro vztah dvou proměnných použijeme lineární regresi (nebo korelaci), vyjadřujeme tím určité předpoklady a hypotézy. Jaké? A je toto vždy splněno v případě vašich dat?
- 2) V práci je korelována velikost genomu (C- i Cx hodnota) a viabilita pylu. Zároveň jde o složitý polyploidní komplex a vliv velikosti genomu a ploidie nelze separovat. Jsou známy jiné studie (v jiných rodech), kde opravdu vyšla závislost viability a velikosti genomu jako takové, bez vlivu plodie? Jak je vztah vysvětlován?
- 3) V diskusi mi chybí krátké shrnutí – co vyšlo podle předpokladů a které výsledky byly naopak neočekávané a které z nich jsou pro vás nejvíc překvapivé?
- 4) V diskusi zmiňujete, že u snědků se příliš nepotvrdil předpoklad, že by sudé ploidie měly mít dobře vyvinutý pyl (v některých případech neměly) a liché by měly mít životaschopnost nižší (někdy měly dobrou). Také píšete, že o příčinách lze jen spekulovat – tak tedy spekulujte: čím by to mohlo být, jaké hypotézy k dalšímu testování lze navrhnout, atd.

Závěr. Předložená práce splňuje nároky kladené na magisterské diplomové práce. Studentka odvedla poměrně dost práce (a to i různé restriktce kvůli epidemii covid-19) a získala data, která jsou výborným základem pro další studium snědků. Práci tedy jednoznačně doporučuji k obhajobě; vzhledem k určitým nedostatkům navrhuji v případě zdařilé obhajoby hodnocení **velmi dobře**.

V Českých Budějovicích, 20. 8. 2021

Mgr. Petr Koutecký, Ph.D.