

Univerzita Karlova

Přírodovědecká fakulta

Studijní program: Biologie (N1501)

Studijní obor: NBOTA (1507T004)



Bc. Hana Mašková

Biosystematická studie okruhu *Carlina vulgaris* ve střední Evropě  
s využitím molekulárních a morfometrických metod

Biosystematic study of the *Carlina vulgaris* complex in Central Europe  
based on molecular and morphometric approaches

Diplomová práce

Vedoucí práce: Mgr. Zdeněk Kaplan, Ph.D.

Konzultant: Mgr. Tomáš Fér, Ph.D.

Praha 2018

## Prohlášení

Prohlašuji, že jsem závěrečnou práci zpracovala samostatně a že jsem uvedla všechny použité informační zdroje a literaturu. Tato práce ani její podstatná část nebyla předložena k získání jiného nebo stejného akademického titulu.

V Praze, 10.8.2018

Podpis: .....

## Poděkování

Děkuji. Děkuji všem, kteří se spolu se mnou podíleli na této práci. Předně svému školiteli Mgr. Zdeňkovi Kaplanovi, Ph.D. a konzultantovi Mgr. Tomášovi Férovi, Ph.D. Pak lidem, kteří ochotně vozili vzorky pupav z celé Evropy, a zaměstnancům Botanického ústavu AV ČR, kteří se o rostliny starali v experimentální zahradě v Průhonicích. Dále Ing. Lence Flaškové za pomoc s přípravou vzorků pro analýzu DNA i personálu sekvenační laboratoře v BioCevu. Děkuji prof. Ing. Václavovi Hlaváčovi, CSc. za vlídné přijetí na půdě ČVUT a jeho studentovi Bc. Ondřejovi Holešovskému za pomoc s programem ToothFinder. Děkuji doc. Mgr. Patrikovi Mrázovi, Ph.D. a RNDr. Zdeňkovi Janovskému, Ph.D. za rady ohledně práce v programu R. A za další rady děkuji neznámým lidem, kteří nezištně odpovídají na dotazy na různých internetových fórech. Děkuji rodině, zejména svému nastávajícímu manželovi Ing. Martinovi Jeřábkovi, za všestrannou podporu. Děkuji všem botanikům, kteří se přede mnou problematice věnovali. A děkuji všem lidem, kterým není lhostejný osud naší přírody a potažmo i botanických lokalit (včetně těch na kterých rostou pupavy Biebersteinovy pravé). Děkuji.

## Abstrakt

Okruh pupavy obecné (*Carlina vulgaris* agg.) představuje v rámci střední Evropy několik ekologicky, morfologicky a chorologicky odlišených linií. V rámci této práce byly studovány vztahy mezi taxony rozeznávanými na území České republiky: pupava obecná pravá (*Carlina vulgaris* subsp. *vulgaris*), pupava Biebersteinova pravá (*C. biebersteinii* subsp. *biebersteinii*), pupava Biebersteinova prostřední (*C. biebersteinii* subsp. *brevibracteata*) a pupava Biebersteinova jesenická (*C. biebersteinii* subsp. *sudetica*).

Molekulární analýzy ukázaly na přítomnost dvou geneticky odlišených skupin, první zahrnuje zástupce reliktních stanovišť a horské rostliny klasifikované jako *C. biebersteinii* subsp. *biebersteinii* a *C. biebersteinii* subsp. *sudetica*, druhá pak rostliny řazené k *C. vulgaris* a *C. biebersteinii* subsp. *brevibracteata*. Diferenciace obou skupin je podpořena i morfometrickými analýzami. Vzájemné vztahy populací uvnitř těchto dvou skupin však zůstávají nejasné.

Jisté je, že české populace *C. biebersteinii* subsp. *biebersteinii* a *C. biebersteinii* subsp. *sudetica* jsou příbuzné horským populacím z Alp a Karpat; jedná se o glaciální relikty zasluhující zvláštní pozornost z hlediska ochrany přírody. Nicméně statut samostatného endemického taxonu pro jesenickou populaci *C. biebersteinii* subsp. *sudetica* není pravděpodobně oprávněný.

Klíčová slova: *Carlina vulgaris* agg., *Carlina biebersteinii*, endemismus, mikroevoluce, postglaciální diferenciace, taxonomie

## Abstract

The *Carlina vulgaris* complex in central Europe includes several lineages defined by their ecology, morphology and distribution. This diploma thesis is focused on relationships between the taxa recognized in the Czech Republic, namely *Carlina vulgaris* subsp. *vulgaris*, *C. biebersteinii* subsp. *biebersteinii*, *C. biebersteinii* subsp. *brevibracteata* and *C. biebersteinii* subsp. *sudetica*.

Molecular analysis revealed two genetically defined groups. One includes samples from relict populations in western Bohemia and from high mountains classified as *C. biebersteinii* subsp. *biebersteinii* and *C. biebersteinii* subsp. *sudetica*. The other is represented by plants classified as *C. vulgaris* and *C. biebersteinii* subsp. *brevibracteata*. This genetic differentiation was also confirmed by morphometric analysis. However, relationships within these two groups remain unclear.

The Czech populations of *Carlina biebersteinii* subsp. *biebersteinii* as well as of *C. biebersteinii* subsp. *sudetica* are closely related to the mountain populations in the Alps and Carpathians. Their occurrence in the Czech Republic is relict and they should be in focus of nature conservation. However, the separate taxonomic position of the claimed endemic *C. biebersteinii* subsp. *sudetica* is probably unjustified.

Key words: *Carlina vulgaris* complex, *Carlina biebersteinii*, endemic, microevolution, postglacial differentiation, taxonomy

# Obsah

1. Úvod	1
2. Teoretický úvod	2
2.1. Změny v rozšíření rostlin a jejich genetické aspekty	2
2.2. Přehled studovaných taxonů	3
2.3. Použité metody	7
2.3.1. Morfometrické analýzy	7
2.3.2. Molekulární analýzy	8
2.3.2.1. NGS metody pro rekonstrukci fylogeneze	8
2.3.2.2. Hyb-Seq	10
3. Materiál a metodika	13
3.1. Vzorky rostlin	13
3.2. Morfometrická analýza	14
3.3. Izolace DNA	17
3.4. Příprava knihovny pro Hyb-Seq	18
3.4.1. Sonikace	18
3.4.2. End repair and A-tailing (opravy délky fragmentů a navázání A)	19
3.4.3. Ligace adaptorů	19
3.4.4. Přečištění produktů	19
3.4.5. Délková selekce DNA	19
3.4.6. PCR obohacení (enrichment)	20
3.4.7. Přečištění PCR produktů	21
3.4.8. „Poolování“ knihovny	21
3.4.9. Hybridizace (obohacení) knihovny	21
3.4.10. Post-capture amplifikace	22
3.5. Vyhodnocení získaných dat	23
3.5.1. Molekulární analýzy	23
3.5.2. Morfometrické analýzy	30
3.5.2.1. Použité soubory dat	30
3.5.2.2. Vlastní analýzy	30
4. Výsledky	33
4.1. Sekvence DNA, vstupní data	33

4.2.	Vlastní fylogenetické analýzy	35
4.3.	Morfometrické analýzy	44
4.3.1.	Test správnosti odečtu znaku počet ostnitých zubů	44
4.3.2.	Vliv rozdílných podmínek na fenotyp rostlin	44
4.3.3.	Morfologické rozdíly mezi geneticky definovanými skupinami	46
4.3.4.	Znaky vhodné k odlišení definovaných skupin	51
4.3.5.	Ekologická diferenciacce	56
4.3.6.	Pozice taxonu <i>Carlina biebersteinii</i> subsp. <i>sudetica</i> v rámci studované skupiny	57
5.	Diskuze	60
5.1.	Možnosti a omezení použitých metod pro rekonstrukci fylogeneze skupiny	60
5.2.	Genetická podobnost a evoluční vztahy	63
5.3.	Morfologická variabilita	63
5.4.	Populace z reliktních lokalit	65
5.5.	Taxonomická pozice a hodnocení <i>C. biebersteinii</i> subsp. <i>brevibracteata</i>	66
6.	Závěr	67
7.	Seznam použité literatury	68
8.	Přílohy	78
8.1.	Seznam vzorků	78
8.2.	Výsledky sekvenace DNA a mapování na referenční sekvence	81
8.3.	Seznam znaků použitých pro morfometrické analýzy	84
8.4.	Skript použitý pro získání vstupních souborů do programu ToothFinder (Corney et al. 2012b)	86
8.5.	Popisné statistiky měřených znaků	88
8.6.	Přehled balíčků a funkcí statistického programu R použitých v této práci	90
8.7.	Výstupy PCA a CDA	91

# 1. Úvod

Střední Evropa, kde se setkává řada migračních tras, je v dlouhodobém měřítku jevištěm řady fylogeografických procesů a vegetačních změn. K nejprostudovanějším patří ty spojené s koncem doby ledové (o čemž vypovídá řada paleobotanických dat i fylogeografických studií, např. Kajtoch et al. 2016, Taberlet et al. 1998). V krajině se začaly více uplatňovat lesní dřeviny, které vytlačily mnohé světlomilné a horské druhy, či izolovaly jejich zbytkové populace v refugiích. Právě vzájemná izolace a malá velikost populací spolu se specifickými podmínkami prostředí může působením genetického driftu urychlit evoluční procesy včetně speciace (Kruckeberg & Rabinowitz 1985). Na druhou stranu oteplení umožnilo migraci řady teplomilných druhů do střední Evropy. V místech, kde došlo k opětovnému překryvu dříve oddělených areálů, pak mohlo docházet k hybridizaci, která může v místním či regionálním měřítku vést k tomu, že se jinde dobře odlišitelné druhy stávají taxonomicky problematickou skupinou (Hewitt 2008).

Výše zmíněné procesy mohly hrát svou úlohu i v okruhu pupavy obecné (*Carlina vulgaris* agg.), který je tak ideální modelovou skupinou pro výzkum evolučních procesů, které se podílely a podílejí na formování druhů středoevropské květeny. V rámci skupiny totiž vznikla pestrá směsice různých ekomorfologicky odlišitelných typů. Jsou zde zástupci ze stanovišť reliktní povahy s možnou návazností na alpské populace (klasifikované jako *Carlina biebersteinii* subsp. *biebersteinii*, v karech Hrubého Jeseníku dokonce jako samostatný poddruh subsp. *sudetica*), dále stepní typy xerothermních trávníků (patřící zřejmě k poddruhu *C. biebersteinii* subsp. *brevibracteata*, jehož moravské lokality jsou na okraji areálu v panonské oblasti), nebo rostliny osidlující převážně antropogenní stanoviště, jako jsou opuštěné lomy, násypy komunikací či výsypky (hodnocené jako *C. vulgaris* s. str.).

Cílem předkládané práce je objasnění evolučních vztahů této skupiny a zhodnocení morfologické variability rozlišených taxonů. Dílčí otázky jsou následující: 1) Jaké genetické skupiny lze v rámci studovaného okruhu *Carlina vulgaris* agg. vylišit? 2) Jaká je genetická diferenciace mezi taxony rozeznávanými v současnosti na území ČR? 3) Jaké znaky lze použít pro morfologické odlišení geneticky definovaných skupin? 4) Jsou na stanovištích refugiálního charakteru v nižších a středních polohách (sutě a hluboce zaříznutá údolí řek v západních Čechách, v údolí Vltavy a v údolích v Praebohemiku) reliktní rostliny blízké *C. biebersteinii*? 5) Nakolik se jesenická *C. biebersteinii* subsp. *sudetica* liší od *C. biebersteinii* subsp. *biebersteinii* z Alp a Karpat?

Vedle biosystematické studie zajímavé v širším kontextu postglaciálních migrací a evoluce reliktních skupin je výstup předkládané práce užitečný i z hlediska ochrany přírody. Přináší údaje o variabilitě (genetické, morfologické) a evoluční historii, které jsou důležité pro efektivní ochranu ohrožených taxonů, jakými jsou všechny tři poddruhy *C. biebersteinii*: subsp. *biebersteinii* je zařazena v Červeném seznamu cévnatých rostlin ČR do kategorie C3, subsp. *brevibracteata* je v kategorii C4b a subsp. *sudetica* v kategorii C1r (Grulich 2017). Zvláště poslední jmenovaný poddruh, jeden z mála doposud neprobádaných endemitů ČR, je fyto geograficky velmi významným prvkem a zasluhuje tak mimořádnou pozornost.

## 2. Teoretický úvod

### 2.1. Změny v rozšíření rostlin a jejich genetické aspekty

Současné areály organismů jsou výsledkem řady procesů. V dlouhodobém měřítku to je ve střední Evropě především střídání dob ledových a meziledových. Se změnami klimatických podmínek dochází k změnám v rozšíření organismů. V rámci Evropy lze z pohledu migrační historie rozdělit dnešní biotu v zásadě do tří skupin (Schmitt 2007). První zahrnuje druhy s mediteránními refugii, u nichž se předpokládá, že během glaciálu přežívaly pouze v jižní Evropě (Iberský, Apeninský a Balkánský poloostrov), a částečně i v jihozápadní Asii. Druhou představují kontinentální druhy, jejichž refugia byla zřejmě v perialpínských oblastech, karpatské pánvi a též na části Balkánského poloostrova; spadají sem i druhy s asijskými refugii (jižní Ural). V případě těchto zmíněných dvou skupin v interglaciálech dochází k opětovné kolonizaci oblastí, které během dob ledových nejsou vhodné pro přežití těchto druhů, na řadě míst se tak setkávají dlouhodobě oddělené linie, může dojít k jejich zkřížení a vzniku hybridních zón. Z oblasti střední Evropy, kde se migrační trasy potkávají, je řada z nich doložena (více viz review Hewitt 1999).

Právě hybridizace je jedním z klíčových evolučních dějů. Na jedné straně může přispět k vyšší genetické variabilitě a tím umožnit vznik nových vlastností, na straně druhé může vyústit ve vznik neplodných hybridů nebo jedinců se sníženou životaschopností (zde se pak uplatňuje disruptivní selekce a opětovný překryv areálů dvou dříve oddělených populací povede k posílení reprodukčních bariér mezi nimi). Situace může být různorodá i v rámci jedné oblasti, kde se linie potkávají, a sice vlivem heterogenity prostředí a různých nároků organismů i jejich disperzních schopností (vznikají tzv. mozaikové hybridní zóny) (Harrison 1993). Jaký vliv má křížení na původní taxony a linie se též liší případ od případu. Některé druhy se mohou vyskytovat sympatricky a příležitostně vytvářet hybridy, aniž by ztrácely svoji genetickou identitu (Steeves et al. 2010), jindy dochází ke smývání rozdílů (vliv introgrese) mezi původními liniemi (Allendorf et al. 2001, Ayres et al. 2003).

Změny klimatu však mohly v řadě případů mít i přesně opačný důsledek, a sice zmenšení areálů (takové druhy tvoří třetí skupinu z pohledu různé postglaciální migrační historie). Jejich areál zahrnoval rozsáhlé stepní periglaciální oblasti; typicky se sem řadí rostliny s alpským a alpsko-arktickým rozšířením (Schmitt 2007), které s oteplením migrovaly za vhodnými podmínkami severněji, případně zůstaly v izolovaných refugiích v horách. Podobnou historii však měla i celá řada dalších druhů, které nebyly schopné čelit nastupující konkurenci dřevin. Jejich refugii se tak, v závislosti na jejich konkrétních nárocích, mohly stát různé typy nelesních stanovišť, skalní výchozy (zvláště v říčních údolích), sutě, strmé svahy s častými sesuvy půdy, hadce, slatiniště, rašelinště, lavinové dráhy nebo kary). I s druhy, které mají centrum rozšíření v Alpách a Karpatech, se tak lze u nás setkat na těchto reliktních stanovištích i v nižších polohách (Kaplan 2017).

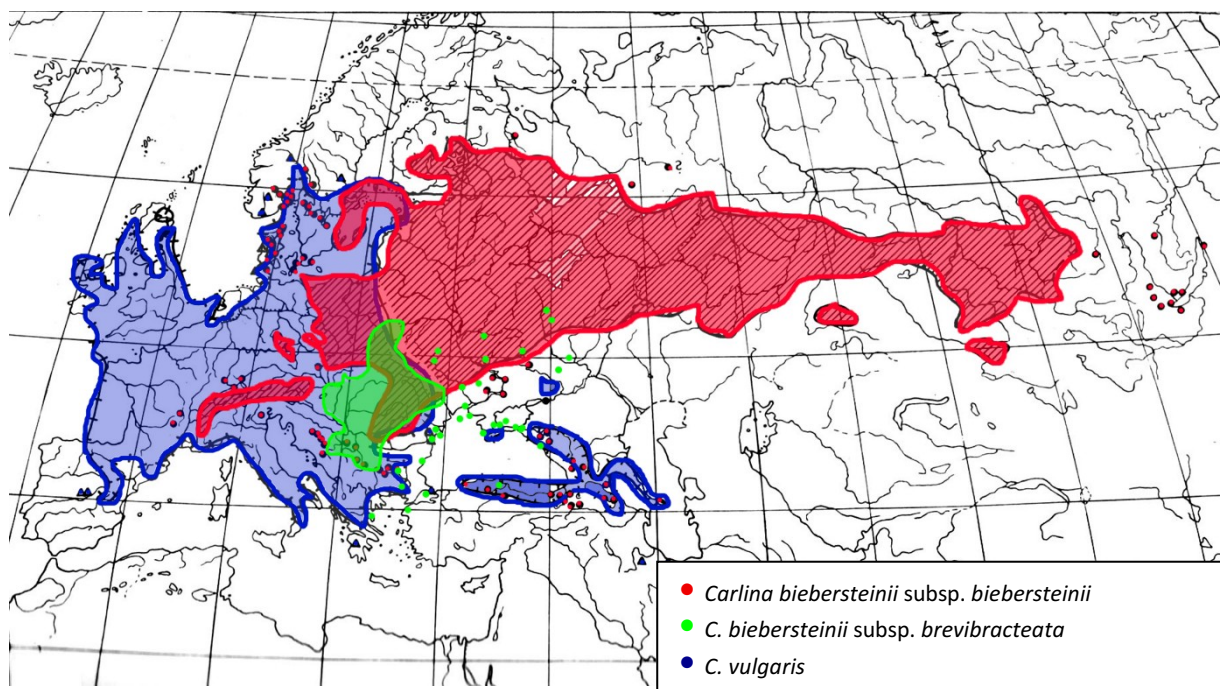
Další druhy však mohly vykazovat disjunktí rozšíření i během glaciálu (ač adaptované k chladu, byly limitované jinými faktory, zejména suchem). Řada evropských pohoří tak má své endemické linie. Zástupci těchto populací mohou se změnou podmínek migrovat do



vyšších či nižších nadmořských výšek pouze v rámci jednoho horského systému či oblasti a zůstat tak izolovány od dalších linií i během glaciálů (Schmitt 2007). Zmíněná fragmentace areálů s sebou přináší omezení genového toku a tím těž zvýšení genetické diferenciace mezi jednotlivými populacemi. Vlivem náhodných mutací, genetickým driftem, ale i adaptací na konkrétní podmínky dané oblasti může dojít k alopatrické speciaci. V malých populacích se navíc uplatňuje efekt hrdla láhve, je zde snížena heterozygotita, náhodné vlivy obecně se zde projevují daleko více, což je případ řady endemických druhů (Kruckeberg & Rabinowitz 1985).

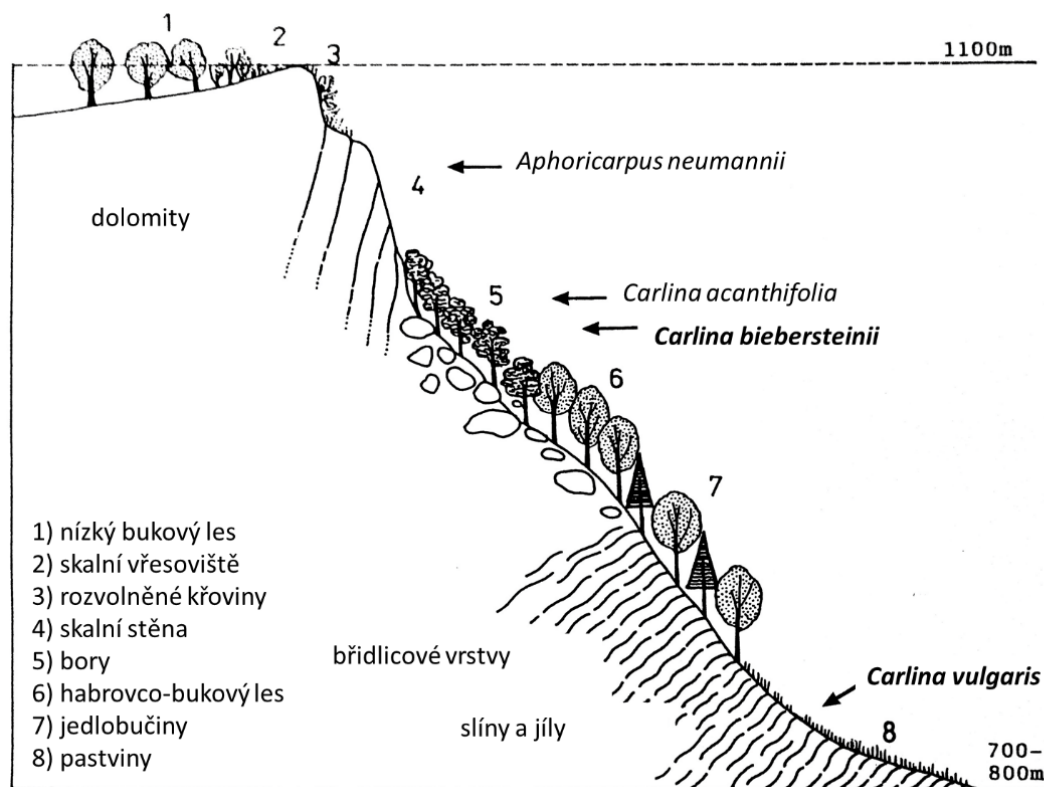
## 2.2. Přehled studovaných taxonů

Rod *Carlina* (česky pupava) je dle recentní fylogenetické studie (Barres et al. 2013) součástí podtribu *Carlininae*, tribu *Cardueae* a podčeledi *Carduoideae* v rámci čeledi *Asteraceae* (*Compositae*). Tato práce se týká skupiny *Carlina vulgaris* agg., která v rámci rodu spadá do sekce *Carlina* a subsekce *Carlina* (Meusel & Kästner 1990). Skupina zahrnuje druhy *C. vulgaris* a *C. biebersteinii*, v rámci kterých je rozeznáváno několik taxonů na úrovni poddruhů a variet. Vedle sarmatských a baltických typů, označovaných jako *C. biebersteinii* subsp. *biebersteinii* var. *fennica* a jihoevropských *C. vulgaris* subsp. *spinosa*, se jedná o taxony, jejichž areály se prolínají ve středoevropském prostoru (viz obr. 1). Je to nominátní poddruh *C. vulgaris* subsp. *vulgaris*, který se vyskytuje od mediteránu přes jihozápadní Evropu po jih Skandinávie, dále nominátní varieta *C. biebersteinii* subsp. *biebersteinii* var. *biebersteinii*, která má rozsáhlý souvislý areál od východní Evropy po Sibiř a dále několik arelů v evropských pohořích (Vogézy, Alpy, Karpaty) a poddruh *C. biebersteinii* subsp. *brevibracteata* známý z ponticko-panonské oblasti (Meusel & Kästner 1994, Meusel 1993). V České republice k předchozím ještě přibývá relativně nedávno popsán taxon *C. biebersteinii* subsp. *sudetica*, endemit Hrubého Jeseníku, znám pouze z Malé a Velké kotliny (Kovanda 2002).



Obr. 1 Rozšíření zájmových taxonů. Převzato (Körner & Meusel 1986, Meusel 1993) a upraveno.

Z výše uvedeného plyne, že se na našem území v rámci studované skupiny můžeme setkat se zástupci s pravděpodobně různou postglaciální historií. *Carlina biebersteinii* subsp. *biebersteinii* a subsp. *sudetica* v evropských pohořích mohou být příkladem glaciálního reliktu, zatímco zbylé dva taxony mohou reprezentovat dvě genetické linie šířící se v postglaciálu z různých refugií a setkávající se ve střední Evropě. Vyloučená není hybridizace v těch případech, kdy se vzhledem k přítomnosti zařízých údolí se strmými svahy poměrně blízko k sobě dostávají i jinak ekologicky dobře oddělené nominální podruhy *C. biebersteinii* subsp. *biebersteinii* a *C. vulgaris* subsp. *vulgaris*. Taková situace byla popsána z balkánského pohoří Igman, kde se zástupci různých vegetačních jednotek, obr. 2, mohou dostat do kontaktu např. prostřednictvím disturbance v podobě laviny, či vlivem lidské činnosti (Meusel & Kästner 1994).



Obr. 2 Schéma vegetačního krytu svahů štítu Igman blízko Sarajeva. *Carlina biebersteinii* (a *C. acanthifolia*, druh podobný *C. acaulis*) roste ve světlých lesích při úpatí skal, *C. vulgaris* níže, na pastvinách. Převzato (Meusel & Kästner 1994) a upraveno.

Přítomnost morfologicky intermediárních jedinců (Kubát 2013) mezi taxony *C. biebersteinii* subsp. *brevibracteata* a *C. vulgaris* subsp. *vulgaris* by naznačovala též možné křížení těchto taxonů a výskyt hybridní zóny, která prochází Českou republikou. Stírající se rozdíly mezi nimi pak mohou být důvodem různorodých pohledů na taxonomii okruhu *C. vulgaris* agg. na našem území. Literární rešerše na toto téma byla předmětem předchozí práce (Mašková 2015), proto jsou zde uvedeny pouze základní znaky používané k odlišení taxonů a krátká charakteristika (formou dichotomického klíče dle monografie rodu *Carlina* (Meusel & Kästner 1994) a článku věnovaného rostlinám okruhu *C. vulgaris* agg. v ČR (Kubát 2013)).

**1a** Přízemní a lodyžní listy peřenolaločné až peřenosečné, úkrojky s nestejně dlouhými ostny. Čepel, zvláště u středních a horních lodyžních listů, je zprohýbaná, na okraji zvlněná až kadeřavá, listy jsou kožovité. Úbory (10–)12–16(–18) mm v průměru (dle Kubáta (2013) až 22 mm), listeny vnějšího zákrovu kratší než úbor (obr. 3c).

..... *Carlina vulgaris* subsp. *vulgaris*

Roste na křovinatých stráních, pastvinách, mezích, v lesních lemech a ve světlejších lesích, dále na antropogenních stanovištích jako jsou lomy a výsypky. Ze zahrnutých taxonů je nejhojnější. Jedná se o velmi variabilní druh. (Autoři Meusel & Kästner (1994) rozlišují 3 různé morfotypy, *planifolia*, *orophila* a *polycephala*, kterým však nepřikládají taxonomickou hodnotu.)

**1b** Přízemní a lodyžní listy ploché, obvykle nedělené, okraj listu pravidelně ostnitý. Pakliže se vytváří ostnité úkrojky, jsou v rovině čepel. .... **2**

**2a** Dolní a střední lodyžní listy úzce kopinaté, horní vejčité kopinaté až vejčité, krátce špičaté, často podle střední žilky složené. Dolní polovina čepel často s několika ostnitými úkrojky. Lodyžní listy pod úborem dlouhé 11–20(–25) mm. Úbory (10–)12–17(–20) mm široké (dle Kubáta širší, až 25 mm). Listeny vnějšího zákrovu kratší než úbor (obr 3b). Rostliny často vzrůstné ((25–)40–100(–125) cm vysoké), zpravidla větvené, s (1–)3–10(–40) úbory (nicméně Kubát (2013) uvádí jen 1–3(–5) úborů).

..... *C. biebersteinii* subsp. *brevibracteata*

Rostliny xerothermních trávníků. Zasahují na jižní Moravu a s morfologicky obdobnými typy se lze setkat i v Českém středohoří (severozápadní hranice areálu tohoto taxonu, který byl popsán z ponticko-panonské oblasti, není dostatečně známa, Meusel & Kästner (1994) citují z území ČR pouze 1 herbářovou položku).

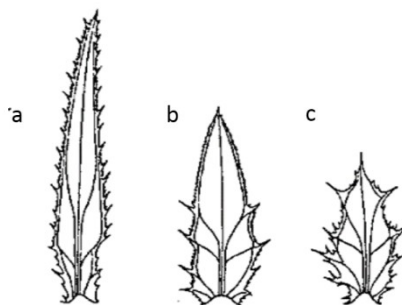
**2b** Přízemní listy přecházejí v postupně se zmenšující listy lodyžní (obr. 4a), které jsou podlouhlé, úzce kopinaté, do špičky pozvolna se zužující, ploché, podle střední žilky neprohnuté. Okraj listu obvykle celistvý, pravidelně ostnitý, případně vtroušeny skupiny ostének, ostnité úkrojky mohou být patrné na čepeli středních a horních lodyžních listů. Úbory větší (15–)19–28 (–45) mm. Listeny vnějšího zákrovu podobné horním lodyžním listům (obr 3a), úzce kopinaté, stejně dlouhé jako úbor nebo delší, rostliny vysoké (20–)30–60(–95) cm .... **3**

**3a** Úbory do 39 mm v průměru –V ČR velmi vzácně (Manětín, údolí Teplé)

..... *C. biebersteinii* subsp. *biebersteinii*

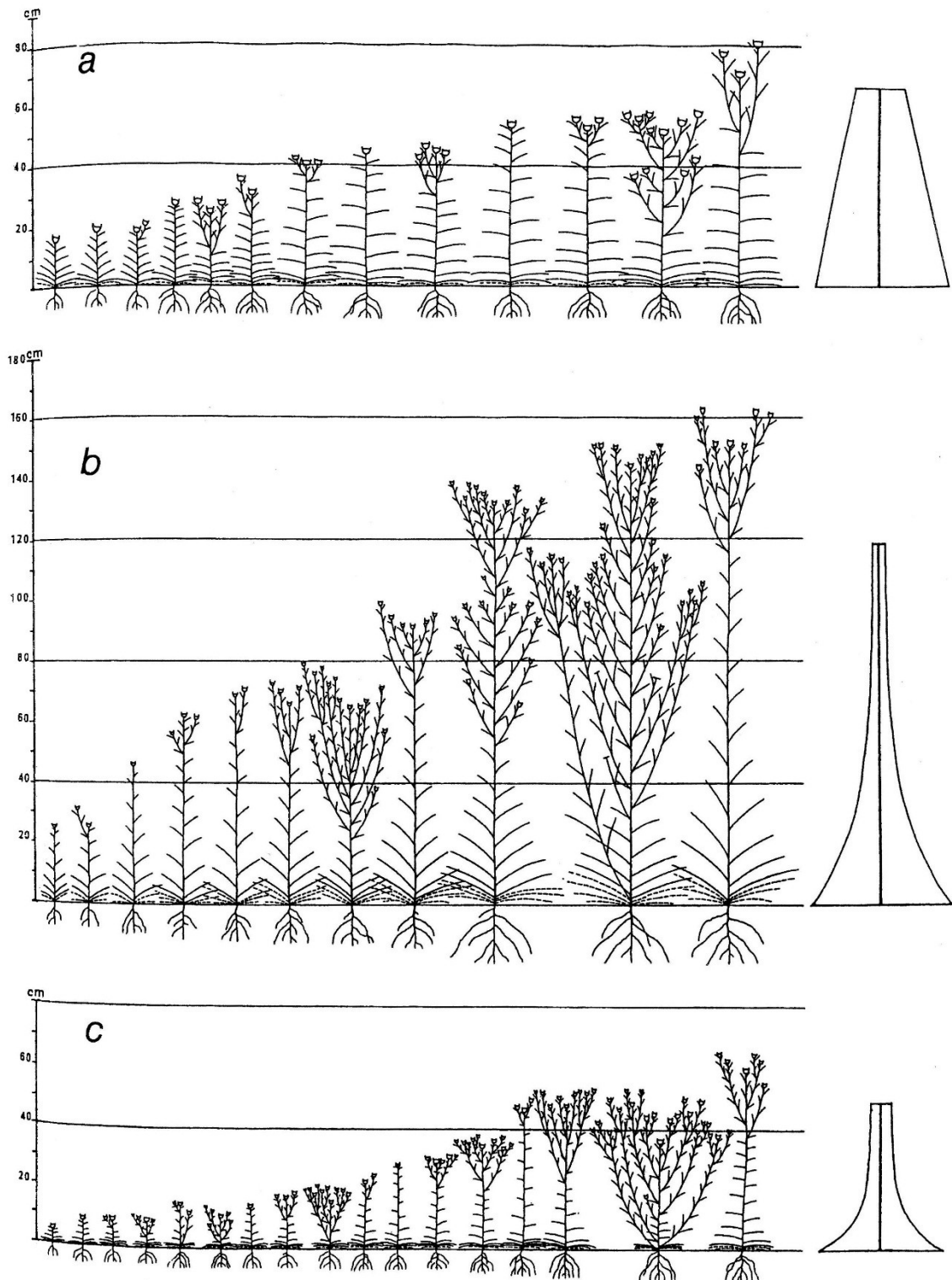
**3b** Úbory o průměru až 45 mm – Endemit Hrubého Jeseníku

..... *C. biebersteinii* subsp. *sudetica*



Obr. 3 Tvar vnějších zákrovních listenů. a) *Carlina biebersteinii* subsp. *biebersteinii*, b) *C. biebersteinii* subsp. *brevibracteata* c) *C. vulgaris* subsp. *vulgaris*. Převzato z Fischer et al. (2008).





Obr. 4 Variabilita v rámci skupiny *Carlina vulgaris* agg. ve vzrůstu a větvení. a) *Carlina biebersteinii* subsp. *biebersteinii*, b) *C. biebersteinii* subsp. *brevibracteata*, c) *C. vulgaris* subsp. *vulgaris*. Vpravo znázorněna silueta olistěné lodyhy (bez bočních větví) pro ukázkou rozdílného charakteru lodyžních listů. Převzato z práce Meusel & Kästner (1994).

## 2.3. Použité metody

### 2.3.1. Morfometrické analýzy

K zhodnocení morfologické variability pozorované v rámci populací studovaných taxonů lze použít různé přístupy. Od 50. let minulého století se za účelem studia tvarů používá tzv. tradiční morfometrika založená na kvantitativních znacích. Typicky jde o počty (např. listů, květů), délky a úhly. Výběr vhodných znaků je zde klíčový, stejně tak jako rozhodnutí ohledně počtu měřených rostlin na populaci. Vedle jiných by měly být zahrnuty také znaky, které jsou používány k odlišení jednotlivých taxonů v klíčích a florách pro pozdější ověření jejich relevance. Při sběru rostlin ke zhodnocení by měla být postihnuta co největší šíře různých stanovišť. Zároveň by měly být vyloučeny znaky, které jsou silně ovlivněny ekologickými podmínkami či takové, které jsou patrné pouze na orgánech v určité fázi vývoje; např. odění často v průběhu stárnutí mizí, není tedy dobrým znakem, pakliže si nejsme jisti, že všechny námi hodnocené listy jsou ve stejné fázi vývoje (Marhold 2011).

Zmiňované kvantitativní znaky lze zpravidla snadno naměřit ručně na živých rostlinách či herbářových položkách. Nicméně pakliže to znak, který chceme měřit, umožní, je vhodné pořídit fotografie nebo scany materiálu a případně využít nástroje na analýzu obrazu. Nápomocné jsou programy typu ImageJ (Schneider et al. 2012), které měření podstatně urychlují, vypisují hodnoty do tabulky po stisknutí klávesové zkratky, umožňují využívat různé doplňky, např. LeafJ (Maloof et al. 2013) atd. Nicméně i tak je nezbytná časově mnohdy náročná práce samostatně s každým souborem. V poslední době se v souvislosti se snahou vyvinout nástroje pro automatickou identifikaci rostlin (více viz kniha Computational Botany (Remagnino et al. 2017) či review autorů Thyagarajan & Kiruba Raji (2018)) nebo pro použití ve funkčně ekologických studiích (review od autorů MacLeod & Steart (2015)), objevují skripty umožňující např. detekci listů z herbářových položek, zcela automatické odečítání jejich délek a šířek a v neposlední řadě i analýzu tvaru okraje listů (Cope et al. 2012, Corney et al. 2012a, Corney et al. 2012b, Jin et al. 2015). Jejich aplikace za účelem získání dat pro morfometrické studie není zatím příliš rozšířená, individuální přístup stále umožňuje podchytit řadu skutečností, které používané algoritmy neodhalí. Na druhou stranu do budoucna jsou perspektivní alternativou, která umožní obohatit tradičně používané znaky o další, třeba pro odlišení taxonů významné, ale prostřednictvím ručního měření těžce postihnutelné.

Získaná data pak jsou zpravidla analyzována prostřednictvím mnohorozměrných statistických metod. V zásadě jsou nejpoužívanější tři typy, a sice shlukovací (cluster analysis), ordinační a diskriminační analýzy. Zatímco první dvě jsou exploratorní, tedy ukazují strukturu v pozorovaných datech, třetí je konfirmační, založená na předem (a priori) stanovených skupinách (např. za pomoci molekulárních analýz), a testuje příslušnost vzorku k té které skupině (Henderson 2006, Marhold 2011).

Dalším z možných přístupů je tzv. geometrická morfometrika zaměřená na analýzu tvaru jako takového (nepracuje s velikostí objektů). Využívá tzv. landmarky, což jsou homologické body identifikovatelné na všech zkoumaných objektech (Cooke & Terhune 2015). Pro

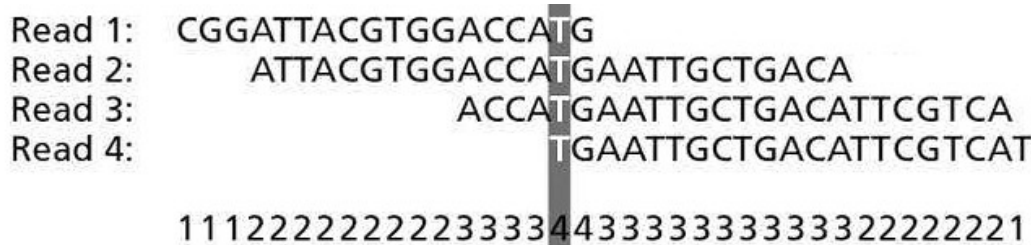
zhodnocení variability v rámci taxonů s dělenými listy, viz např. studii věnovanou rodu *Quercus* (Viscosi et al. 2009), je tento postup vhodnější než výše zmíněná tradiční morfometrika. Oproti ní ale neposkytuje jasné implikace v podobě nalezení vhodných znaků pro odlišení taxonů, proto nebyl v rámci této práce použit, ač by mohl do budoucna být vhodným doplněním ostatních analýz.

## 2.3.2. Molekulární analýzy

### 2.3.2.1. NGS metody pro rekonstrukci fylogeneze

Molekulárních metod použitelných k odlišení blízce příbuzných druhů je celá řada. Na jedné straně se lze zaměřit na srovnávání variabilních úseků genomu, jako jsou nekódující oblasti chloroplastové DNA (trnL-trnF) nebo ribozomální jaderná DNA (rDNA); na straně druhé lze využít přístupů, které redukují komplexitu variability v celkové DNA prostřednictvím tzv. arbitrárních primerů nebo restričních enzymů a poskytují tak vzhled do celkové variability napříč celým genomem (např. tradiční metody AFLP, RFLP, RAPD). Pro řešení otázek systematiky i populační a ochranářské genetiky se stále častěji uplatňují metody tzv. Next-generation sequencing (NGS), více viz review Soltis et al. (2013). V porovnání se sekvenováním tzv. první generace (Sanger et al. 1977) je nejmarkantnějším rozdílem množství takto získaných dat (Soltis et al. 2013). Snadno lze osekvenovat celé genomy, nicméně stále se jedná o analýzy spojené s vysokými náklady. Navíc pro systematické účely není znalost celého genomu nezbytná, stačí srovnávat vybrané dostatečně variabilní úseky. Cílem metod pro odvození fylogeneze na základě NGS dat je tedy redukce komplexity genomu.

Možnosti, jak k této problematice přistupovat, představuje např. review Lemmon & Lemmon (2013). Jednou z nich je tzv. mělké, povrchní sekvenování (genome skimming) (Cronn et al. 2012). Princip vychází ze skutečnosti, že NGS přístupy poskytují vedle vlastní sekvence i údaj o kvantitě, tedy v kolika získaných úsecích DNA (sequencing reads, sekvenačních readech) se báze na určité pozici vyskytla (pro ilustraci viz obr. 5). Tato tzv. hloubka (depth, coverage) markantně narůstá v případě DNA vyskytující se ve více kopiích, jako je ribozomální, chloroplastová, či mitochondriální DNA. Pro získání sekvencí takových částí genomu tedy postačí celkově menší počet readů (a tedy i menší počet sekvenačních cyklů spojený s nižšími náklady na vlastní sekvenaci).



Obr. 5 Hloubka (depth), či lépe pokrytí (coverage) vyjadřuje zastoupení nukleotidů na stejné pozici v rámci souboru získaných sekvencí (tzv. readů). Ilustrační obrázek převzat z Wikimedia.org, Genomics Education Programme, CC BY 2.0.

([https://commons.wikimedia.org/wiki/File:Read\\_read\\_length\\_and\\_read\\_depth\\_to\\_achieve\\_a\\_read\\_depth\\_of\\_4.jpg](https://commons.wikimedia.org/wiki/File:Read_read_length_and_read_depth_to_achieve_a_read_depth_of_4.jpg)).

Další možností je redukce genomové komplexity ve vzorku (eliminace nadbytečných částí genomu), prostřednictvím tzv. obohacení (enrichment). Podstatných je zde několik faktorů: specificita (rozpoznání cílených úseků), senzitivita (tedy efektivnost namnožení cíleného úseku) a uniformita (či naopak variabilita) napříč cílenými oblastmi (Cronn et al. 2012). Svou roli při výběru postupu v neposlední řadě hraje cena, množství DNA nezbytné pro reakci, či snadná aplikace dané metodiky (Mamanova et al. 2010). Nejčastěji se lze setkat s následujícími:

- a) Využití restrikčních enzymů (podobně jako u tradičních metod AFLP, RFLP zmíněných výše). Příkladem mohou být metody RRLs (reduced-representation libraries) a RAD-seq. RRLs využívá tzv. frequent-cutting enzymy a selekce dle délky takto získaných úseků, RAD-seq tzv. less-frequently cutting enzymy a polymorfismus je v tomto případě detekován na úrovni koncových úseků daných sekvencí (Soltis et al. 2013).
- b) Sekvence transkriptomu (tzv. RNA-seq). Z organismu získaná RNA je během této metody reverzně transkribována do DNA (tzv. cDNA, complementary DNA), která je následně sekvenována prostřednictvím NGS postupů. Získá se tak velké množství sekvencí exonů a nekódujících oblastí (odvozených z nekódující RNA), které mohou velmi dobře posloužit pro účely fylogenetických studií napříč různými taxonomickými úrovněmi. Na druhou stranu, je tato metoda poměrně náročná – je nezbytné použít čerstvý materiál, musí se vždy jednat o stejný typ tkáně (a jedinci, z kterých byla odebrána, musí být ve stejné fázi vývojového cyklu), navíc je třeba při sekvenování dosáhnout poměrně vysoké hloubky, „sequencing depth“ (v protipólu k mělkému sekvenování zmíněném výše se lze v dané souvislosti setkat s termínem hluboké sekvenování, tzv. deep sequencing), aby byla podchycena i RNA ze slabě exprimovaných genů. Též přiřazení (alignment) úseků transkriptomu může být problematické díky alternativnímu sestřihu (alternative splicing; vede ke vzniku izoforem odvozených z různých kombinací exonů) (Lemmon & Lemmon 2013).
- c) PCR obohacení. Jeho výhodou je vysoká specificita, reprodukovatelnost a sekvenční uniformita. Vyžaduje však konzervovaná místa k nasednutí primerů (priming sites). Této metody může být použito pouze pro amplifikaci regionů s poměrně nízkou mutační rychlostí. Též zde zůstávají chyby spojené s PCR (PCR rekombinace, PCR bias), obtížně se pracuje při sekvenování více vzorků naráz (tzv. multiplex), problémy mohou být na úrovni reakce kvůli nízké kvalitě vzorku, či inhibici kontaminacemi (Cronn et al. 2012), primery použité souběžně v rámci jednoho multiplexu mohou spolu vzájemně interagovat, což vede k vysokému zastoupení nespecifických amplifikací. Moderní postupy však umožňují proběhnutí PCR reakcí v rámci jedné zkumavky separátně (roztok je v takovém případě emulsi z oddělených kapének, např. RainStorm, RainDance Technologies). Přes nesporné zlepšení výsledků ve srovnání s běžnou PCR i zde zůstává nezbytnost vyvinout specifické primery, též cena je v takovém případě relativně vysoká (Mamanova et al. 2010)
- d) Obohacení prostřednictvím MIP (Molecular Inversion Probes). Tento postup je založen na použití jednovláknových prób, které jsou po nasednutí na cílené sekvence



vzájemně propojeny ligázou. Úseky, které nejsou takto ošetřeny, jsou v dalším kroku odstraněny (Mamanova et al. 2010)

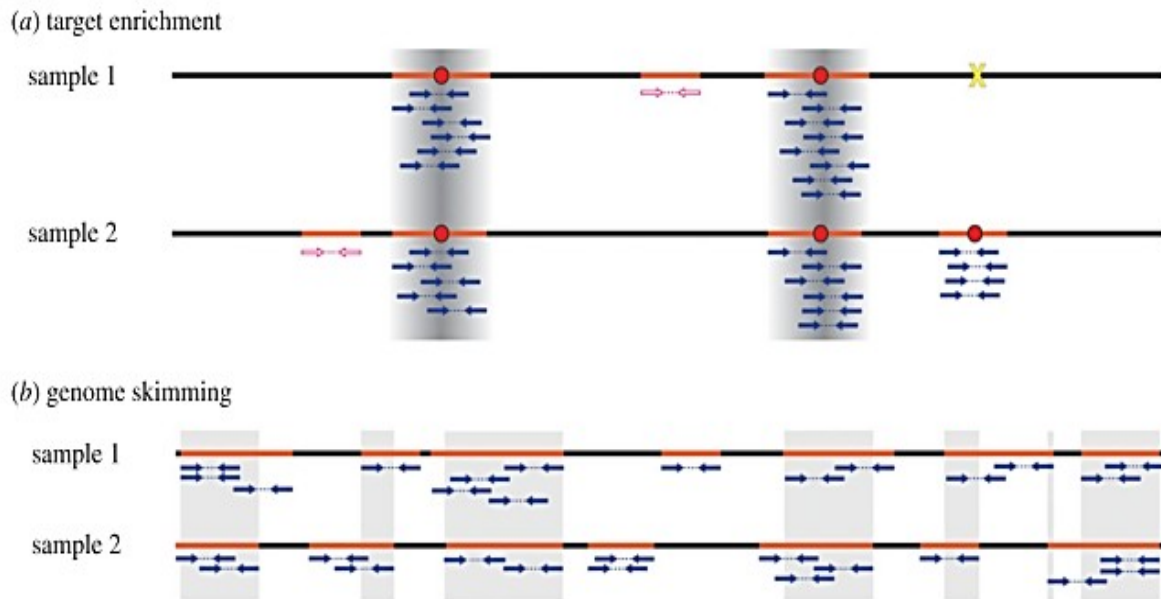
- e) Obohacení na základě hybridizace (tzv. hybridization-based enrichment). Tento přístup využívá značené oligonukleotidové sekvence o délce 60–120 bp, které jsou komplementární k požadovaným úsekům DNA. Po vzájemné hybridizaci dochází k selekci podle zmíněných značených sekvencí, takto získaný materiál je dále namnožen PCR reakcí a sekvenován (Lemmon & Lemmon 2013). V současné době se za tímto účelem používají nejčastěji biotinylem značené RNA próby (tzv. baits). Vzniklé RNA-DNA komplexy cílených úseků jsou pak zachytávány na streptavidinem potažené magnetické částice, od zbylého roztoku je lze pak snadno separovat pomocí magnetu (Gnirke et al. 2009). Cíl obohacení může být různý, k metodám patří např. AE (anchored enrichment) (Lemmon et al. 2012), resp. UCE (ultraconserved element enrichment) (Faircloth et al. 2012), či exon enrichment, používaný zejména v případě lidského genomu (Clark et al. 2013). Cílenou oblastí může být též organelová DNA (Hamilton & Buell 2012, Briggs et al. 2009). Nicméně vyvinutí těchto tzv. baits je bioinformaticky poměrně komplikované a náročné (viz např. program Sondovač (Schmickl et al. 2016), MarkerMiner (Chamala et al. 2015). Pro některé skupiny organismů však takové próby byly již navrženy (Čeled' složnokvětých (*Compositae*) mezi ně patří, proto byl v rámci této práce postup využit a je podrobněji popsán dále).

#### 2.3.2.2. Hyb-Seq

Metoda zvaná Hyb-Seq (Weitemier et al. 2014) použitá v této práci, kombinuje dva z výše zmíněných přístupů. A sice mělké sekvenování (genome skimming; Straub et al. 2012) s obohacením na základě hybridizace (targeted enrichments; Gnirke et al. 2009). Jejich srovnání je na obr. 6. Umožňuje tedy získat jak sekvence organelové či ribozomální DNA (hojně zastoupené přímo ve vyzolované DNA), tak předem vybraných single-copy jaderných exonů a přilehlých oblastí (počet jejich kopií ve vzorku je cíleně navýšen prostřednictvím obohacení na základě hybridizace).

Klíčovou otázkou zde ale je, na jaké úseky cílit, resp. jaké použít próby. Ty mohou být pro cílené obohacení navrženy na základě porovnávání vzorků z veřejně dostupných databází (ať již se jedná o projekty zaměřené na sekvenaci transkriptomu, viz onekp.com, či kompletních rostlinných genomů, <https://phytozome.jgi.doe.gov>), problematické však je nalézt úseky s optimálními vlastnostmi. Je nutné nalézt vhodný kompromis mezi univerzálními lokusy, které jsou málo variabilní na jedné straně (low-variation universal loci) a variabilnějšími, které však nemusí být přítomny u všech zkoumaných druhů, na straně druhé (Cronn et al. 2012). Pro nemodelové druhy tak vyžaduje nalezení vhodných hybridizačních prób poměrně komplikovaný postup, při kterém jsou data z mělkého sekvenování jednoho vzorku porovnávána s referenčním transkriptomem dále upravována prostřednictvím řady kroků (filtrování vzájemně příliš podobných úseků, délkově nevyhovujících sekvencí, či organelární DNA, atd.) (Schmickl et al. 2016, Weitemier et al. 2014).





Obr. 6 Schematické srovnání cíleného obohacení (a) a mělkého sekvenování (b). Cílené obohacení (a) je zaměřeno na homologní úseky kolem místa, kde nasedá hybridizační próba (červené kolečko, srovnávané úseky DNA vyznačeny šedivě). Postrádá-li sekvence odpovídající místo k nasednutí próby (žlutý křížek), není daný úsek amplifikován (a tak není zahrnut do následné analýzy). Též je osekvenována část necílených oblastí (označené růžovými šipkami); zejména se tak děje v případě DNA, která je ve vzorku výrazně zastoupena (např. organelová DNA, high-copy ribosomální DNA). Mělké sekvenování (genome skimming, b) slouží k získání náhodných fragmentů. Po jejich poskládání do kontigů (assembly) lze dohledat homologní sekvence pro dané vzorky (opět vyznačeny šedě). Získané úseky DNA (sequencing reads) v obou případech vyznačeny modrými šipkami. Převzato (Hollingsworth et al. 2016).

Další možností je použití konzervativních úseků genomu. Podaří-li se najít takové, které sousedí s mnohem variabilnějšími regiony, lze právě z nich odvodit hybridizační próby použitelné i u vzdáleně příbuzných druhů (Faircloth et al. 2012). Takové próby založené na tzv. UCEs (ultraconserved elements) byly úspěšně aplikovány v případě řešení fylogeneze řady skupin, ať již se jednalo o ptáky (McCormack et al. 2016), blanokřídlý hmyz (Faircloth et al. 2015) či pavoukovce (Starrett et al. 2017). V případě rostlin se za stejným účelem používá tzv. COS, conserved orthologous set (použití UCEs není vhodné vzhledem k možnému horizontálnímu přenosu DNA z organel i častým polyploidizačním událostem (Weitemier et al. 2014)). Soubor těchto sekvencí odvozený z ESTs<sup>1</sup> různých zástupců čeledi *Compositae*, posloužil již řadě fylogenetických studií celé čeledi (Chapman et al. 2007, Mandel et al. 2017, Mandel et al. 2015). Je komerčně dostupný v rámci produktů firmy Arbor Biosciences (stejně jako výše zmíněné UCEs), pod obchodní značkou MyBaits®<sup>2</sup>. Proto byl také v této práci za účelem obohacení využit.

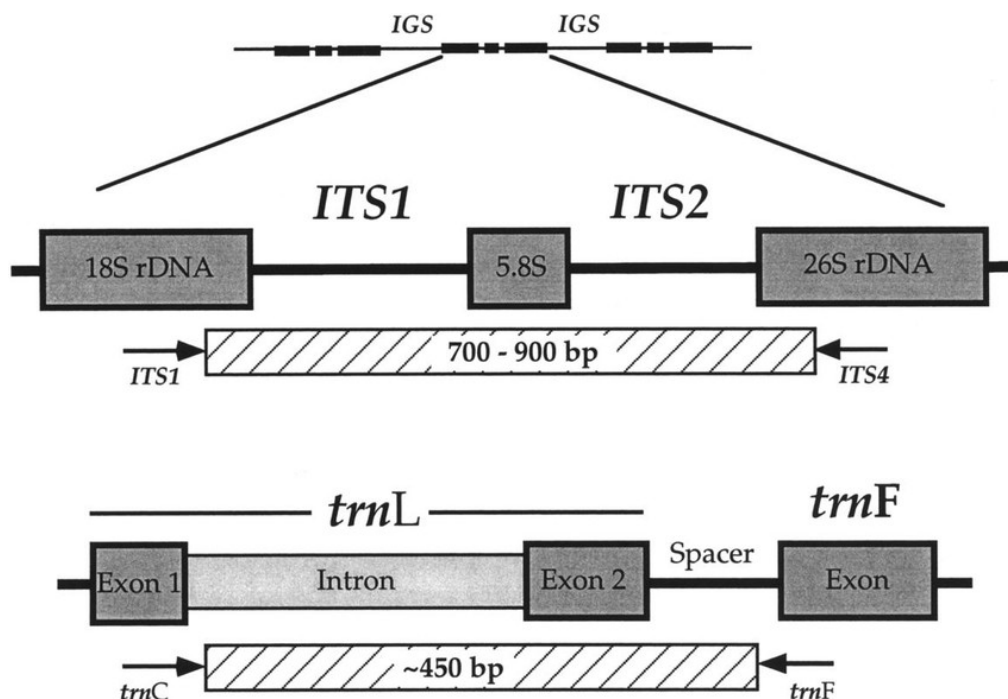
Jak již bylo zmíněno výše, prostřednictvím metody Hyb-Seq lze získat též informativní data z necílených, ale hojně zastoupených, úseků DNA. Prostřednictvím mapování na vhodnou referenci tak lze odvodit sekvence chloroplastové DNA či cistronů jaderné ribosomální DNA (rDNA) (Weitemier et al. 2014). V případě obou se jedná o molekulární markery, které

<sup>1</sup> ESTs (z anglického Expressed sequence tags) jsou krátké úseky komplementární DNA (vzniká reverzní transkripce buněčné mRNA). Od 90. let se používají k vymapování genů, jejich soubory jsou dostupné pro řadu organismů (Rudd 2003).

<sup>2</sup> viz <http://www.arborbiosci.com/products/uces/>

mohou při srovnání s cílenými úseky jaderné DNA přinést další informace o evoluční historii.

Chloroplastová DNA je děděna u krytosemenných rostlin po maternální linii a nedochází v jejím případě k rekombinaci. Lze ji proto v kombinaci s daty z jaderné DNA, která je děděna biparetálně, využít např. k detekci rodičovských taxonů v případě hybridizace. Časté uplatnění má i ve fylogeografických studiích, nejčastěji je za tím účelem používán úsek *trnL-trnF* (jeho struktura viz obr. 7) (Olmstead & Palmer 1994). Též oblast cistronu jaderné ribozomální DNA má časté uplatnění ve fylogenetických studiích (Álvarez & Wendel 2003). Sestává ze sekvencí, které se v genomu v rámci tandemových repetitivních sekvencí několikrát opakují. Tvoří je konzervativní oblasti kódující malou a velkou podjednotku ribozomu a variabilnější oblast sestávající ze dvou ITS sekvencí (intergenic spacers) mezi nimi. Vzhledem k zmíněné konzervativnosti sekvencí 18S a 26S napříč organismy jsou primery pro amplifikaci ITS snadno dostupné, navíc, jak již bylo řečeno výše, je podíl rDNA ve vyzolovaných vzorcích DNA výrazný. Proto se často používají k odvození fylogeneze v případě nemodelových organismů (Álvarez & Wendel 2003, Shaw et al. 2002). Při studii fylogeneze rodu *Carlina* (Wahrmund et al. 2010) však ITS jako marker nemohl být použit, neboť nebyl dostatečně variabilní, autoři této práce využívali ETS (external transcribed spacer, na obr. 7 není znázorněn, je součástí IGS (intergenic spacers), tedy variabilní netranskribované oblasti mezi opakujícími se 18S-5.8S-26S sekvencemi).



Obr. 7 úseky DNA nejčastěji používané ve fylogenetických a fylogeografických studiích. Nahoře schéma úseku 18S-5.8S-26S cistronu rDNA (popis v textu), dole úsek *trnL-trnF* chloroplastové DNA. Převzato (Shaw et al. 2002).

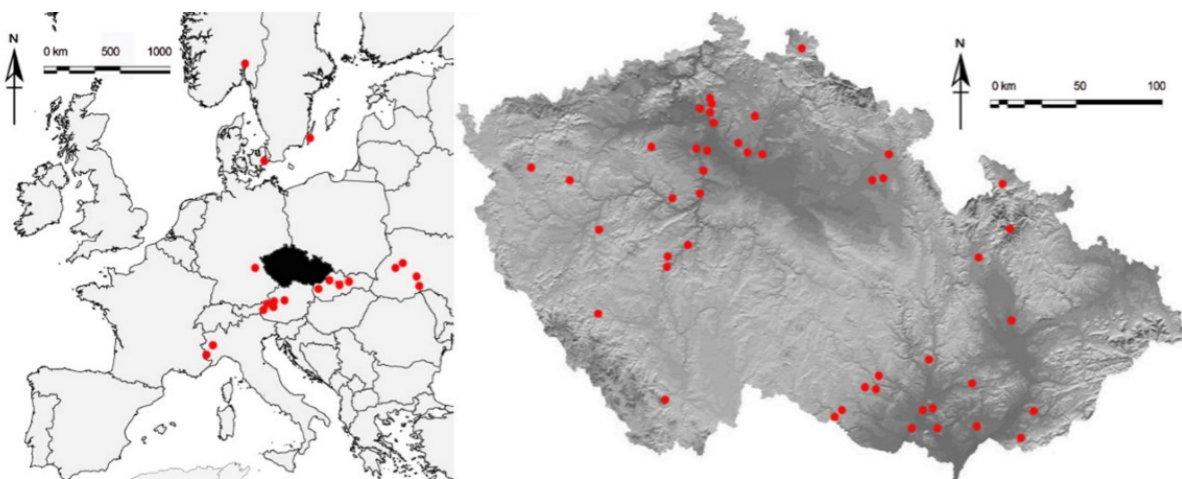
Po laboratorní stránce analýza zahrnuje vlastní izolaci vzorků DNA, přípravu knihovny pro vybranou sekvenční platformu (v rámci této práce byla použita Illumina) a hybridizaci získaných vzorků s výše zmíněnými baits. Výstupy z následné sekvenace jsou pak

zpracovávají prostřednictvím bioinformatických nástrojů. Je nezbytné a) efektivně uspořádat získané sekvence do samostatných lokusů, b) následně lokusy „zalignovat“ pro úspěšné odvození fylogeneze a c) optimálně rozšířit informativní oblast kódujících sekvencí (na které bylo cíleno) i o mnohem variabilnější, vzdálenější introny (Johnson et al. 2016). Vlastní analýza sestává z několika kroků. Za účelem jejího snadného průběhu jsou sestavovány tzv. pipeline, soubory skriptů, které umožňují propojení řady různých programů a jejich výstupů mezi sebou do jednoho rámce. Pro analýzu obohacených dat existují např. PHYLUCE (Faircloth 2016) nebo alignreads.py (Straub et al. 2011), avšak tyto programy nejsou zcela vhodné pro Hyb-Seq data. Přímo za účelem jejich analýzy byly vytvořeny např. soubory skriptů HybPiper (Johnson et al. 2016) a HybPhyloMaker (Fér & Schmickl 2018).

### 3. Materiál a metodika

#### 3.1. Vzorky rostlin

Pro fylogenetickou analýzu bylo získáno 77 vzorků ze 73 populací (přehled viz tab. I v příloze). V naprosté většině případů jeden vzorek reprezentoval jednu populaci. Do analýz byly dále zahrnuty 3 vzorky původem z jesenické populace endemického poddruhu *C. biebersteinii* subsp. *biebersteinii* z různých let (zde označeny jako vzorky 1, 63, 64), též 3 vzorky původem z lokality Březí, v Dunajovických kopcích, kde se jednalo a) o rostliny pěstované po několik generací v Botanické zahradě PřF UK (populace 10), b) rostliny ze semen zde odebraných a dopěstovaných v experimentální zahradě BÚ AV ČR v Průhoncích (populace 42) a c) rostliny z původní populace rostoucí na dané lokalitě (populace 30). 45 vzorků pochází z ČR, zatímco ostatní převážně z Alp a Karpat. Pro srovnání bylo zahrnuto několik vzorků ze Slovenska, Ukrajiny a Skandinávie.



Obr. 8 Lokalizace sběrů vzorků v Evropě a České republice. Vyneseno na podkladovou mapu Evropy (ze stránek <https://d-maps.com/>) a mapu ČR (z programu DIVA-GIS (Hijmans et al. 2001)).

Lokality v rámci celé České republiky byly vybrány podle dostupných mapových podkladů na florabase.cz (Daníhelka et al. 2009) tak, aby byla podchycena co největší variabilita stanovišť. Na každé lokalitě byl odebrán vzorek listů z jedné rostliny (optimálně v horní části

lodyhy) a byl uložen do sáčku se silikagelem pro následnou izolaci DNA. Celkem z 32 lokalit z ČR, Slovenska a Rakouska byly dále sebrány rostliny pro morfometrické analýzy. Optimálně bylo 10–15 rostlin změřeno přímo na místě a dalších 10–15 sterilních růžic bylo převezeno do experimentální zahrady za účelem alespoň částečného odfiltrování vlivu fenotypové plasticity. V průběhu následujícího roku byly dopěstovány do květu a pak též měřeny. Na ostatních lokalitách byly populace příliš chudé nebo rostliny po přesazení do zahrady nepřežimovaly, a proto nemohly být pro morfometrické vyhodnocení použity. Herbářové doklady jsou uloženy na Botanickém ústavu AV ČR v Průhonicích.

Rostliny z Malé kotliny v Hrubém Jeseníku, která je na taxonomické úrovni hodnocena jako samostatný endemický poddruh *C. biebersteinii* subsp. *sudetica*, byly získány díky laskavé spolupráci s CHKO Jeseníky a dopěstovány též v experimentální zahradě BÚ AV ČR.

### 3.2. Morfometrická analýza

Na rostlinách byly měřeny znaky na lodyze, listech i úbořech (soupis všech měřených znaků viz v tab. III v příloze). Šířka lodyhy ve výšce cca 3 cm nad bází (měřena pomocí posuvného plastového měřítka), výška rostliny (po terminální úbor i celková) a počet listů byly zaznamenány přímo v terénu. Dále byly pořízeny fotografie lodyhy pro pozdější změření úhlu, pod kterým od ní odstávají listy (obr. 9a), a fotografie příčného řezu listem (proveden žiletkou) ke stanovení úhlu příčného prohnutí listu zhruba v polovině jeho délky (fotografie byla pořízena kolmo na rovinu řezu, obr. 9b). Vlastní hodnoty úhlů byly odečítány v programu ImageJ (Schneider et al. 2012).



Obr. 9 Příklady fotografií, z kterých byly v programu ImageJ (Schneider et al. 2012) odečítány úhly. a) úhel mezi listem a lodyhou (uhel), b) úhel příčného prohnutí listu na příčném řezu listem (proh1-3).

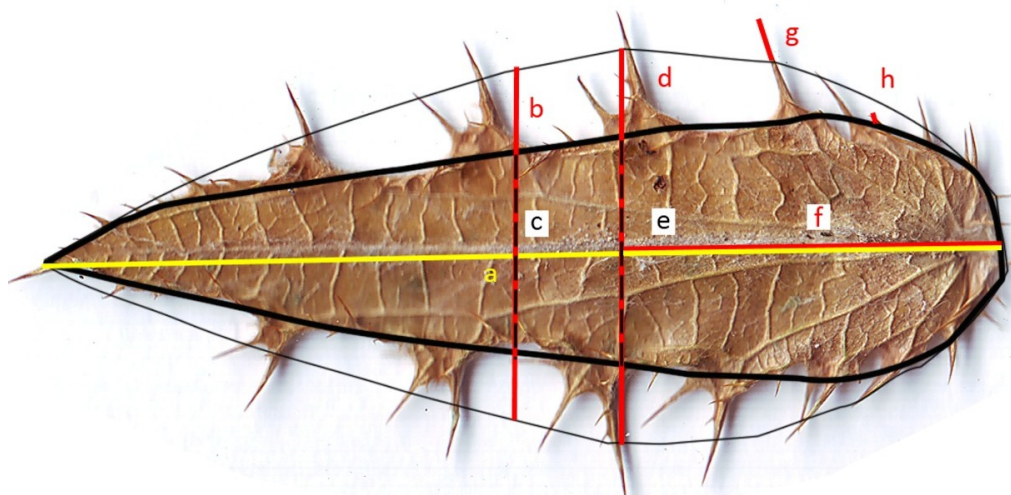
Z každé rostliny bylo odebráno 9 listů pro následující morfometrické analýzy; vždy se jednalo o 3 listy na bázi lodyhy, 3 listy ze střední části lodyhy a 3 listy pod terminálním úbořem. Dále byly z každé rostliny sebrány i 3 horním lodyžním listům nejbližší listy vnějšího zákrovu úboru. Po usušení v novinovém papíře pod přiměřeným zatížením byly



listy i listeny naskenovány do formátu .tif při rozlišení 600 dpi (příklad naskenovaných listů viz na obr. 10). Měřené znaky jsou znázorněny na obr. 11, jejich hodnoty byly opět odečítány pomocí programu ImageJ.

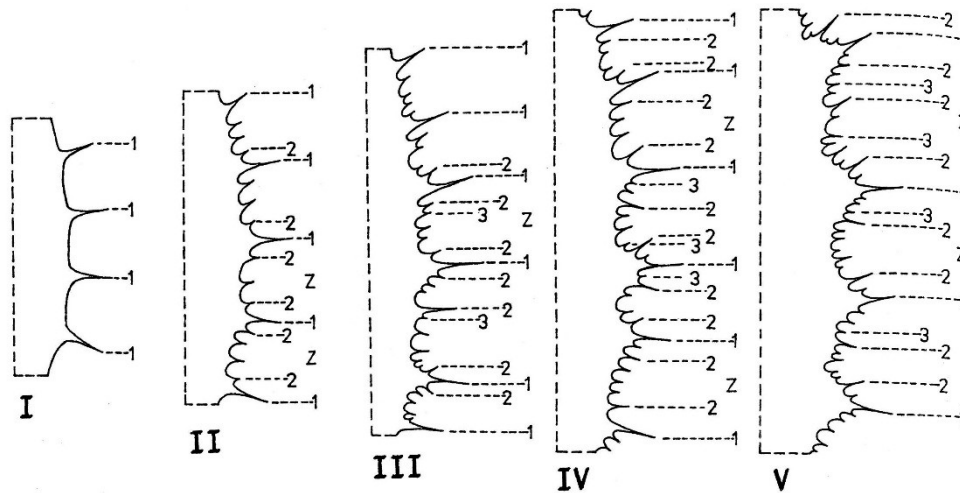


Obr. 10 Příklad naskenovaných listů z jedné rostliny, které sloužily k odečtení hodnot měřených znaků (na následujícím obrázku).



Obr. 11 Znaky měřené na listech v programu ImageJ. a) délka listu, b) šířka listu v polovině jeho délky, c) šířka listu v zárezu mezi úkrojky v polovině délky listu, d) šířka listu v nejširší části, e) šířka listu v zárezech mezi úkrojky v nejširší části listu, f) vzdálenost mezi nejširší částí listu a bází listu, g) délka ostnu prvního řádu (viz obr. 12), h) délka ostnů nejnižšího řádu (zpravidla mezi úkrojky listu nebo na okraji nečleněných listů, odpovídá označení Z na obr. 12).

Z diskrétních znaků byl zjišťován počet ostének a ostnitých zubů. Od původního záměru počítat též ostnité úkrojky listu (laloky) bylo nakonec upuštěno z důvodu jejich obtížné definice pro potřeby této práce vyplývající z plynulého přechodu mezi zuby a úkrojky listu (viz obr. 12). Namísto toho byla „laločnatost“ listu vyjádřena prostřednictvím „indexu hloubky zářezu“. Ten byl definován jako poměr šířky listu (v nejširší části) a rozdílu šířek v nejširší části a v zářezech mezi úkrojky (na obrázku výše odpovídá  $d/(d-e)$ ).



Obr. 12 Vrstující členitost okraje listové čepele u druhu *Carlina vulgaris*. I jednotlivé zuby, II Sdružené zuby (1) a zoubky (2), III-V peřenolaločný okraj s úkrojky a zuby a zoubky několika řádů (1-3). Z – samostatné ostny mezi úkrojky. Převzato z práce Meusel & Kästner (1990).

Vzhledem k možným chybám daným subjektivním hodnocením znaků počet ostnitých zubů na listu z dolní/střední/horní části lodyhy a počet ostnitých zubů na vnějších zákrovnicích listenech, a též k časové náročnosti spojené s vizuálním odečtem těchto hodnot, bylo přistoupeno k použití programu ToothFinder (Corney et al. 2012b; ke stažení na stránkách <https://github.com/dcorney/ToothFinder>). Vstupní soubory ve formátu .mat byly získány z obrázků ve formátu .tif pomocí skriptu segmentation\_one.py (Holešovský, nepublikováno) s drobnými úpravami (popsané v příloze). Též skript ToothFinder.m, který byl používán v prostředí GNU Octave (Eaton et al. 2016), bylo pro potřeby této práce nutné upravit. Nejpodstatnější změnou, ke které bylo přistoupeno, je skutečnost, že pro vlastní analýzu okrajů byly použity přímo souřadnice ze vstupního souboru (původně je skript ToothFinder.m využíván pouze pro zmíněnou detekci oblasti zájmu, pro vlastní počítání zubů je získává de novo pomocí funkce Contour\_following.m (<https://www.mathworks.com/matlabcentral/fileexchange/14947-contour-following>), která je sice možná přesnější, avšak selhávala v případě dlouhých úzkých listů).

Pro ověření, zdali je možné program ToothFinder za účelem zjištění počtu ostnitých zubů na listech v rámci této práce skutečně využít, a jsou-li jeho výsledky srovnatelné s vizuálním odečtem, byl použit vzorek 83 rostlin (u každé bylo provedeno měření na 3 dolních, 3 středních a 3 horních lodyžních listech a 3 listenech vnějšího zákrovu). Počet ostnitých zubů byl odečten z naskenovaných obrázků vizuálně v programu ImageJ. Stejně obrázky listů byly

pak analyzovány prostřednictvím programu ToothFinder. Dále bylo pracováno s průměrnou hodnotou pro 3 listy, případně listeny.

Z porovnání datových souborů bylo patrné, že počty zubů získané oběma přístupy spolu silně korelují (více viz Výsledky) a lze tedy program ToothFinder v rámci práce za účelem zjištění počtu ostrnitých zubů používat. Počty ostrnitých zubů na všech měřených listech tedy byly získány tímto algoritmem (s jehož použitím jednak odpadl vliv chyb daných subjektivním hodnocením znaků, jednak se tak proces podstatně urychlil).

Statistické zpracování morfometrických dat proběhlo v programu R (R Development Core Team 2018). Blíže je popsáno v kapitole 3.5.2. Pro přehled použitých funkcí viz tab. V příloze.

### 3.3. Izolace DNA

Listy pro izolaci DNA byly odebírány z horní části lodyhy. Vysušeny a skladovány byly v silikagelu. S ohledem na množství sekundárních metabolitů obsažených ve vzorcích byla zvolena izolace DNA za použití sorbitolu. S drobnými úpravami byl využit pracovní postup podle Štorchové et al. (2000).

Materiál byl rozdrčen ve 2 ml zkumavkách za pomoci wolframkarbidových kuliček v třepače Retsch Mixer Mill 400 (do každé zkumavky byly vloženy 2 kuličky, drcení probíhalo 3 min s frekvencí 30 třepů za minutu). Po vizuální kontrole materiálu byl v případě nedostatečného rozmělnění čas drcení prodloužen na 5 min. Dále byla přidána špetka polyvinylpyrrolidoinu pro odstranění sekundárních metabolitů, 1300  $\mu$ l extrakčního pufru a 4  $\mu$ l RNAsy A (10 mg/ml, Thermo). Po 20 minutách inkubace při laboratorní teplotě následovala centrifugace (5 min, 7 000 rpm, centrifuga Eppendorf 5415R). Dalším krokem bylo odstranění supernatantu a přidání 300  $\mu$ l extrakčního pufru. Po promíchání v ruce byl přidán lysis pufr (též 300  $\mu$ l) a vzorky inkubovány 15 minut na termobloku (teplota 65 °C, 300 rpm).

Další postup spočívá v přečištění směsí chloroform:isoamylalkohol v poměru 24:1. Ke každému vzorku bylo přidáno 600  $\mu$ l zmiňované směsi a zkumavky byly dále minimálně 3 minuty promíchávány překlápěním (v ruce). Následovala centrifugace (10 min, 9 000 rpm) a odebrání horní vodné fáze. Tato byla umístěna do nových 2 ml zkumavek a přečištěna ještě jednou. Po opětovném odebrání vodné fáze (ideálně 600  $\mu$ l), byla tato přepipetována do zkumavek o objemu 1,5 ml a byly přidány 2/3 objemu (tj. 400  $\mu$ l, v případě, že byl výtěžek z nějakého důvodu nižší, bylo množství upraveno na odpovídající poměr) vymraženého isopropanolu. Po promíchání v ruce byly ponechány vzorky půl hodiny v mrazáku při teplotě -20 °C.

Tekutá složka byla opět oddělena prostřednictvím centrifugace (15 min, 13 000 rpm, centrifuga předem chlazená na teplotu 4 °C) a odlita. K peletu na dně zkumavky bylo přidáno 700  $\mu$ l 80 % etanolu o laboratorní teplotě. Zkumavky byly pak několikrát promíchány překlápěním v ruce. Po následné centrifugaci (1 min, 13 000 rpm) byla opět slita tekutá složka. Vzniklé peletky byly dosušeny na termobloku (10 min, 60 °C); po vizuální kontrole byl případně čas prodloužen na 15 min. Posledním krokem extrakce DNA je



rozpuštění vzniklé peletky v 50  $\mu$ l TE pufru, opět na termobloku (10 min, 60 °C, 400 rpm); v případě, že se peletka hůře rozpouštěla, byl čas též prodloužen o několik minut, v případě některých vzorků bylo nutné přidat více TE pufru.

Z důvodu omezeného rozsahu měření koncentrací DNA na přístroji Qubit (do 100 ng/ $\mu$ l), byla tato předem orientačně zjišťována prostřednictvím přístroje Nanodrop 1000 (Thermo). Přesahovala-li hodnota 90 ng/ $\mu$ l, byly vzorky ředěny (10 krát) a až pak byla stanovena jejich přesná koncentrace (dle návodu dodávaného k přístroji Qubit 2.0 Fluorometer (Thermo), byl použit 1  $\mu$ l vzorku DNA na 199  $\mu$ l reakčního roztoku (Qubit buffer a reagent, podrobnosti viz na stránkách výrobce

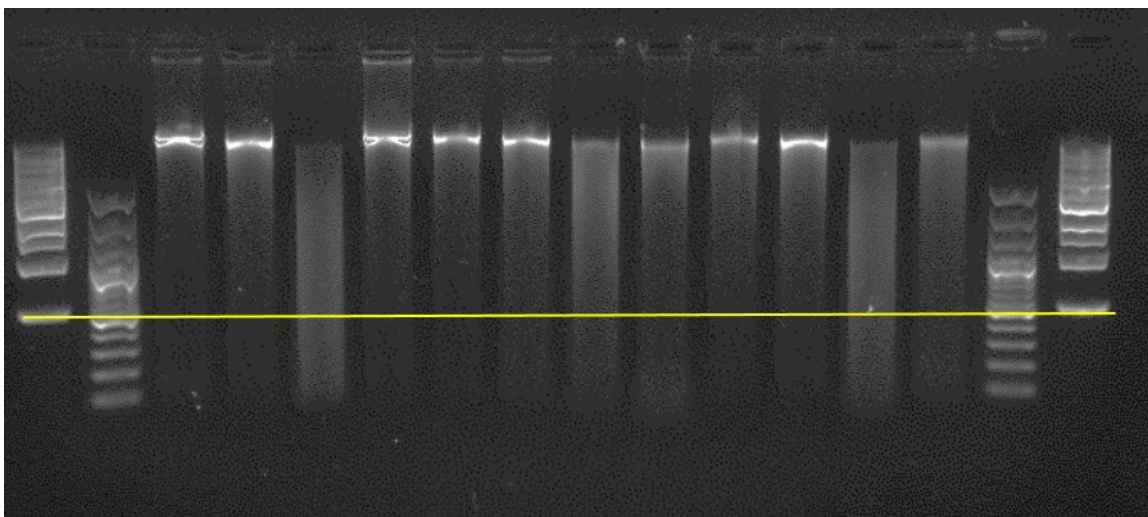
<https://www.thermofisher.com/content/dam/LifeTech/migration/en/filelibrary/cell-tissue-analysis/qubit-all-file-types.par.0519.file.dat/qubit-2-fluorometer-user-manual.pdf>).

### 3.4. Příprava knihovny pro Hyb-Seq

#### 3.4.1. Sonikace

K fragmentaci DNA byl použit ultrasonikátor Covaris M220. Tato část laboratorní práce probíhala v Laboratoři sekvenace DNA Přírodovědecké fakulty v centru BioCev (<http://www.biocev.eu/corefacilit/omics-genomika/>). Pro vlastní sonikaci je potřeba 500 až 1000 ng DNA. Proto byla DNA ředěna (dle předem naměřených koncentrací) na výslednou hodnotu 800 ng (v 62  $\mu$ l). Z toho 2  $\mu$ l roztoku byly využity pro kontrolu délky získaných fragmentů na 0,8 % agarózovém gelu s barvou Gel Red. Na gel byly naneseny 2  $\mu$ l DNA spolu s 1  $\mu$ l barviva 6x Orange LD; do krajních jamek 0,7  $\mu$ l žebříku 1 kb a 0,4  $\mu$ l žebříku 100 bp.

Dle fotografie gelu (její příklad viz na obr. 13) bylo rozhodnuto, které vzorky je třeba sonikovat a které nikoliv (tj. již obsahují dostatečné množství požadovaných fragmentů o délce 500 až 600 bp díky degradaci DNA v použitém materiálu).



Obr. 13 0.8% TAE agarózový gel s barvou Gel Red a nanesenými vzorky před vlastní sonikací. 3. a 11. vzorek zprava není třeba sonikovat, obsahuje již degradovanou DNA o požadované délce úseků 500 až 600 bp (hranice cca 500 bp je vyznačena žlutou čarou).



Pro vlastní sonikaci pak byly vzorky přeneseny do speciálních zkumavek microTUBE AFA Fiber Snap-Cap 6x16mm (o objemu 130 µl). Dále bylo postupováno dle návodu výrobce přístroje Covaris M220 (více viz [http://covarisinc.com/wp-content/uploads/pn\\_010157.pdf](http://covarisinc.com/wp-content/uploads/pn_010157.pdf)). Na základě doporučení se jeví jako optimální pro získání fragmentů dané délky 500 až 600 bp následující nastavení: 50W, 2% duty factor, 200 cyklů, t 6 °C po dobu 95 sekund.

Po vlastní sonikaci byly vzorky přeneseny zpět do původní zkumavky. Délka fragmentů DNA byla opět zkontrolována na agarózovém gelu (obdobně jako výše). Pakliže většina vzorku odpovídala delším úsekům, než byly požadovány, byla sonikace zopakována.

#### 3.4.2. End repair and A-tailing (opravy délky fragmentů a navázání A)

K vlastní přípravě knihovny pro sekvenaci na platformě Illumina byl použit komerční kit NEBNext Ultra DNA Library Prep Kit for Illumina (E7370). Bylo postupováno dle návodu výrobce (viz <https://www.neb.com/~media/Catalog/All-Products/23EF8BC783C749D9B664D90C8C72FA73/Datacards%20or%20Manuals/manualE7370.pdf>)

Prvním krokem je úprava konců DNA fragmentů, aby byla obě vlákna stejně dlouhá, a následné navázání adeninů. Výchozí roztok DNA o objemu 55,5 µl se zde míchá se směsí obsahující tzv. End Prep enzyme Mix a End Repair Reaction Buffer. Kapalina je vazká, proto je nezbytné promíchání pomocí pipety. Následuje inkubace v cycleru (30 min, 20 °C a poté 30 min, 65 °C).

#### 3.4.3. Ligace adaptorů

Následuje ligace adaptorů, v našem případě byly použity adaptory pro multiplex (NEBNext Adaptor for Illumina kit E7335S). Směs adaptorů, ligázy (Blunt/TA Ligase Master Mix) a roztoku označovaného jako Ligation Enhancer je přidána ke vzorkům DNA, opět musí dojít k důkladnému promíchání pomocí špičky. Dále se zkumavky inkubují v cycleru při 20 °C 15 min. NEBNext adaptory vázané k DNA tvoří smyčku, předchází se tak vzniků dimerů mezi adaptory a zvyšuje se tak účinnost ligace. Ve středu sekvence adaptoru je určil, v následujícím kroku (přidání tzv. USER enzymu, opět z kitu E7335S) dojde k jeho vyštěpení, uvolní se tak konce adaptorů pro následné PCR obohacení.

#### 3.4.4. Přečištění produktů

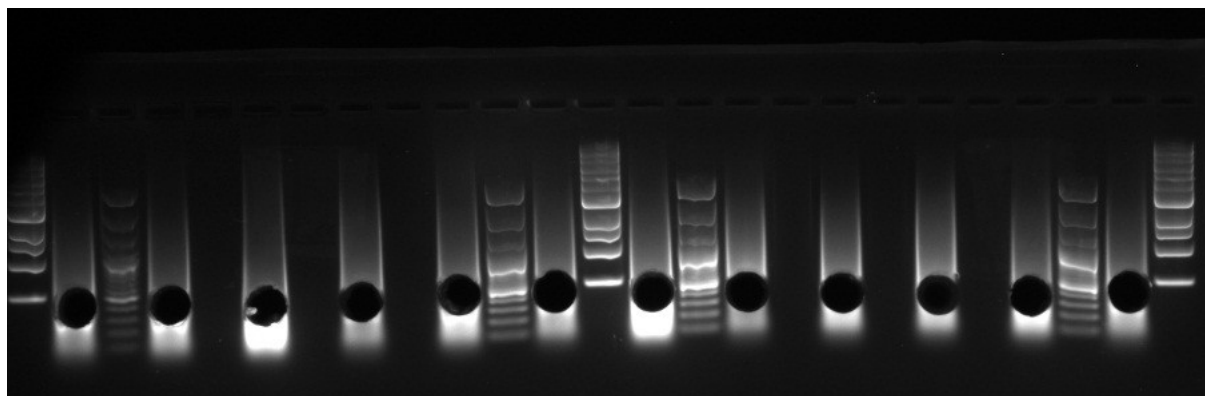
Objem jednoho vzorku (po provedení kroků uvedených výše) byl cca 83,5 µl. Pro další krok, kterým je délková selekce prostřednictvím agarózového gelu, je nezbytné celý vzorek nanést na gel a jeho objem proto musí být zredukován. Toho bylo v našem případě docíleno pomocí kitu QIAquick PCR Purification Kit (Qiagen, postup dle návodu výrobce, <https://www.qiagen.com/nl/resources/resourcedetail?id=3987caa6-ef28-4abd-927e-d5759d986658&lang=en>). Eluce produktu byla provedena do 30 µl ddH<sub>2</sub>O.

#### 3.4.5. Délková selekce DNA

K selekci námi požadovaných úseků o délce 500 až 600 bp byla použita gelová elektroforéza na 1% agarózovém gelu (200 ml low EDTA/TAE pufru, 2 g agarózy, 2 µl barvy Gel Red). K

vizualizaci délky fragmentů byla ke vzorkům přidána barva 6× Orange DNA loading dye (7 μl na 30 μl roztoku DNA z předchozího kroku). Vzorky na gel byly nanášeny ob jamku, aby se snížila pravděpodobnost vzájemné kontaminace. Do krajních a prostředních jamek byly nanášeny žebříky (3 μl 1 kb Fast Ladder a 2 μl 100 bp). Gelová elektroforéza probíhala při 150 V cca 1,5 h (doba ovlivňuje „roztažení“ úseků DNA napříč gelem, čím je kratší, tím blíže jsou úseky podobné délky u sebe).

Pod UV světlem byla pomocí alobalu s rovným okrajem přiloženého k pruhům žebříků lokalizována oblast s úseky o požadované délce. Ty byly následně z gelu vyřezány prostřednictvím obrácené špičky (1 ml, Eppendorf), příklad gelu po vyřezání terčů viz na obr. 14. Takto získané terčíky byly přeneseny do předem zvážených 2 ml zkumavek (Eppendorf). Jako záloha byly použity i produkty těsně pod vyřezávanou oblastí (cca 300 až 500 bp), opět umístěné do 2 ml předem zvážených zkumavek; tyto byly dále skladovány v mrazáku pro použití v případě problémů s první sadou vzorků.



Obr. 14 1% agarózový gel (200 ml low EDTA/TAE pufru, 2 g agarózy, 2 μl barvy Gel Red) se vzorky po sonikaci a vyřezání terčů s DNA o patřičné délce.

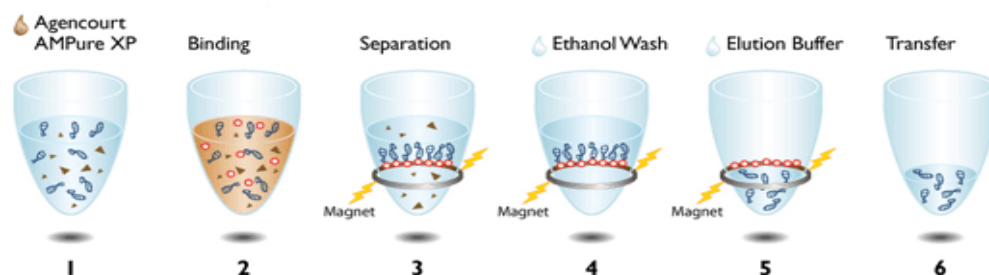
Zkumavky s gelem byly opět zváženy a dopočten rozdíl, neb je pro následující přečištění nezbytné znát hmotnost gelu. Přečištění proběhlo za použití kitu Qiagen QIAquick Extraction Kit. Postup byl proveden dle protokolu výrobce, pouze rozpouštění gelu v pufru probíhalo v termomixeru (450 rpm, 25 °C) za občasného promíchání vzorků překlápěním v ruce. Eluce vzorků DNA byla provedena do 18 μl ddH<sub>2</sub>O.

#### 3.4.6. PCR obohacení (enrichment)

Pro PCR obohacení byl zvolen Master mix – Q5 Hot Start HiFi PCR, postup byl proveden dle protokolu výrobce. Ke každému vzorku byl tedy přidán jiný indexační primer (byly používány 2 sady indexačních primerů pro sekvenování více vzorků zároveň (multiplex) a sice NEBNext Multiplex Oligos for Illumina (Index Primers Set 1) E7335 a (Index Primers Set 2) E7500). Dále mastermix získaný smícháním NEBNext Q5 HotStart HiFi PCR Master Mix-u a univerzálního PCR primeru. V rámci PCR cyklu proběhlo 8 cyklů 98 °C 10 sec, 65 °C 1 min 15 sec (jejich počet závisí na počáteční koncentraci DNA, je-li nižší než 500 ng, doporučuje se počet cyklů zvýšit, dochází pak ale k vyšší chybovosti).

### 3.4.7. Přečištění PCR produktů

K přečištění PCR produktů byl využit systém Agencourt AMPure XP (Beckman), v rámci kterého je DNA patřičné délky vyvázána na magnetické kuličky (viz obr. 15 krok 2). Ty jsou následně zachyceny v jamkách kapkovací destičky díky magnetickému stojánku (krok 3). Po odsátí původního roztoku a následném přečištění 70% etanolem (krok 4), jsou vzorky umístěny mimo stojánek, kde dojde k eluci do 40  $\mu$ l ddH<sub>2</sub>O. Opětovným postavením na magnet jsou z roztoku vyvázány čisté kuličky (krok 5) a roztok přečištěné DNA může být přemístěn do zkumavky (krok 6).



Obr. 15 Schéma průběhu přečištění PCR produktů prostřednictvím Agencourt AMPure XP (Beckman). Popis v textu. Převzato ze stránek výrobce <http://www.beckman.com/Media/LifeSciences/Nucleic-Acid-Sample-Preparation/DNA-Extraction-and-Clean-Up/PCR-Purification/agencourt-process-overview.gif>

Délku fragmentů, které budou z roztoku DNA odstraněny se v případě použití Agencourt AMPure XP lze ovlivnit prostřednictvím poměru v jakém bude roztok kuliček přidáván k DNA; čím je tento poměr větší, tím kratší úseky budou zachovány. Přečištění proto bylo provedeno za použití poměru 0,7 : 1 (pro odstranění úseků 100-200 bp). Pro získání produktů o vyšší čistotě byl postup zopakován dvakrát (první eluce byla do 40  $\mu$ l ddH<sub>2</sub>O, druhá do 20  $\mu$ l).

Kvalita přečištění byla dále ověřena na 1% TAE agarózovém gelu (0,6 g agarózy, 60 ml pufru) s barvou Gel Red (1  $\mu$ l). Nanášeny byly 2  $\mu$ l DNA spolu s 1  $\mu$ l barvy 6 $\times$  Orange LD a žebříky 100bp (0,4  $\mu$ l) a 1 kb (0,7  $\mu$ l). Pakliže na gelu byly patrné dva proužky (spodní odpovídá adaptorům z PCR), bylo nezbytné přečištění ještě jednou zopakovat.

### 3.4.8. „Poolování“ knihovny

Všechny vzorky jsou v následujícím kroku přeneseny do jedné tzv. Low-bind zkumavky (tato má speciálně upravený povrch, aby na něm neulpívala DNA). Tento proces je označován jako pooling („poolování“). Vzhledem k tomu, že je nezbytné, aby v celkovém objemu bylo od každého vzorku stejné množství, předchází vlastnímu „poolování“ měření koncentrace na Qubit. Vzorky jsou pak smíchány dohromady, ideálně by obsah DNA ve zkumavce měl činit více než 500 ng (ačkoliv je možné pracovat i s nižším obsahem, hraniční hodnotou je 100 ng).

### 3.4.9. Hybridizace (obohacení) knihovny

Pro vlastní hybridizaci je potřeba objem roztoku ve zkumavce zredukovat na 6  $\mu$ l. Za tím účelem je používán koncentrátor.

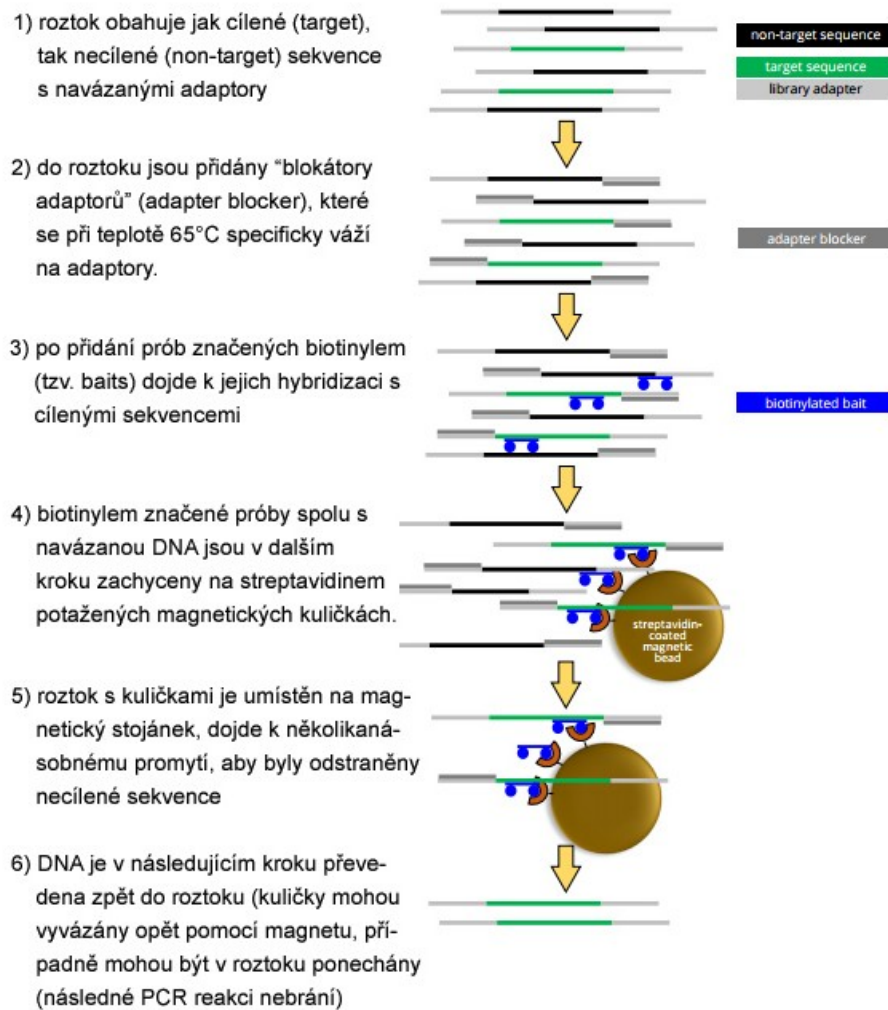
Použitá metoda obohacení vychází z protokolu (Gnirke et al. 2009), její schéma je na obr. 16, k provedení byly použity komerčně vyráběné próby MYBaits (MYcroarray). Jedná se o soubor COS navržený pro čeleď Compositae (CompCOS) základě ESTs druhů *Carthamus tinctorius* L. (světlice barvířská, ~20 000 ESTs), *Helianthus annuus* L. (slunečnice roční, ~70 000 ESTs), a *Lactuca sativa* L. (locika setá, ~70 000 ESTs). Obsahuje celkově 9678 hybridizačních prób, které cílí na 1061 ortologních genů (Mandel et al. 2014). V případě prvních vzorků analyzovaných v rámci pilotní studie byl použit za účelem obohacení laboratorní kit verze 2. Všechny ostatní vzorky (celkem 72) byly připraveny prostřednictvím novější verze kitu (verze 3), změny oproti předchozí verzi jsou detailně popsány na straně 6 protokolu výrobce, podle kterého bylo postupováno (viz <http://www.mycroarray.com/pdf/MYbaits-manual-v3.pdf>). Kromě změny poměrů vstupních objemů vzorků a přidávaných látek do reakcí, došlo ke změně v použitých pufrách, zatímco ve verzi 2 se k promývání streptavidinových kuliček používají dva typy pufrů, v novějším protokolu pouze 1)

#### 3.4.10. Post-capture amplifikace

Posledním krokem je amplifikace, tedy namnožení cílených sekvencí prostřednictvím PCR. Za tímto účelem byla použita KAPA Hifi HotStart polymeráza (bližší informace viz na stránkách výrobce <https://www.kapabiosystems.com/assets/KAPA-HiFi-HotStart-TDS-KR0369-v8.16.pdf>). Byly použity forward i reverse primery pro získání sekvencí v obou směrech.

Do 0,2 ml Low-bind zkumavky byl namíchán mix (Nuclease-free H<sub>2</sub>O 8,5 až 18,5 µl (dle objemu roztoku DNA), KAPA Hifi HotStart ReadyMix 25 µl, PCR primer mix 1,5 µl a obohacená knihovna 5 až 15 µl). Následoval PCR cyklus a po něm přečištění prostřednictvím komerčního kitu QIAquick PCR Purification Kit (Qiagen, obdobné jako výše).

Takto připravená knihovna byla odeslána k sekvenaci do sekvenační laboratoře v BioCevu (<http://www.biocev.eu/corefacilit/omics-genomika/> na sekvenátoru Illumina MiSeq, s kitem MiSeq Reagent Kit v2 pro 300 cyklů (výrobní číslo MS-102-2002), 2x150 bp paired-end reads).



Obr. 16 Schematické znázornění průběhu obohacení (obrázek převzat z protokolu výrobce <http://www.mycroarray.com/pdf/MYbaits-manual-v3.pdf>, upraveno)

## 3.5. Vyhodnocení získaných dat

### 3.5.1. Molekulární analýzy

Molekulární data obdržená ze sekvenační laboratoře (ve formátu FASTQ) byla analyzována zejména za použití skriptů HybPhyloMaker (Fér & Schmickl 2018; volně dostupné na adrese <https://github.com/tomas-fer/HybPhyloMaker>). Pro výpočetní operace byla využita gridová infrastruktura MetaCentrum (CESNET 2018), na úložišti MetaCentra jsou data také uložena (/storage/projects/botany-uk/Asteraceae/).

Prvním krokem je filtrování surových dat obdržených ze sekvenační laboratoře (označovaných jako raw reads). Jsou vyhledány kontrolní sekvence fágové DNA přidávané ke knihovně pro zvýšení komplexity, tzv. PhiX reads (za použití nástroje Bowtie 2 (Langmead & Salzberg 2012)) a následně odstraněny (SAMtools (Li & Durbin 2009)). Též jsou vymazány sekvence adaptorů a sekvence nízké kvality (Trimmomatic (Bolger et al. 2014)), případně sekvence příliš krátké, dále jsou odstraněny duplikáty (FastUniq (Xu et al. 2012)).



Takto získané úseky DNA jsou namapovány na tzv. pseudoreferenci. Z možných nástrojů, které je možné za tímto účelem použít, byl pro tento krok zvolen BWA (Li & Durbin 2009), ostatní jsou vhodné spíše pro kratší sekvence (Hatem et al. 2013). Pseudoreferenci program vytváří ze sekvencí práb použitých pro cílené obohacení během přípravy NGS knihovny, mezi které vkládá určitý počet prázdných pozic (N). V případě této práce byl využit původní soubor hybridizačních sond zmíněný výše, COS pro čeleď složnokvětých (*Compositae*) (Mandel et al. 2014), mezi které bylo vloženo 400 N. Ze souborů vzniklých mapováním (tzv. BAM souborů) byla vytvořena konsenzuální reference s použitím pravidla „majority rule“ přičemž byly zahrnuty pouze pozice s pokryvností nejméně 4 (např. Kamneva et al. (2017)). K tomu byl využit program kindel (Bede 2017). Při mapování vzniklé konsenzuální sekvence byly pro další analýzu rozděleny na jednotlivé exony. Ty byly dále porovnány se sekvencemi z původních souborů hybridizačních sond (za použití programu BLAT (Kent 2002)) k identifikaci cílových sekvencí. Byly vyřazeny všechny sekvence, které se lišily o více než 10 % od referenčního souboru.

Následovalo vzájemné přiřazení odpovídajících si úseků DNA (tedy alignment) programem MAFFT (Kato & Standley 2013). Po tomto kroku lze, v případě, že je předpokládána přítomnost paralogních genů, využít skript HybPhyloMaker4a2\_NoHet.sh, který umožňuje nalezení lokusů se zvýšeným zastoupením heterozygotních míst (jejich počet lze nastavit). Následně mohou být takové exony vyfiltrovány ze souboru pro další analýzy (viz manuál k programu HybPhyloMaker [https://github.com/tomas-fer/HybPhyloMaker/blob/master/docs/HybPhyloMaker\\_manual1.6.4.pdf](https://github.com/tomas-fer/HybPhyloMaker/blob/master/docs/HybPhyloMaker_manual1.6.4.pdf), str. 19). Dalším krokem je filtrování na základě chybějících dat. V tomto kroku je pro správné odvození fylogeneze nezbytné rozhodnout, jaké množství chybějících dat ponechaných v datovém souboru poskytne zároveň dostatečné množství genů a zároveň relevantní vztahy v rámci konstruovaného fylogramu. Proto byla analýza provedena několikrát s různým nastavením filtrování chybějících dat.

Dalším parametrem, který ovlivňuje, jaké sekvence budou dále použity, je „speciespresence“, který vyjadřuje požadované procentuální zastoupení vzorků v jednotlivých genech. V případě blízce příbuzných jedinců, je doporučeno (Fér, ústní sdělení) používat striktnější podmínky (90 – 100 %).

Dále se vytváří genové stromy, za tím účelem se využívá program RAxML 8.0.0 (Stamatakis 2014a) s modelem GTRGAMMA, se 100 opakováními bootstrapu; souhrnné tabulky a histogramy jsou vizualizovány prostřednictvím programu R (R Development Core Team, 2018), za tím účelem jsou použity skripty plotting\_correlations.R, tree\_props.R ([https://github.com/marekborowiec/good\\_genes](https://github.com/marekborowiec/good_genes), Borowiec et al. 2015) a treepropsPlot.R implementované též v rámci skriptů HybPhyloMaker.

Vzniklé genové stromy pak jsou základem pro výpočet tzv. druhových stromů. Za tímto účelem lze využít různé metody. V zásadě se používají dva typy modelů. První jsou založené na konkatenaci, sestavují alignované sekvence různých částí genomu do jedné matice, která následně slouží k odvození výsledné druhové fylogeneze prostřednictvím tradičních fylogenetických metod (Yu et al. 2011), např. pomocí maximum likelihood

implementované v programech EXaML (Kozlov et al. 2015), FastTree (Price et al. 2010). Druhou skupinou metod tvoří postupy založené na koalescentním modelu (Liu et al. 2009; Yu et al. 2011). Při jejich aplikaci je odvozován druhový strom na základě kombinace genových fylogenezí (MP-EST, ASTRAL) (Warnow 2015).

V rámci skriptů nástroje HybPhyloMaker jsou implementovány jak koalescenční metody ASTRAL (Mirarab et al. 2014) a ASTRID (Vachaspati & Warnow 2015), tzv. supertree metody MRL (Nguyen et al. 2012), nebo BUCKy založený na Bayesovské analýze korekondance (Larget et al. 2010), tak konkatenční přístupy využívající maximální věrohodnosti (program EXaML (Kozlov et al. 2015), či FastTree (Price et al. 2010)).

Pro porovnání byl při analýze molekulárních dat použit i alternativní přístup, který nabízí soubor skriptů HybPiper (Johnson et al. 2016), který umožňuje mimo jiné práci s introny. Výsledky obou programů by měly být vzájemně srovnatelné, byť se jejich vstupní předpoklady liší (Fér & Schmickl 2018).

HybPiper v porovnání s programem HybPhyloMaker nezahrnuje nezbytnou vstupní úpravu a filtrování dat. Autoři skriptů doporučují provést tyto kroky automaticky v Illumina BaseSpace a následně v programu Trimmomatic (Bolger et al. 2014). Dalším krokem je vymapování na referenční sekvenční. HybPiper využívá zkonkatenované geny cílených exonů (každý samostatně, HybPhyloMaker pro tento krok používá pseudoreferenci). K přiřazení sekvencí dochází, stejně jako ve výše zmíněném HybPhyloMaker-u, za pomoci programu BWA (Li and Durbin 2009) (v případě že jsou použity nukleotidové sekvenční), anebo programem BLASTX (Camacho et al. 2009) (pro peptidové). V tomto kroku se pracuje se samostatnými ready, nejsou složeny do sekvencí odpovídajících jednotlivým vzorkům (čímž se přístup liší od dalších používaných pipeline, např. (Faircloth et al. 2015), umožňuje tak dále pracovat s introny). Dalším krokem v rámci skriptů HybPiper je de novo assembly pro každý gen zvlášť. Analyzované sekvenční jsou uspořádány do kontigů pomocí programu SPAdes (Bankevich et al. 2012). Následně jsou seřazeny podle sekvenční hloubky a zalignovány podle reference. K získání kódujících sekvencí z kontigů, které obsahují i introny, je použit program Exonerate (Slater & Birney 2005); každý z kontigů je porovnán se srovnávací sekvencí na základě modelu „protein2genome“ (v případě předchozího použití BWA je proteinová reference vytvořena přímou translací v rámci jednoho z nástrojů balíku Biopython (Cock et al. 2009)). Výstupem je alignment z nepřekrývajících se úseků, které odpovídají minimálně z 60 % referenci (dle výchozího nastavení, tato hodnota může být změněna).

V další fázi skript `exonerate_hits.py` vybírá kontigy, které jsou od sebe vzdáleny alespoň 85 % délky původní sekvenční. Dalším kritériem pro výběr je sekvenční hloubka a případně podobnost s referencí. Skript nazvaný `intronerate.py` naopak cílí na introny, pracuje s nimi podobně jako předchozí skript s exony. Pro každý lokus vygeneruje zvlášť 2 sekvenční, a sice sekvenci „superkontigu“, která obsahuje jak exony, tak introny a poté sekvenci sestávající pouze z intronů. Použití analýzy intronových sekvencí může být užitečné zvláště v případě rychlých radiací a u blízce příbuzných jedinců (např. v rámci druhových komplexů), kdy

nemusí být pro odhalení divergence větví sekvence exonů dostatečně variabilní (Folk et al. 2015), což je i případ studované skupiny.

Pro další analýzu dat je v rámci nástroje HybPiper dostupný skript `retrieve_sequence.py`, který vytváří soubor sekvencí ve FASTA formátu. Získaný výstup lze pak použít pro zalignment genů v programu MAFFT (Katoh & Standley 2013) a dále pak k tvorbě genových stromů.

Co se týče aplikace skriptů HybPiper a HybPhyloMaker na HybSeq data, jak již bylo řečeno výše, měly by být jejich výstupy srovnatelné. Nicméně liší se v několika hlavních bodech, což znemožňuje při použití identických dat nastavit stejné podmínky (Fér & Schmickl 2018). Předně HybPhyloMaker využívá pro assembly pomocný referenční genom (zde zvaný pseudoreference) získaný ze sekvencí hybridizačních sond oddělených prázdnými pozicemi, zatímco HybPiper de novo assembly přístup. Druhý jmenovaný by mohl být pro typ hybridizačních sond, jaké byly v této práci použity, teoreticky vhodnější (Fér, ústní sdělení).

Dalším rozdílem je způsob jakým je v rámci těchto programů řešena otázka paralogních genů. V případě skriptů HybPiper dochází k detekci možných paralogů během vytváření kontigů programem SPAdes. Za paralogy jsou považovány takové geny, pro které bylo získáno více kontigů o délce 75 % délky referenční proteinové sekvence. V rámci skriptů HybPhyloMaker lze vyfiltrovat možné paralogy také, a sice na základě počtu přítomných heterozygotních míst (za použití skriptu `HybPhyloMaker4a2_NoHet.sh`). Též je v rámci skriptů programu HybPhyloMaker možné v případě nízké variability vzorků a přítomnosti heterozygotních míst, přistoupit k použití ambiguitních kódů NC-IUPAC při tvorbě konsenzuální sekvence (za různých nastavení prahových hodnot viz dále) za účelem podchycení alelické variability pomocí programu ConsensusFixer (Töpfer n.d., volně dostupný na adrese <https://github.com/cbg-ethz/ConsensusFixer>).

Další rozdíl spočívá v přístupu k možným sekvencím intronů, při použití skriptů HybPhyloMaker jsou sice nalezeny během assembly, nicméně dále se s nimi nepracuje, zatímco HybPiper umožňuje pracovat i s těmito, oproti cíleným exonům podstatně variabilnějšími, sekvencemi.

Oba zmíněné programy byly použity pro analýzu cílených exonů (soubor COS pro *Asteraceae* (Mandel et al. 2014)).

Vzhledem ke skutečnosti, že v rámci mladé skupiny, jakou je *Carlina vulgaris* agg., lze předpokládat přítomnost různých evolučních procesů, jako je incomplete lineage sorting (ILS, více viz diskuze), hybridizace atd., jejichž vliv lze však těžko odhadnout, byly v případě skriptů HybPhyloMaker využity tři přístupy a sice:

a) tvorba konsenzuální sekvence (v rámci skriptu `HybPhyloMaker2_read_mapping.sh`) za pomoci programu kindel (Bede 2017; volně dostupný na adrese <https://github.com/bede/kindel>). S minimální pokryvností 4 (minimum site coverage, v rámci skriptů HybPhyloMaker-u nastavení parametru "mincov", tedy místo nukleotidů, které na dané pozici jsou v méně než 4 readech, se vkládá do výsledné sekvence znak pro chybějící bázi "N") a s nastavením prahové hodnoty pro "majority consensus" pro případ



heterozygotních míst na 0.8 (“majority threshold for consensus calling“, parametr „majthres“, do konsenzuální sekvence je zařazena báze, která je zastoupena ve vstupních readech z více jak 80 %); analýza všech získaných exonů (pouze s vyfiltrováním chybějících dat s nastavením “missingpercent” = 70, “speciespresence” = 90),

b) tvorba konsenzuální sekvence za pomoci programu kindel (s obdobným nastavením jako výše); další analýza s odfiltrováním možných paralogů (tedy lokusů se zastoupením heterozygotních míst, v tomto případě vyšším než 4, menší hodnoty odpovídají spíše chybám, ke kterým došlo při mapování, než reálné skutečnosti a nebo alelické variabilitě),

c) tvorba konsenzuální sekvence za pomoci programu ConsensusFixer (Töpfer n.d.) (více možných variant v alignmentu označuje pomocí kódů pro ambiguitní báze) s nastavením parametru “plurality” = 0,3 (odpovídá minimálnímu zastoupení odlišné báze v sekvencích, tedy alespoň ve 3 z 10 původních sekvencí musela být zastoupena odlišná báze, aby byla pozice obsazena zástupným kódem); další analýza zahrnovala odfiltrování možných paralogů obdobně jako výše.

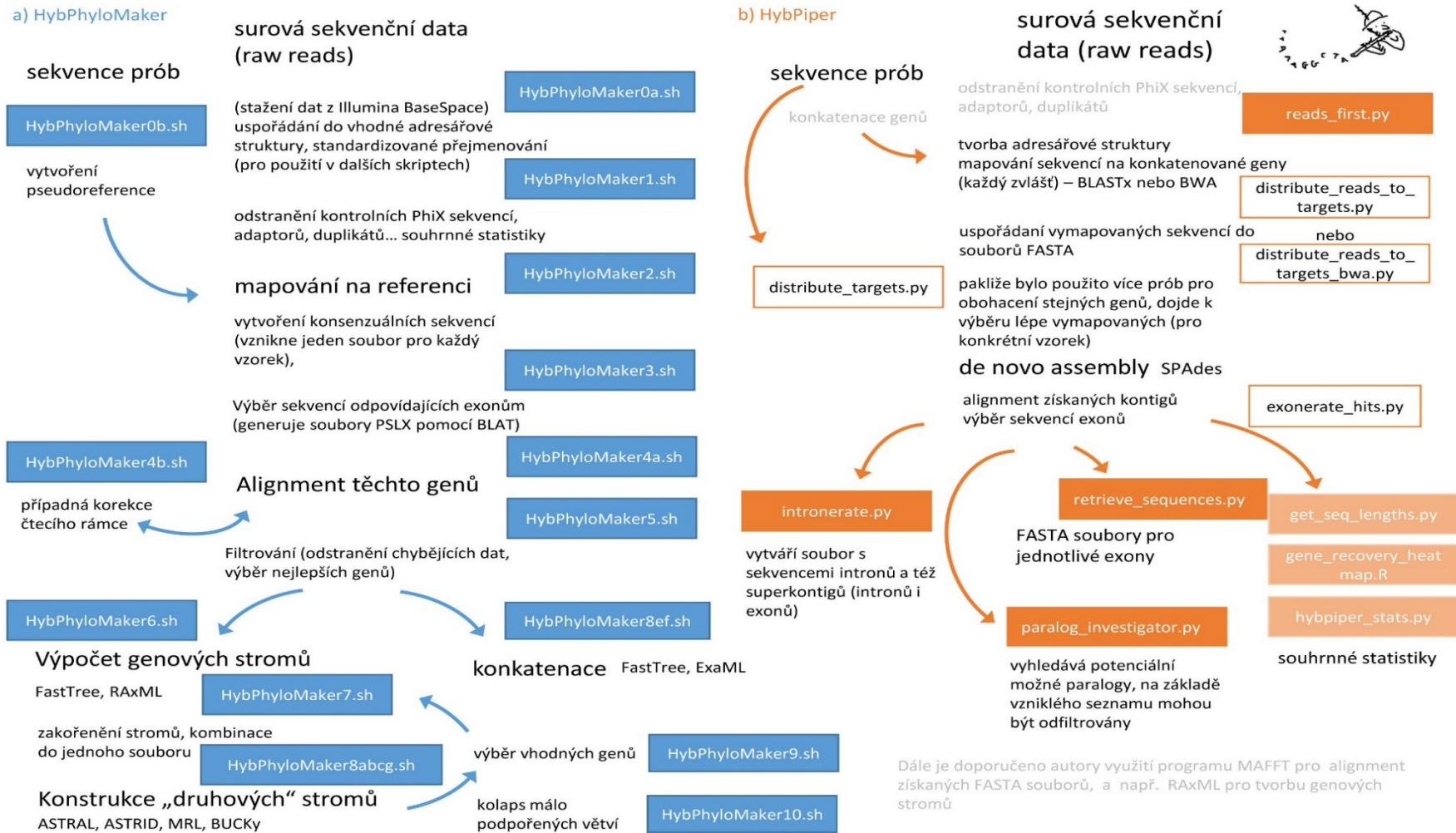
Pro práci s introny byl použit program HybPiper. Za tím účelem byla sestavena reference ze sekvencí práb cílících pouze na exony bez heterozygotních míst (byly nalezeny obdobně jako v případě analýz v HybPhyloMaker-u) za účelem redukce počtu možných paralogů. Se získanými daty (se souborem, který sestával ze supercontigů, tedy exonů a intronů spojený dohromady) byl dále proveden alignment v programu MAFFT (Katoh & Standley 2013) (jak doporučují tvůrci programu HybPiper). Vzniklé alignmenty byly zkonkaténovány pomocí programu AMAS (Borowiec 2016).

Pro mapování chloroplastové DNA, opět za použití programu HybPhyloMaker, byl jako referenční sekvence použit celý chloroplastový genom *Carthamus tinctorius* získaný z GenBank (NC\_030783.1, volně dostupný z adresy [https://www.ncbi.nlm.nih.gov/nuccore/NC\\_030783.1/](https://www.ncbi.nlm.nih.gov/nuccore/NC_030783.1/)). Nicméně podařilo se tak získat pouze omezené procento cpDNA sekvencí, nedostatečné pro odvození relevantní fylogeneze, proto nebylo k další analýze přistoupeno (pro podrobnosti viz tab. II v příloze).

Pro získání rDNA byla sestavena reference z 675 bazí úseku ITS1-28S *Carlina vulgaris*, (získaná z GenBank, no. kf301217, volně dostupná z adresy <https://www.ncbi.nlm.nih.gov/nuccore/KF301217.1/>). K této sekvenci byly v programu Geneious přiřazeny sekvence vzorku 38 prostřednictvím opakovaného (iterativního) mapování (v prvním kroku se mapují sekvenční ready přímo na referenci, k takto vzniklým úsekům se přidávají v dalších krocích další, které se s původními částečně překrývají). Vznikla tak referenční sekvence dlouhá 6921 bp (viz elektronické přílohy práce, soubor *Carlina-vulgaris\_rDNAref.fasta*). Po namapování na tuto referenci v programu BWA (Li & Durbin 2009) s nastavením parametrů “mincov” = 4, “majthres” = 0,9 (vysvětlení výše), byla prostřednictvím programu ConsensusFixer (Töpfer n.d.) vytvořena konsenzuální sekvence s ambiguitními znaky NC-IUPAC v místech, kde byla detekována přítomnost různých bazí v rámci readů jednoho vzorku (a sice s nastavením -plurality 0.3 -mcc 4 (vysvětlení užití parametru “plurality” viz výše, “mcc” je obdobou parametru “mincov” ). Získané sekvence byly zalignovány v programu MAFFT (Katoh & Standley 2013).

V konkatenovaných souborech ve formátu FASTA byla následně odstraněna prázdná místa v programu trimAl v1.2 (Capella-Gutiérrez et al. 2009) (parametr -nogaps) s defaultním nastavením, za účelem další redukce možného biasu. Ke konstrukci fylogenetických stromů byl ve všech případech použit program RAxML 8.0.0 (Stamatakis 2014) s modelem GTRGAMMA, se 100 opakováními bootstrapu. Pro vizualizaci vzájemných vztahů mezi vzorky, kde lze očekávat dopady retikulární evoluce, je však vhodnějším zobrazením sítí. Fylogenetické sítě za použití neighbour-net metody byly generovány programem SplitsTree v4.10 (Huson & Bryant 2006) na základě neupravené P-vzdálenosti (uncorrected P-distance) a s defaultním nastavením, uloženy byly ve formátu .nex. V případě souboru, v kterém byly zahrnuty znaky pro ambiguitní pozice bylo použito nastavení "Average" (pro popis parametru viz manuál <http://ab.inf.uni-tuebingen.de/data/software/splitstree4/download/manual.pdf>).

Stromy byly zakořeny prostřednictvím skriptu retree v rámci program PHYLIP 3.6 (<http://evolution.genetics.washington.edu/phylip/phylip.html>, Felsenstein 1993). Výsledky pak byly vizualizovány v programu R (R Development Core Team 2018). Fylogenetické stromy prostřednictvím funkcí "read.tree" a "plot.phylo" z balíčku "ape 5.1 (Paradis et al. 2004)". Fylogenetické sítě prostřednictvím funkcí "read.nexus.networx" a "plot.networx" z balíčku "phangorn 2. 4. 0 (Schliep 2011)". Pro soupis použitých balíčků a funkcí viz tab. V v příloze.



Obr. 17 Znárodnění jednotlivých dílčích kroků používaných „pipeline“. a) HybPhyloMaker (Fér and Schmickl 2018) umožňuje vedle práce se surovými daty (raw reads) a tvorby výstupních FASTA souborů získaných exonů i následující konstrukci fylogenetických stromů, b) HybPiper (Johnson et al. 2016) naproti tomu umožňuje práci s introny a paralogy. Hlavní rozdíl v přístupech obou programů tkví ve způsobu vmapování a assembly získaných sekvencí DNA, HybPhyloMaker využívá pseudoreferenci vzniklou ze sekvencí práb tzv. reference-guide assembly, zatímco HybPiper přistupuje k vmapování každého exonu zvlášť a skládá je až v dalších krocích prostřednictvím de novo assembly (šedivě odlišený text označuje kroky, které musí být provedeny mimo program HybPiper). Bližší popis v textu. Sestaveno na základě schémat dostupných na stránkách [https://github.com/tomas-fer/HybPhyloMaker/blob/master/docs/HybPhyloMaker\\_workflow.png](https://github.com/tomas-fer/HybPhyloMaker/blob/master/docs/HybPhyloMaker_workflow.png) a <https://github.com/mossmatters/HybPiper/wiki>

## 3.5.2. Morfometrické analýzy

### 3.5.2.1. Použité soubory dat

Pro morfometrické analýzy byly sbírány kvetoucí rostliny. Avšak v případě rostlin z osluněných stanovišť (včetně těch pěstovaných v experimentální zahradě) byly v době květu většinou zaschlé listy z dolní části lodyhy, někdy i ze střední. Dalším problémem byly často poškozené terminální úbory (ať již okusem zvěří, nebo hmyzem - larvy některých zástupců se vyvíjí přímo v úborech, ty jsou pak podstatně menší než v případě zdravých rostlin a ne vždy se rozvíjí do květu). V neposlední řadě pak lze zmínit svazčitost (fasciaci), která byla zejména v sezonách let 2015 a 2016 přítomna u řady rostlin v zahradě. Šlo pravděpodobně o důsledek virového onemocnění; ač některé populace byly zcela zdravé, značná část pěstovaných rostlin více či méně vykazovala znaky poškození tohoto typu. Též byly u řady rostlin pozorovány příznaky napadení rzemi. Vzhledem k výše zmíněným skutečnostem nebylo možné odečíst všechny stanovené znaky, řada měřených rostlin proto byla vyřazena a následné analýzy byly proto prováděny na 2 různých datových souborech (oba jsou součástí elektronických příloh této práce, označeny jako 197samples.csv a 401samples.csv).

První zahrnoval všechny měřené znaky, jejich soupis viz tab. III v příloze (celkem se jednalo pouze o 197 rostlin, z toho 145 z terénu a 52 dopěstovaných v experimentální zahradě).

Druhý ze souborů dat pak představoval výběr rostlin a měřených znaků, které vykazovaly nejmenší podíl chybějících hodnot. Byly vynechány některé znaky na listech, které nebylo možné odečíst na seschlých nebo jinak poškozených listů (uhel<sup>3</sup>, proh1, proh2, proh3, počty ostnitých zubů (numostd, numostp, numosth), a z nich počítané poměrové proměnné (ostdol, ostpro, osthor)), dále znaky v květenství (vnz a stz), které nebyly měřené v případě odkvétajících a poškozených úborů. Celkem se jednalo o soubor 401 rostlin (93 ze zahrady a 308 z terénu).

Na základě exploratorní analýzy dat byly z datového souboru vyloučeny rostliny pěstované v Botanické zahradě PřF UK (populace 10); důvodem byl výrazně odlišný vzrůst daný atypickými podmínkami, zároveň nízký počet těchto vzorků. Též bylo přistoupeno k hodnocení zvláště datových souborů rostlin měřených v experimentální zahradě a z terénu (vzhledem k posunu v ordinačním prostoru PCA, viz výsledky). Rozdělení hodnot znaků v rámci souborů dat bylo testováno na normalitu Shapirovým-Wilkovým testem. V případě, že rozdělení nebylo normálové, bylo přistoupeno k logaritmické nebo odmocninové transformaci takových proměnných.

### 3.5.2.2. Vlastní analýzy

Pro posouzení vlivu rozdílných podmínek na fenotyp rostlin byly porovnány datové soubory znaků naměřené na rostlinách pěstovaných v experimentální zahradě a na rostlinách sbíraných v terénu. Za tím účelem byla provedena na datovém souboru zahrnujícím všechny znaky (počet vzorků 197) analýza hlavních komponent (PCA). Pro vlastní výpočet PCA byla použita funkce dudi.pca z balíčku ade 4 1. 7. 11 (Dray & Dufour

---

<sup>3</sup> Pro vysvětlení zkratk znaků viz tab. III v příloze.

2007) s výchozím nastavením (standardizace dat na průměr rovný nule, škálování směrodatnou odchylkou), vyřazeny byly silně korelované znaky (korelační koeficient vyšší jak 0,95). Dopočtení procentuálních příspěvků k jednotlivým osám pak bylo provedeno funkcí PCA z balíčku FactoMineR 1.41 (Lê et al. 2008).

Dále byla porovnávána morfologická variabilita mezi předem definovanými skupinami. Skupiny vylišené na základě molekulární analýzy byly pouze dvě, a sice skupina pupavy Biebersteinovy pravé (včetně rostlin z reliktních biotopů západních Čech a jesenického endemita, subsp. *sudetica*) na jedné straně a všechny ostatní rostliny *Carlina vulgaris* agg. zahrnuté v této práci na straně druhé (více viz v kapitole 4.2, níže). Nicméně v rámci druhé ze skupin lze na základě pozorování z terénu i experimentální zahrady odlišit několik morfotypů, minimálně podruhy *Carlina vulgaris* subsp. *vulgaris* a *C. biebersteinii* subsp. *brevibracteata* jsou v typické podobě dobře odlišitelné, zvláště při pěstování ve stejných podmínkách. Existují však mezi nimi přechodné typy. Další skupinou, která by se mohla od ostatních lišit, jsou, vzhledem k povaze jejich stanovišť, potenciálně reliktní populace z hluboce zaříznutých říčních údolí. Proto pro porovnání byla v následujících analýzách znázorněna též příslušnost k těmto pracovním skupinám (jejich soupis je v tabulce níže).

pracovní skupina	b.	vzorky
rostliny z Malé kotliny ( <i>Carlina biebersteinii</i> subsp. <i>sudetica</i> )	●	1, 63, 64
rostliny reliktních stanovišť v západních Čechách (Chlumská hora u Manětína, údolí Teplé)	●	19, 36
horské rostliny z Alp	●	3, 4, 22, 24, 38
horské rostliny z Karpat	●	40, 41, 45, 57
rostliny klasifikované jako <i>Carlina vulgaris</i> subsp. <i>vulgaris</i> (obvykle nižšího vzrůstu, z lomů a jiných mělkých půd)	●	6, 12, 13, 14, 15, 16, 23, 25, 32A, 33, 49, 51, 52, 54, 55, 56, 60, 69, 70, 71, 72, 78
rostliny hluboce zaříznutých údolí řek (Vltava, Dyje, Jihlava, Rokytná)	●	5, 31, 50, 59, 73, 74
rostliny klasifikované jako <i>Carlina biebersteinii</i> subsp. <i>brevibracteata</i> (převážně z xerothermních trávníků)	●	7, 8, 9, 10, 11, 17, 18, 20, 28, 29, 30, 37, 39, 42, 58, 64, 65, 66, 67, 68
přechodné a nejasné typy mezi <i>Carlina vulgaris</i> subsp. <i>vulgaris</i> a <i>C. biebersteinii</i> subsp. <i>brevibracteata</i>	●	2, 21, 26, 27, 34, 46, 47, 48, 75, 76, 77
(nedoložené vzorky)	●	32B, 35, 53, 61, 62, 79

Tab. 1 Přehled pracovních skupin definovaných na základě místa původu, morfologie a ekologie rostlin. b. = barva použitá v následujících grafech a schématech k znázornění skupin, vzorky = označení vzorků, které byly do jednotlivých skupin přiřazeny (jejich soupis viz v tab. I v příloze).

Pro morfologické porovnání zmíněných skupin byl použit datový soubor představující více vzorků (401), neboť zahrnoval více populací. Analyzovány byly zvláště rostliny pěstované a měřené v experimentální zahradě a rostliny z terénu. Na oba datové soubory byla aplikována PCA; jak na úrovni jedinců (viz obr. 28), tak na úrovni populací (viz obr. 29). Ověření přítomnosti testovaných skupin v rámci této analýzy bylo provedeno na základě randomizačního testu (funkce `ade4::rtest`). Na tentýž datový soubor byla použita i



shlukovací analýza, za tím účelem byla z dat sestavena matice vzdáleností (funkce `vegdist` z balíčku `vegan` 2. 5. 2 (Oksanen et al. 2018)), použita byla manhattanská vzdálenost, data byla standardizována směrodatnou odchylkou a centrována (funkce „`decostand`“, která v rámci balíčku `vegan` slouží k přípravě dat pro další analýzu). Vlastní dendrogramy byly sestaveny za pomoci funkce `hclust`, dále vizualizovány prostřednictvím funkcí z balíčku `dendextend` 1. 8. 0. (Galili 2015).

Diskriminační analýzy pak zahrnovaly kanonickou diskriminační analýzu (provedena na stejných maticích dat jako předchozí PCA funkcí „`diskrimin`“ z balíčku `ade4`), dále lineární diskriminační analýzu (provedena funkcí „`lda`“ z balíčku `MASS` 7. 3. 49 (Venables & Ripley 2002)) s křížovou validací (leave-one-out cross-validation, soubor bez jednoho z objektů je použit jako tréninkový, na jeho základě je odvozeno klasifikační kritérium, které je zpětně aplikováno na objekt, který byl vynechán) a případně kvadratickou diskriminační analýzu (`MASS::qda`). Vhodné diskriminátory, znaky k odlišení předem definovaných skupin, byly odvozeny jednak na základě příspěvků vysvětlené variability ke kanonické ose, jednak za pomoci Wilkova kritéria  $\lambda$  (funkce `greedy.wilks` z balíčku `klaR` 0. 6. 14 (Weihs et al. 2005)). (Z datového souboru byly vyřazeny znaky s vyšším korelačním koeficientem, než 0,9. Vedle snahy omezit vliv vzájemné korelace znaků, byla jejich redukce nezbytná vzhledem k malému počtu vzorků, testováno tedy bylo pouze 31 znaků. Hodnota Wilkovo  $\lambda$  vyjadřuje diskriminační sílu navrženého modelu bez zahrnutí testovaného znaku. Test začíná se znakem, který k rozdělení skupin přispívá největší měrou, v dalším kroku je tento odstraněn, pokračuje, dokud p-hodnota dosáhne předem stanovené významnosti (tj. znaky přispívají k rozdělení do skupin), zde byla použita defaultní hladina 0,02. Ze zmíněných 31 znaků splňovalo kritérium 13 viz tab. 4.

Jedním z faktorů, na který je v souvislosti s morfometrickou analýzou třeba brát zřetel, je možnost vlivu různých makroekologických podmínek na fenotyp rostlin. Jejich sledování však nebylo předmětem této práce a ani nebyly na lokalitách podrobně zaznamenávány. Nicméně jedním ze základních ukazatelů různých podmínek prostředí je nadmořská výška. Vzorky byly proto rozděleny do 3 arbitrárních skupin (< 350 m n. m., 350–500 m n. m., > 500 m n. m.) a toto zařazení vyznačeno na ordinačním diagramu analýzy PCA (obr. 35). Dále byla pro podrobnější srovnání morfologické a stanovištní variability použita klimatická data získána z databáze WorldClim (Fick & Hijmans 2017) prostřednictvím funkce „`getData`“ z balíčku `raster` 2. 6. 7. (Hijmans 2017). Vedle zmíněné nadmořské výšky tak byly získány informace o průměrných ročních srážkách a průměrné teplotě v nejteplejším a nejstudenějším měsíci. Na výslednou matici byla opět aplikována PCA (`ade4::dudi.pca`) s defaultním nastavením. Pro porovnání podobnosti této matice a předchozí odvozené z morfometrických dat byl použit jednak mantelův test (`vegan::mantel`), jednak koinerční analýza (`ade4::coiner`), v rámci které je počítán i RV koeficient.

## 4. Výsledky

### 4.1. Sekvence DNA, vstupní data

Sekvence DNA probíhala ve 4 separátních bžích. První byl proveden v rámci pilotní studie, kdy bylo testováno, zdali při použití hybridizačních prb navržených pro celou čeeď *Compositae* dojde k dostatečnému obohacení DNA studovaného taxonu (tato knihovna obsahovala vedle vzorků ze skupiny *Carlina vulgaris* agg. též DNA zástupců rodů *Hieracium*, *Lasiocephalus* a *Tripleurospermum*). Kromě jednoho vzorku, v jehož případě došlo díky nešetrné manipulaci během přípravy k snížení koncentrace, bylo procento získaných cílených úseků DNA dostačující. Proto bylo pristoupeno k přípravě dalších knihoven (celkem tři po 24 vzorcích). Oproti zmíněnému běhu byla použita novější verze kitu pro obohacení (viz kapitola 3.4.9.). Celkem byly získány sekvence 77 vzorků (nepoužitelný vzorek z pilotní studie byl osekvenován znovu v rámci jedné z dalších knihoven). Přehled vzorků viz tab. I v příloze.

Množství získaných readů na vzorek se pohybovalo od 478 966 (u vzorku 40, podobně nízké hodnoty byly též u vzorků 39 a 20) po 2 669 331 (u vzorku 17), průměrně bylo z jednoho vzorku osekvenováno 1 233 877 readů. V následném filtrování dat prostřednictvím programu Trimmomatic (Bolger et al. 2014) bylo odstraněno průměrně 2,52 % sekvencí (od 0,76 u vzorku 75 po 7,41 % u vzorku 33). V dalším kroku byly odstraněny duplikáty, jednalo se v průměru o 30 % sekvencí readů (nejméně to bylo u vzorku 33, 7,25 %, nejvíce u vzorku 40 58,58 %). Celkem u 20 vzorků tvořily duplikáty více jak polovinu získaných readů. Podrobnosti viz v tab. II v příloze.

Dále proběhlo mapování takto filtrovaných dat na referenční sekvence. V případě cílených exonů bylo zastoupení vymapované DNA mezi ready v průměru 30 %. Nejmenší podíl této cílené DNA, kolem 10 %, obsahovaly vzorky, které byly analyzovány v rámci pilotní studie a byly obohaceny starší verzí laboratorního kitu. Největší podíl pak byl v DNA vzorku 75 (43,5 %), největší absolutní počet u vzorku 4 (446701 readů, 30,71 % z původních 1454716), nejmenší u vzorku 40 (67434 readů, tedy 34,73 % z původních 468795).

U necílených úseků (cpDNA a rDNA) bylo zastoupení vymapovaných úseků podstatně nižší. V případě cpDNA představovalo průměrně pouhých 0,91 % (více než 1 % chloroplastové DNA bylo zastoupeno u 18 vzorků (nejvyšší hodnoty, mezi 1,7 až 3,88 %) vykazovaly opět vzorky připravované starší verzí laboratorního kitu v rámci pilotní studie). K získání dostatečně dlouhých sekvencí cpDNA pro odvození relevantních fylogenetických vztahů je třeba minimálně 10 000 readů (Fér, ústní sdělení), takových bylo v případě této práce pouhých 12, proto bylo od následné analýzy těchto dat ustoupeno. Na referenci z rDNA cistronu se pak namapovalo 355 (u vzorku 51) až 9965 readů (vzorek 17 z pilotní analýzy), tato DNA ve vzorcích byla zastoupena s podílem od 0,06 % (u zmíněného vzorku 51) po 0,92 % (u vzorku 65), s průměrem 0,27 %.

V případě analýzy cílené DNA je důležitým krokem filtrování dat. Základními parametry, které používá program HybPhyloMaker za tímto účelem, je přítomnost chybějících dat v sekvencích získaných exonů (missing percent) a procento vzorků, ve kterých se tyto sekvence

vyskytují (species presence), jak již je popsáno výše. Dále je možné porovnat exony prostřednictvím souhrnných tabulek generovaných programem AMAS (Borowiec 2016). Ukazují další parametry (délka alignmentu daného exonu, průměrné bootstrapové hodnoty, průměrné délky větví genových stromů). Z původních cílených 1061 orthologních exonů se podařilo u všech vzorků získat pouze 236, které obsahovaly méně než 50 % chybějících dat. V případě použití méně striktních podmínek pro filtrování se pohyboval počet získaných exonů mezi 450 – 512. Při filtrování dat na základě délky alignmentu exonů, byl počet redukován přibližně na polovinu původních sekvencí, které splňovaly dané podmínky (optimálně dlouhý alignment 200 bází). Chyby v odvozené fylogenezi by tak mohly být způsobeny nedostatečným množstvím vstupních dat, proto bylo od tohoto kritéria dále ustoupeno. V tab. 2 jsou příklady použitého filtrování dat (výsledné topologie byly inkongruentní, srovnatelné více méně s následujícími v kapitole Výsledky, v rámci práce nejsou zobrazeny).

Pro další práci bylo použito nastavení parametrů „missing percent“ = 70, „species presence“ = 90. Do další analýzy tak bylo zahrnuto 450 exonů s průměrnou délkou alignmentu 248 bází, případně 348 exonů s průměrnou délkou alignmentu 268 bp v případě odstranění heterozygotních lokusů a při počítání s ambiguitními bazemi 172 exonů o průměrné délce alignmentů 186 bp.

Co se týče aplikace programu HybPiper na datový soubor, z 534 vybraných exonů s nižším zastoupením heterozygotních míst, se vymapovalo 453. I tak byly některé detekovány prostřednictvím skriptu paralog\_investigator.py jako paralogy, z dalších analýzy byly odstraněny, výsledný počet superkontigů tak byl 294 sekvencí s průměrnou délkou alignmentu 2115 bp.

<b>filtrování dat</b>	<b>missing percent</b>	<b>species presence</b>	<b>exony</b>	<b>délka alignmentu</b>	<b>bootstrap</b>	<b>délka větví</b>
odstraněny sekvence s více jak 70 % chybějících dat a přítomné u méně než 75 % vzorků (defaultní nastavení programu)	70	75	480	234	0,456	0,00035
odstraněny sekvence s více jak 80 % chybějících dat a přítomné u méně než 80 % vzorků	80	80	505	246	0,445	0,00023
odstraněny sekvence s více jak 90 % chybějících dat a přítomné u méně než 75 % vzorků	90	75	521	263	0,421	0,00063
odstraněny sekvence s více jak 50 % chybějících dat	50	100	236	245	0,52	0,00062
odstraněny sekvence s více jak 50 % chybějících dat, dále odstraněny sekvence o délce 200 bází a kratší	50	100	128	332	0,56	0,00069

**Tab. 2** Příklady použitého filtrování dat. Popis parametrů „missing percent“ a „species presence“ viz v textu. **exony** = počet exonů získaných na základě zmíněných nastavení, **délka alignmentu** = průměrná délka alignmentu exonů (počet bází), **bootstrap** = průměrná bootstrapová podpora v genových stromech, **délka větví** = průměrná délka větví v genových stromech



## 4.2. Vlastní fylogenetické analýzy

Jak již bylo uvedeno v kapitole metodika (3.5.1.), fylogenetické stromy a sítě byly konstruovány z dostupných dat jednak na základě cílených exonových sekvencí jaderné DNA (a to 3 různými přístupy v rámci programu HybPhyloMaker, a programem HybPiper, který umožňuje zahrnout do analýz i přilehlé úseky intronů), jednak na základě vypapované cistronové rDNA. Přehled použitých datových souborů viz tab. 3.

data	konk. sekvence	trimAl	var. místa	% var. míst	bootstrap	podp. nody	délka větví	obr.
exony, filtrování chybějících dat	111125	64877	1579	2,43	25	5	0,019	18
exony, filtrování chybějících dat a odstranění heterozygotních lokusů	51095	44551	1093	2,45	21	3	0,020	19
exony, použity ambiguitní báze, filtrování chybějících dat a odstranění heterozygotních lokusů	59103	31973	170	0,5	24	8	1,05E-04	20 a 21 a
superkontigy (výstup z programu HybPiper)	619675	135508	3496	2,57	39	14	8,2E-04	20 b 21 b
rDNA	6951	6926*	143	2,06	24	6	5,23E-05	22

**Tab. 3** Datové soubory použité pro odvození fylogeneze. konk. sekvence = výsledná délka konkatenované sekvence, trimAl = délka konk. sekvence po odstranění prázdných pozic v programu trimAl (Capella-Gutiérrez et al. 2009), var. místa = počet variabilních míst v rámci sekvence, % var. míst = procento variabilních míst, bootstrap = průměrné bootstrapové podpory výsledných fylogenetických stromů, délka větví = průměrná délka větví výsledných fylogenetických stromů, obr. = označení obrázků, které reprezentují výstupy z jednotlivých datových souborů. \*V případě datového souboru rDNA byla provedena úprava konkatenované sekvence manuálně v programu Tablet

Získané fylogenetické stromy ukazují dvě základní evoluční linie. První zahrnuje vzorky příslušící k taxonům *Carlina biebersteinii* subsp. *biebersteinii* a *C. biebersteinii* subsp. *sudetica*, tedy ty původem z Alp a Karpat (označené čísly 3, 22, 24, 38, 40, 41, 45, 57), reliktních lokalit v západních Čechách (19, 36) a z Malé kotliny (1, 63, 64). Druhou pak reprezentují všechny ostatní vzorky analyzované v rámci této studie, tedy ty řazené k *C. vulgaris*, *C. biebersteinii* subsp. *brevibracteata* a přechodné morfotypy mezi typickými rostlinami těchto taxonů. Co se týče rozdělení vzorků do těchto dvou skupin, je obdobné takřka ve všech získaných fylogenetických stromech a sítích (s bootstrapovou podporou od 85 v rámci stromu rDNA po 100 ve fylogenezi odvozené na základě superkontigů); výjimku tvoří několik vzorků (4, 5, 13, 66 a 77), v jejichž případě se příslušnost do zmíněných skupin různí ve fylogenetických stromech konstruovaných na základě exonů nDNA a rDNA (více viz diskuze).

Vzájemné vztahy uvnitř těchto dvou linií však zůstávají nejasné: větvení ve vzniklých fylogenetických stromech vykazují nízkou bootstrapovou podporu, jednotlivé stromy

získané na základě různého přístupu k vstupním datům se různí. Nejvýraznější je tato skutečnost u vzorků, které pocházejí ze stejných populací (vzorky 1, 63 a 64 z Malé kotliny a částečně i vzorky 10, 30 a 42 z lokality Březí v Dunajovických kopcích). V rámci fylogenetických stromů, v jejichž případě nebylo pracováno s ambiguitními bazemi (obr. 18, 19), jsou vzorky z Malé Kotliny součástí různých větví v rámci skupiny *C. biebersteinii* subsp. *biebersteinii*. V případě prvního (pouze vyfiltrovaná chybějící data, obr. 18) vychází jako sesterský vzorku 64 vzorek 66, nicméně bootstrapové hodnoty jsou obecně nízké. V rámci fylogenetické sítě odvozené ze stejných dat však jsou sesterské vzorky 63 a 64, vzorek 66 je na bázi vlastního shluku reprezentujícího *C. biebersteinii* subsp. *biebersteinii*. (Jedná se o vzorek, reprezentující rostliny z Ukrajiny diametrálně odlišné morfologie i ekologie, jeho pozice je i v rámci ostatních fylogenetických stromů problematická. Obdobně v dalším z nich (obr. 19a) je sesterský vzorku 63, v dalších se nachází na různých pozicích v rámci skupiny *C. biebersteinii* subsp. *biebersteinii* včetně subsp. *sudetica*). Vzorky 1, 63 a 64 tvoří jednu linii v případě jednoho z fylogenetických stromů (obr. 20 a), a dvou fylogenetických sítí (obr. 21). Též v případě dalších vzorků, které byly pořízeny z jedné populace (byť se jednalo o rostliny pěstované po několik generací v odlišných podmínkách, dala by se u nich očekávat blízká příbuznost), se zde netvoří skupina, vzorky 10 a 42 sice jsou v rámci několika topologií u sebe, vzorek 30 však stojí v pozici poměrně vzdálené. Reálné vztahy mezi populacemi (není-li zřejmý ani vztah v rámci jedné populace), tedy nelze z těchto dat odvozovat.

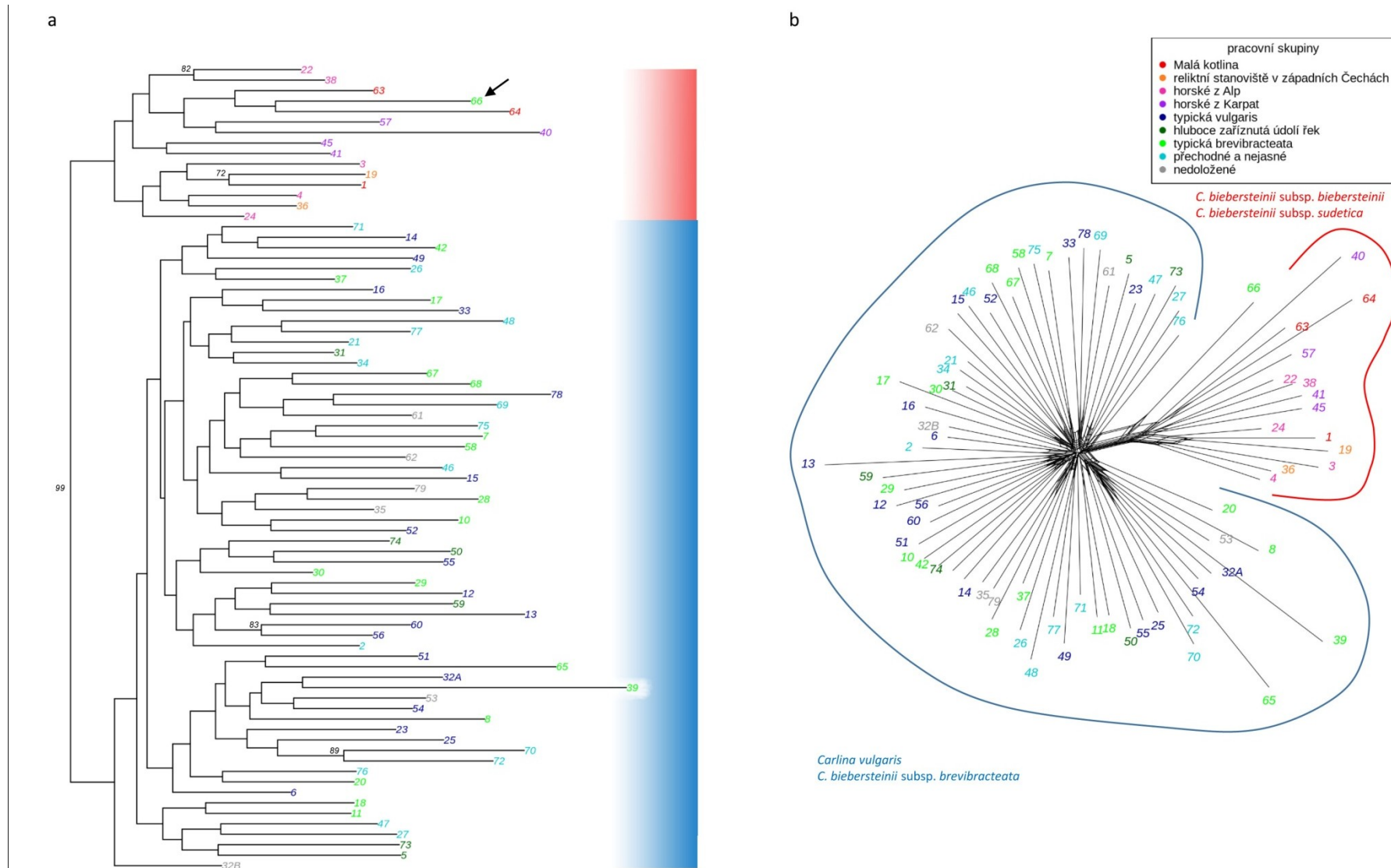
Fylogenetický strom konstruovaný na základě cistronu rDNA (obr. 22) představuje poněkud odlišnou topologii, ani zde však nelze předpokládat reflexi reálných vztahů na nižších úrovních, než jsou již zmíněné dvě hlavní skupiny. V rámci větve zahrnující vzorky 1, 63, 64 je umístěn vzorek 57, což je rostlina z Pienin. Zvláštní pak zde je především pozice vzorků 5 a 13, z nichž první reprezentuje rostlinu hluboce zaříznutých údolí řek, druhý pak je typickou *C. vulgaris*. Oba jsou v rámci fylogenetického stromu i sítě uvnitř shluku *C. biebersteinii* subsp. *biebersteinii*. Další nezvyklou pozici představuje vzorek 77 (přechodný typ), který je na bázi druhé skupiny, navíc oddělen dlouhou větví vypovídající o značné genetické vzdálenosti. Pak je zde diskutabilní i pozice vzorku 4, což jsou horské rostliny z Alp spadající v ostatních stromech jednoznačně do skupiny *C. biebersteinii* subsp. *biebersteinii* (i když v rámci fylogenetického stromu z dat exonů s vyfiltrovanými heterozygotními lokusy a obsahujícími kódy pro ambiguitní báze (obr. 20a) je v bazální pozici). Naproti tomu, v případě vzorku 66, který byl problematický v stromech konstruovaných na základě cílených exonů, spadá tento ve fylogenetickém stromě i síti z rDNA dat do skupiny ostatních vzorků *C. vulgaris* s. str. a *C. biebersteinii* subsp. *brevibracteata* (v jedné větvi se vzorky 58 a 37, tedy morfologicky podobným rostlinám z území ČR).

Jak je již výše řečeno, na základě získaných výsledků nelze spolehlivě vyvozovat přesné vztahy mezi jednotlivými populacemi. Na grafických výstupech jsou barevně odlišeny pracovní skupiny (viz tab. 1), v rámci kterých by mohla být předpokládána příbuznost na základě morfologické a ekologické podobnosti takových populací. V rámci získaných fylogenetických stromů a sítí však tato skutečnost není patrná. Ač v některých případech lze pozorovat, že morfologicko-ekologicky podobné typy sdílí společné větve stromu, těžko lze

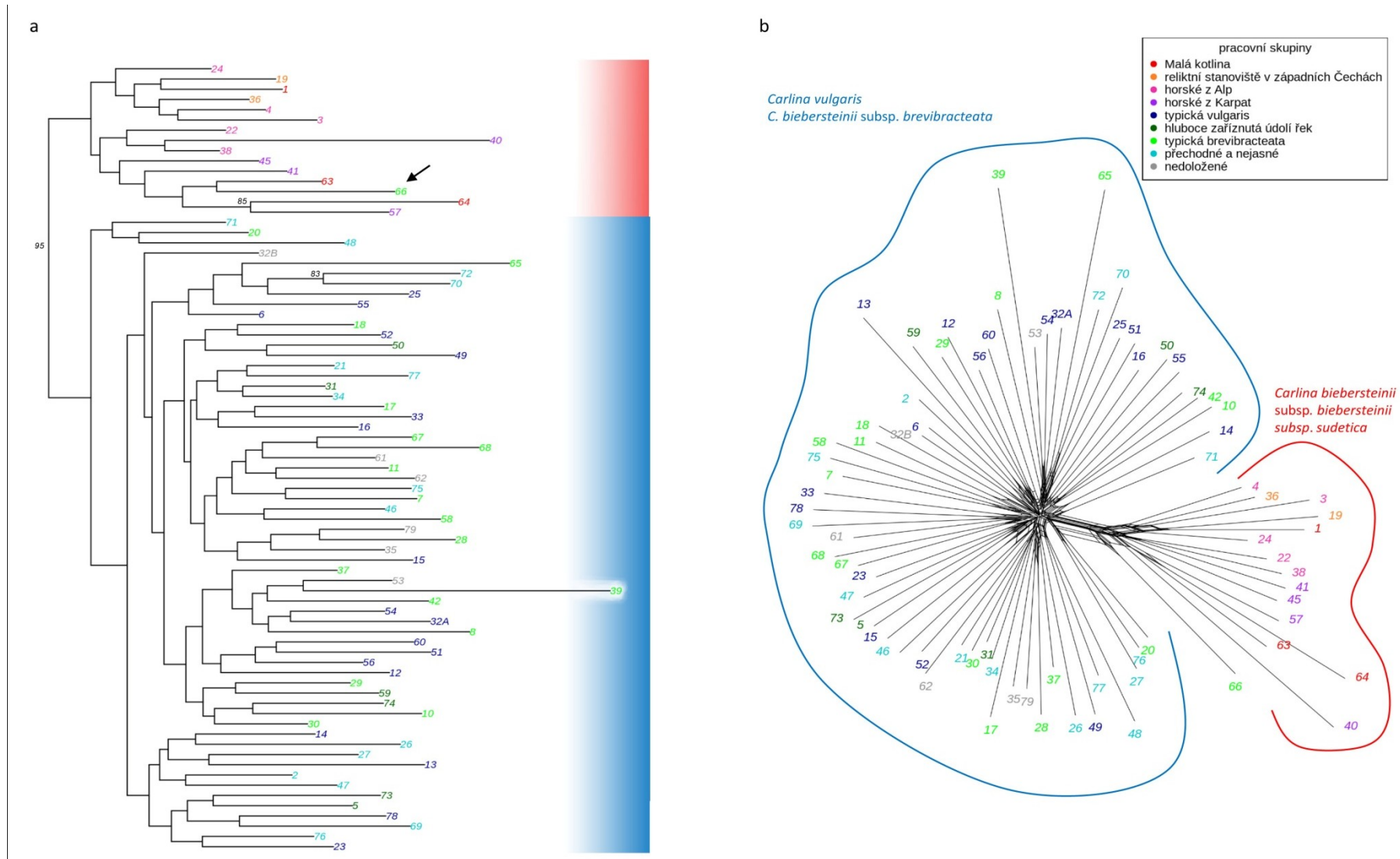
taková data interpretovat jako vzájemnou příbuznost. Výjimku snad tvoří skupina alpských vzorků *C. biebersteinii* s. str. (vzorky 38, 24, 22) v rámci některých z fylogenetických sítí (na obr. 19b a 21a). Jinak je přítomnost takových skupin ve fylogenetických stromech a sítích pravděpodobně artefaktem, např. příbuznost vzorků 60, 51, 56 a 12 na obr. 19 je již spíše náhodná, ač se jedná o rostliny s podobným morfotypem, pocházejí však z geograficky odlehlých lokalit (vápencový lom v Bavorsku, italské Alpy, Pieniny a podhůří Jeseníků). Obdobně v dalším ze stromů vychází v rámci jedné větve již zmíněný vzorek 51 v blízkosti vzorků 52, 54, 15, 13 a 6, které reprezentují zástupce ze Skandinávie, Českého krasu, Mělnicka a Příbramska.

V rámci stromu rDNA (obr. 22) lze pozorovat takových skupin několik, jednak větve reprezentující morfotyp *Carlina biebersteinii* subsp. *brevibracteata* (vzorky 68, 39, 11, 20), jednak takové odpovídající morfologicky *C. vulgaris* subsp. *vulgaris* (vzorky 15, 78, 25, 32A, (případně se zahrnutím přechodného typu 69), 23 a další skupinu tvoří vzorky 6, 54, 60, 46). Opět však nelze usuzovat, že by zde byly podchyceny reálné příbuzenecké vztahy, neboť tyto vzorky pocházejí ze značně vzdálených oblastí. Vzorek 68 pochází z Ukrajiny, 39 z Hodonínska a vzorky 11 a 20 z Královehradecka. V další zmíněné skupině pak jsou vzorky z Českého krasu (15), Plzeňska (78), Rakouska (spolková země Salzburg, 25) či severních Čech (32A). Obdobná situace je i v druhé takové skupině, která pojí dohromady vzorky pocházející ze středních Čech (6, 46), Dánska (54) a Německa (60).

Navíc jsou vzájemné pozice ve fylogenetických stromech odvozených na základě různých přístupů značně rozdílné (a s velmi nízkými podporami bootstrapu). Málokdy se objevují v rámci stromů stejné topologie. Nejodlišnější je od všech ostatních právě fylogenetický strom odvozený z rDNA, jeho porovnání s jedním z předchozích stromů ukazuje obr. 23.

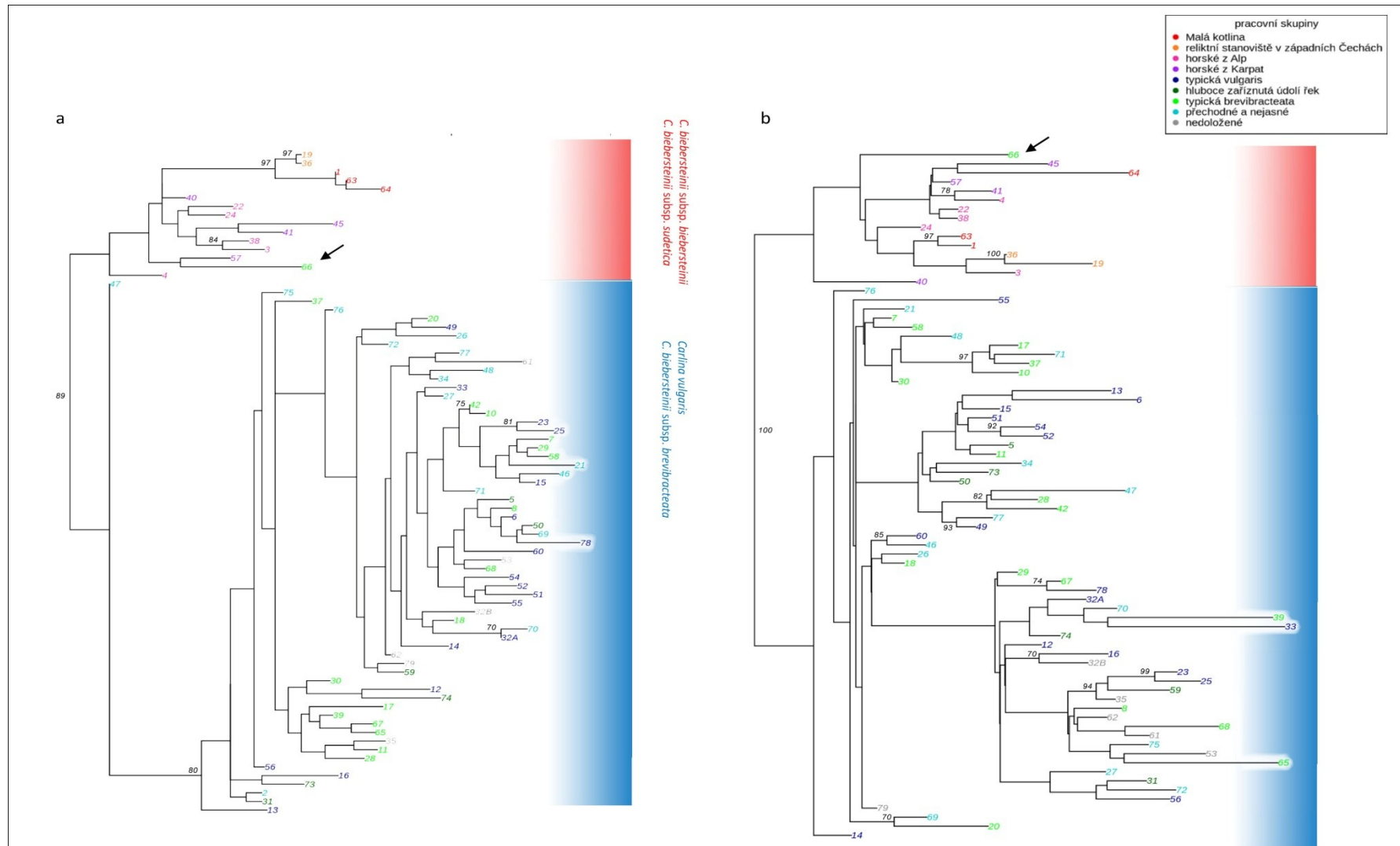


Obr. 18 a) Fylogenetický strom získaný programem RAxML (v příslušných nodech jsou vyznačeny bootstrapové podpory 70 a vyšší) na základě analýzy konkaténovaného souboru, v kterém pouze byla odfiltrována chybějící data (majority consensus tree). b) Síť odvozená ze stejných dat programem SplitsTree. Barvy v označení vzorků odpovídají přiřazení do pracovních skupin (viz tab. 1). Barevně jsou vyznačeny i dvě základní skupiny: *C. vulgaris* subsp. *vulgaris* a *C. biebersteinii* subsp. *brevibracteata* jsou označeny modře a *C. biebersteinii* subsp. *biebersteinii* včetně subsp. *sudetica* označena červeně. Šipkou je označena pozice kontroverzního vzorku 66.



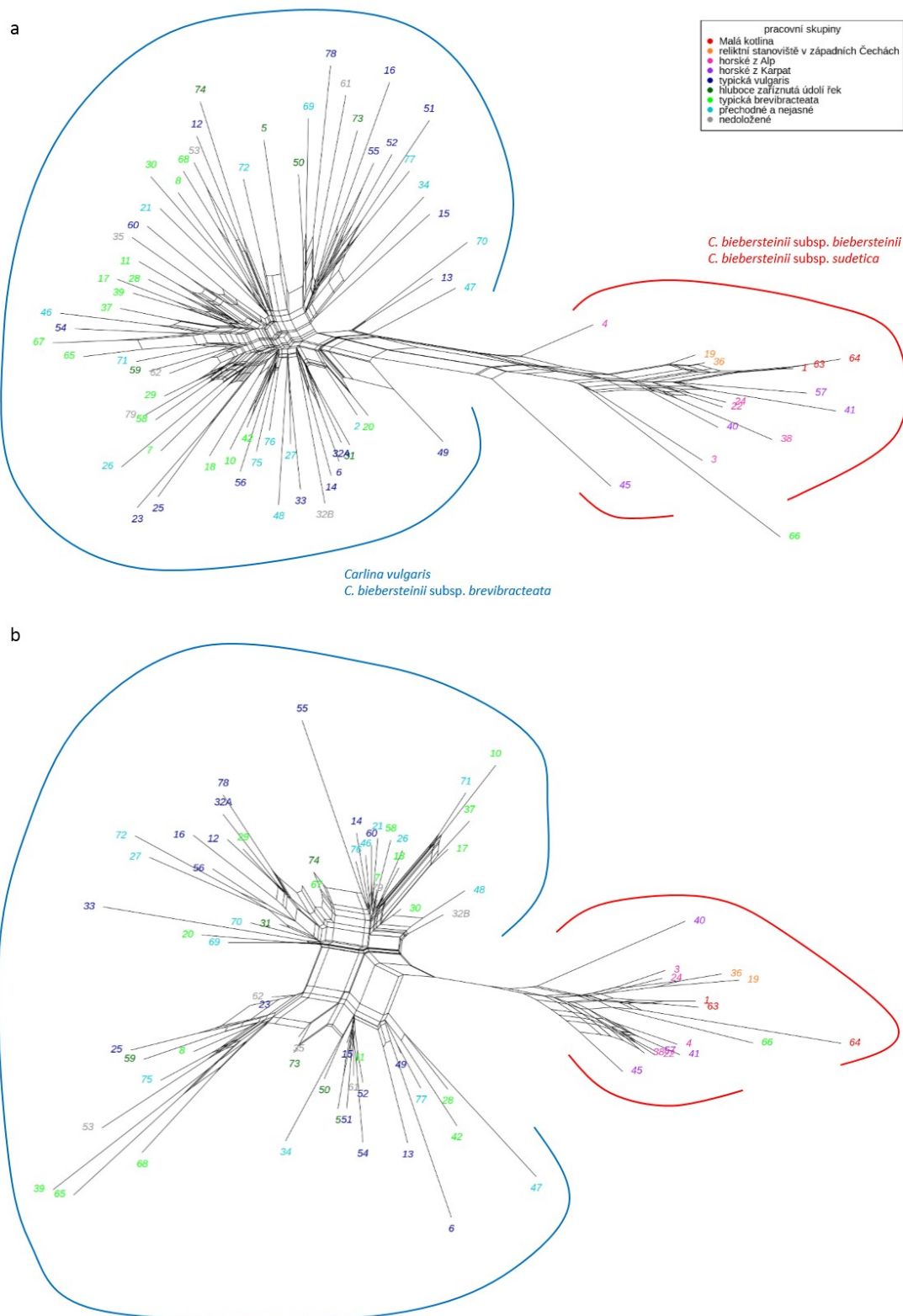
Obr. 19 a) fylogenetický strom získaný programem RAxML (v příslušných nodech jsou vyznačeny bootstrapové podpory 70 a vyšší) na základě analýzy konkatenaného souboru, v kterém byly vyfiltrovane lokusy (omezen počet heterozygotních míst). b) Síť odvozená ze stejných dat programem SplitsTree. Barvy v označení vzorků odpovídají přiřazení do pracovních skupin (viz tab. 1). Barevně jsou vyznačeny i dvě základní skupiny: *C. vulgaris* subsp. *vulgaris* a *C. biebersteinii* subsp. *brevibracteata* jsou označeny modře a *C. biebersteinii* subsp. *biebersteinii* včetně subsp. *sudetica* označena červeně. Šipkou je označena pozice kontroverzního vzorku 66.



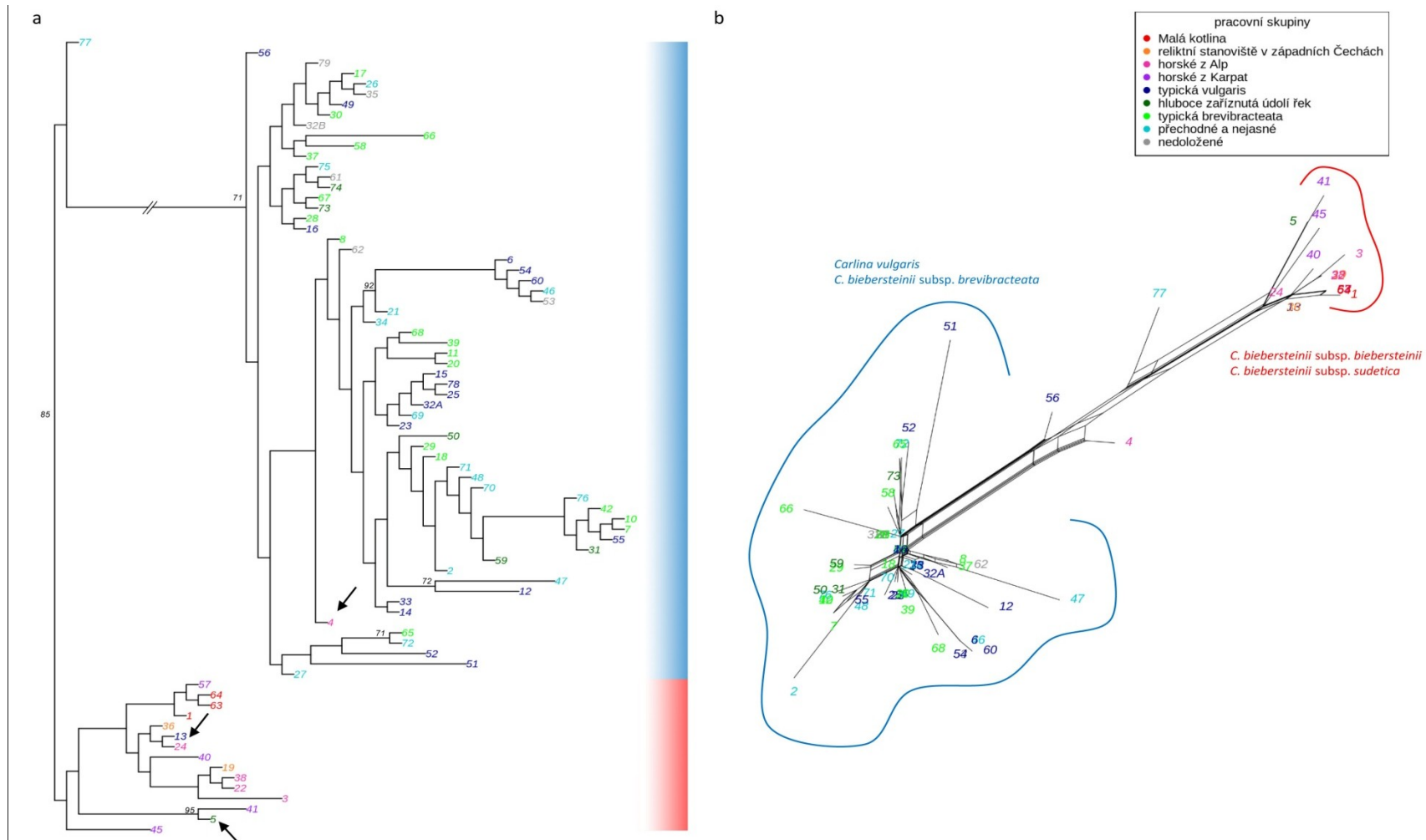


Obr. 20 Fylogenetické stromy získané programem RAxML (v příslušných nodech jsou vyznačeny bootstrapové podpory 70 a vyšší). a) na základě analýzy supercontigů z programu HybPiper, b) konkatenovaného souboru, v kterém byly vyfiltrovane lokusy (omezen počet heterozygotních míst) a zahrnuty ambigitní báze. Barvy v označení vzorků odpovídají přiřazení do pracovních skupin (viz tab. 1). Barevně jsou vyznačeny i dvě základní skupiny: *C. vulgaris* subsp. *vulgaris* a *C. biebersteinii* subsp. *brevibracteata* jsou označeny modře a *C. biebersteinii* subsp. *biebersteinii* včetně subsp. *sudetica* označena červeně. Šipkou je označena pozice kontroverzního vzorku 66.

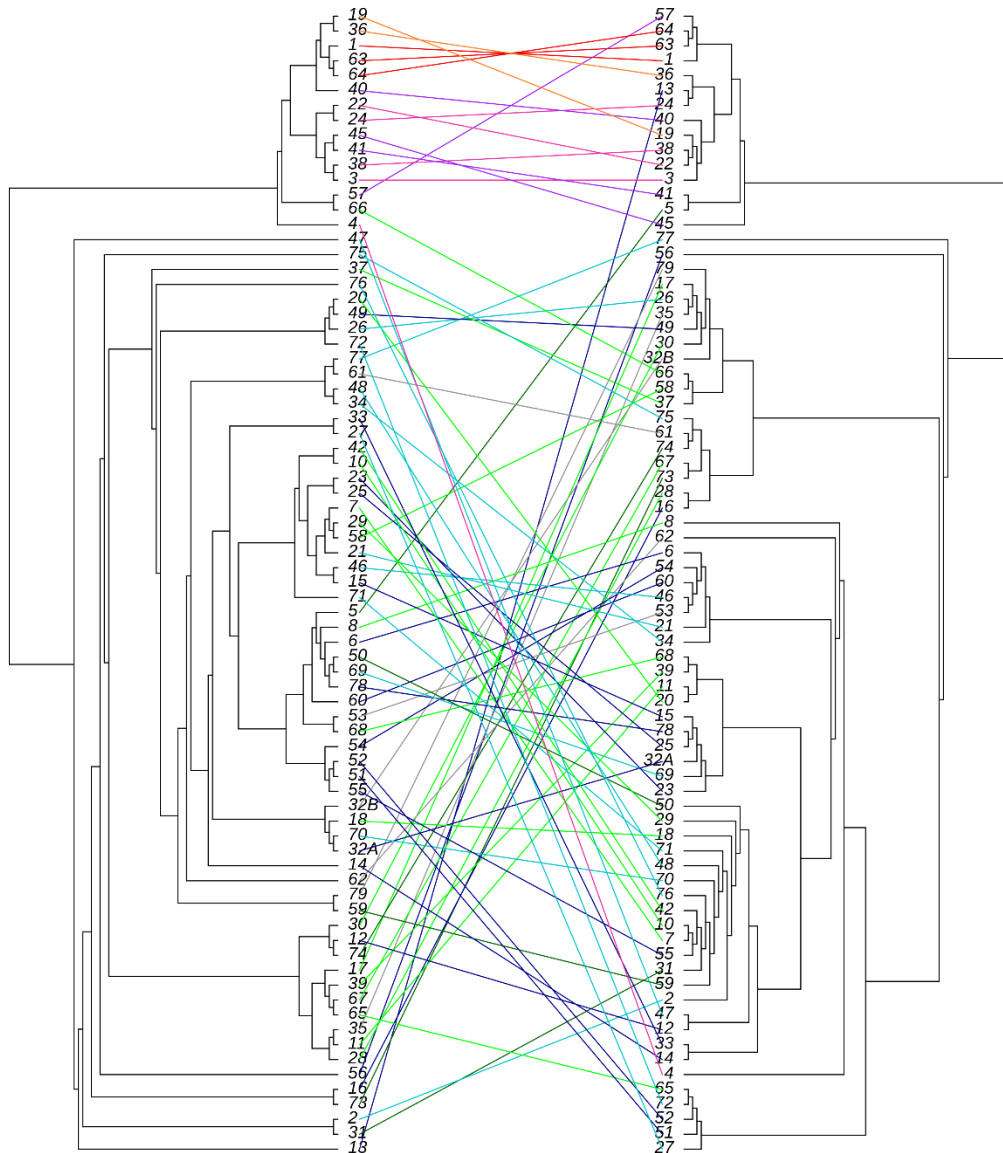




Obr. 21 Fylogenetické sítě získané programem SplitsTree a) na základě analýzy konkatenovaného souboru, v kterém byly vyfiltrovány lokusy (omezen počet heterozygotních míst) a zahrnuty ambiguitní báze, b) na základě analýzy superkontigů (z programu HybPiper). Barvy v označení vzorků odpovídají přiřazení do pracovních skupin (viz tab. 1). Barevně jsou vyznačeny i dvě základní skupiny: *C. vulgaris* subsp. *vulgaris* a *C. biebersteinii* subsp. *brevibracteata* jsou označeny modře a *C. biebersteinii* subsp. *biebersteinii* včetně subsp. *sudetica* označena červeně.



Obr. 22 a) fylogenetický strom získaný programem RAxML (v příslušných nodech jsou vyznačeny bootstrapové podpory 70 a vyšší) na základě analýzy rDNA. b) Síť odvozená ze stejných dat programem SplitsTree. Barvy v označení vzorků odpovídají přiřazení do pracovních skupin (viz tab. 1). Patrné jsou dvě základní skupiny: *C. vulgaris* subsp. *vulgaris* a *C. biebersteinii* subsp. *brevibracteata* jsou označeny modře a *C. biebersteinii* subsp. *biebersteinii* včetně subsp. *sudetica* označena červeně. Šipky označují pozice vzorků 5, 13 a 4, které nekorespondují v rámci pozorovaných skupin s předchozími stromy.



Obr. 23 Porovnání topologií fylogenetických stromů (RAxML) získaných na základě cílených genů z jaderné DNA (vlevo, pro bootstrapové podpory viz obr. 20 a) a rDNA (obr. 22 a). Barvy čar odpovídají přiřazení vzorků do pracovních skupin (viz tab. 1).

### 4.3. Morfometrické analýzy

#### 4.3.1. Test správnosti odečtu znaku počet ostnitých zubů

Při porovnání datových souborů počtů ostnitých zubů prostřednictvím vizuálního odečtu a s použitím programu ToothFinder se ukázalo, že při prvním ze zmíněných přístupů bylo detekováno ostnitých zubů více. U dolních listů, které jsou nejdelší, dosahoval jejich počet až 410 (při střední hodnotě 232, nejnižší hodnota byla 116). Naproti tomu algoritmus našel v případech dolních listů nejvýše 325 ostnitých zubů. Porovnání počtu ostnitých zubů získaných vizuálním hodnocením a rozlišených programem ToothFinder pro jednotlivé kategorie měřených listů je na obr. 24.

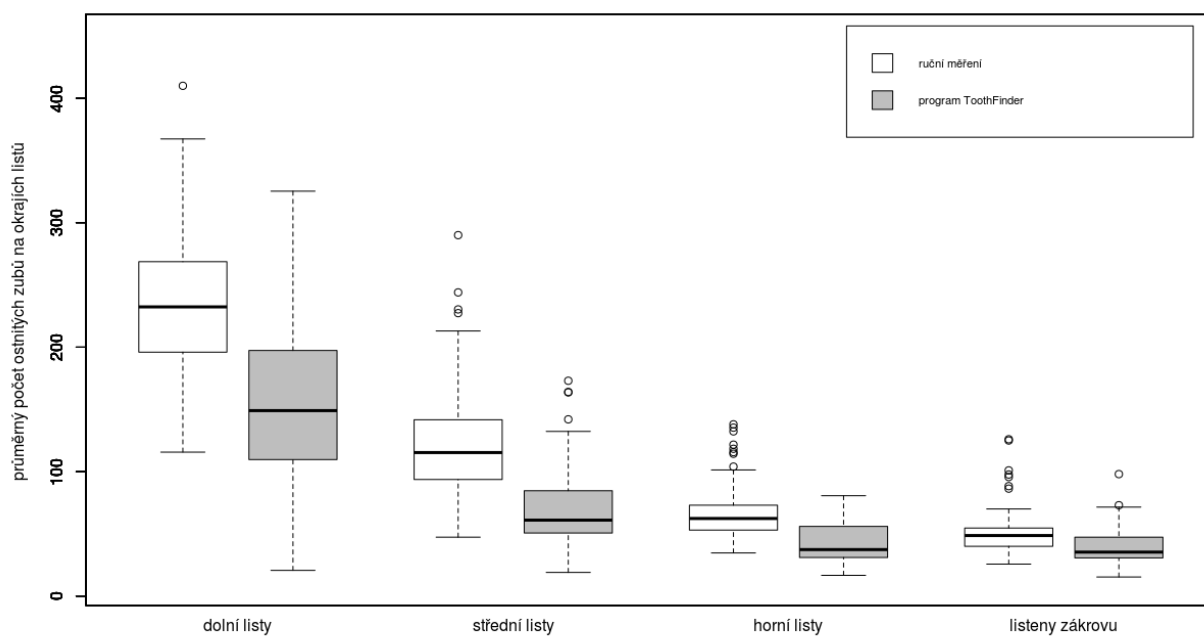
Nižší hodnoty zjištěné programem ToothFinder jsou zde dány arbitrárním nastavením toho, jak velký výstupek je označován jako zub. Stejně tak je velikost výstupku detekovaného jako zub omezena shora, což může vést k podhodnocení v některých jiných případech.

Každopádně jsou tak všechny listy hodnoceny na základě stejných kritérií (odpadá subjektivní rozhodování ohledně toho, jak velké výstupky v rámci naskenovaného listu počítat jako zub). Na druhou stranu však je výsledek ovlivněn chybami. Algoritmus není schopen rozeznat osténkaté zuby, které během lisování byly přiklopeny na čepel listu. Zároveň v případě přeložených, či jinak zdeformovaných listů představuje nesmyslné výsledky (nejnižší hodnota pro spodní listy byla 20 nalezených zubů). Trend v rámci zkušebního datasetu je však jasný. Počty zubů spolu signifikantně korelují ( $r = 0.839$ ,  $p$ -hodnota  $< 2.2 \cdot 10^{-16}$ ) (viz též obr. 25) podobný výsledek, jaký představili I tvůrci programu, v jejich případě šlo o porovnání počtu naměřených počtů zubů na 50 listech rodu *Tillia*,  $r = 0.873$ ,  $p$ -hodnota  $< 2.2 \cdot 10^{-16}$ ). Počty zubů na všech měřených listech pro následné analýzy proto byly zjišťovány za pomoci programu ToothFinder, tedy jednotnou metodikou s jasnou definicí toho, jak velké objekty v rámci okraje listu jsou do výsledného počtu zahrnuty. Následující analýzy nebyly proto ovlivněny subjektivním rozhodnutím v tomto směru.

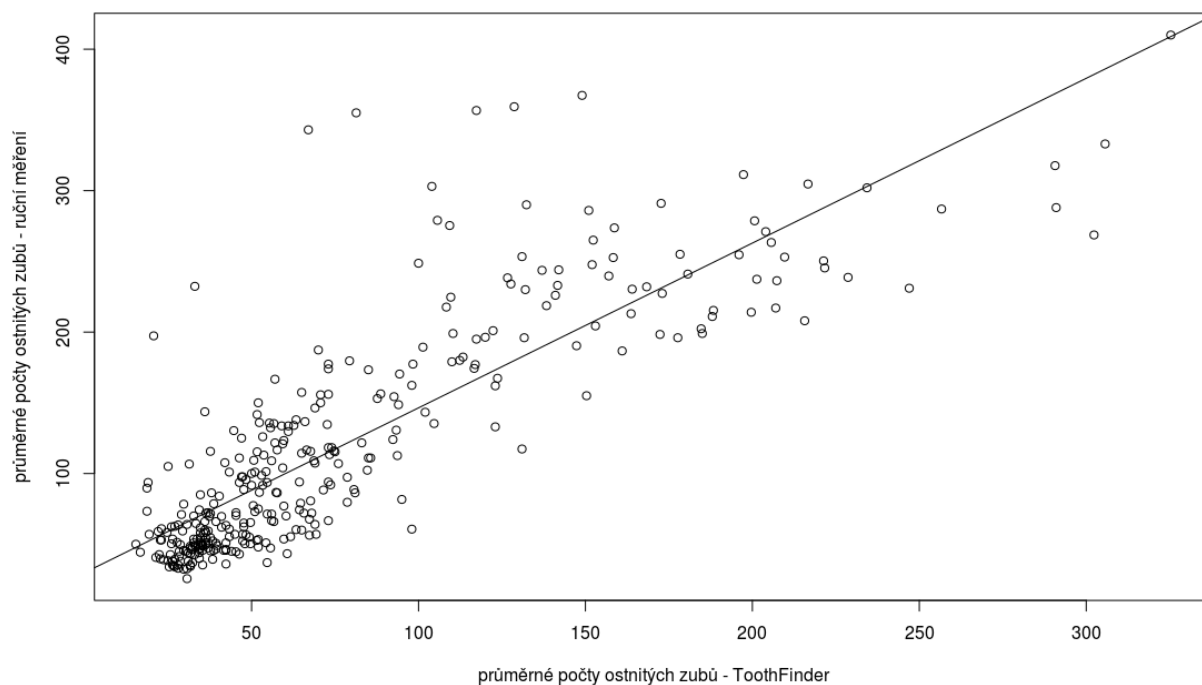
#### 4.3.2. Vliv rozdílných podmínek na fenotyp rostlin

Za účelem posouzení vlivu rozdílných podmínek na fenotyp rostlin byla provedena analýza hlavních komponent (PCA), která ukázala na posun v ordinačním prostoru daný rozdílným původem měřených rostlin (centroid pro rostliny z kultivace má souřadnice ( $x = 2.76$ ,  $y = -0.96$ ), z terénu ( $x = -0.99$ ,  $y = 0.34$ ), viz obr. 26). Posun v rámci ordinačního prostoru lze tedy vyjádřit vektorem, který s osou  $x$  svírá úhel  $-19.15^\circ$ . Je tedy nejbližší vektoru znaku poměr výšek "pomvys" ( $-19.86^\circ$ ). Pěstování v zahradě, kromě tohoto znaku, má vliv i na další, s tímto korelované, jako je průměr terminálního úboru (terub), délka listů z dolní části lodyhy (deldol) a ze střední části lodyhy (delpro), či na opačně orientovaný znak ostnitost horních listů (osthor), znázornění vektorů znaků viz na obr. 27.

Tyto výsledky lze interpretovat tak, že zmíněné znaky jsou plastičtější než ty znaky, jejichž směrový vektor nekoresponduje s posunem, který byl dán změnami podmínek (přesazení do experimentální zahrady).



Obr. 24 Krabicové diagramy pro porovnání počtu ostnitých zubů získaných ručním měřením (bílá) a programem ToothFinder (šedá).



Obr. 25 Porovnání datových souborů počtu ostnitých zubů získaných pomocí ručního odečtu hodnot (osa y) a za pomoci programu ToothFinder (osa x), jedná se o 340 průměrných hodnot vždy pro 3 měřené listy/listeny. (Proloženy přímkou za pomoci metody nejmenších čtverců, intercept 29.93, odlehle hodnoty vypovídají o chybách daných nepřesnostmi algoritmu.)



### 4.3.3. Morfologické rozdíly mezi geneticky definovanými skupinami

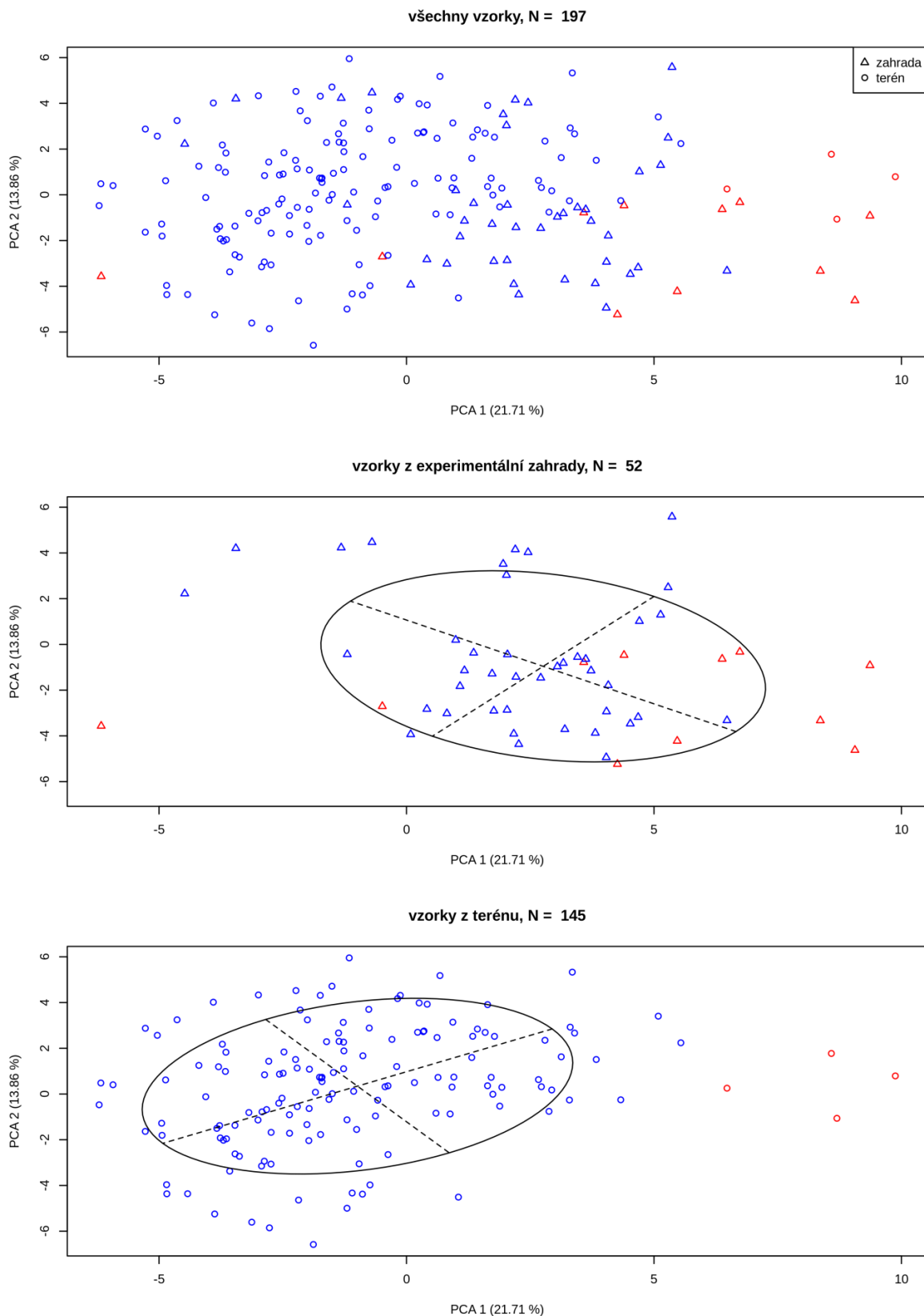
Jak již bylo uvedeno výše, byly porovnávány jednak skupiny definované na základě analýzy molekulárních dat, jednak na základě subjektivního přiřazení do skupin dle morfologických a ekologických charakteristik. Za tím účelem byl použit soubor dat zahrnující 401 jedinců s vybranými znaky (obsahoval vzorky z více populací).

PCA aplikovaná na úrovni jedinců ukazuje na morfologickou variabilitu v rámci populací, oblaky bodů v ordinačním diagramu se do jisté míry překrývají již při prezentaci pouhých dvou skupin, viz obr. 28) a procento vysvětlené variability prvními dvěma komponenty je 40,29 % pro vzorky z kultivace v experimentální zahradě a 41,35 % pro vzorky z terénu. Výstupy z PCA pro populační průměry ukazují strukturu v souboru dat lépe, představují jednak více vysvětlené variability první a druhou osou (50,62 % pro vzorky z experimentální zahrady, 54,24 % pro vzorky z terénu), jednak ostřeji ohraničené skupiny (viz obr. 29).

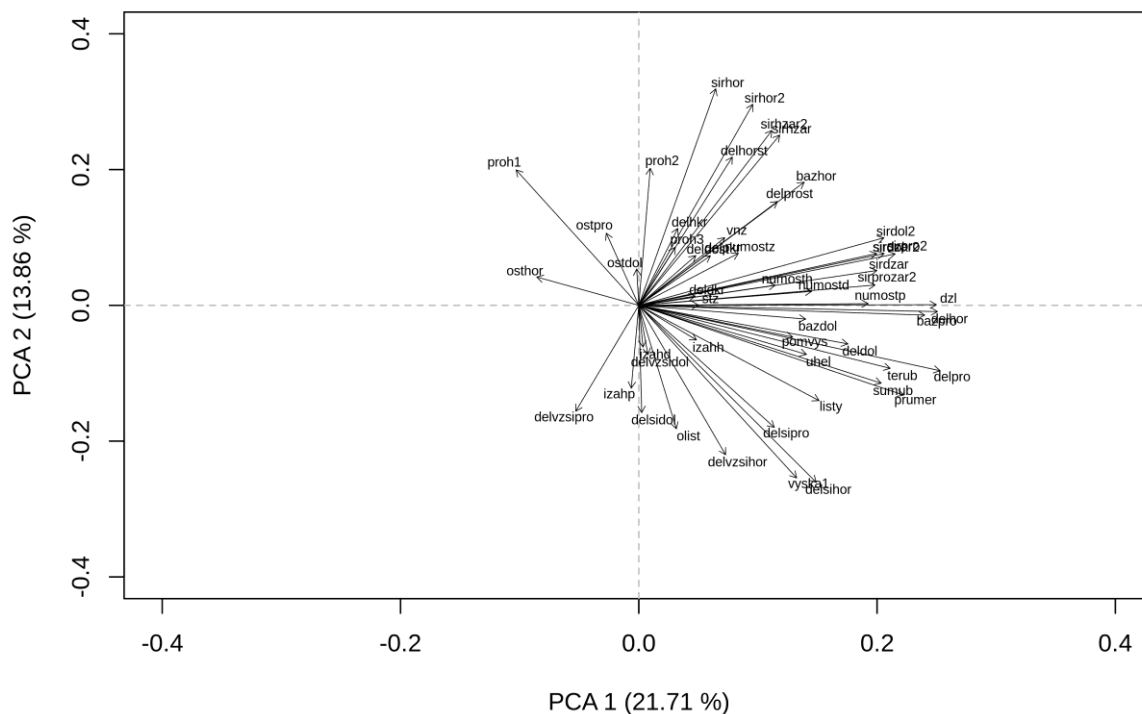
Přítomnost oddělených skupin v ordinačním prostoru PCA byla potvrzena prostřednictvím randomizačního testu (funkce `ade4::rtest`), dle kterého k celkové inerci (suma odchylek od náhodného vztahu proměnných a objektů) přispívá příslušnost do dané skupiny 11 % v případě jednotlivých vzorků (u všech podsouborů vyšel test obdobně s rozdílem v setinách procenta) a kolem 20 % v případě počítání s populačními průměry (18 % pro rostliny z experimentální zahrady, 21 % pro rostliny z terénu), s p-hodnotou 0,001.

Dále byla na data aplikovaná shlukovací analýza, opět na dva zmíněné podsoubory separátně (obr. 30 a 31). Ukazuje, obdobně jako předchozí PCA, překryv populací. V případě vzorků z uniformních podmínek experimentální zahrady představuje vzniklý dendrogram dvě odlišené skupiny. První zahrnuje vzorky pupavy Biebersteinovy pravé a několik jedinců z populací 18 a 11, v druhé pak jsou jedinci z ostatních populací a jedna rostlina z alpské populace 4 (zbylí zástupci této populace tvoří samostatný shluk v rámci první skupiny).

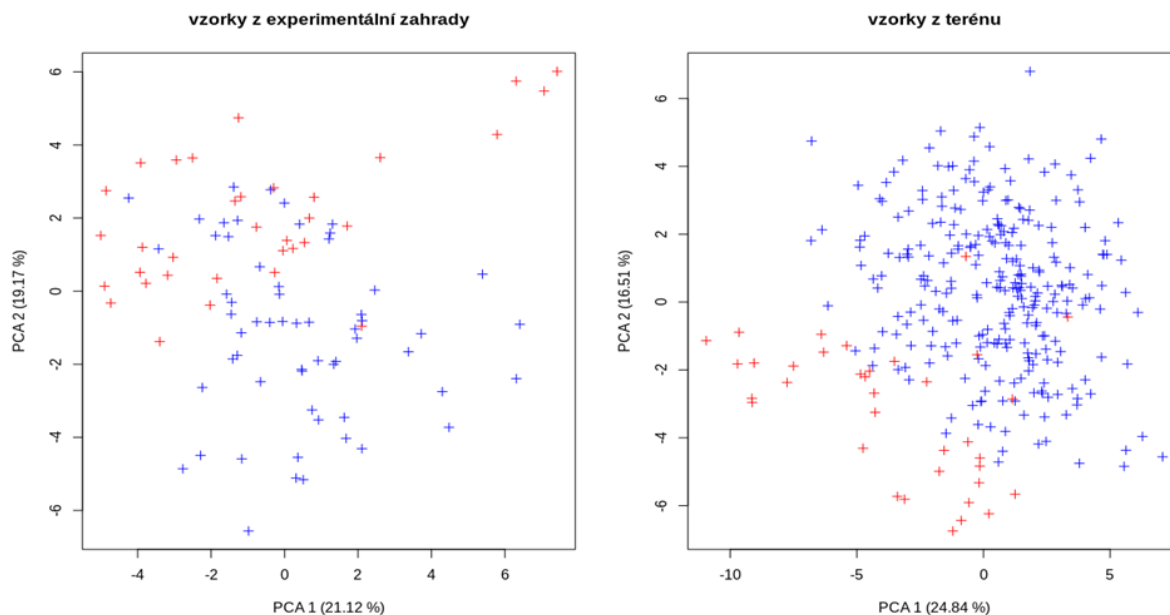
Výsledky shlukovací analýzy v případě vzorků z terénu jsou ovlivněny podmínkami prostředí, v kterých rostliny rostly. V rámci tohoto dendrogramu nejsou zástupci označení jako *Carlina biebersteinii* s. str. v rámci jedné skupiny, nýbrž ve dvou větvích, z nichž jedna sdružuje jedince z populací 22, 24 a několik z populace 4, druhou pak reprezentují jedinci z populace 3 a opět několik zástupců z populace 4. Další čtyři rostliny z této alpské populace 4 pak jsou na různých místech dendrogramu mezi ostatními zástupci *Carlina vulgaris* agg. Tyto výsledky vypovídají pravděpodobně o heterogenitě prostředí, která se v rámci různých lokalit vyskytovala.



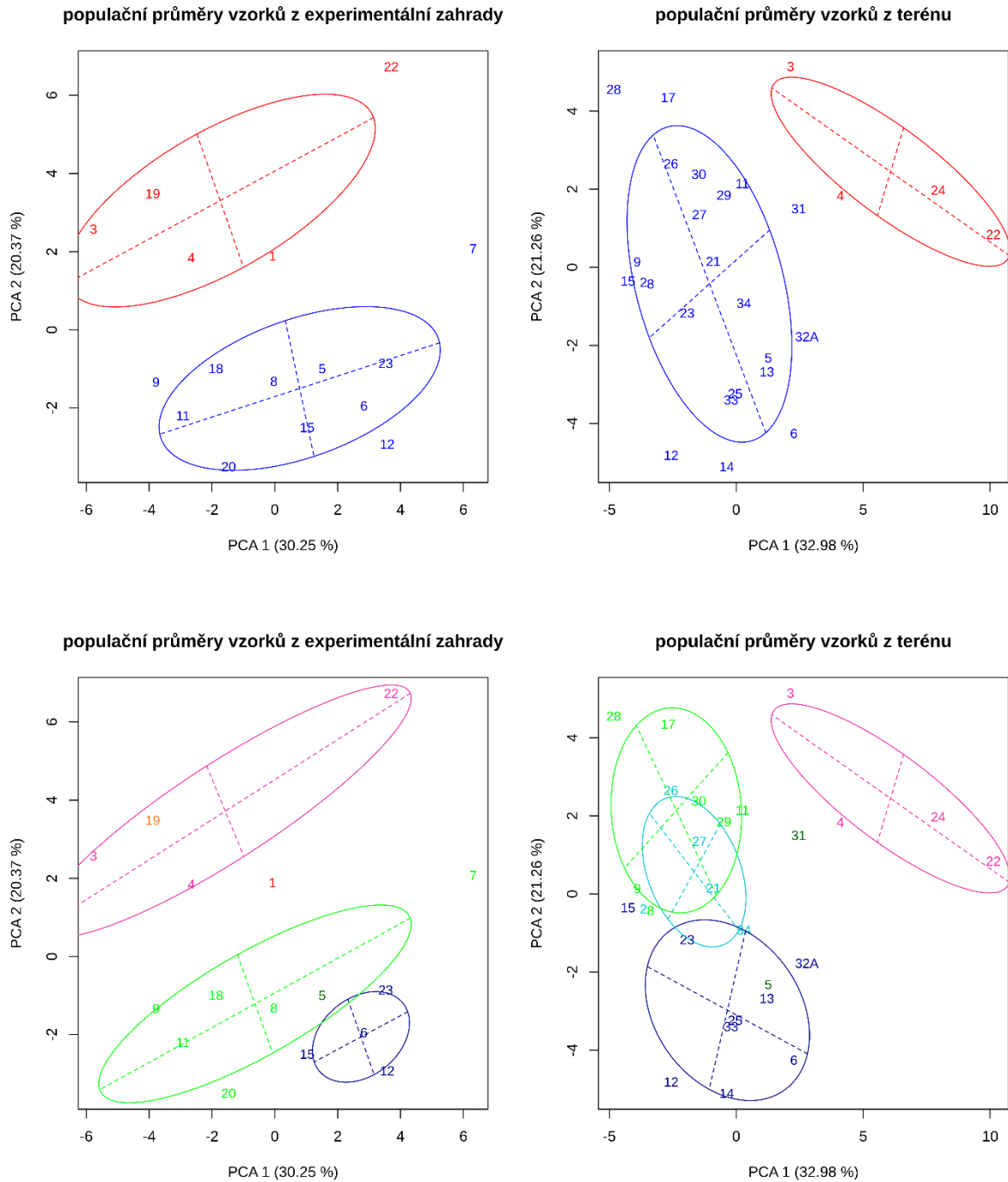
Obr. 26 Ordinační diagram PCA pro všechny měřené znaky (celkem 197 rostlin). Pro názornost je níže uveden tentýž graf jednou s promítnutím vzorků pouze z experimentální zahrady, podruhé pouze z terénu s vyznačenými centroidy. Elipsy setrvačnosti (angl. inertia ellipses) vymezují prostor kolem centroidu, kde se nachází 2/3 vzorků z příslušné skupiny. Modře skupina *Carlina vulgaris* a *C. biebersteinii* subsp. *brevibracteata*, červeně skupina *Carlina biebersteinii* subsp. *biebersteinii* včetně subsp. *sudetica*.



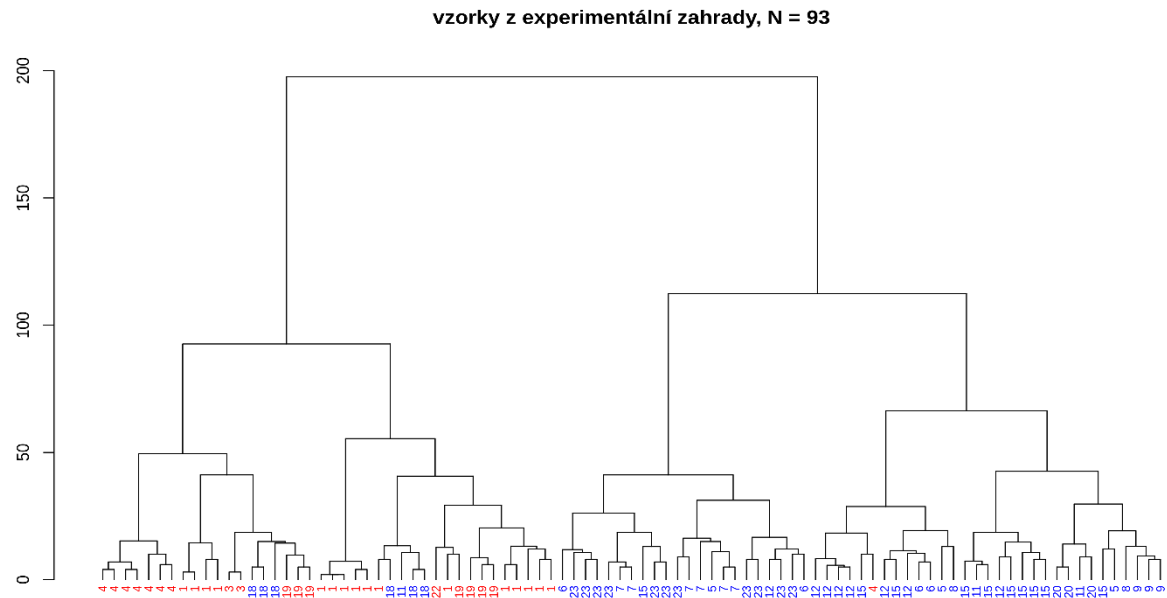
Obr. 27 Graf komponentních vah PCA pro všechny měřené znaky (na 197 rostlinách). Rozdíl mezi centroidy vzorků z kultivace v experimentální zahradě a z terénu (viz obr. 25 výše) lze vyjádřit vektorem odpovídajícím zhruba směru znaku „pomvys“, resp. opačnému směru vektoru znaku „osthor“ v protilehlém kvadrátu.



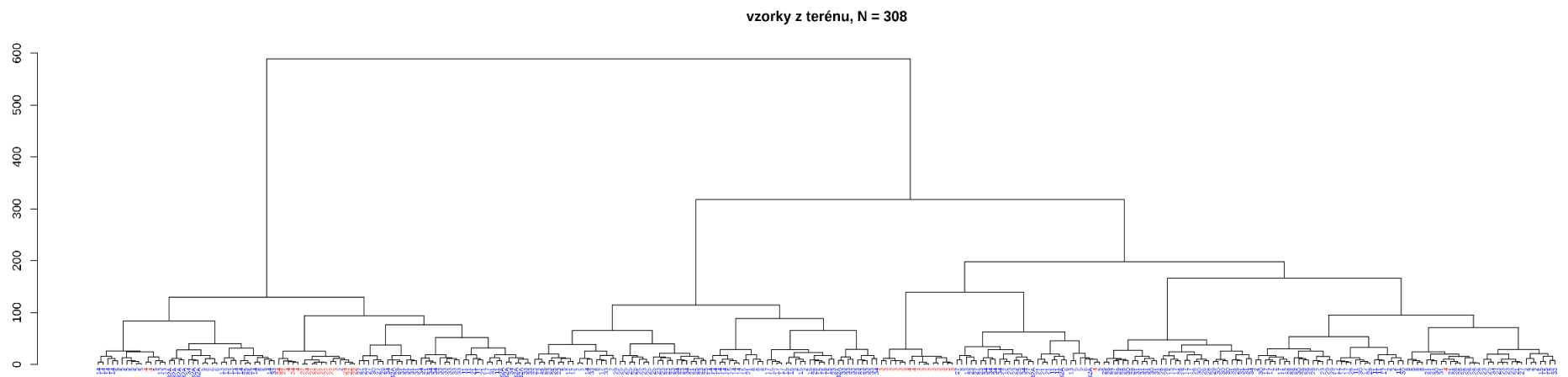
Obr. 28 Ordinační diagram PCA pro datový soubor vybraných znaků. Zleva: pro vzorky z experimentální zahrady (95 jedinců), pro vzorky z terénu (308 jedinců). Modře skupina *Carlina vulgaris* a *C. biebersteinii* subsp. *brevibracteata*, červeně skupina *Carlina biebersteinii* subsp. *biebersteinii* včetně subsp. *sudetica*.



Obr. 29 Ordinační diagram PCA pro datový soubor vybraných znaků, použity populační průměry. Zleva: pro vzorky z experimentální zahrady (15 populací), pro vzorky z terénu (27 populací). Nahoře vyznačená příslušnost ke geneticky definovaným skupinám (modře skupina *Carlina vulgaris* a *C. biebersteinii* subsp. *brevibracteata*, červeně skupina *Carlina biebersteinii* subsp. *biebersteinii* včetně subsp. *sudetica*). Dole tentýž graf znázorňující pracovní skupiny dle tab. 1.



Obr. 30 Dendrogram reprezentující výsledky shlukovací analýzy (použita Wardova metoda) pro vzorky pěstované a měřené v experimentální zahradě. Modře skupina *Carlina vulgaris* a *C. biebersteinii* subsp. *brevibracteata*, červeně skupina *Carlina biebersteinii* subsp. *biebersteinii* včetně subsp. *sudetica*.



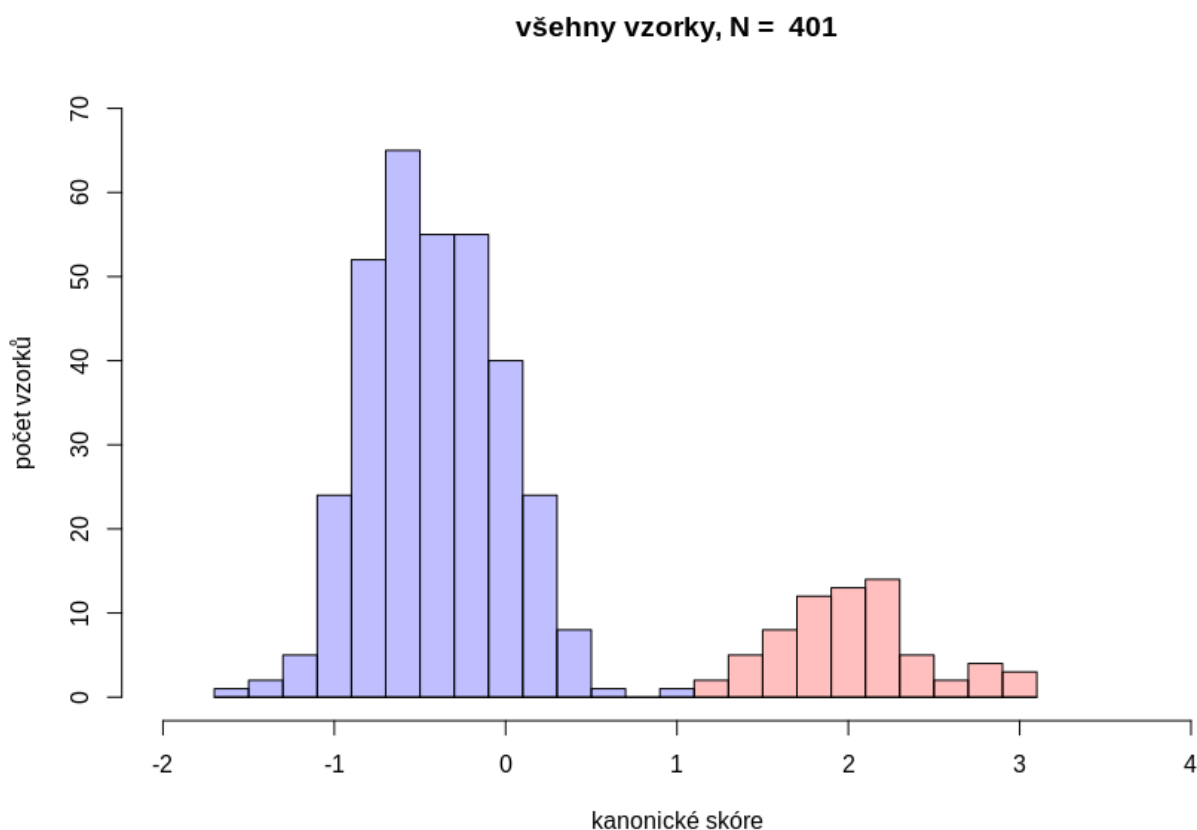
Obr. 31 Dendrogram reprezentující výsledky shlukovací analýzy (použita Wardova metoda) pro vzorky měřené v terénu. Modře skupina *Carlina vulgaris* a *C. biebersteinii* subsp. *brevibracteata*, červeně skupina *Carlina biebersteinii* subsp. *biebersteinii* včetně subsp. *sudetica*.



#### 4.3.4. Znaky vhodné k odlišení definovaných skupin

Za účelem nalezení vhodných znaků pro odlišení skupin byla provedena diskriminační analýza, opět separátně pro oba zmíněné datové soubory. Dvě definované skupiny prostřednictvím této analýzy byly distinktně odděleny ve všech případech. Obr. 32 reprezentuje bimodální rozdělení vzorků dle kanonického skóre z diskriminační analýzy pro dataset 401 vzorků (rostlin pěstovaných v experimentální zahradě i měřených v terénu). Ukazuje, že na základě měřených znaků dojde k rozdělení do dvou skupin i v případě, kdy jsou do souboru zahrnuty rostliny z různých stanovišť. Z toho vyplývá, že vliv různých podmínek na fenotyp rostlin tedy není tak výrazný, aby smýval rozdíly v měřených znacích mezi danými skupinami).

Celkové strukturní koeficienty představující míru korelace znaku a kanonické osy byly nejvyšší v případě znaku délka vnějších zákrovních listenů (dzl) v rámci všech testovaných souborů. Dále pak znaky vypovídající o rozměrech prostředních listů, a sice znak poměr délky a šířky listu ze střední části lodyhy v nejširším místě (delsipro), v případě datového souboru s omezeným počtem znaků (401 rostlin), a znak délka prostředního listu (delpro), v případě analýzy všech znaků (197 rostlin). Dalším znakem, který měl v případě všech provedených diskriminačních analýz nejvyšší hodnoty celkového strukturního koeficientu je



Obr. 32 Frekvenční histogram diskriminační analýzy provedené na datovém souboru vybraných znaků dohromady pro vzorky z uniformních podmínek experimentální zahrady i z terénu (401 rostlin). Modře skupina *Carlina vulgaris* a *C. biebersteinii* subsp. *brevibracteata*, červeně skupina *Carlina biebersteinii* subsp. *biebersteinii* včetně subsp. *sudetica*.

délka horního listu (delhor), jedná se však o znak významně korelovaný s již zmíněným znakem délka vnějších zákrovních listenů (dzl) (což vedlo k jeho vyřazení z analýzy v případě datového souboru rostlin pěstovaných v experimentální zahradě, v případě ostatních nebyl korelační koeficient vyšší než stanovená hodnota 0,95, ale blížil se k ní). Vzájemný blízký vztah těchto dvou znaků, stejně jako v případě některých dalších, které byly vzájemně značně korelované, ukazuje dendrogram na obr. 33.

V porovnání s procentuálním příspěvkem k vysvětlené variabilitě os PCA je zřejmé, že zmíněné znaky měly značný podíl na vysvětlené variabilitě i v této analýze, kde k nim přistupoval ještě znak vzdálenost mezi nejširší částí listu ze střední části lodyhy a jeho bázi (bazpro) v datových souborech pro vybrané znaky, a dále vzdálenost mezi nejširší částí listu z dolní části lodyhy a jeho bázi (deldol) a výška po vrchol terminálního úboru (vyska1) v případě vzorků pěstovaných na zahradě a všech použitých znaků. Přehledové tabulky, tab. VI a VII, s hodnotami pro jednotlivé znaky jsou uvedeny v příloze.

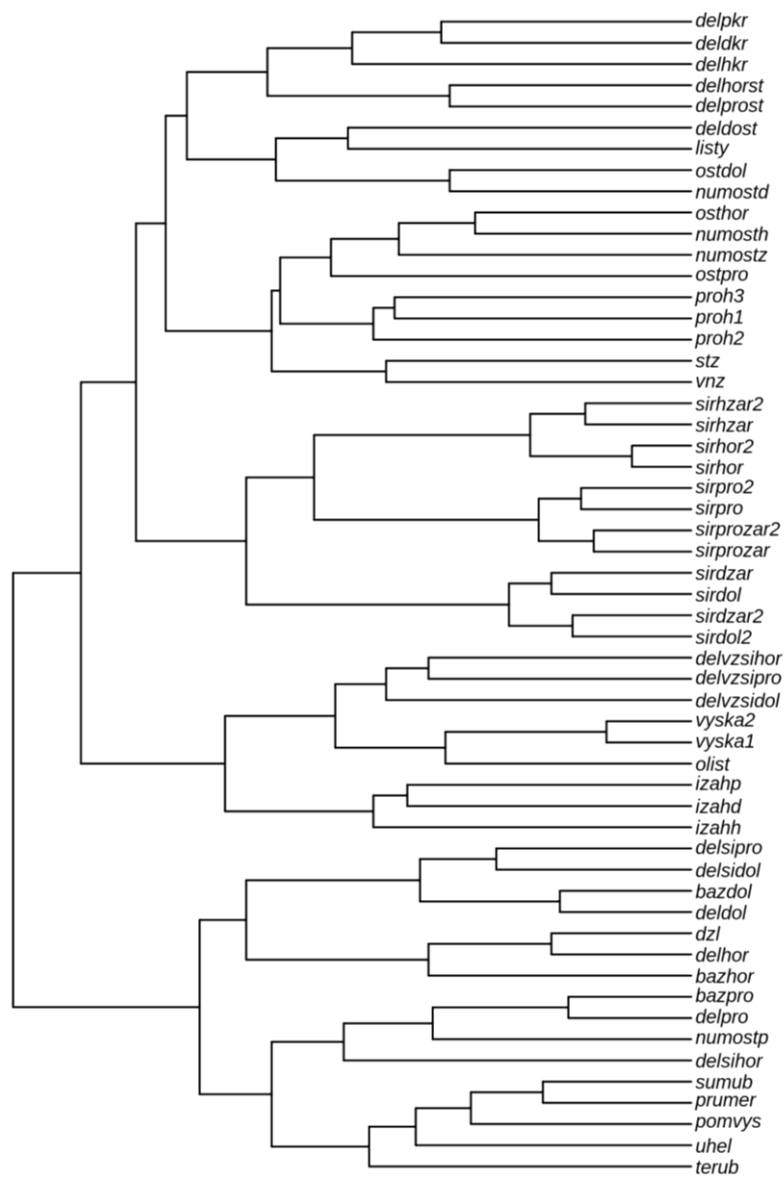
Dále bylo za účelem ověření vlivu zmíněných znaků na diskriminaci objektů do dvou skupin aplikováno testové kritérium Wilkovo lambda. Výsledky tohoto testu ukazují 13 znaků z původních 31, které signifikantně přispívají k rozdělení do skupin, výsledky testu viz v tab 4.

	znaky	Wilkovo lambda	F-test	p-hodnota	F-test (part)	p-hodnota (part)
1	dzl	0.45	479.02	2.49E-70	479.02	2.49E-70
2	delsipro	0.32	417.04	2.18E-98	161.90	< 2.2 E-160
3	bazdol	0.27	352.41	1.63E-111	72.76	3.33E-16
4	izahd	0.24	319.09	2.02E-122	60.55	6.26E-14
5	olist	0.22	276.47	1.44E-126	25.86	5.67E-07
6	sumub	0.21	244.10	2.36E-129	19.05	1.63E-05
7	prumer	0.20	227.23	6.77E-134	27.51	2.56E-07
8	bazpro	0.19	203.43	1.96E-134	8.09	4.68E-03
9	delvzsidol	0.19	183.98	1.28E-134	6.31	1.24E-02
10	terub	0.19	168.46	7.93E-135	6.31	1.24E-02
11	delvzsihor	0.19	155.10	9.55E-135	4.86	2.80E-02
12	bazhor	0.18	143.71	1.50E-134	4.23	4.03E-02
13	vyska1	0.18	133.57	4.19E-134	2.99	8.46E-02

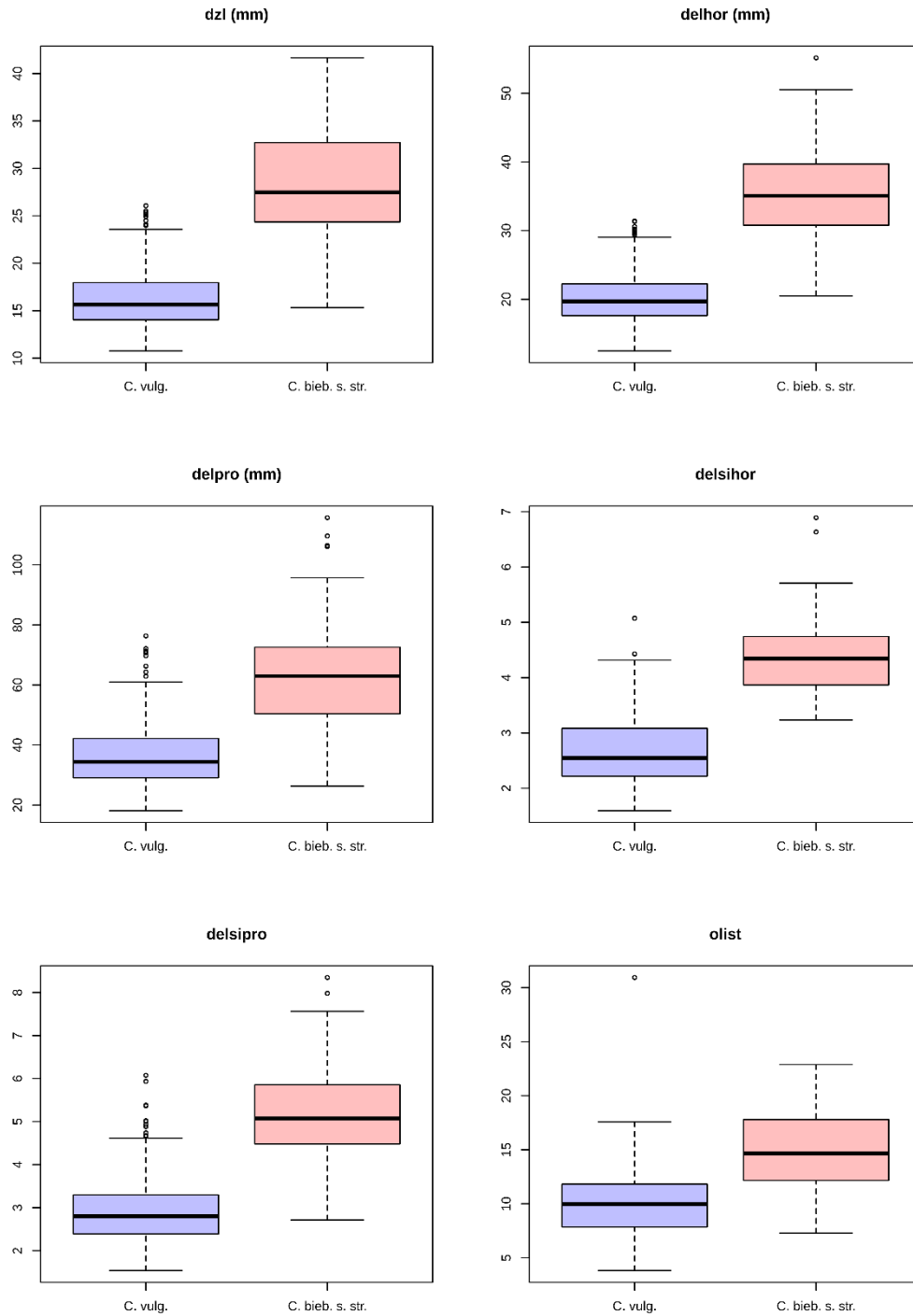
Tab. 4 Wilkovo kritérium lambda pro diskriminaci 401 objektů do 2 skupin. F-test = výsledek F-testu pro daný model, F-test (part) = výsledek F-testu pro unikátní příspěvek dané proměnné k diskriminaci, p-hodnota a p-hodnota (part) = statistická významnost F-testů.

Mezi relativně dobré diskriminátory byly tedy tímto testem vedle zmíněných znaků délka vnějších zákrovních listenů (dzl) a poměr délky a šířky listu ze střední části lodyhy v nejširším místě (delsipro) zařazeny též znaky vzdálenost mezi nejširší částí listu z dolní části lodyhy a jeho bázi (bazdol), index hloubky zářezu u listů z dolní části lodyhy (izahd) či délka internodií (olist).

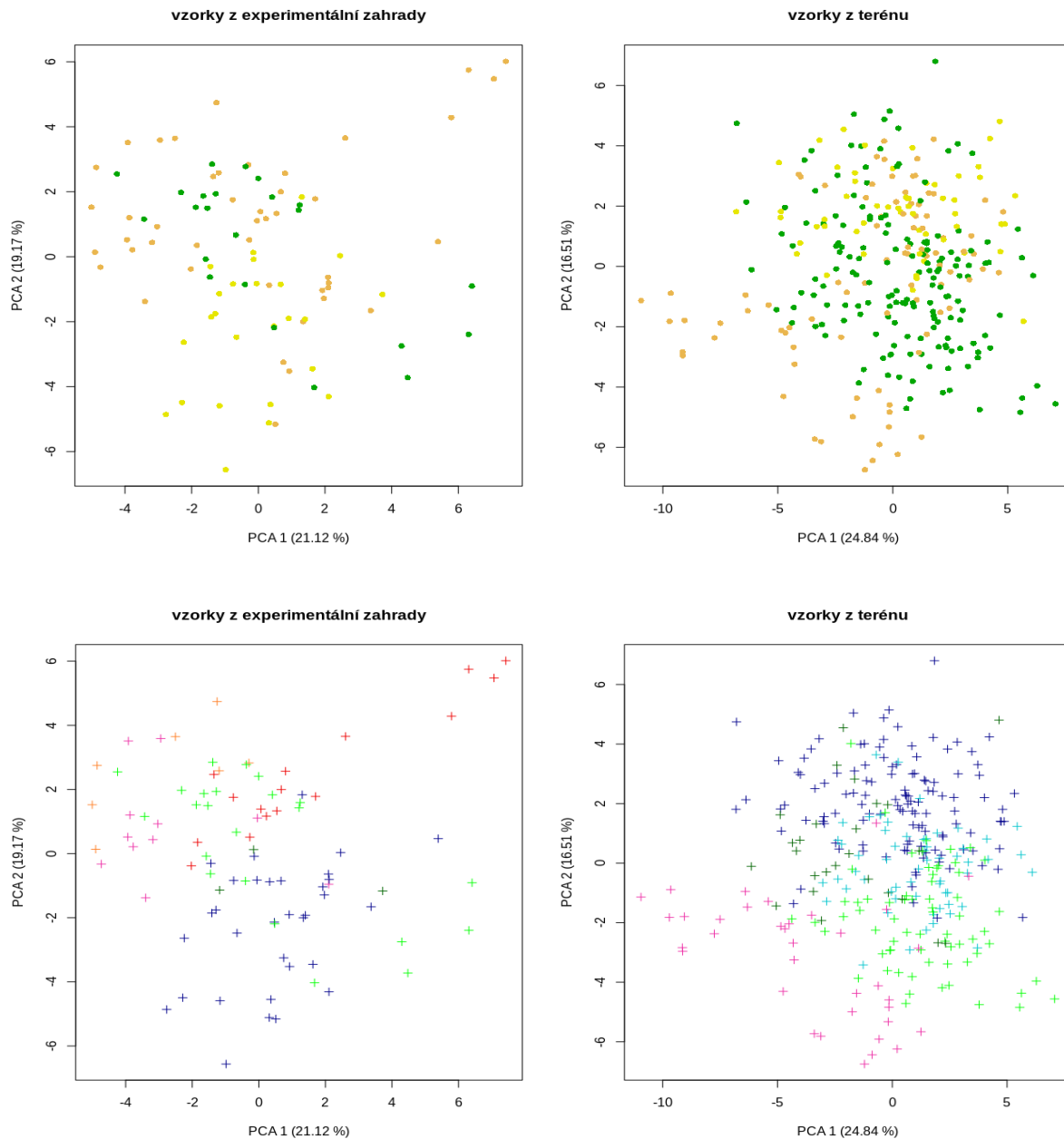
Účinnost klasifikačního kriteriá (souboru znaků) pro rozdělení do dvou skupin pak byla testována lineární diskriminační analýzou s křížovou validací, která jasně ukázala na rozdělení vzorků do dvou zmíněných skupin ve všech podsouborech. Úspěšnost diskriminační funkce pro přiřazení objektů do dvou zmíněných skupin se pohybovala mezi 93 a 100 %. Vzhledem ke skutečnosti, že data neměla mnohorozměrné normálové rozdělení (testováno pomocí balíčku MVN 5.5 (Korkmaz et al. 2014)), byla použita i kvadratická diskriminační analýza (bohužel v případě podsouborů nebyl pro danou analýzu dostatečně velký počet pozorování, proto byla aplikována pouze na soubor zahrnujícího všech 401 vzorků). Zde byla úspěšnost nižší v případě skupiny *Carlina biebersteinii* s. str., činila 76 % (chybně bylo klasifikováno 16 z 68 jedinců).



Obr. 33 Vizualizace vztahů mezi použitými znaky pomocí dendrogramu (výstup shlukovací analýzy (Wardova metoda) aplikované na transponovanou matici všech znaků a 197 objektů). Popis jednotlivých znaků viz v tab. III v příloze.



Obr. 34 Krabicové diagramy některých významných znaků (podle celkových strukturních koeficientů CDA, viz v tab. VI a VII v příloze, soupis souhrnných statistik pro všechny znaky viz v tab. IV tamtéž). Modře skupina *Carlina vulgaris* subsp. *vulgaris* a *C. biebersteinii* subsp. *brevibracteata*, červeně skupina *Carlina biebersteinii* subsp. *biebersteinii* včetně subsp. *sudetica*.



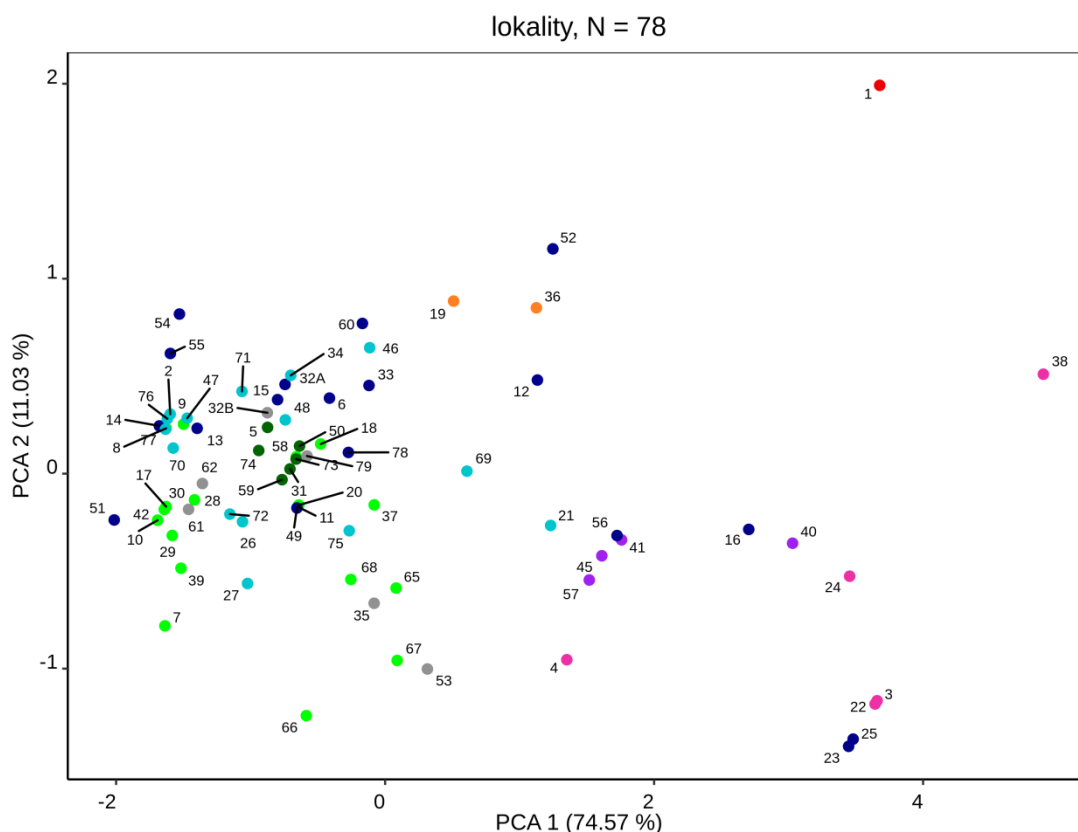
Obr. 35 Nahoře: PCA (výstup analýzy 401 vzorků) s vyznačením nadmořských výšek, z kterých rostliny pochází. Vzorky z populací z nižších poloh (do 350 m n. m.) označeny žlutou barvou, středních poloh (350 – 500 m n. m.) zelenou, a podhorských až vysokohorských (500 m n. m. a výše) béžovou. Dole: Pro srovnání stejný ordinační diagram s označením vzorků podle příslušnosti do pracovních skupin (jejich popis viz tab. 1).



#### 4.3.5. Ekologická diference

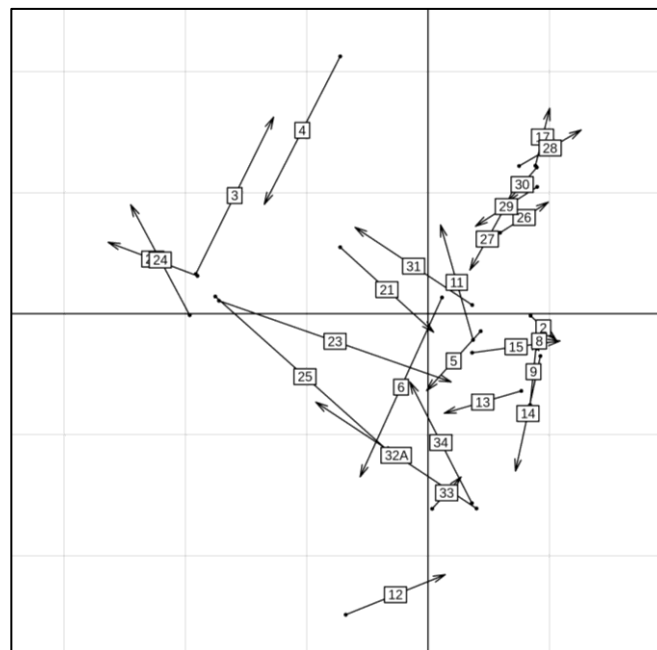
Na PCA diagramu (obr. 35) s vyznačením 3 arbitrárních skupin dle nadmořské výšky se její výrazný vliv neprojevil (v porovnání s příslušností k pracovním skupinám, která je na stejném grafu znázorněna níže). Lze vidět, že rostliny klasifikovatelné jako *C. biebersteinii* subsp. *brevibracteata* osidlují nadmořské výšky středních poloh (350 – 500 m n. m.), zatímco *C. biebersteinii* subsp. *biebersteinii* roste ve vyšších nadmořských výškách. Rostliny klasifikované jako *C. vulgaris* subsp. *vulgaris* pak mohou růst v celém rozsahu nadmořských výšek.

Dále byla pro srovnání morfologické a stanovištní variability použita klimatická data dostupná v rámci databáze WorldClim (Fick & Hijmans 2017). Graf PCA, který reprezentuje pozici sbíraných populací v ordinačním prostoru těchto několika proměnných prostředí, ukazuje podobný vzor, jako PCA výše ohledně populací *Carlina vulgaris* s. str., které v rámci ordinačního prostoru mají největší rozptyl. Specifickou pozici mimo ostatní skupiny zde zaujímá populace z Malé Kotliny (na obr. 36 označena číslem 1), další odlehlý bod představuje alpská populace 38. Jinak je zde patrná skupina odpovídající jedincům *Carlina biebersteinii* subsp. *brevibracteata* (světle zelené body) a s ní se překrývající skupina představující přechodné typy (světle modré body).



Obr. 36 Ordinační diagram PCA pro vybrané ekologické faktory (nadmořská výška, průměrné roční srážky, průměrná teplota v nejteplejším a nejchladnějším měsíci) pro všechny lokality sběru vzorků. Barevné označení odpovídá pracovním skupinám v tab. 1.

Ač jsou data převzatá z obecných modelů pouze nepříliš přesnou aproximací a reálné podmínky odráží pouze omezeně, byla za účelem porovnání vlivu prostředí na fenotyp rostlin testována podobnost distančních matic z těchto dat a z populačních průměrů morfometrických znaků. Za tím účelem byl použit Mantelův test (vegan::mantel), který ukazuje na relativně slabou korelaci (korelační koeficient  $r = 0,36$ ,  $p$ -hodnota 0.005). Obdobně RV koeficient (vícerozměrné zobecnění Pearsonova korelačního koeficientu) s hodnotou 0,34. Oba ukazují na vzájemný vztah mezi zmíněnými maticemi. Ten byl dále testován koinerční analýzou. Jejím výstupem je graf na obr. 37, který ukazuje, že největší rozdíly (zde reprezentované délkou šipek) mezi porovnávanými maticemi vzdáleností představují jednak alpské populace 3 a 4, jednak opět populace přiřazené ke *Carlina vulgaris* subsp. *vulgaris* (6, 23, 25, 32A), případně populace přechodná (34). Pozice populací 17, 28, 29, 30, které zde reprezentují *C. biebersteinii* subsp. *brevibracteata*, jsou si naopak v rámci obou matic podobné.



Obr. 37 normované skóre objektů v ordinačním prostoru koinerčních os. Začátky šipek reprezentují pozici v prostoru danou na základě první matice (tedy podmínek prostředí), konce pak pozici dle druhé matice (morfologické znaky). Délka šipek odpovídá shodě mezi projekcemi.

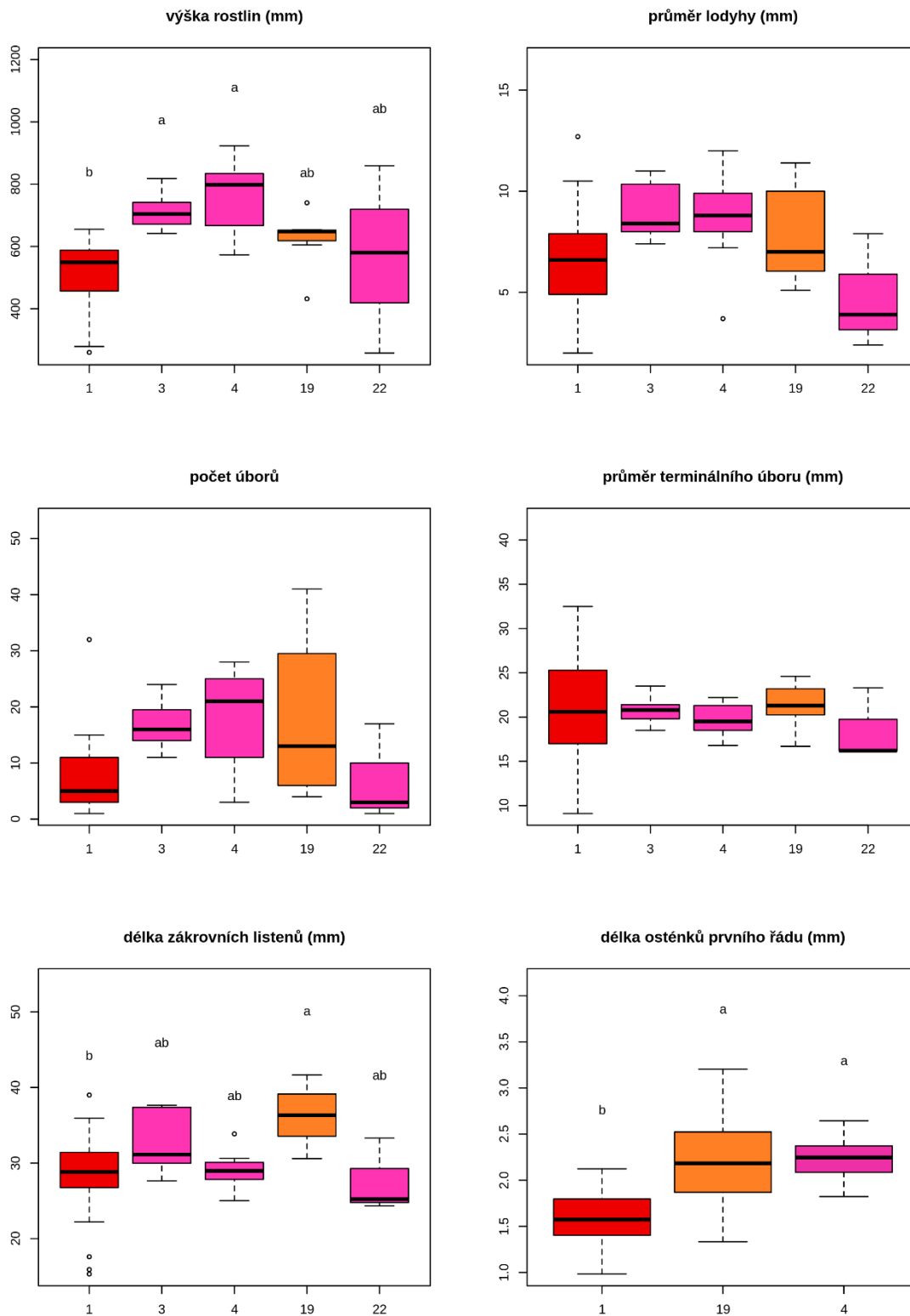
#### 4.3.6. Pozice taxonu *Carlina biebersteinii* subsp. *sudetica* v rámci studované skupiny

Taxon popsáný z Jeseníků, jeden z endemitů ČR, je dle autora popisu reprezentován statnými rostlinami s výškou až 55 cm a průměrem lodyhy až 6,5 mm a počtem úborů 1–2(–3). Od ostatních zástupců se liší zejména velikostí terminálního úboru (42 mm v průměru bez zákrovu), délkou ostnů na listech (které tvoří pouze 2 řády, první reprezentují 2,4 mm dlouhé osténky, druhý 0,8 mm) a dále tvarem zákrovních listenů (úzce kopinaté, dlouhé až 52 mm a až 9 mm široké). Vnitřní zákrov pak je dlouhý 16–19 mm.

Vzhledem k tomu, že na poslední existující lokalitě taxonu roste jen okolo deseti rostlin, z nichž obvykle kvete jen jedna nebo pár, byly v rámci této práce znaky měřeny pouze na rostlinách pěstovaných v experimentální zahradě; jednalo se celkem o 19 rostlin. Ty byly

porovány s 20 rostlinami ze tří alpských populací (8 zástupců z lokality 3, 9 z lokality 4 a 3 z lokality 22) a se zástupci reliktní západočeské populace 19 (9 jedinců), též pěstovaných v experimentální zahradě. Ač je datový soubor značně omezený a postihuje pouze několik lokalit, lze na jeho základě polemizovat se zmíněným popisem. Rostliny subsp. *sudetica* pěstované v experimentální zahradě (jak podotýká i sám autor popisu Kovanda (2002)) mají vyšší počet úborů (v našem případě průměrně 8), s čímž však souvisí i menší velikost úborů. V případě terminálních úborů u subsp. *sudetica* to bylo průměrně  $20,3 \pm 6$  mm (průměr největšího terminálního úboru měřený v rámci této populace dosahoval 32 mm). Pro ostatní populace se průměrná velikost úboru pohybovala též kolem 20 mm (u populace 19 to bylo  $21,36 \pm 2,7$  mm, u alpských rostlin  $19,98 \pm 2,1$  mm). Od toho se odvíjí i kratší listeny vnějšího zákrovu (v rámci srovnávaných souborů byla jejich délka u jesenické populace nejmenší ( $27,98 \pm 6,4$  mm), v případě alpských populací pak průměrně  $30,37 \pm 4,1$  mm, nejdelší listeny byly zaznamenány u rostlin z populace 19 ( $36,28 \pm 4,1$  mm), pro srovnání populačních průměrů viz krabicový diagram na obr. 38. I v případě délek ostnů jsou v případě subsp. *sudetica* kratší ( $1,57 \pm 0,3$  mm) než v dalších ze srovnávaných rostlin ( $2,18 \pm 0,5$  mm v případě populace 19 a  $2,24 \pm 0,2$  mm). (Zde byla použita data z dolních, středních i horních listů. Vzhledem k tomu, že autor popisu nespécifikoval, jakých lodyžních listů se údaj o délce ostnů týká, nelze jeho popis se získanými daty dobře porovnat. Navíc jsou dostupná data v případě alpských rostlin pouze pro populaci 4).

Z výše uvedeného vyplývá, že znaky charakterizující *C. biebersteinii* subsp. *sudetica* nejsou, navzdory tvrzení uvedeném v popisu (Kovanda 2002), stabilní při změně podmínek prostředí. Ač se tato populace signifikantně liší např. od populací 3 a 4 absolutní výškou rostlin, v porovnání s dalšími (populace 19 a 22) již rozdíl není výrazný. Podobná je situace v případě délky zákrovních listenů, kde se odlišují jedinci subsp. *sudetica* od zástupců populace 19 a průměrné hodnoty alpských populací leží na pomezí mezi nimi (viz krabicové diagramy na obr. 38). Lze tedy tvrdit, že rámci těchto znaků rostliny z Malé kotliny v Hrubém Jeseníku výrazně nevybočují z variability, jaká je mezi ostatními populacemi *C. biebersteinii* z Alp. Výraznou diferenciaci neukazují ani molekulární data získaná v rámci této práce. Bohužel nepřinesla, vedle potvrzení blízké příbuznosti zmíněné subsp. *sudetica*, rostlin z reliktních lokalit v západních Čechách a horských zástupců *C. biebersteinii*, jednoznačnou představu o jejich vzájemných vztazích. Nejbližší příbuzní našich populací, mohou být, jak v Alpách, tak v Karpatech. Každopádně dle dostupných dat jsou součástí jedné vývojové linie.



Obr. 38 Krabicové diagramy pro porovnání vybraných znaků *C. biebersteinii* subsp. *sudetica* (1), *C. biebersteinii* subsp. *biebersteinii* z lokality v západních Čechách (19) a z Alp (3,4,22). Písmena označují skupiny, které se signifikantně liší dle Tukeyovy metody mnohonásobného porovnání (barvy odpovídají označení populací v tab. 1).

## 5. Diskuze

### 5.1. Možnosti a omezení použitých metod pro rekonstrukci fylogeneze skupiny

Před vlastní rekonstrukcí fylogeneze byla data nejprve filtrována na základě různých kritérií. Tím se výrazně zredukoval počet dat (např. duplikáty tvořily více než 50 % získaných readů v případě 20 vzorků). V dalších krocích pak byla eliminována chybějící data. Je zde nutné čelit dilematu, zdali pracovat s větším množstvím méně kvalitních dat, či naopak s menším podsouborem, zato však dlouhých alignmentů s menším podílem chybějících hodnot. Počet genů zahrnutých do analýzy může značně ovlivnit přesnost získané fylogeneze. Na příklad při použití 5 genů činí rozdíl mezi topologiemi odvozenými na jejich základě různými metodami 60 %, u 500 genů to je 10 % (Skopalíková 2017). Navíc je množství genů, které je pro rekonstrukci fylogeneze nezbytné, ovlivněno také její taxonomickou úrovní. Pro vztahy mezi druhy a poddruhy je nutné zahrnout do analýz vyšší počet genů, než jaké stačí pro oddělení hlavních linií v rámci studované skupiny (Skopalíková 2017, Ai & Kang 2015). Např. ve studii vnitrorodových vztahů rodu *Curcuma* bylo pro zařazení druhů do správné skupiny v rámci podrodu potřeba až 200 genů (Skopalíková 2017). Jaký počet by byl optimální na nižší úrovni (a mezi podstatně mladšími taxony, jako jsou zástupci skupiny *Carlina vulgaris* agg.) je těžké odhadnout. V rámci studie (Ai & Kang 2015), jejímž cílem bylo nalézt nejmenší počet genů nutný pro odvození relevantní fylogeneze, přítomnost jen jedné rychle radiující větve, výrazně hůře podpořeného tripletu, vedla ke zvýšení potřebného počtu genů ze 75 na 175. Bude-li takových víc, což lze v našem případě předpokládat, nemusí být počet exonů roven 450 vůbec dostačující.

Použitím většího množství genů s větším množstvím chybějících dat je také rizikem. Vedle jejich samotného počtu může mít vliv i jejich rozložení v rámci dostupných alignmentů (Bleidorn 2017). Chybějící data nejsou jediným problémem, na úrovni alignmentu, svoji roli může hrát celá řádka různých dalších jevů, více viz tomuto tématu věnované review Philippe et al. (2011). Část z nich by bylo možné řešit filtrováním dat podle více kritérií, jako je délka alignmentu, délka větví jednotlivých genových stromů (čím jsou kratší, tím větší je vliv rekombinace (Philippe et al. 2011)), nebo jejich podpora (Borowiec et al. 2015). Nicméně za použití těchto filtrování relevance odvozených topologií nebyla výrazně vyšší než bez jeho aplikace, vzájemně byly stromy značně inkongruentní (částečně i proto, že počet alignmentů, které splňovaly optimální podmínky, byl omezený). Důsledkem filtrování dat může být na jedné straně zlepšení podpory některých, před jeho aplikací nepodpořených větví, na druhé straně však v některých případech může vést i k odstranění informativních úseků a tím k většímu biasu (Herrando-Moraira et al. 2018, Streicher et al. 2018).

Dalším z faktorů, na které je třeba brát zřetel, i když byly použity próby cílené na obohacení low-copy genů, je v případě recentně divergujících taxonů přítomnost paralogních genů. Základní skripty HybPhyloMaker-u vychází z tzv. „adjustable majority“, počítá s nejčastěji se vyskytujícími se sekvencemi. Obecně se má za to, že takové jsou vždy ortologní (Fér & Schmickl 2018). Nicméně nemusí to platit absolutně. Proto byly exony filtrovány podle



přítomnosti heterozygotních míst (rozšiřující skript HybPhyloMaker4a2 ve verzi 1.6.4, pro bližší popis viz manuál [https://github.com/tomas-fer/HybPhyloMaker/blob/master/docs/HybPhyloMaker\\_manual1.6.4.pdf](https://github.com/tomas-fer/HybPhyloMaker/blob/master/docs/HybPhyloMaker_manual1.6.4.pdf), str. 13). Za účelem přesnějšího filtrování a detekce alelické variability se pak pracovalo se sekvencemi, které zahrnovaly ambiguitní znaky NC-IUPAC. Tento postup přinesl alespoň v případě vzorků z jedné populace topologii odpovídající očekávání (viz dále), přesto nelze vyvozovat závěry o ostatních vztazích.

Na úrovni blízké příbuzných druhů probíhá řada jevů, které ztěžují, ba přímo znemožňují jejich jednoznačné taxonomické vymezení. Patří mezi ně hybridizace, tzv. incomplete lineage sorting, či mezigenomový přenos (intergenomic transfer), organizace genomu (genome organisation). V obecnější rovině pak mohou úspěšnost taxonomického řešení ovlivnit i důsledky selekce; nezanedbatelnou roli hraje také demografická a fylogeografická struktura zkoumaných populací. Tradičně používané modely pro odvození fylogeneze však většinou předpokládají dichotomické větvení a hierarchické vztahy mezi liniemi zkoumaných organismů, zanedbávají ancestrální polymorfismus a dopady různých typů rozmnožovacího systému na vnitro- i mezidruhovou genetickou strukturu (Naciri & Linder 2015).

Východiskem pro rekonstrukci takové retikulární evoluce, může být použití modelů vytvářejících evoluční sítě namísto stromů (Huson & Bryant 2006; Linder & Rieseberg 2004). Proto byla v rámci této práce za účelem vizualizace vztahů mezi skupinami vedle programu RAXML (Stamatakis 2014) ke konstrukci fylogenetických stromů na základě metody maximum likelihood, použita také metoda neighbour-net v rámci programu SplitsTree (Huson & Bryant 2006). Pro tento typ dat je zřejmě vhodnějším způsobem prezentace, příkladem může být pozice problematického vzorku 66 na bázi skupiny *Carlina biebersteinii* subsp. *biebersteinii* ve fylogenetické síti v porovnání s pozicí, kterou prezentuje fylogenetický strom RAXML (zmíněný vzorek vmezeřen mezi vzorky 63 a 64). Dalším takovým případem je vzorek 4, který v rámci fylogenetického stromu rDNA vytváří větev mezi vzorky skupiny *C. vulgaris* subsp. *vulgaris* a *C. biebersteinii* subsp. *brevibracteata*.

Avšak ani fylogenetické sítě počítané metodou neighbour-net nejsou uspokojivým řešením, nepočítají s efektem ILS, předpokládají pouze vliv hybridizace (Nakhleh 2011), což může vést k mylným výsledkům. Obecně s tímto efektem pracují modely založené na koalescenci, ty však pro změnu vychází z předpokladu, že po speciaci mezi vzorky nedošlo k vzájemnému genovému toku; nelze je tedy aplikovat na data, kde lze hybridizaci předpokládat (Kubatko et al. 2009), jako jsou tato. Navíc získané alignmenty byly spíše krátké (průměrně 248 bp), což omezuje úspěšnost koalescentních metod (Xi et al. 2015).

Vedle příčin uvedených výše (malý počet získaných genů pro odvození vztahů na vnitrodruhové úrovni, jakou recentně vzniklé taxony představují a/nebo nevhodně zvolené parametry pro výběr těchto genů, jejich malá variabilita, rozsáhlá vzájemná hybridizace, ILS a jiné evoluční procesy) je nutné též zvažovat možné chyby při vlastním procesu přípravy vzorků (chybovost PCR). Je zde patrný vliv rozdílné verze použitého laboratorního kitu pro obohacení DNA, vedle skutečnosti, že vzorky analyzované v rámci pilotní studie měly vyšší koncentraci i podstatně vyšší zastoupení cpDNA a rDNA, sdílejí pravděpodobně i část stejných dat (ať již chybějící data, nebo naopak více ambiguitních bází jako důsledek většího

pokrytí). V důsledku se pak objevuje v několika fylogenetických stromech a sítích artefakt v podobě sesterských vzorků 17 a 33 a větev prezentující vzájemnou příbuznost vzorků 1, 19 a 3, které byly analyzovány právě během pilotního běhu. Lze spekulovat o vlivu knihovny, v rámci které vzorky byly připravovány, na výsledky i v ostatních případech. Bohužel však vzorky nebyly do knihoven rozřazeny náhodně (příprava probíhala postupně, vzorky získané z jedné sběrné cesty, tedy z blízkých lokalit, často byly sekvenovány společně), proto jej nelze dobře testovat. Počet variabilních míst ve výsledném konkatenovaném souboru po filtrování navíc činil pouze půl procenta. Co z této nízké variability lze brát za reálný základ fylogenetických vztahů a co je šumem v datech, ať již jakéhokoli původu, je otázkou, každopádně i změna v několika málo nukleotidech se za takových podmínek může markantně projevit (což je zřejmě i důvod diametrálně odlišných stromů za různého filtrování dat).

Diskutabilní je i použití univerzálních hybridizačních prób, které sice přinášejí dostatečně rozlišené topologie na mezidruhové úrovni (Chau et al. 2018), dosavadní studie prováděné na úrovni poddruhů až populací stavěly na próbách specificky navržených pro daný taxon (Nicholls et al. 2015; Villaverde et al. 2018). Obecně tento přístup, ač přináší úskalí spojená s nalezením vhodných cílených úseků, poskytuje vyšší počet vymapovaných sekvencí apod., je považován za vhodnější (Kadlec et al. 2017), i když výsledky na základě univerzálních prób mohou být srovnatelné a pro mezidruhovou úroveň zcela postačující (Chau et al. 2018). Nicméně, i příklady studií, kde sice oproti dosud používaným markerům přineslo použití Hyb-Seq s univerzálními próbami podstatně lépe vyřešené stromy, ale přesto neposkytlo dostatečně variabilní sekvence pro odvození vztahů na nižších úrovních, nalézt lze, např. fylogeneze rodu *Sarracenia* (Stephens et al. 2015), nebo recentně publikovaná práce řešící vztahy v rámci komplikované skupiny 4 rodů z podčeledi *Cardueae* (Herrando-Moraira et al. 2018). Autoři jednak srovnávali použití pipeline PHYLUCE a HybPiper pro analýzu dat, jednak různé metody filtrování. Ač jejich analýzy přinesli dobře vyřešené topologie zástupců rodů *Arctium*, *Cousinia* a *Saussurea*, v případě čtvrtého, *Jurinea*, byla výsledná fylogeneze tvořena překvapivě krátkými větvemi s nízkou podporou. Metoda tedy má široké uplatnění napříč různými taxonomickými úrovněmi, nicméně má své limity a úspěšnost její aplikace se liší druh od druhu v závislosti na jeho evoluční historii. Možným řešením dle autorů studie je (vedle zvýšení počtu vzorků) jednak použití metod „statistical binning“ (Bayzid et al. 2015) (umožňují eliminovat efekt krátkých alignmentů pro konstrukci genových stromů v rámci koalescentních metod), jednak použití variabilnějších částí genomu (intronů, či superkontigů), které však má své opodstatnění zejména v případě, že není očekáván výrazný efekt ILS. V rámci této práce byl datový soubor ze superkontigů použit, přinesl sice lépe podpořenou topologii (s průměrným bootstrapem 39, s podporou vyšší jak 70 bylo 14 větví), nicméně přesto nereflektující reálné vztahy v přírodě lépe než předchozí. Dalším z možných řešení pak je použití některé z jiných metod NGS, které poskytují rozlišení na vnitrodruhové a populační úrovni jako je RAD-sekvenování (Davey et al. 2010).

Jak již bylo zmíněno v kapitole Výsledky, vzorky původem z jedné populace (1, 63, 64), jejichž postavení ve vzájemné blízkosti by mělo odpovídat reálným vztahům, tvoří jednu linii pouze v jednom fylogenetickém stromu (viz obr. 20a). Ohledně jejich příbuznosti by

neměl být pochyb vzhledem ke skutečnosti, že se jedná o vzorky z reliktní lokality v Malé Kotlině, kde se bude jednat o populaci geneticky značně uniformní vzhledem k její izolovanosti i nízkému počtu jedinců. Předpokládaný vztah je prezentován v rámci fylogenetické sítě a stromu odvozených na základě dat s vyfiltrovanými heterozygotními lokusy při použití ambiguitních kódů pro víceznačné pozice (pak je též v rámci fylogenetické sítě odvozené na základě superkontigů, obr. 21b, v rámci stromu RAxML ze stejných dat však již není tato skupina patrná, obr. 20b). Na základě toho lze snad usuzovat, že z použitých přístupů k sestavení konsenzuální sekvence a filtrování dat, poskytuje výše zmíněný přístup nejlepší představu o vzájemných vztazích i v případě dalších vzorků. Přesto však zdaleka neodpovídá realitě v přírodě, minimálně některé z dalších pracovních skupin by v takovém případě byly v rámci fylogenetického stromu či sítě pozorované.

Rozdíly mezi fylogenetickými stromy odvozenými na základě jaderných low-copy genů a ribozomální DNA (viz obr. 23) jsou obecně dány odlišnou evoluční historií. Jednak je to již zmíněným jevem „incomplete lineage sorting“, jednak tím, že rDNA často podléhá tzv. concerted evolution (Feliner and Rosselló 2007). V rámci této práce fylogeneze odvozená na základě rDNA nepřináší v porovnání s exonovou DNA lepší výsledky (i zde je průměrná bootstrapová hodnota 24, průměrná délka větvi je však o řád menší). Naopak představuje několik pozic vzorků, které vybočují i z předchozích dvou skupin, jinak dobře podpořených ve fylogenetických stromech odvozených na základě jaderné DNA. To je dáno zřejmě nízkou variabilitou (v rámci obou datových souborů).

## 5.2. Genetická podobnost a evoluční vztahy

Získané stromy nejsou vzájemně kongruentní v řadě větvení. Procentuální vyjádření Robinsonovy-Fouldovy vzdálenosti (RF) mezi jejich topologiemi dosahovalo od 79 % (porovnání souborů s a bez vyfiltrovaných heterozygotních lokusů) po 100 (rDNA v porovnání s ostatními). Výsledky tohoto typu, v souvislosti s tím, že použitá metoda neposkytla dostatečné rozlišení na úrovni jednotlivých populací, ukazují na komplikovanost vzájemných vztahů v rámci studované skupiny. Navzdory tomu lze ve všech fylogenetických stromech a sítích (kromě pozic několika sporných vzorků) pozorovat jednoznačně odlišené dvě hlavní genetické linie, které odpovídají *C. vulgaris* subsp. *vulgaris* a *C. biebersteinii* subsp. *brevibracteata* na jedné straně a *C. biebersteinii* subsp. *biebersteinii* a subsp. *sudetica* na straně druhé. I tato informace je nesmírně cenná v kontextu dosavadních sporných názorů na taxonomické řešení skupiny *Carlina vulgaris* agg. na území ČR (Kovanda 2004, Kubát 2013), jak je blíže popsáno v dalších kapitolách.

## 5.3. Morfologická variabilita

Morfologické znaky pro zhodnocení variability v rámci studovaných taxonů byly odečítány jak na rostlinách přímo v terénu, tak na rostlinách, které byly přesazeny ve stadiu přizemních růžic do experimentální zahrady a dopěstovány do květu. Většinou nakvétaly v následující sezoně. Tento postup byl aplikován za účelem eliminace vlivu podmínek prostředí na fenotyp rostlin. O znacích, které se s přesazením měnily nejčastěji, vypovídá posun v ordinačním diagramu PCA (obr. 26) a příslušný graf komponentních vah (obr. 27).

Nejvíce byly ovlivněny znaky poměr výšek (pomvys), úhel, který svírají listy s lodyhou (uhel), průměr terminálního úboru (terub), znaky počet úborů na rostlině (sumub) a průměr lodyhy ve výšce 3 cm nad zemí (prumer). První ze zmíněných je dán pravděpodobně tím, že rostliny byly pěstovány na plném slunci s vyloučením kompetice dalších druhů. Posun v dalších ze znaků vypovídá o vyšší fitness rostlin, kterou zřejmě jedinci v zahradě díky pravidelné závlivce a další péči vykazovaly. Většina z nich byla nezávisle na lokalitě původu bohatě větvená s velkým počtem úborů. Úhel, který svírají listy s lodyhou (uhel), byl tímto faktorem také značně ovlivněn, neboť se nezdá stávalo zejména u rostlin *C. vulgaris* subsp. *vulgaris*, které se větily již od báze, že v paždí všech středních lodyžních listů vyrážela další větev s květenstvím. Tato pozorování jsou v souladu s tím, že ve zmíněných znacích jsou tyto rostliny značně plastické (viz obr. 4) i s výsledky studie zaměřené na fenotypovou plasticitu *C. vulgaris* (Becker 2005), která vliv znaků vypovídajících o fitness rostlin porovnává s ostatními, které s ní přímo nesouvisí. Plasticita ve znacích jako jsou počet úborů a jejich velikost může být adaptivní (v čím větším rozsahu dokáží rostliny využít dostupné zdroje, tím lépe budou přežívat v různorodých podmínkách prostředí) a rostliny ze stejných lokalit proto mohou být ve zmíněných znacích značně variabilní. Z toho také plynou nesrovnalosti v popisech. Rostliny z Českého středohoří (pakliže se nejedná přímo o důsledky hybridizace *C. biebersteinii* subsp. *brevibracteata* a *C. vulgaris* s. str.), mohou být na okraji svého areálu, tudíž v suboptimálních podmínkách, drobnější. Údaj, že rostliny *C. biebersteinii* subsp. *brevibracteata* mají (1)3-5 úborů (Kubát 2013) tak vypovídá spíše o obdobných ekologických parametrech lokalit, než příslušnosti rostlin k některému z taxonů.

Značná je variabilita uvnitř i mezi populacemi (více viz práce Berg et al. 2005), kterou ukazují i překrývající se shluky bodů na ordinačním diagramu PCA na základě datového souboru pro jednotlivé jedince (obr. 28, obr. 35 dole), výsledky shlukovací analýzy zejména v případě rostlin z terénu (obr. 30) i přímá pozorování (např. *C. vulgaris* rostoucí ve stínu může mít ploché listy ze střední části lodyhy, v extrémních případech s podvinutým okrajem). Přesto je řada znaků stabilní a rostliny na jejich základě lze dobře klasifikovat minimálně do dvou skupin definovaných na základě molekulárních analýz (viz obr. 32 a výsledky CDA v tab. VI a VII v příloze). K nejlepším diskriminačním znakům, které dobře odlišují dvě zmíněné genetické linie, patří v souladu s popisy v monografii rodu (Meusel & Kästner 1994) i s recentně publikovanou determinační literaturou (Fischer et al. 2008, Jäger 2011) znaky délka listenů vnějšího zákrovu (dzl) a znaky popisující tvar horních a středních lodyžních listů (blíže popsáno v kapitole 4.3.4.).

Další znaky by byly použitelné pouze v porovnání rostlin z identických podmínek (příkladem může být právě znak počet úborů na rostlině (sumub), jehož celkový strukturní koeficient CDA je v případě rostlin ze zahrady relativně vysoký -0,47 v porovnání s analýzou, do které byly zahrnuty rostliny všechny, kde je jeho hodnota pouhých 0,02). Také byla pozorována fenologická diferenciaci. Rostliny klasifikované jako *C. biebersteinii* subsp. *biebersteinii* rozkvétaly o několik týdnů dříve, než druhá skupina.

Rozdíly mezi jedinci klasifikovatelnými jako *C. vulgaris* subsp. *vulgaris* a *C. biebersteinii* subsp. *brevibracteata* byly v terénu i v experimentální zahradě dobře patrné. První taxon reprezentovaly nižší rostliny, větvené od báze, druhý spíše vyšší „metlicovité“ rostliny



(odpovídá zhruba druhému habituálnímu typu zprava na obr. 4). Morfometrické analýzy na rozlišení těchto taxonů však nebyly do práce zahrnuty z důvodu absence jejich vymezení na základě molekulárních výsledků. I zde se projevoval částečný fenologický posun: rostliny *C. vulgaris* subsp. *vulgaris* zpravidla rozkvétaly o týden nebo dva později než rostliny *C. biebersteinii* subsp. *brevibracteata*.

#### 5.4. Populace z reliktních lokalit

Populace ze dvou západočeských lokalit (z Chlumské hory u Manětína (19) a z údolí Teplé (36)) svou výjimečnou morfologií vždy poutaly pozornost botaniků. Zmiňuje se o nich již Čelakovský (1888), který je ztotožnil s alpskou *C. vulgaris* var. *longifolia* (dnes synonymum jména *biebersteinii* subsp. *biebersteinii*). Názory na jejich taxonomickou identitu se však v průběhu uplynulých desetiletí různily. Někteří autoři (např. Hendrych 1986) se domnívali, že ztotožnění s vysokohorskými populacemi je mylné a tamější rostliny přiřadili k taxonu *C. vulgaris* subsp. *intermedia* (dnes synonymum jména *C. biebersteinii* subsp. *brevibracteata*). Poddruh *C. biebersteinii* subsp. *biebersteinii* (v originále *C. vulgaris* subsp. *stricta*) podle tohoto pojetí na území ČR reprezentovaly pouze rostliny z karů Hrubého Jeseníku. Tato populace však byla následně popsána jako samostatný poddruh *C. biebersteinii* subsp. *sudetica* (Kovanda 2002). V Květeně ČR (Kovanda 2004) byl taxon *C. biebersteinii* subsp. *biebersteinii* vymezen příliš široce a kromě zástupců zmíněných reliktních západočeských populací zahrnoval i populace ze svahů hluboce zaříznutých údolí řek a xerothermních trávníků.

Navzdory různým dříve nastíněným hypotézám přinesly molekulární i morfometrické analýzy provedené v této práci důkaz o příbuznosti rostlin z reliktních lokalit v západních Čechách (19 a 36) s rostlinami sbíranými v Alpách a Karpatech (vzorky 3, 4, 22, 38, 40, 41, 45 a 57). Všechny tyto vzorky se shlukují do jedné skupiny a lze je považovat za taxonomicky identické. Tato práce tedy jako první přesvědčivě prokázala, že tyto západočeské populace patří k převážně horskému taxonu *C. biebersteinii* subsp. *biebersteinii*.

Do stejného shluku patří i *Carlina biebersteinii* subsp. *sudetica*. Jedná se o poddruh původně vylišený pouze na základě morfologických znaků, které jsou však značně ovlivněny specifickými podmínkami prostředí. 20 rostlin původem ze zmíněné jesenické populace, které byly pěstovány v experimentální zahradě a měřeny v rámci této práce, byly v klíčových znacích srovnatelné se zástupci *C. biebersteinii* subsp. *biebersteinii* pěstovanými v obdobných podmínkách (viz krabicové diagramy na obr. 38).

Výše uvedené výsledky ukazují, že *Carlina biebersteinii* subsp. *biebersteinii* představuje jeden z taxonů, které mají ve střední Evropě centrum rozšíření v Alpách a Karpatech, ale zároveň našly refugia v nižších polohách; dalšími příklady mohou být *Calamagrostis varia* a *Tofieldia calyculata* (Kaplan 2017). Jedná se o populace unikátní svou postglaciální historií, přežívající dnes na izolovaných lokalitách. Není to pouze jesenická populace, ať již je, či není hodnocena jako samostatný poddruh, ale i další výskyty v údolí Teplé a na Chlumské hoře u Manětína, které si zaslouží zvláštní pozornost. V současné době je v červeném seznamu (Grulich 2017) vedle *C. biebersteinii* subsp. *sudetica* v kategorii C1 (kriticky ohrožený druh) veden taxon *C. biebersteinii* subsp. *biebersteinii* v kategorii C3 jako ohrožený druh, což odráží přetrvávající



nesrovnalosti v taxonomickém hodnocení skupiny a chybné vymezení *C. biebersteinii* subsp. *biebersteinii* v Květeně ČR; Kovanda 2004). Příloha II vyhlášky č. 395/1992 Sb. se seznamem zvláště chráněných druhů rostlin neobsahuje žádný z pojednávaných taxonů. V obou případech je tedy nezbytná revize, i když většina lokalit není přímo ohrožena lidskou činností. Rizika, která rostliny ohrožují, jsou dána spíše malou velikostí populací a tím i nízkou genetickou variabilitou, což se může fatálně projevit při náhlých klimatických výkyvech apod.

Na dalších lokalitách reliktního charakteru (hluboká údolí řek se strmými svahy ve středních Čechách a na jihozápadní Moravě) nebyla přítomnost *C. biebersteinii* subsp. *biebersteinii* prokázána. Na studovaných lokalitách byly zjištěny jen populace rostlin, které geneticky patří do skupiny s *C. vulgaris* subsp. *vulgaris* a *C. biebersteinii* subsp. *brevibracteata*. Tyto (i mnohé jiné) výskyty jsou zřejmě pozůstatkem tradičního hospodaření v krajině, neboť *C. vulgaris* je jedním z druhů favorizovaných pastvou (Becker 2005). V zastínění, které je dáno zapojováním porostů, pak vytváří morfotypy (snad morfotyp *planifolia*), jejichž identita může být bez znalosti biotopu a možnosti srovnání s dalšími rostlinami matoucí.

### 5.5. Taxonomická pozice a hodnocení *C. biebersteinii* subsp. *brevibracteata*

Ohledně rostlin morfologicky hodnocených jako *C. biebersteinii* subsp. *brevibracteata* nelze na základě dostupných dat vyvozovat mnoho závěrů. Na základě molekulárních analýz netvoří v rámci skupiny, kterou sdílí s *C. vulgaris*, samostatné linie, což může být důsledek opakované hybridizace a introgrese. Přítomnost intermediárních populací mezi pupavou obecnou pravou (*Carlina vulgaris* subsp. *vulgaris*) a pupavou Biebersteinovou prostřední (*Carlina biebersteinii* subsp. *brevibracteata*) by tomu i nasvědčovala.

Nicméně nabízí se otázky, jak výrazná je genetická diferenciace v dalších územích, kde je tento taxon rozlišován, a zdali je vskutku oprávněné ho hodnotit jako poddruh v rámci druhu *C. biebersteinii*. Dle teorie vzniku a šíření studovaných taxonů, kterou nastiňuje Meusel (1993), došlo k oddělení linie *C. biebersteinii* (včetně subsp. *brevibracteata*) od mediteránních typů *C. vulgaris* v severní Anatolii za formování hor během pleistocénu. Jedna ze zde vzniklých linií (předkové dnešní *C. biebersteinii* subsp. *biebersteinii*) se pak během teplejších období mohla šířit napříč kontinentální Eurasií až do západní Sibiř, zatímco dnešní *C. biebersteinii* subsp. *brevibracteata* je možným potomkem původních populací v jižní části areálu (oblast Černého moře a Kavkaz), který se rozšířil až po poslední době ledové v ponticko-panonské oblasti. Ve střední Evropě pak dochází k setkání všech linií (*C. vulgaris* sem domigrovala z jihu, zatímco *C. biebersteinii* subsp. *biebersteinii* zde zůstala v oblastech bezlesí, které se reliktně zachovaly ve Francouzském středohoří, Vogézách, Alpách a Karpatech). Molekulární data získaná v rámci této práce však ukazují provázané vztahy mezi *C. vulgaris* a *C. biebersteinii* subsp. *brevibracteata*, zatímco *C. biebersteinii* subsp. *biebersteinii* tvoří samostatnou skupinu, což by v protikladu k nastíněné historii mohlo svědčit o tom, že se tyto dvě větve oddělily dříve, než vznikl poddruh *C. biebersteinii* subsp. *brevibracteata*. Pro potvrzení takových domněnek by však bylo zapotřebí jednak vyšší rozlišení v rámci analýz, jednak zahrnutí vzorků z celého areálu skupiny.

## 6. Závěr

Na základě provedené molekulární analýzy jsou v rámci studovaných vzorků dvě jasně odlišitelné skupiny, první odpovídá pupavě Biebersteinovy pravé (*Carlina biebersteinii* subsp. *biebersteinii* a subsp. *sudetica*) a druhá sdružuje všechny ostatní vzorky zahrnuté do této studie (nominální podruh *C. vulgaris*, *C. biebersteinii* subsp. *brevibracteata* a přechodné typy). Vztahy v rámci obou skupin však nejsou dobře postihnutečné, ať již kvůli limitům použité metody a/nebo možné hybridizaci a dalším evolučním jevům.

Pro morfologické odlišení zmíněných dvou skupin je, v souladu s recentně publikovanou determinační literaturou (např. Fischer 2008, obr. 3), na základě diskriminační analýzy jakož i Wilkova koeficientu lambda nejvýznamnější znak délka vnějších zákrovních listenů (dzl). A to jak v souborech vzorků dat naměřených na rostlinách pěstovaných v experimentální zahradě, tak v případě rostlin přímo z původních lokalit. Dále se jedná o tvar horních a středních lodyžních listů (znaky jako je poměr délky a šířky listu ze střední části lodyhy v nejširším místě (delsipro) či samotná délka prostředního a horního listu (delhor, delpro)

Lokality reliktního charakteru v západních Čechách hostí rostliny, které lze přiřadit k alpským populacím pupavy Biebersteinovy pravé (*C. biebersteinii* subsp. *biebersteinii*). Dokazují to jak molekulární tak morfometrické analýzy provedené v rámci této práce. V případě dalších stanovišť obecně považovaných za reliktní (údolí řek Vltavy, Dyje, Jihlavy a Rokytné) se však jedná o rostliny spadající do druhé z nalezených skupin. Jejich původ je pravděpodobně mladší, spíše než o postglaciální relikty se jedná o pozůstatky pastevní krajiny přežívající ve světlých lesích na strmých svazích říčních údolí.

*Carlina biebersteinii* subsp. *sudetica* spadá geneticky mezi ostatní zástupce *C. biebersteinii* subsp. *biebersteinii*. I když pozice této větve v rámci skupiny není na základě dostupných dat zcela jednoznačná (nejbližší příbuzní této populace mohou být jak na další z reliktních lokalit (Chlumská hora u Manětína), tak v Karpatech). Výrazná není ani její morfologická diference (shlukovací analýza aplikovaná na data odečtená na rostlinách pěstovaných v experimentální zahradě přiřadila část jedinců původem z této populace k zástupcům z reliktní populace na Chlumské hoře u Manětína, další část k rostlinám z alpských lokalit). Obdobné byly výsledky porovnání populačních průměrů pro hodnoty klíčových znaků. Oprávněnost statutu samostatného podruhu je tedy na tomto místě více než diskutabilní. Pozornost z hlediska ochrany přírody si vedle této populace z Hrubého Jeseníku zaslouží i další zástupci pupavy Biebersteinovy pravé (*C. biebersteinii* subsp. *biebersteinii*), jejichž izolované lokality na našem území s největší pravděpodobností představují reliktní výskyty většího rozšíření během posledního glaciálu nebo raného postglaciálu.

## 7. Seznam použité literatury

- Ai B. & Kang M. (2015):** How many genes are needed to resolve phylogenetic incongruence? – *Evol. Bioinform.* 11: 185–188.
- Allendorf F. W., Leary R. F., Spruell P. & Wenburg J. K. (2001):** The problems with hybrids: setting conservation guidelines. – *Trends Ecol. Evol. (Amst.)* 16: 613–622.
- Álvarez I. & Wendel J. F. (2003):** Ribosomal ITS sequences and plant phylogenetic inference. – *Mol. Phylogenet. Evol.* 29: 417–434
- Ayres D. R., Strong D. R. & Baye P. (2003):** *Spartina foliosa* (Poaceae) – a common species on the road to rarity? – *Madroño* 50: 209–13.
- Bankevich A., Nurk S., Antipov D., Gurevich A. A., Dvorkin M., Kulikov A. S., Lesin V. M., Nikolenko S. I., Pham S., Prjibelski A. D., Pyshkin A. V., Sirotkin A. V., Vyahhi N., Tesler G., Alekseyev M. A., Pevzner P. A. (2012):** SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing. – *J. Comput. Biol.* 19:455–77.
- Barres L., Sanmartín I., Anderson C. L., Susanna A., Buerki S., Galbany-Casals M. & Vilatersana R. (2013):** Reconstructing the evolution and biogeographic history of tribe Cardueae (Compositae) – *Am. J. Bot.* 100: 867–82.
- Bayzid M. S., Mirarab S., Boussau B., Warnow T. (2015):** Weighted statistical binning: enabling statistically consistent genome-scale phylogenetic analyses – *PLoS One* 10: e0129183
- Becker U. (2005):** Population biology of *Carlina vulgaris* and *Hypochoeris radicata* in fragmented european grasslands. (dizertační práce) – Philipps-Universität Marburg, Marburg/Lahn
- Bede C. (2017):** Kindel: Indel-Aware Consensus Calling. – URL: <https://github.com/bede/kindel> (navštíveno 6. srpna 2018)
- Berg H., Becker U. & Matthies D. (2005):** Phenotypic plasticity in *Carlina vulgaris*: effects of geographical origin, population size, and population isolation. – *Oecologia* 143: 220–231.
- Bleidorn, Ch. (2017):** Sources of error and incongruence in phylogenomic analyses. – In Bleidorn Ch., *Phylogenomics*: 173–193, Springer, Cham.
- Bolger A. M., Lohse M. & Usadel B. (2014):** Trimmomatic: a flexible trimmer for Illumina sequence data. – *Bioinformatics* 30: 2114–2120.
- Borowiec M. L. (2016):** AMAS: a fast tool for alignment manipulation and computing of summary statistics. – *PeerJ* 4: e1660
- Borowiec M. L., Lee E. K., Chiu J. C. & Plachetzki D. C. (2015):** Extracting phylogenetic signal and accounting for bias in whole-genome data sets supports the Ctenophora as sister to remaining Metazoa. – *BMC Genomics* 16: 987

- Briggs A. W., Good J. M., Green R. E., Krause J., Maricic T., Stenzel U., Lalueza-Fox C., Rudan P., Brajkovic D., Kucan Z., Gusic I., Schmitz R., Doronichev V. B., Golovanova L. V., de la Rasilla M., Fortea J., Rosas A. & Pääbo S. (2009):** Targeted retrieval and analysis of five neandertal mtDNA Genomes. – *Science* 325: 318–321.
- Camacho C., Coulouris G., Avagyan V., Ma N., Papadopoulos J., Bealer K., Madden T. L. (2009):** BLAST+: architecture and applications." – *BMC Bioinformatics* 10: 421.
- Capella-Gutiérrez S., Silla-Martínez J. M. & Gabaldón T. (2009):** a tool for automated alignment trimming in large-scale phylogenetic analyses. – *Bioinformatics*. 25: 1972–1973.
- CESNET, z. s. p. o (2018):** MetaCentrum VO - virtuální organizace pro celou akademickou obec – URL: <https://metavo.metacentrum.cz/> (navštíveno: 7. srpna 2018)
- Clark M. J., Chen R. & Snyder M. (2013):** Exome sequencing by targeted enrichment. – In Ausubel, F. M. (ed.), *Curr Protoc Mol Biol*. 102: 7.12.1–7.12.21, John Wiley & Sons, New York.
- Cock P. J., Antao T., Chang J. T., Chapman B. A., Cox C. J., Dalke A., Friedberg I., Hamelryck T., Kauff F., Wilczynski B. & de Hoon M. J. (2009):** Biopython: freely available Python tools for computational molecular biology and bioinformatics. – *Bioinformatics* 25: 1422–1423.
- Cooke S. B. & Terhune C. E. (2015):** Form, function, and geometric morphometrics. – *Anat Rec (Hoboken)* 298:5–28.
- Cope J. S., Corney D., Clark J. Y., Remagnino P. & Wilkin P. (2012):** Plant species identification using digital morphometrics: a review. – *Expert Syst Appl* 39: 7562–7573.
- Corney D., Tang H. L., Clark J. Y., Hu Y. & Jin J. (2012a):** Automating digital Leaf measurement: the tooth, the whole tooth, and nothing but the tooth – *PLoS One* 7: e42112.
- Corney D., Clark J. Y., Tang H. L., & Wilkin P. (2012b):** Automatic extraction of leaf characters from herbarium specimens. – *Taxon*, 61: 231–244.
- Cronn R., Knaus B. J., Liston A., Maughan P. J., Parks M., Syring J. V. & Udall J. (2012):** Targeted enrichment strategies for next-generation plant biology. – *Am. J. Bot.* 99: 291–311
- Čelakovský L. (1888):** Resultate der botanischen Durchforschung Böhmens im Jahre 1886. – *Sitzungsber. Königl. Böhm. Ges. Wiss. Prag, Math.-Naturwiss. Cl.*, 1887: 193–194.
- Davey J. W., Davey J. L., Blaxter M. L. & Blaxter M. W. (2010):** RADSeq: next-generation population genetics. – *Brief. Funct. Genomics* 9: 416–423.
- Danihelka J., Petřík P. & Wild J. (2009):** Databanka flóry České republiky – URL: <http://florabase.cz/databanka/index.php> (navštíveno 9. srpna 2018).
- Dray, S. & Dufour A. B. (2007):** The ade4 package: implementing the duality diagram for ecologists. – *J. Stat. Softw.* 22: 1–20.

- Eaton J. W., Bateman D., Hauberg S. & Wehbring R. (2016):** GNU Octave 4.2.0 Manual: A high-level interactive language for numerical computations. – URL: <http://www.gnu.org/software/octave/doc/interpreter> (navštíveno 7. srpna 2018).
- Faircloth B. C., McCormack J. E., Crawford N. G., Harvey M. G., Brumfield R. T., Glenn T. C. (2012):** Ultraconserved elements anchor thousands of genetic markers spanning multiple evolutionary timescales. – *Syst. Biol.* 61: 717–726.
- Faircloth B. C. (2016):** PHYLUCE is a software package for the analysis of conserved genomic loci – *Bioinformatics* 32: 786–788.
- Faircloth B. C., Branstetter M. G., White N. D. & Brady S. G. (2015):** Target enrichment of ultraconserved elements from arthropods provides a genomic perspective on relationships among Hymenoptera – *Mol. Ecol. Resour.* 15: 489–501.
- Feliner G. N. & Rosselló J. A. (2007):** Better the devil you know? Guidelines for insightful utilization of nrDNA ITS in species-level evolutionary studies in plants – *Mol. Phylogenet. Evol.* 44: 911–919
- Felsenstein J. (1993):** PHYLIP (Phylogeny Inference Package) version 3.697 – URL: <http://evolution.genetics.washington.edu/phylip.html> (navštíveno 7. srpna 2018)
- Fér T. & Schmickl R. E. (2018):** HybPhyloMaker: target enrichment data analysis from raw reads to species trees. – *Evol. Bioinform.* 14: 117693431774261.
- Fick S. E., & Hijmans R. J. (2017):** WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. – *Int. J. Climatol.* 37: 4302–4315.
- Fischer M. A. (ed.) (2008):** Exkursionsflora Für Österreich, Liechtenstein und Südtirol. – Biologiezentrum der Oberösterreichischen Landesmuseen, Linz.
- Folk R. A., Mandel J. R. & Freudenstein J. V. (2015):** A protocol for targeted enrichment of intron-containing sequence markers for recent radiations: A phylogenomic example from *Heuchera* (*Saxifragaceae*). – *Appl. Plant Sci.* 3:1500039.
- Galili T. (2015):** dendextend: an R package for visualizing, adjusting and comparing trees of hierarchical clustering. – *Bioinformatics* 31: 3718–3720.
- Gnirke A., Melnikov A., Maguire J., Rogov P., Le Proust E. M., Brockman W., Fennell T., Giannoukos G., Fisher S., Russ C., Gabriel S., Jaffe D. B., Lander E. S. & Nusbaum C. (2009):** Solution hybrid selection with ultra-long oligonucleotides for massively parallel targeted sequencing. – *Nat. Biotechnol.* 27: 182–189.
- Grulich V. (2017):** Červený seznam cévnatých rostlin ČR. – *Příroda* 35: 75–133.
- Hamilton J. P. & Buell C. R. (2012):** Advances in plant genome sequencing. – *Plant J.* 70: 177–190.
- Harrison R. G. (1993):** Hybrid zones and the evolutionary process. – Oxford Univ. Press, New York.
- Hatem A., Bozdağ D., Toland A. E. & Çatalyürek Ü. V. (2013):** Benchmarking short sequence mapping tools. – *BMC Bioinformatics* 14: 184



- Henderson A. (2006):** Traditional morphometrics in plant systematics and its role in palm systematics. – *Bot. J. Linn. Soc.* 151: 103–111.
- Hendrych R. (1986):** Polozapomenuté a nové nálezy z květeny ČSR. – *Zpr. Čs. Bot. Společ.* 21: 45–55.
- Herrando-Moraira S. & Cardueae Radiations Group (2018):** Exploring data processing strategies in NGS target enrichment to disentangle radiations in the tribe *Cardueae* (*Compositae*) – *Mol Phylogenet Evol.* 128: 69–87.
- Hewitt G. M. (1999):** Post-glacial re-colonization of European biota. – *Biol. J. Linn. Soc.* 68: 87–112.
- Hewitt G. M. (2008):** Speciation, hybrid zones and phylogeography – or seeing genes in space and time. – *Mol. Ecol.* 10: 537–49.
- Hijmans R. J., Cruz M., Rojas E. & Guarino L. (2001):** DIVA-GIS 1.4. A geographic information system for the management and analysis of genetic resources data. – International Potato Center and International Plant Genetic Resources Institute, Lima.
- Hijmans R. J. (2017):** raster: geographic data analysis and modeling. – URL: <https://cran.r-project.org/package=raster> (navštíveno 7. srpna 2018).
- Hollingsworth P. M., Li D. Z., van der Bank M. & Twyford A. D. (2016):** Telling plant species apart with DNA: from barcodes to genomes. – *Phil. Trans. R. Soc. Lond. B. Biol. Sci.* 371(1702).
- Huson D. H. & Bryant D. (2006):** Application of phylogenetic networks in evolutionary studies. – *Mol. Biol. Evol.* 23: 254–267.
- Chamala S., García N., Godden G. T., Krishnakumar V., Jordon-Thaden I. E., De Smet R., Barbazuk W. B., Soltis D. E. & Soltis P. S. (2015):** MarkerMiner 1.0: A new application for phylogenetic marker development using angiosperm transcriptomes. – *Appl Plant Sci* 3:1400115.
- Chapman M. A., Chang J., Weisman D., Kesseli R. V., Burke J. M. (2007):** Universal markers for comparative mapping and phylogenetic analysis in the *Asteraceae* (*Compositae*). – *Theor. Appl. Genet.* 115: 747–755.
- Chau J. H., Rahfeldt W. A., Olmstead R. G. (2018):** Comparison of taxon-specific versus general locus sets for targeted sequence capture in plant phylogenomics. – *Appl. Plant. Sci.* 6: e1032.
- Jäger E. J. (2011):** Exkursionsflora von Deutschland 4. – Spektrum Akademischer Verlag, Heidelberg
- Jin T., Hou X., Li P., Zhou F. (2015):** A novel method of automatic plant species identification using sparse representation of leaf tooth features. – *PLoS One* 10: e0139482.
- Johnson M. G., Gardner E. M., Liu Y., Medina R., Goffinet B., Shaw A. J., Zerega N. J., Wickett N. J. (2016):** HybPiper: Extracting coding sequence and introns for phylogenetics from high-throughput sequencing reads using target enrichment. – *Appl. Plant. Sci.* 4: 1600016.

- Kadlec M., Bellstedt D. U., Le Maitre N. C., Pirie M. D. (2017):** Targeted NGS for species level phylogenomics: "made to measure" or "one size fits all"? – *PeerJ* 5: e3569.
- Kajtoch Ł., Cieślak E., Varga Z., Paul W., Mazur M. A., Sramkó G. & Kubisz D. (2016):** Phylogeographic patterns of steppe species in eastern Central Europe: A review and the implications for conservation. – *Biodivers. Conserv.* 25: 2309–39.
- Kamneva O. K., Syring J., Liston A. & Rosenberg N. A. (2017):** Evaluating allopolyploid origins in strawberries (*Fragaria*) using haplotypes generated from target capture sequencing – *BMC Evol. Biol.* 17: 180
- Kaplan Z. (2017):** Flora and phytogeography of the Czech Republic. – In: Chytrý M., Danihelka J., Kaplan Z., Pyšek P. (eds), *Flora and vegetation of the Czech Republic*: 89–163, Springer, Cham.
- Katoh K. & Standley D. M. (2013):** MAFFT multiple sequence alignment software version 7: improvements in performance and usability. – *Mol. Biol. Evol.* 30: 772–780
- Kent W. J. (2002):** BLAT – the BLAST-like alignment tool. – *Genome Res.* 12: 656–664.
- Kovanda M. (2002):** Observations on *Carlina biebersteinii*. – *Thaiszia-J. Bot* 12: 75–82.
- Kovanda M. (2004):** *Carlina* L. – pupava. – In: Slavík B., Štěpánková J. & Štěpánek J. (eds), *Květena České republiky [Flora of the Czech Republic]* 7: 356–361.
- Kozlov A. M., Andre A. J. & Stamatakis A. (2015):** ExaML version 3: a tool for phylogenomic analyses on supercomputers – *Bioinformatics* 31: 2577–2579.
- Korkmaz S., Goksuluk D. & Zararsiz G. (2014):** MVN: An R package for assessing multivariate normality. – *The R Journal* 6: 151–162.
- Körner C. & Meusel H. (1986):** Zur ökophysiologischen und ökogeographischen Differenzierung nah verwandter *Carlina*-Arten. – *Flora* 178: 209–232.
- Kruckeberg A. R. & Rabinowitz D. (1985):** Biological aspects of endemism in higher plants. – *Annu. Rev. Ecol. Syst.* 16: 447–79
- Kubát K. (2013):** *Carlina vulgaris* agg. v České Republice. – *Severočes. Přír.* 44: 1–5.
- Kubatko L. S., Carstens B. C., Knowles L. L. (2009):** STEM: species tree estimation using maximum likelihood for gene trees under coalescence. – *Bioinformatics* 25: 971–973.
- Kuhn M. (2018):** caret: Classification and regression training. – URL: <https://cran.r-project.org/web/packages/caret/index.html> (navštíveno 7. srpna 2018)
- Langmead B. & Salzberg S. L. (2012):** Fast gapped-read alignment with Bowtie 2. – *Nat. Methods.* 9: 357–59
- Larget B. R., Kotha S. K., Dewey C. N. & Ané C. (2010):** BUCKy: gene tree/species tree reconciliation with Bayesian concordance analysis. – *Bioinformatics* 26: 2910–2911.
- Lê S., Josse J. & Husson F. (2008):** FactoMineR: a package for multivariate analysis. – *J. Stat. Soft.* 25: 1–18.

- Lemmon A. R., Emme S. A. & Lemmon E. M. (2012):** Anchored hybrid enrichment for massively high-throughput phylogenomics. – *Syst. Biol.* 61: 727–744.
- Lemmon E. M. & Lemmon A. R. (2013):** High-throughput genomic data in systematics and phylogenetics. – *Annu. Rev. Ecol. Evol. Syst.* 44: 99–121.
- Li H. & Durbin R. (2009):** Fast and accurate short read alignment with burrows-wheeler transform. – *Bioinformatics* 25: 1754–1760.
- Linder C. R. & Rieseberg L. H. (2004):** Reconstructing patterns of reticulate evolution in plants. – *Am. J. Bot.* 91: 1700–1708.
- Liu L., Yu L., Kubatko L., Pearl D. K. & Edwards S. V. (2009):** Coalescent methods for estimating phylogenetic trees – *Mol. Phylogenet. Evol.* 53: 320–328.
- MacLeod N. & Steart D. (2015):** Automated leaf physiognomic character identification from digital images – *Paleobiology* 41: 528–553.
- Malooof J., Kazunari N., Mumbach M. R. & Palmer Ch. M. (2013):** LeafJ: an ImageJ plugin for semi-automated leaf shape measurement – *J. Vis. Exp.* 71: e50028.
- Mamanova L., Coffey A. J., Scott C. E., Kozarewa I., Turner E. H., Kumar A., Howard E., Shendure J. & Turner D. J. (2010):** Target-enrichment strategies for next-generation sequencing – *Nat. Methods* 7: 111–118.
- Mandel J. R., Dikow R. B., Funk V. A., Masalia R. R., Staton S. E., Kozik A., Michelmore R. W., Rieseberg L. H. & Burke J. M. (2014).** A target enrichment method for gathering phylogenetic information from hundreds of loci: An example from the *Compositae*. – *Appl. Plant Sci.* 2: 1300085.
- Mandel J. R., Barker M. S., Bayer R. J., Dikow R. B., Gao T.-G., Jones K. E., Keele S., Kilian N., Ma H., Siniscalchi C. M. & Susanna A. (2017):** The *Compositae* tree of life in the age of phylogenomics. – *J. Syst. Evol.* 55: 405–410.
- Mandel J. R., Dikow R. B. & Funk V. A. (2015):** Using phylogenomics to resolve megafamilies: An example from *Compositae*. – *J. Syst. Evol.* 53: 391–402.
- Marhold K. (2011):** Multivariate morphometrics and its application to monography at specific and infraspecific levels. – In: Stuessy T. F. & Lack H. W., *Monographic plant systematics: Fundamental assessment of plant biodiversity*: 73-99, Ganter, Ruggell.
- Mašková H. (2015):** Diferenciace v rámci okruhu *Carlina vulgaris* agg. ve střední Evropě. – Bakalářská práce, Univerzita Karlova v Praze, Praha.
- McCormack J. E., Tsai W. L. & Faircloth B. C. (2016):** Sequence capture of ultraconserved elements from bird museum specimens. – *Mol. Ecol. Resour.* 16: 1189–1203
- Meusel H. (1993):** Arealanalysen an *Carlina biebersteinii* BERNH. EX. HORNEM. – Versuch Einer Ökogeographischen Sippencharakteristik – *Flora* 188: 73–84.
- Meusel H. & Kästner A. (1990):** Lebensgeschichte der Gold- und Silberdisteln. Monographie der mediterran-mitteleuropäischen Compositen-Gattung *Carlina*. Bd. 1. – Österreichische Akademie der Wissenschaften, Wien.

- Meusel H. & Kästner A. (1994):** Lebensgeschichte der Gold- und Silberdisteln. Monographie der mediterran-mitteleuropäischen Compositen-Gattung *Carlina*. Bd. 2. – Österreichische Akademie der Wissenschaften, Wien.
- Mirarab S., Reaz R., Bayzid M. S., Zimmermann T., Swenson M. S. & Warnow T. (2014):** ASTRAL: genome-scale coalescent-based species tree estimation. – *Bioinformatics* 30: i541–i548.
- Naciri Y. & Linder H. P. (2015):** Species delimitation and relationships: the dance of the seven veils. – *Taxon* 64: 3–16.
- Nakhleh L. (2011):** Evolutionary phylogenetic networks: models and issues. – In: Heath L. S. & Ramakrishnan N. (eds), *Problem Solving Handbook in Computational Biology and Bioinformatics*: 125–158, Springer US, Boston
- Nguyen N., Mirarab S. & Warnow T. (2012):** MRL and SuperFine+MRL: new supertree methods – *Algorithms Mol. Biol.* 7: 3.
- Nicholls J. A., Pennington R. T., Koenen E. J., Hughes C. E., Hearn J., Bunnefeld L., Dexter K. G., Stone G. N. & Kidner C. A. (2015):** Using targeted enrichment of nuclear genes to increase phylogenetic resolution in the neotropical rain forest genus *Inga* (*Leguminosae: Mimosoideae*). – *Front. Plant Sci.* 6: 710.
- Oksanen J., Blanchet F.G., Friendly M., Kindt R., Legendre P., McGlenn D., Minchin P. R., O'Hara, R. B., Simpson G. L., Solymos P., Stevens M. H., Scoecs E. & Wagner H. (2018):** Vegan: community ecology package – URL: <https://cran.r-project.org/package=vegan> (navštíveno 7. srpna 2018).
- Olmstead R. G. & Palmer J. D. (1994):** Chloroplast DNA systematics: a review of methods and data analysis. – *Am. J. Bot.* 81: 1205 - 1224.
- Paradis E., Claude J. & Strimmer K. (2004):** APE: Analyses of phylogenetics and evolution in R Language. – *Bioinformatics* 20: 289–90.
- Pebesma E. J. & Bivand R. S. (2005):** Classes and Methods for Spatial Data in R. – *R News* 5: 9–13.
- Philippe H., Brinkmann H., Lavrov D. V., Littlewood D. T., Manuel M., Wörheide G. & Baurain D. (2011):** Resolving difficult phylogenetic questions: why more sequences are not enough. – *PLoS Biol.* 9: e1000602.
- Price M. N., Dehal P. S. & Arkin A. P. (2010):** FastTree 2 – approximately maximum-likelihood trees for large alignments – *PLoS One* 5: e9490.
- R Development Core Team (2018):** R: a language and environment for statistical computing. – URL: <https://www.r-project.org/> (navštíveno 7. srpna 2018).
- Remagnino P., Mayo S. J., Wilkin P., Cope J. & Kirkup D. (2017):** *Computational Botany: Methods for Automated Species Identification*. – Springer, Berlin.
- Revell L. J. (2012):** Phytools: an R package for phylogenetic comparative biology (and other things). – *Methods Ecol. Evol.* 3: 217–23.

- Rudd S. (2003):** Expressed sequence tags: alternative or complement to whole genome sequences? – *Trends Plant Sci.* 8: 321–329.
- Sanger F., Nicklen S. & Coulson A. R. (1977):** DNA sequencing with chain-terminating inhibitors. – *Proc. Natl. Acad. Sci. U.S.A.* 74: 5463–5467.
- Shaw A. J., McDaniel S. F., Werner O. & Ros R. M. (2002):** Phylogeography and Phylodemography – *Bryologist* 105: 373–383.
- Schliep K. P. (2011):** Phangorn: phylogenetic analysis in R. – *Bioinformatics* 27: 592–593.
- Schmickl R., Liston A., Zeisek V., Oberlander K., Weitemier K., Straub S. C., Cronn R. C., Dreyer L. L. & Suda J. (2016):** Phylogenetic marker development for target enrichment from transcriptome and genome skim data: the pipeline and its application in southern African *Oxalis* (*Oxalidaceae*) – *Mol. Ecol. Resour.* 16: 1124–1135.
- Schmitt T. (2007):** Molecular biogeography of Europe: Pleistocene cycles and postglacial trends. – *Front. Zool.* 4: 11.
- Schneider C. A., Rasband W. S. & Eliceiri K. W. (2012):** NIH Image to ImageJ: 25 years of image analysis. – *Nat. Methods* 9: 671–675.
- Skopalíková J. (2017):** Využití metody Hyb-Seq pro rekonstrukci retikulární vnitrorodové fylogeneze: příklad z polyploidního rodu *Curcuma* L. (*Zingiberaceae*). – Diplomová práce, Univerzita Karlova, Praha.
- Slater G. S. & Birney E. (2005):** Automated generation of heuristics for biological sequence comparison. – *BMC Bioinformatics* 6: 31.
- Soltis D. E., Gitzendanner M. A., Stull G., Chester M., Chanderbali A., Chamala S., Jordon-Thaden I., Soltis P. S., Schnable P. S. & Barbazuk W. B. (2013):** The potential of genomics in plant systematics. – *Taxon* 62: 886–898.
- Stamatakis A. (2014):** RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies – *Bioinformatics* 30: 1312–1313.
- Starrett J., Derkarabetian S., Hedin M., Bryson R. W. Jr., McCormack J. E. & Faircloth B. C. (2017):** High phylogenetic utility of an ultraconserved element probe set designed for Arachnida. – *Mol. Ecol. Resour.* 17: 812–823.
- Steeves T. E., Maloney R. F., Hale M. L., Tylianakis J. M. & Gemmell (2010):** Genetic analyses reveal hybridization but no hybrid swarm in one of the world's rarest birds. – *Mol. Ecol.* 19: 5090–5100.
- Stephens J. D., Rogers W. L., Heyduk K., Cruse-Sanders J. M., Determann R. O., Glenn T. C. & Malmberg R. L. (2015):** Resolving phylogenetic relationships of the recently radiated carnivorous plant genus *Sarracenia* using target enrichment. – *Mol. Phylogenet. Evol.* 85: 76–87.
- Straub S. C., Fishbein M., Livshultz T., Foster Z., Parks M., Weitemier K., Cronn R. C. & Liston A. (2011):** Building a model: developing genomic resources for common milkweed (*Asclepias syriaca*) with low coverage genome sequencing. – *BMC Genomics* 12: 211.



- Straub S. C., Parks M., Weitemier K., Fishbein M., Cronn R. C. & Liston A. (2012):** Navigating the tip of the genomic iceberg: Next-generation sequencing for plant systematics. – *Am. J. Bot.* 99: 349–364.
- Streicher J. W., Miller E. C., Guerrero P. C., Correa C., Ortiz J. C., Crawford A. J., Pie M. R. & Wiens J. J. (2018):** Evaluating Methods for Phylogenomic Analyses, and a New Phylogeny for a Major Frog Clade (*Hylloidea*) Based on 2214 Loci. – *Mol. Phylogenet. Evol.* 119: 128–43.
- Štorchová H., Hrdličková R., Chrtěk J., Tetera M., Fitze D. & Fehrer J. (2000):** An improved method of DNA isolation from plants collected in the field and conserved in saturated NaCl/CTAB solution. – *Taxon* 49: 79–84.
- Taberlet P., Fumagalli L., Wust-Saucy A.-G. & Cosson J.-F. (1998):** Comparative phylogeography and postglacial colonization routes in Europe. – *Molec. Ecol.* 7: 453 - 464.
- Thyagarajan K. K. & Kiruba Raji I. (2018):** A review of visual descriptors and classification techniques used in leaf species identification. – *Arch. Computat. Methods Eng.* 1–28.
- Töpfer A. (n. d.):** ConsensusFixer. – URL: <https://github.com/cbg-ethz/ConsensusFixer> (navštíveno 7. srpna 2018).
- Vachaspati P. & Warnow T. (2015):** ASTRID: accurate species trees from internode distances. – *BMC Genomics* 16 (Suppl 10): S3.
- Venables W. N. and Ripley B. D. (2002):** Modern applied statistics with S. – Springer US, New York.
- Villaverde T., Pokorny L., Olsson S., Rincón-Barrado M., Johnson M. G., Gardner E. M., Wickett N. J., Molero J., Riina R. & Sanmartín I. (2018):** Bridging the micro- and macroevolutionary levels in phylogenomics: Hyb-Seq solves relationships from populations to species and above. – *New Phytol.*
- Viscosi V., Lepais O., Gerber S. & Fortini P. (2009):** Leaf morphological analyses in four European oak species (*Quercus*) and their hybrids: a comparison of traditional and geometric morphometric methods. – *Plant Biosyst.* 143: 564–574.
- Wahrmund U., Heklau H., Röser M., Kästner A., Vitek, E., Ehrendorfer, F. & Hagen, K. (2010):** A molecular phylogeny reveals frequent changes of growth form in *Carlina* (*Asteraceae*). – *Taxon*, 59: 367–378.
- Warnow T. (2015):** Concatenation analyses in the presence of incomplete lineage sorting. – *PLoS Curr.* 7: ecurrents.currents.tol.8d41ac0f13d1abedf4c4a59f5d17b1f7.
- Weihls C., Ligges U., Luebke K. & Raabe N. (2005):** klaR Analyzing German Business Cycles. – In: Baier D., Decker R. & Schmidt-Thieme L. (eds), *Data Analysis and Decision Support*: 335–343, Springer, Berlin.
- Weitemier K., Straub S. C., Cronn R. C., Fishbein M., Schmickl R., McDonnell A. & Liston A. (2014):** Hyb-Seq: Combining target enrichment and genome skimming for plant phylogenomics. – *Appl. Plant Sci.* 2: apps.1400042.

- Xi Z., Liu L. & Davis Ch. C. (2015):** Genes with minimal phylogenetic information are problematic for coalescent analyses when gene tree estimation is biased. – *Mol. Phylogenet. Evol.* 92: 63–71.
- Xu H., Luo X., Qian J., Pang X., Song J., Qian G., Chen J. & Chen S. (2012):** FastUniq: a fast de novo duplicates removal tool for paired short reads. – *PLoS One* 7: e52249.
- Yu Y., Than C., Degnan J. H. & Nakhleh L. (2011):** Coalescent histories on phylogenetic networks and detection of hybridization despite incomplete lineage sorting. *Syst. Biol.* 60: 138–49

## 8. Přílohy

### 8.1. Seznam vzorků

vz.	místo původu, lokalita	souřadnice (WGS 84)		n. v.	col.
1	rostliny získané z rušené kultivace; původ: CZ, okr. Bruntál: Karlov pod Pradědem: Pupavová skalka v Malé kotlině	50°02'22.2"N	17°12'33.2"E	1260	
2	CZ, okr. Hlavní město Praha: Malá Chuchle, PP Homolka: stepní společenstva na jižním svahu vrchu nad bývalým vápencovým lomem	50°0'55.76"N	14°22'32.74"E	280	HM
3	A, Salzburg: okr. Hallein: Osterhomgruppe: trávníky s <i>Calamagrostis varia</i> na strmých svazích nad lesní cestou v údolí řeky Aubach	47°36'22.6"N	13°17'44.9"E	680	ZK
4	A, Oberösterreich: okr. Kirchdorf an der Krems: trávníky s <i>Calamagrostis varia</i> a <i>Carex flacca</i> na strmých svazích nad lesní cestou, západní svah hory Rammelspitz	47°52'08.5"N	14°14'17.9"E	570	ZK
5	CZ, okr. Příbram: Cholín: řídký listnatý les na skalnaté, ZJZ orientované stráni nad levým břehem vodní nádrže Slapy	49°43'23.4"N	14°19'32.1"E	350	ZK+HM
6	CZ, okr. Příbram: Nepřejov: xerothermní trávník s keří v lesním lemu na kraji koňské pastviny na svahu údolí	49°38'21.1"N	14°10'21.6"E	480	ZK+HM
7	SK, okr. Senica: Sekule: xerothermní trávník v lemu borového lesa na písčité v soustavě zatopených pískoven	48°37'06.3"N	17°00'00.6"E	150	ZK
8	CZ, okr. Mladá Boleslav: Kropáčova Vrutice: xerothermní trávníky s keří na stráni	50°19'58.1"N	14°42'08.1"E	230	ZK
9	CZ, okr. Litoměřice: Čakovice: xerothermní trávníky na bílé stráni obklopené křovinami a lesem na JZ svahu návříší Na Vinici	50°29'00.7"N	14°22'54.2"E	210	ZK
10	rostliny pěstované v Botanické zahradě PŘF UK; původ: CZ, okr. Břeclav: Březi: Dunajovické kopce, starší kamenitý úhor na svahu	48°50'15"N	16°33'42"E	240	
11	CZ, okr. Hradec Králové: Jeníkovice: křovinami zarůstající xerothermní trávník mezi polní cestou a lesem Osík	50°13'25.2"N	15°58'54.0"E	260	ZK
12	CZ, okr. Jeseník: Vápenná: xerothermní trávník postupně zarůstající dřevinami (bříza, osika, jiva, smrk) na střední terase opuštěného vápencového lomu	50°16'57.1"N	17°05'25.0"E	470	ZK
13	CZ, okr. Mělník: Nelahozeves: řídký trávník na hraně těžebního prostoru pískovny Uhý	50°16'22.36"N	14°17'47.79"E	230	HM
14	CZ, okr. Mělník: Dřínov, PP Dřínovská stráž: xerothermní trávníky na jižním svahu Dřínovského vrchu	50°15'59.62"N	14°23'16.15"E	220	HM
15	CZ, okr. Beroun: Sv. Jan pod Skalou, NPR Karlštejn: dno bývalého vápencového lomu	49°58'19.21"N	14°8'39.91"E	380	HM
16	SK, okr. Liptovský Mikuláš: Liptovský Ján, NP Nízke Tatry: východní svah hory Biela Marušová, při kraji cesty	49° 0'46.39"N	19°40'24.35"E	740	HM
17	CZ, okr. Břeclav: Pouzdřany, NPR Pouzdřanská step - Kolby: xerothermní trávníky na svazích Uherčické nové hory	48°56'43.72"N	16°38'38.48"E	270	HM
18	CZ, okr. Šumperk: Hrabová: nekosený trávník u cesty do starého opuštěného lomu	49°51'12.0"N	16°57'08.7"E	310	LP
19	CZ, okr. Karlovy Vary: Manětín: prudká suťová stráž v řídkém lese na JZ svahu vrchu Chlumská hora	50°00'19.3"N	13°12'16.6"E	590	ZK
20	CZ, okr. Hradec Králové: Jeníkovice: malý zbytek křovinami zarůstajícího xerothermního trávníku v lese Osík	50°13'31.2"N	15°59'05.6"E	270	ZK
21	CZ, okr. Český Krumlov: Vyšný: vápencový lom na východním úpatí Výřího vrchu	48°49'55.15"N	14°17'55.89"E	580	HM
22	A, Salzburg: okr. Hallein: sutě na úpatí vápencových skal Lueg Palfen v údolí řeky Salzach	47°34'01.3"N	13°10'18.1"E	670	ZK+IK

vz.	místo původu, lokalita	souřadnice (WGS 84)		n. v.	col.
23	A, Salzburg: okr. Sankt Johann im Pongau: xerothermní trávníky podél cesty do vápencového lomu	47°31'24.5"N	13°10'11.7"E	560	ZK+IK
24	A, Salzburg: okr. Zell am See: vápencové sutě na strmých svazích na jižním úpatí hory Lärfeldkogel	47°34'48.2"N	12°39'57.8"E	770	ZK+IK
25	A, Salzburg: okr. Sankt Johann im Pongau: řídký les na vápencové suti na strmých svazích v údolí řeky Salzbach	47°30'36.4"N	13°10'30.0"E	580	ZK+IK
26	CZ, okr. Vyškov: Nevojice: xerothermní trávníky zv. Malhotky na svahu pod krajem Milonického lesa	49°08'55.8"N	17°03'14.9"E	270	ZK+HM
27	CZ, okr. Uherské Hradiště: Uherský Brod-Havříce: xerothermní trávníky na S břehu jezírka v opuštěné cihelně	49°01'35.1"N	17°36'46.1"E	220	ZK+HM
28	CZ, okr. Břeclav: Kurdějov: xerothermní trávníky v dolní části svahu 720 m S od vrcholu Kamenného vrchu	48°58'10.7"N	16°45'00.2"E	310	ZK+HM
29	CZ, okr. Břeclav: Přítluky: xerothermní trávníky podél cesty vedoucí na Přítluckou horu	48°51'37.1"N	16°46'06.6"E	220	ZK+HM
30	CZ, okr. Břeclav: Březí: Dunajovické kopce, řídké xerothermní trávníky na kamenitém svahu	48°50'22.2"N	16°33'47.4"E	260	ZK+HM
31	CZ, okr. Třebíč: Šemíkovice: xerothermní trávník ve světlíně v boru na svahu nad levým břehem říčky Rokytne	49°02'42.5"N	16°08'05.0"E	340	ZK+HM
32A	CZ, okr. Litoměřice: Tetčíněves: světlý bor (v porostu <i>Brachypodium pinnatum</i> ), ve svahu nad lesní cestou	50°33'29.921"N	14°21'36.09"E	290	HM
32B	CZ, okr. Litoměřice: Hoštka: kosený trávník u silnice mezi obcemi Hoštka a Malešov	50°30'3.618"N	14°20'4.524"E	220	HM
33	CZ, okr. Liberec: Raspenava, PR Vápenný vrch: dno bývalého lomu a světliny v dubohabřině na jeho horní hraně	50°53'35.600"N	15°8'13.600"E	380	HM
34	CZ, okr. Litoměřice: Julčín: bylinný podrost květnatého boru Na Čemč	50°32'50.6"N	14°22'29.4"E	310	ZK
35	SK, okr. Ilava: Krivoklát: PP Babiná, stráně na JZ úbočí vrchu	49°02'02"N	18°10'37"E	340	HM
36	CZ, okr. Sokolov: Nová ves, Louka: PR Údolí Teplé, železniční násypy	50°2'57.072"N	12°49'31.48"E	610	HM
37	CZ, okr. Náchod: Česká Skalice: xerothermní mez nad silnicí pod lesem Rousín	50°22'24.3"N	16°02'41.2"E	290	FK
38	A, Salzburg: okr. Zell am See: NP Hohe Tauern, příležitostně přepásaný trávník u lesní cesty na SV svahu hory Spielbergkogel	47°14'54.9"N	12°28'50.8"E	1130	ZK+IK
39	CZ, okr. Hodonín: Hodonín: les Důbrava, travnatá světlina v dubové výsadbě S od silnice Hodonín – Mutěnice	48°53'01.0"N	17°06'21.2"E	200	ZK+HM
40	SK, okr. Ružomberok: Liptovská Osada: trávník na prudkém vápencovém svahu nad lesní cestou v Teplé dolině	48°56'46.1"N	19°13'36.1"E	720	ZK
41	SK, okr. Banská Bystrica: Homý Hamanec: řídký trávník na prudkém suťovém vápencovém svahu	48°49'21.2"N	19°01'13.6"E	700	ZK
42	rostliny pěstované ze semen z Botanické zahrady PřF UK; původ: CZ, okr. Břeclav: Březí: Dunajovické kopce, starší kamenitý úhor na svahu	48°50'15"N	16°33'42"E	240	
45	SK, okr. Martin: Folkušová: řídký trávník na prudkém suťovém svahu pod vápencovou skálou v lese v Gaderské dolině	48°57'13.5"N	18°58'59.6"E	660	ZK
46	CZ, okr. Kladno: Bílichov: okraj lesa u světlé lesní cesty 180 m ZSZ–Z od Třetího rybníka na Zlonickém potoce 1,3 km ZJZ od obce.	50°14'57.2"N	13°53'36.2"E	430	ZK
47	CZ, okr. Litoměřice: Stračín: zbytky xerothermních trávníků přerostlých křovinami a nízkými stromy	50°26'58.6"N	14°24'36.4"E	230	ZK
48	CZ, okr. Litoměřice: Svářeň: xerothermní trávníky na bílé stráni obklopené křovinami na Z svahu mělkého údolíčka	50°30'02.9"N	14°18'53.5"E	200	ZK
49	CZ, okr. Hradec Králové: Hradec Králové-Svinary: trávník mezi železniční tratí a motokrosovou dráhou	50°13'13.2"N	15°54'28.2"E	260	ZK
50	CZ, okr. Příbram: Solenice: světlina v borovém lese	49°36'13.3"N	14°10'38.3"E	360	ZK
51	I, kraj Piemonte: prov. Torino: Superga: stará lesní cesta poblíž silnice "Strada Palucco"	45°04'19.8"N	7°47'08.0"E	520	MŠ
52	I, kraj Piemonte: prov. Cuneo: Dronero, Tetti: okraj louky při cestě "Borgata Biut"	44°27'21.7"N	7°18'15.9"E	830	MŠ
53	N, Oslo (ostrov Malmøya): vápnomilná vegetace při jižním pobřeží ostrova	59°51'44.179"N	10°45'17.705"E	30	FKo
54	DK, Zealand (ostrov Sjælland: Holtug: útesy Stevns Klint: staré křídové lomy, skály a sutě	55°20'25.872"N	12°26'38.724"E	10	FKo

vz.	místo původu, lokalita	souřadnice (WGS 84)		n. v.	col.
55	SWE, Kalmar; okr. Mörbylånga (ostrov Öland): alvar mezi cestou Resmo-Stenåsa a jezerem Möcklemossen	56°31'42.744"N	16°31'17.076"E	20	FKo
56	SK, Pieniny; okr. Stará Lubovňa: Lesnica: relativně strmý a suchý okraj světlé lesní cesty	49°23'49.3" N	20°27'31.2" E	660	VZ
57	SK, Pieniny; okr. Stará Lubovňa: Lesnica: velmi prudký suchý svah pod skálou nad cestou	49°24'52.8" N	20°27'06.2" E	550	VZ
58	CZ, okr. Znojmo: Čížov: travnatý kosený břeh velkého Čížovského rybníka	48°52'59.040"N	15°53'8.280"E	400	HM
59	CZ, okr. Znojmo: Čížov: skalnatá stráň v ohybu cesty (u hraničního přechodu Hardegg)	48°51'17.160"N	15°52'10.860"E	350	HM
60	D, Bayern; okr. Oberpfalz: Zogenreuth, bývalý vápencový lom v rámci vojenského prostoru Grafenwöhr	49°42'0.490"N	11°40'0.590"E	530	HM
61	CZ, okr. Břeclav: Uherčice: zarůstající xerothermní stráň "Na Ramholcích"	48°57'17.5"N	16°40'50.9"E	280	HM
62	CZ, okr. Břeclav: Borkovany: zarůstající xerothermní stráň	49°01.4"N	16°48.14"E	300	HM
63	CZ, okr. Bruntál: Karlov pod Pradědem: Pupavová skalka v Malé kotlině	50°02'22,2"N	17°12'33,2"E	1260	ZK+LB
64	CZ, okr. Bruntál: Karlov pod Pradědem: Pupavová skalka v Malé kotlině	50°02'22,2"N	17°12'33,2"E	1260	ZK+LB
65	UA, reg. Lviv: Perelisky: suchý trávník na kopci při levé straně cesty z obce na Litovyshe	49°57'07.2"N	25°10'39.4"E	320	FK+MD
66	UA, reg. Chmelnick: Kamjanec-Podilskyj: skalní step na prudkých svazích nad řekou Smotrych	48°34'31.4"N	26°38'28.3"E	160	FK+MD
67	UA, reg. Chmelnick: Ivakhnivtsi: suchý trávník ( <i>Cirsio-Brachypodium</i> ) a skalní výchozy vápence	49°04'29.4"N	26°20'10.5"E	300	FK+MD
68	UA, reg. Lviv: Pid'jarkiv: bývalý úhor, nyní zarostlý suchým trávníkem	49°43'44"N	24°20'15"E	290	FK+MD
69	CZ, okr. Klatovy: Čepice (PR Čepičná): světlý bor s válečkou prapořitou na J úpatí vrchu Chanovc	49°16'18.418"N	13°36'12.548"E	530	HM
70	CZ, okr. Mladá Boleslav: Benátky nad Jizerou: kraj borového lesa	50°17'16.139"N	14°52'16.497"E	220	HM
71	CZ, okr. Mladá Boleslav: Bělá pod Bezdězem: bor ve svahu naproti železniční trati	50°29'49.820"N	14°46'52.482"E	300	HM
72	CZ, okr. Olomouc: Krčmaň: PP U Strejčkova lomu: xerothermní trávníky v bývalém těžebním prostoru	49°31'24.521"N	17°19'38.232"E	230	HM
73	CZ, okr. Třebíč: Přesovice: hřbet nad meandrem řeky Rokytne; v úvozu cesty z Vilímova Mlýna na vrch Cibuluška	49°1'54.120"N	16°5'12.971"E	360	HM
74	CZ, okr. Třebíč: Dukovany: okraj NPR Mohelenská hadcová step, bor na skalnatém svahu nad pravým břehem Jihlavy poblíž hráze Mohelenské přehrady	49°6'6.291"N	16°10'51.615"E	370	HM
75	CZ, okr. Hodonín: Velká nad Veličkou: NPR Zahrady pod Hájem, podrost řídkého borového háje	48°53'6.568"N	17°31'59.773"E	380	HM
76	CZ, okr. Praha-západ: Roztoky u Prahy (PR Údolí Únětického potoka): roztroušeně na kamenité stráni nad Trojanovým mlýnem	50°8'51.660"N	14°22'32.640"E	270	HM
77	CZ, okr. Mladá Boleslav: Kochánky: bílá stráň 1 km SSZ od obce.	50°17'16,8"N	14°46'25,6"E	230	ZK
78	CZ, okr. Rokycany: Ejovice: lesní lemy podél světlé lesní cesty nad S břehem vodní nádrže Ejovice 750 m SSZ od obce.	49°45'27,6"N	13°30'37,1"E	340	ZK
79	CZ, okr. Brno-venkov: Kanice: NPR Hádecká planinka, xerothermní trávníky nad lomovou stěnou	49°13'15.2"N	16°40'18.6"E	360	HM

Tab. I Seznam vzorků a lokalit, z kterých byly tyto vzorky odebrány. vz. = označení vzorku, n. v. = nadmořská výška v m n. m., písmena ve sloupci „col.“ označují původce sběrů: HM – H. Mašková, ZK – Z. Kaplan, FK – F. Krahulec, FKo – F. Kolář, VZ – V. Zeisek, MŠ – M. Štech, LP – L. Paštová, IK – I. Kaplanová, LB – L. Bureš, MD – M. Duchoslav.



## 8.2. Výsledky sekvenace DNA a mapování na referenční sekvence

vz.	Nr. reads	filtrování dle kvality	odf. (%)	odstranění duplikátů	dupl.(%)	Nr.paired reads	Nr. for. reads	Nr. rev. reads	Nr.map. nDNA	nDNA (%)	Nr.map. cpDNA	cpDNA (%)	Nr. reads po filtr.	Nr.map. rDNA	rDNA (%)
1	2053703	1912033	6.9	1757918	8.06	822022	106263	7420	190931	10.86	29919	1.70	1644048	4973	0.30
2	1583846	1551605	2.04	741085	52.24	356833	24191	3022	256626	34.63	3804	0.51	713666	1960	0.27
3	2267449	2136232	5.79	1887859	11.63	891756	94852	9304	182870	9.69	46124	2.44	1783513	5868	0.33
4	2024892	1952532	3.58	1454716	25.50	697035	54369	5829	446701	30.71	6880	0.47	1394070	1710	0.12
5	660448	652912	1.15	590912	9.50	292047	5944	704	204847	34.67	5845	0.99	584094	802	0.14
6	1404576	1356579	3.42	1055890	22.17	507772	35815	4192	361920	34.28	6236	0.59	1015546	2322	0.23
7	1026198	999174	2.64	491720	50.79	234218	21223	1887	171747	34.93	3039	0.62	468436	1809	0.39
8	1078328	1052494	2.4	504677	52.05	241620	18621	2713	178642	35.40	3583	0.71	483241	1954	0.40
10	1269738	1245818	1.89	565373	54.62	272546	17633	2437	200422	35.45	8773	1.55	545093	1841	0.34
11	700284	691987	1.19	630899	8.83	311726	6424	835	234674	37.20	3480	0.55	623452	613	0.10
12	1178354	1152506	2.2	535029	53.58	256572	19848	1898	184924	34.56	4502	0.84	513144	3267	0.64
13	617196	601870	2.49	294907	51.00	140913	11706	1314	95290	32.31	4815	1.63	281826	1585	0.56
14	1894960	1863455	1.67	785118	57.87	379195	22576	3929	283886	36.16	5649	0.72	758390	1927	0.25
15	613234	606663	1.08	552065	9.00	273017	5175	676	192425	34.86	1126	0.20	546034	913	0.17
16	1442246	1390173	3.62	1051796	24.34	504169	37518	5583	374660	35.62	5356	0.51	1008339	3074	0.30
17	2669331	2520489	5.58	2245644	10.90	1064319	105753	11046	205090	9.13	165590	7.37	2128640	9965	0.47
18	1569468	1538236	1.99	738174	52.01	356077	23145	2637	244578	33.13	4410	0.60	712154	2648	0.37
19	2137060	1995706	6.62	1838653	7.87	862782	104558	8356	164330	8.94	71419	3.88	1725566	3508	0.20
20	547674	542287	0.99	497306	8.29	246210	4126	615	186133	37.43	3470	0.70	492420	665	0.14
21	1312544	1267755	3.42	987289	22.12	474848	33528	3763	322102	32.62	3378	0.34	949696	1357	0.14
22	753944	746986	0.93	672449	9.98	333051	5291	837	277732	41.30	3848	0.57	666102	703	0.11
23	647276	641237	0.94	579295	9.66	286891	4655	684	230207	39.74	2770	0.48	573782	511	0.09
24	1518258	1461625	3.74	1090595	25.38	521603	42968	4133	338014	30.99	6324	0.58	1043207	2571	0.25
25	711066	704852	0.88	645026	8.49	319624	4736	812	265473	41.16	2112	0.33	639248	1433	0.22
26	1157534	1123977	2.9	651469	42.04	311553	24571	3610	223610	34.32	6967	1.07	623106	855	0.14
27	556710	551421	0.96	505914	8.25	250524	4003	668	213765	42.25	3558	0.70	501048	1339	0.27
28	1946110	1890799	2.85	1098816	41.89	526006	40448	6063	358222	32.60	13268	1.21	1052012	2897	0.28
29	1431140	1400128	2.17	634467	54.69	304091	23552	2546	217002	34.20	14021	2.21	608183	1842	0.30
30	1921424	1846452	3.91	1359505	26.37	648253	57481	5179	409379	30.11	11914	0.88	1296509	5615	0.43
31	1338880	1287055	3.88	994499	22.73	475506	39818	3441	316039	31.78	8454	0.85	951012	2442	0.26

vz.	Nr. reads	filtrování dle kvality	odf. (%)	odstranění duplikátů	dupl.(%)	Nr.paired reads	Nr. for. reads	Nr. rev. reads	Nr.map. nDNA	nDNA (%)	Nr.map. cpDNA	cpDNA (%)	Nr. reads po filtr.	Nr.map. rDNA	rDNA (%)
32A	1843432	1811362	1.74	784496	56.69	378531	22925	4295	298240	38.02	5803	0.74	757063	3175	0.42
32B	1992568	1925967	3.35	1372230	28.75	658056	50152	5607	444159	32.37	6317	0.46	1316115	1477	0.11
33	1630382	1509626	7.41	1400207	7.25	651124	92117	5711	144840	10.34	42637	3.04	1302250	4841	0.37
34	1380230	1316842	4.6	976075	25.88	461297	49475	3775	303516	31.10	4223	0.43	922594	1979	0.21
35	1672828	1620170	3.15	1124890	30.57	540318	39077	4893	363793	32.34	3330	0.30	1080636	2312	0.21
36	1323314	1268512	4.15	915138	27.86	434561	42201	3619	262031	28.63	12459	1.36	869122	1397	0.16
37	1418642	1366131	3.71	915994	32.95	436076	39544	4093	271393	29.63	5613	0.61	872152	1800	0.21
38	775508	768807	0.87	687812	10.54	340772	5083	940	293988	42.74	5256	0.76	681545	969	0.14
39	479056	466177	2.69	241760	48.14	115267	9826	1339	87422	36.16	1706	0.71	230534	633	0.27
40	478966	468795	2.13	194175	58.58	92806	7571	922	67434	34.73	1801	0.93	185612	437	0.24
41	701066	691817	1.32	629852	8.96	310782	7366	721	229317	36.41	3401	0.54	621564	546	0.09
42	1502556	1476402	1.75	702415	52.42	340282	18629	2957	262604	37.39	7958	1.13	680565	2250	0.33
45	661424	655331	0.93	597123	8.88	295795	4620	701	227639	38.12	2493	0.42	591590	410	0.07
46	1410128	1348379	4.38	952389	29.37	450016	48831	3354	229985	24.15	17805	1.87	900033	2028	0.23
47	690344	683777	0.96	603208	11.78	298638	5063	720	230683	38.24	1236	0.20	597276	900	0.15
48	1691408	1628702	3.71	941569	42.19	444542	47291	5029	249159	26.46	4087	0.43	889085	1122	0.13
49	1599140	1544396	3.43	911405	40.99	432687	40914	4936	260467	28.58	5768	0.63	865376	2206	0.25
50	1631150	1598745	1.99	782756	51.04	377856	23994	2787	264266	33.76	4378	0.56	755712	2145	0.28
51	1231424	1207924	1.91	579300	52.04	279699	17244	2446	219881	37.96	2797	0.48	559398	355	0.06
52	1196468	1149903	3.9	819530	28.73	390239	35439	3442	250394	30.55	2605	0.32	780479	2130	0.27
53	1803648	1774764	1.61	801939	54.81	388897	19951	3889	319015	39.78	3798	0.47	777794	4486	0.58
54	1911464	1868678	2.24	901340	51.77	433157	30905	3851	333230	36.97	15404	1.71	866315	6183	0.71
55	1310092	1284383	1.97	622337	51.55	300348	19019	2452	202423	32.53	13728	2.21	600696	3423	0.57
56	1241474	1211413	2.43	627885	48.17	301294	23263	1832	199513	31.78	3903	0.62	602590	805	0.13
57	1688524	1654753	2.01	782285	52.72	377078	24666	3173	257332	32.89	3039	0.39	754160	5781	0.77
58	1278398	1228639	3.9	726929	40.83	342494	38161	3640	188457	25.93	8024	1.10	684988	2006	0.29
59	1309002	1281851	2.08	582266	54.58	279426	20527	2716	198440	34.08	5190	0.89	558854	1289	0.23
60	1456360	1422914	2.3	722893	49.20	347302	25777	2263	240870	33.32	2196	0.30	694604	927	0.13
61	1330070	1278601	3.87	877572	31.36	417248	39107	3768	281358	32.06	5400	0.62	834498	1836	0.22
62	1479406	1425747	3.63	916351	35.73	435690	40406	4287	294851	32.18	7660	0.84	871380	1704	0.20
63	617128	610680	1.05	544796	10.79	269485	4911	717	201029	36.90	2427	0.45	538970	736	0.14
64	779500	773106	0.83	711423	7.98	352697	4596	1024	289190	40.65	447	0.06	705397	1169	0.17

vz.	Nr. reads	filtrování dle kvality	odf. (%)	odstranění duplikátů	dupl.(%)	Nr.paired reads	Nr. for. reads	Nr. rev. reads	Nr.map. nDNA	nDNA (%)	Nr.map. cpDNA	cpDNA (%)	Nr. reads po filtr.	Nr.map. rDNA	rDNA (%)
65	713164	690185	3.23	395870	42.64	187998	17855	1888	136093	34.38	949	0.24	375997	3468	0.92
66	706660	700224	0.92	600194	14.29	297127	4774	938	220058	36.66	2783	0.46	594255	1130	0.19
67	727062	720355	0.93	590452	18.03	292161	5075	888	217083	36.77	3776	0.64	584322	814	0.14
68	607874	601544	1.05	524013	12.89	259246	4692	670	197601	37.71	3329	0.64	518492	785	0.15
69	808472	801427	0.88	629610	21.44	311614	5283	956	250139	39.73	3964	0.63	623228	1258	0.20
70	696150	689501	0.96	580987	15.74	287469	4949	918	220256	37.91	4163	0.72	574942	1572	0.27
71	1158938	1117660	3.57	772014	30.93	368662	31031	3437	239893	31.07	8407	1.09	737324	2572	0.35
72	699076	691758	1.05	579941	16.16	286626	5547	925	214215	36.94	2502	0.43	573252	2163	0.38
73	489260	484271	1.02	394825	18.47	195266	3609	580	159237	40.33	2076	0.53	390533	630	0.16
74	1606954	1578617	1.77	762752	51.68	369668	20318	2871	259581	34.03	2550	0.33	739336	1274	0.17
75	674674	669570	0.76	580211	13.35	287720	3815	763	252388	43.50	3338	0.58	575440	1339	0.23
76	702376	696423	0.85	607867	12.72	301204	4426	863	233091	38.35	8290	1.36	602408	1158	0.19
77	1433322	1386563	3.27	855261	38.32	407986	34586	4547	255676	29.89	3528	0.41	815972	1191	0.15
78	602778	597105	0.95	512413	14.18	253622	4372	665	188344	36.76	1476	0.29	507245	369	0.07
79	1832780	1780385	2.86	1254889	29.52	605526	37575	5892	420122	33.48	1845	0.15	1211054	5744	0.47
<b>průměr</b>	<b>1233876.77</b>	<b>1197583.38</b>	<b>2.52</b>	<b>805593.56</b>	<b>30.33</b>	<b>387651.09</b>	<b>27108.43</b>	<b>2971.86</b>	<b>246271.91</b>	<b>33.18</b>	<b>9435.05</b>	<b>0.91</b>	<b>775302.91</b>	<b>2102.51</b>	<b>0.27</b>

Tab. II Výsledky sekvenace DNA a mapování na referenční sekvence. vz. = označení vzorku (viz tab. I výše). Nr. reads = počet získaných sekvencí (readů, reads) po odstranění PhiX, filtrování dle kvality = počet readů po filtrování prostřednictvím programu Trimmomatic (Bolger et al. 2014), odf.(%) = procento takto odfiltrovaných readů, odstranění duplikátů = počet readů po odstranění duplikátů, dupl.(%) = procento odstraněných duplikátů, Nr. paired reads = počet obousměrně osekvenovaných „reads“, Nr. for. reads a Nr. rev. reads = počet nespárovaných readů ve „forward“ a „reverse“ směru. Nr. map. nDNA a nDNA (%) = počet readů (resp. procento) namapovaných na referenci z cílených exonů, Nr. map. cpDNA a cpDNA (%) = počet readů (resp. procento) namapovaných na referenci chloroplastové DNA, Nr. reads. po filtr. = počet readů po vyfiltrování pro mapování na referenci rDNA. Nr. map. rDNA a rDNA = počet readů (resp. procento) namapovaných readů na referenci rDNA.

### 8.3. Seznam znaků použitých pro morfometrické analýzy

	<b>zkratka</b>	<b>znak</b>
1	vyska1	výška po vrchol terminálního úboru (mm)
2	vyska2	absolutní výška celé rostliny po vrchol nejhořejšího úboru (mm)
3	prumer	průměr lodyhy ve výšce 3 cm nad zemí (mm)
4	listy	počet listů na lodyze
5	uhel	úhel, který svírají listy s lodyhou, průměr ze tří listů ze střední části lodyhy
6	proh1	příčné prohnutí listu z dolní části lodyhy*
7	proh2	příčné prohnutí listu ze střední části lodyhy*
8	proh3	příčné prohnutí listu z horní části lodyhy*
9	terub	průměr terminálního úboru (mm)
10	sumub	počet úborů na rostlině
11	vnz	délka vnitřních (jazykovitých) zákrovních listenů (mm)*
12	stz	délka středních zákrovních listenů (mm)*
13	deldol	délka listů z dolní části lodyhy* (mm)
14	sirdol	šířka listů z dolní části lodyhy v polovině jejich délky (mm)*
15	sirdzar	šířka listů z dolní části lodyhy v zářezích mezi úkrojky (laloky) v polovině jejich délky (mm)*
16	sirdol2	šířka listů z dolní části lodyhy v jejich nejširší části (mm)*
17	sirdzar2	šířka listů z dolní části lodyhy v zářezích mezi úkrojky (laloky) (mm)*
18	bazdol	vzdálenost mezi nejširší částí listu z dolní části lodyhy a jeho bází (mm)*
19	numostd	počet ostnitých zubů na listu z dolní části lodyhy*
20	deldost	délka ostnů prvního řádu na listu z dolní části lodyhy (mm)*
21	deldkr	délka ostnů nejvyššího řádu na listu z dolní části lodyhy (mm)*
22	delpro	délka listů ze střední části lodyhy (mm)*
23	sirpro	šířka listů ze střední části lodyhy v polovině jejich délky (mm) *
24	sirpzar	šířka listů ze střední části lodyhy v zářezích mezi úkrojky (laloky) v polovině jejich délky (mm) *
25	sirpro2	šířka listů ze střední části lodyhy v jejich nejširším místě (mm) *
26	sirpzar2	šířka listů ze střední části lodyhy v zářezích mezi úkrojky (laloky) (mm)*
27	bazpro	vzdálenost mezi nejširší částí listu ze střední části lodyhy a jeho bází (mm)*
28	numostp	počet ostnitých zubů na listu ze střední části lodyhy*
29	delprost	délka ostnů prvního řádu na listu ze střední části lodyhy (mm) *
30	delpkr	délka ostnů nejvyššího řádu na listu ze střední části lodyhy (mm) *
31	delhor	délka listů z horní části lodyhy (mm) *
32	sirhor	šířka listů z horní části lodyhy v polovině jejich délky (mm) *
33	sirhzar	šířka listů z horní části lodyhy v zářezích mezi úkrojky (laloky) v polovině jejich délky (mm) *
34	sirhor2	šířka listů z horní části lodyhy v jejich nejširším místě (mm) *
35	sirhzar2	šířka listů z horní části lodyhy v zářezích mezi úkrojky (laloky) (mm) *
36	bazhor	vzdálenost mezi nejširší částí listu z horní části lodyhy a jeho bází (mm) *
37	numosth	počet ostének na listu z horní části lodyhy*

38	delhorst	délka ostnů prvního řádu na listu z horní části lodyhy (mm) *
39	delhkr	délka ostnů nejvyššího řádu na listu z horní části lodyhy (mm) *
40	dzl	délka vnějších zákrovních listenů (mm) *
41	numostz	počet ostnitých zubů na vnějších zákrovních listenech*
42	delostz	délka ostnů prvního řádu na vnějším zákrovním listenu*

Poměrové znaky:

43	pomvys	poměr výšek (celá rostlina/po terminální úbor)
44	olist	délka internodií (odpovídá podílu výšky rostliny po terminální úbor (vyska1) a počtu listů)
45	delsidol	poměr délky a šířky listu z dolní části lodyhy (v nejširším místě) *
46	delvzsivol	poměr délky listu z dolní části lodyhy a vzdálenosti jeho nejširšího místa od báze*
47	delsipro	poměr délky a šířky listu ze střední části lodyhy (v nejširším místě) *
48	delvzsipro	poměr délky listu ze střední části lodyhy a vzdálenosti jeho nejširšího místa od báze*
49	delsihor	poměr délky a šířky listu z horní části lodyhy (v nejširším místě) *
50	delvzsihor	poměr délky listu z horní části lodyhy a vzdálenosti jeho nejširšího místa od báze*
51	ostdol	poměr počtu ostének a délky listu z dolní části lodyhy*
52	ostpro	poměr počtu ostének a délky listu ze střední části lodyhy*
53	osthor	poměr počtu ostének a délky listu z horní části lodyhy*
54	izahd	index hloubky zářezu** u listů z dolní části lodyhy*
55	izahp	index hloubky zářezu** u listů ze střední části lodyhy*
56	izahh	index hloubky zářezu** u listů z horní části lodyhy*

Tab. III Seznam znaků (a jejich zkratk), které byly použity v rámci morfometrické analýzy

\* V případě znaků na listech a listenech (znaky 6–8 a 11–42, poměrové znaky 45–56) se jedná o aritmetický průměr hodnot z měření tří listů/listenů z jedné rostliny.

\*\* „Index hloubky zářezu“ (znak 54–56) byl definován jako poměr šířky listu (v nejširší části) a rozdílu šířek v nejširší části a v zářezích mezi úkrojky (laloky) (zde např. sirdol2/(sirdol2-sirdzar2)).



## 8.4. Skript použitý pro získání vstupních souborů do programu ToothFinder (Corney et al. 2012b)

```
#segment one.py - create .mat files for ToothFinder (Corney et al. 2012,
https://github.com/dcorney/ToothFinder)
#Written in 2018 by Ondřej Holešovský (Czech Technical University in Prague) and Hana Mašková
#(Charles University)
#(This code is released to the public domain. https://en.wikipedia.org/wiki/Public\_domain)

from skimage import data
from skimage import filter
import skimage
import numpy as np
from skimage import segmentation
from skimage.feature import corner_harris, corner_subpix, corner_peaks
import matplotlib.pyplot as plt
import scipy.io
import cv2
from PIL import Image

def threshold(image):
    # Otsu thresholding.
    # Use just the green channel for thresholding.
    green = np.mean(img, axis=2)
    val = skimage.filter.threshold_otsu(green)
    mask = green < val
    return mask

def biggest_contour(contours):
    ml = len(contours[0])
    contour = contours[0]
    for c in contours:
        if len(c) > ml:
            ml = len(c)
            contour = c
    return contour

def corners(img, mask):
    coords = corner_peaks(corner_harris(mask), min_distance=7)
    fig, ax = plt.subplots(figsize=(7, 7))
    ax.imshow(img)
    ax.scatter(coords[:,1], coords[:,0])
    plt.show()

import glob
import os

os.chdir("/directory/to/folder_with_TIFs/") #the directory containing my .tif files
for filename in glob.glob("*.tif"): #iterates over all files in the directory ending in .tif

#remove black spaces (was create as output from ImageJ straightening)
im = Image.open(filename)
data = np.array(im)
r1, g1, b1 = 0, 0, 0 # Original value
r2, g2, b2 = 255, 255, 255 # Value that we want to replace it with

red, green, blue = data[:, :, 0], data[:, :, 1], data[:, :, 2]
mask = (red == r1) & (green == g1) & (blue == b1)
data[:, :, :3][mask] = [r2, g2, b2]

im = Image.fromarray(data)
im.save(filename)

#add white frame (for a case of some objects on the margin of image)
img = skimage.io.imread(filename)
top, bottom, right, left = [100]*4
img = cv2.copyMakeBorder(img, top, bottom, left, right, cv2.BORDER_CONSTANT, value=[255, 255,
255])
```



Obr. I Skript `segment_one.py` (Holešovský, nepublikováno) s úpravami, které zahrnují doplnění bílého rámečku (pro případ, kdy do výběru zasahovala část sousedního listu) a odstranění černých ploch (ty vznikly v případě některých obrázků v programu ImageJ při použití funkce „Straighten“ při „narovnání“ prohnutých listů za účelem snadnějšího odečtu jejich délky – příklad před a po této úpravě viz obr. a výše. (Pokračování na další straně).

```

# Threshold the image (Otsu thresholding).
# Output a binary mask (background/foreground).
mask = threshold(img)
# Retrieve the boundaries between background/foreground.
boundaries = skimage.segmentation.find_boundaries(mask)
# Find connected contours among the boundaries.
contours = skimage.measure.find_contours(boundaries, 0)
# Return the biggest contour (the largest number of points).
contour = biggest_contour(contours)

# Plot the image and the biggest contour.
#fig, ax = plt.subplots()
#ax.imshow(img)
#ax.plot(contour[:, 1], contour[:, 0], linewidth=1)

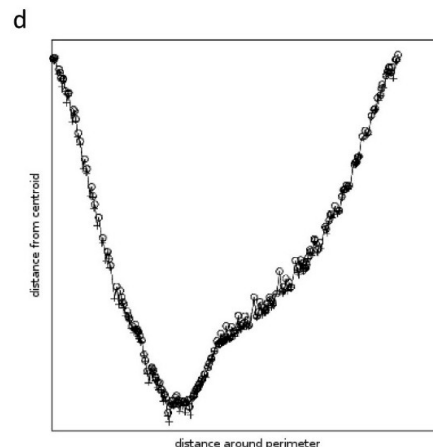
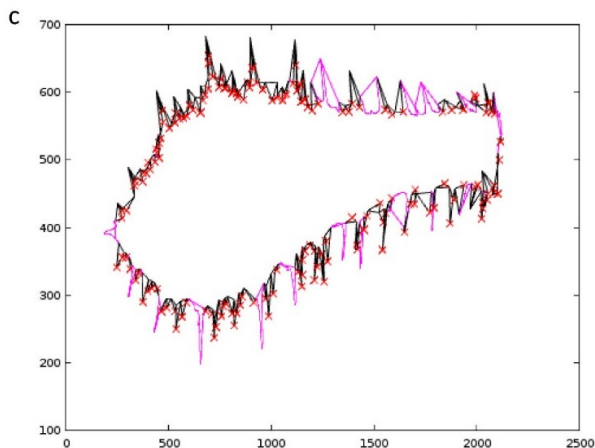
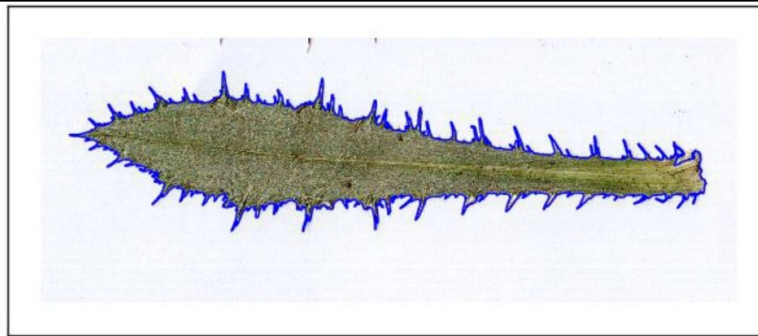
#ax.axis('image')
#ax.set_xticks([])
#ax.set_yticks([])
#plt.show()

#contour = np.int32(contour)
#cr = mask * 0
#cr[contour[:, 0], contour[:, 1]] = 1

    scipy.io.savemat(( filename.rsplit( ".", 1 )[ 0 ] )+ ".mat", dict(x=contour[:, 0],
y=contour[:, 1], imExt=img))

```

b



V rámci skriptu je proveden výběr objektů na základě prahování (tresholding), a pak jsou nalezeny jejich hranice. Do souboru .mat jsou zapsány pouze souřadnice okrajů největšího objektu v obrázku (listu). Jejich vizualizaci viz na obr. b. Takto získané soubory byly dále zpracovány programem ToothFinder. Obr. c – grafické znázornění výstupu, křížky označují nalezené zuby a příslušnou konturu v systému souřadnic; křivka na obr. d představuje průběh algoritmu (na ose x je postup po obvodu listu, osa y odpovídá vzdálenosti od centroidu listu, zub je lokálním maximem funkce (označen kolečkem), příslušný zářez (sinus) křížkem). Podrobný popis viz (Corney et al. 2012b). Počet zubů nalezený algoritmem v tomto konkrétním případě byl 185.

## 8.5. Popisné statistiky měřených znaků

znak	<i>Carlina vulgaris</i> & <i>C. biebersteinii</i> subsp. <i>brevibracteata</i>						<i>C. biebersteinii</i> subsp. <i>biebersteinii</i> <i>C. biebersteinii</i> subsp. <i>sudetica</i>					
	min.	5. p.	průměr ± sd.	95. p.	max.		min.	5. p.	průměr ± sd.	95. p.	max.	
vyska1	111.00	227.50	<b>415.24</b> ± 148.13	691.00	1065		258.00	341.05	<b>579.33</b> ± 143.23	858.55	933	
vyska2	111.00	232.50	<b>449.74</b> ± 162.77	744.50	1144.		258.00	341.05	<b>613.17</b> ± 161.80	923.55	1000	
listy	15.00	24.50	<b>42.31</b> ± 11.08	60.00	83.00		13.00	23.45	<b>40.94</b> ± 12.47	62.55	71.00	
terub	8.10	10.05	<b>14.24</b> ± 3.03	19.35	25.40		9.10	13.00	<b>19.00</b> ± 3.94	25.37	32.50	
sumub	1.00	1.00	<b>9.12</b> ± 16.31	45.00	100.00		1.00	1.00	<b>7.50</b> ± 8.82	27.55	41.00	
deldol	27.40	61.18	<b>101.16</b> ± 28.29	150.66	199.47		13.70	44.45	<b>117.25</b> ± 43.91	188.87	211.72	
sirdol	5.76	7.96	<b>13.15</b> ± 3.72	20.28	26.81		1.26	4.51	<b>13.06</b> ± 5.51	22.00	26.25	
sirdzar	3.83	5.15	<b>8.45</b> ± 2.49	12.61	19.75		0.95	3.94	<b>9.87</b> ± 3.88	16.21	20.12	
sirdol2	6.52	9.12	<b>15.08</b> ± 4.20	23.00	35.14		1.48	4.80	<b>14.36</b> ± 6.07	24.29	27.65	
sirdzar2	3.79	6.14	<b>9.83</b> ± 3.08	14.42	38.17		1.20	4.04	<b>10.67</b> ± 4.21	17.45	21.18	
bazdol	12.60	29.68	<b>59.91</b> ± 20.01	97.08	127.23		8.41	20.45	<b>64.74</b> ± 28.90	114.31	138.07	
delpro	18.12	23.19	<b>36.70</b> ± 10.35	56.45	76.33		6.08	35.07	<b>62.73</b> ± 19.55	101.39	115.69	
sirpro	4.78	7.09	<b>11.82</b> ± 3.30	18.73	23.57		0.92	5.47	<b>10.97</b> ± 3.90	16.44	18.64	
sirpzar	3.34	4.57	<b>7.37</b> ± 2.15	11.55	16.38		0.65	4.61	<b>7.68</b> ± 2.38	11.66	12.97	
sirpro2	5.41	8.06	<b>13.00</b> ± 3.56	20.34	28.95		0.99	6.86	<b>12.31</b> ± 4.07	19.10	20.69	
sirpzar2	3.58	5.19	<b>8.34</b> ± 2.43	12.71	20.69		0.70	5.34	<b>8.58</b> ± 2.68	13.40	16.57	
bazpro	5.90	7.69	<b>14.16</b> ± 5.64	24.54	45.11		2.38	10.62	<b>22.90</b> ± 10.24	40.71	47.40	
delhor	12.48	14.81	<b>20.21</b> ± 3.79	27.15	31.39		2.84	21.38	<b>34.25</b> ± 8.59	44.66	55.16	
sirhor	3.22	4.39	<b>7.02</b> ± 1.75	10.39	13.46		0.67	4.19	<b>6.61</b> ± 2.04	10.13	11.05	
sirhzar	2.00	2.51	<b>4.00</b> ± 1.03	5.89	7.52		0.37	3.16	<b>4.26</b> ± 1.10	6.07	6.43	
sirhor2	3.87	5.20	<b>7.86</b> ± 1.82	11.27	15.37		0.84	5.08	<b>7.96</b> ± 2.12	11.20	12.17	
sirhzar2	2.01	2.93	<b>4.41</b> ± 1.10	6.54	8.24		0.46	3.22	<b>4.98</b> ± 1.34	7.20	7.87	
bazhor	3.36	4.76	<b>7.64</b> ± 2.45	12.25	21.75		0.95	5.71	<b>10.18</b> ± 2.98	15.32	17.76	
delhorst	0.99	1.35	<b>1.79</b> ± 0.34	2.43	3.43		0.20	1.44	<b>1.82</b> ± 0.36	2.20	2.38	
delhkr	0.11	0.15	<b>0.24</b> ± 0.13	0.56	0.89		0.02	0.14	<b>0.23</b> ± 0.09	0.42	0.52	
dzl	10.76	11.94	<b>16.16</b> ± 3.05	21.94	26.07		2.28	17.69	<b>27.72</b> ± 6.66	37.89	41.65	
delostz	1.08	1.39	<b>1.84</b> ± 0.35	2.38	3.74		0.22	1.43	<b>1.97</b> ± 0.35	2.36	2.45	
pomvys	0.91	1.00	<b>1.11</b> ± 0.59	1.29	11.64		1.00	1.00	<b>1.06</b> ± 0.06	1.17	1.26	
olist	3.81	5.72	<b>10.00</b> ± 3.00	14.67	30.94		7.28	10.17	<b>14.83</b> ± 3.59	20.25	22.88	
delvzsídol	1.28	1.42	<b>1.74</b> ± 0.26	2.28	3.16		1.45	1.58	<b>1.89</b> ± 0.32	2.54	3.42	
delsídol	2.83	4.08	<b>6.97</b> ± 2.05	10.61	13.87		3.87	5.32	<b>8.60</b> ± 1.97	11.43	13.23	
delsipro	1.54	1.95	<b>2.91</b> ± 0.75	4.41	6.08		2.71	3.72	<b>5.27</b> ± 1.14	7.43	8.35	
delvzsipro	1.60	1.87	<b>2.72</b> ± 0.51	3.57	4.58		1.93	2.13	<b>2.94</b> ± 0.59	3.96	4.18	
delsihor	1.59	1.84	<b>2.66</b> ± 0.60	3.81	5.07		3.24	3.32	<b>4.36</b> ± 0.77	5.55	6.89	
delvzsíhor	1.42	1.82	<b>2.79</b> ± 0.64	4.12	5.19		2.22	2.66	<b>3.45</b> ± 0.72	5.05	6.44	
prumer	1.10	2.20	<b>4.23</b> ± 2.09	8.70	14.30		2.00	2.47	<b>5.83</b> ± 2.61	11.00	12.70	
deldost	0.77	1.19	<b>2.01</b> ± 1.80	2.98	31.76		0.86	1.01	<b>1.91</b> ± 0.54	2.80	3.23	
deldkr	0.12	0.15	<b>0.25</b> ± 0.14	0.57	1.12		0.11	0.13	<b>0.46</b> ± 1.99	0.41	16.87	
delprost	1.10	1.36	<b>1.91</b> ± 0.41	2.66	3.62		1.09	1.28	<b>2.05</b> ± 0.47	2.74	3.06	
delpkr	0.13	0.15	<b>0.25</b> ± 0.14	0.52	1.32		0.13	0.14	<b>0.24</b> ± 0.10	0.43	0.69	
izahd	-0.57	2.11	<b>2.98</b> ± 0.66	4.07	5.66		2.30	2.60	<b>4.58</b> ± 1.55	7.59	9.49	
izahp	1.86	2.16	<b>2.89</b> ± 0.64	3.70	10.29		-3.23	2.30	<b>3.50</b> ± 1.37	5.50	8.88	
izahh	1.55	1.85	<b>2.34</b> ± 0.34	2.99	3.59		1.92	2.05	<b>2.76</b> ± 0.50	3.63	4.27	
uhel	28.67	31.79	<b>60.76</b> ± 22.38	101.10	120.19		18.10	20.70	<b>59.82</b> ± 29.69	108.20	108.50	
proh1	82.17	96.66	<b>131.39</b> ± 20.47	165.76	194.03		71.63	71.74	<b>107.44</b> ± 30.66	160.22	174.32	
proh2	48.63	92.75	<b>129.28</b> ± 23.45	163.99	192.39		85.61	95.78	<b>135.92</b> ± 26.04	170.09	176.64	

znak	<i>Carlina vulgaris</i> & <i>C. biebersteinii</i> subsp. <i>brevibracteata</i>						<i>C. biebersteinii</i> subsp. <i>biebersteinii</i> <i>C. biebersteinii</i> subsp. <i>sudetica</i>					
	min.	5. p.	průměr ± sd.	95. p.	max.		min.	5. p.	průměr ± sd.	95. p.	max.	
proh3	92.59	114.43	<b>149.11</b> ± 21.09	181.67	205.53		115.05	123.84	<b>156.10</b> ± 21.09	184.75	190.07	
numostd	20.67	56.45	<b>147.02</b> ± 59.38	246.68	351.33		109.67	113.87	<b>220.73</b> ± 82.55	339.83	373.67	
numostp	19.00	39.02	<b>69.19</b> ± 24.93	116.98	150.33		78.67	88.23	<b>136.33</b> ± 27.69	170.43	173.00	
numosth	18.67	26.05	<b>41.67</b> ± 17.18	60.62	175.33		55.67	56.37	<b>76.71</b> ± 16.95	104.10	113.67	
numostz	15.33	24.00	<b>37.70</b> ± 13.72	53.27	144.00		17.60	24.15	<b>44.91</b> ± 16.01	68.10	70.67	
ostdol	0.23	0.55	<b>1.51</b> ± 0.61	2.62	3.40		0.92	1.16	<b>1.81</b> ± 0.57	2.75	2.76	
ostpro	0.31	1.00	<b>1.86</b> ± 0.55	2.71	4.04		1.14	1.25	<b>1.80</b> ± 0.36	2.35	2.54	
osthor	0.93	1.27	<b>2.02</b> ± 0.84	2.96	8.77		1.03	1.13	<b>1.98</b> ± 0.60	2.86	3.28	
vnz	10.37	13.42	<b>16.86</b> ± 1.82	19.23	21.73		14.87	15.26	<b>17.46</b> ± 1.88	20.52	20.57	
stz	5.14	6.78	<b>8.69</b> ± 1.17	10.60	12.25		6.78	6.98	<b>9.01</b> ± 1.06	11.32	12.87	

Tab. IV Souhrnná statistika pro měřené znaky použité v morfometrických analýzách pro dvě molekulárně odlišené skupiny. Prvních 34 znaků bylo měřeno celkem na 401 rostlinách (331 z první skupiny, 70 z druhé), zbylých 13 (odděleny tučnou čarou) bylo měřeno na 197 rostlinách (182 z první skupiny, 15 z druhé). min. = nejnižší naměřená hodnota pro daný znak, 5. p. = 5. percentil, sd. = směrodatná odchylka, 95. p. = 95. percentil, max. = nejvyšší naměřená hodnota pro daný znak. Vysvětlení zkratk znaků viz v tab. III výše.

## 8.6. Přehled balíčků a funkcí statistického programu R použitých v této práci

balíček	funkce	použití
ape 5.1 (Paradis et al. 2004)	read.tree()	načte soubory ve formátu .tre případně výstupy programu RAxML
	plot.phylo()	vizualizuje fylogenetické stromy
	identify.phylo()	vypisuje číslo nodu po kliknutí do fylogramu, pro nalezení uzlu, v kterém byly stromy zakořeněny
	cophyloplot()	porovnává 2 topologie
phangorn 2.4.0 (Schliep 2011)	read.nexus.networx()	načte soubory generované programem SplitsTree
	plot.networx()	vizualizace fylogenetických sítí
phytools 0.6.59 (Revell 2012)	multiRF()	počítá Robinson-Fouldova vzdálenost pro porovnání topologií stromů
vegan 2.5.2 (Oksanen et al. 2018)	vegdist()	výpočet distančních matic
	decostand()	standardizace dat
MVN 5.5 (Korkmaz et al. 2014)	mvn()	testy mnohorozměrného normálového rozdělení
MASS 7.3.49 (Venables & Ripley 2002)	lda()	výpočet lineární a kvadratické diskriminační analýzy
	qda()	
dendextend 1.8.0 (Galili 2015)	labels_colors()	grafická úprava dendrogramů
caret 6.0.80 (Kuhn 2018)	findCorrelation()	pro detekci silně korelovaných znaků
	findLinearCombos()	pro detekci lineárních kombinací znaků
ade4 1.7.11 (Dray & Dufour 2007)	dudi.pca()	výpočet PCA
	rtest()	test inercií
	discrimin()	výpočet DCA
	randtest()	randomizační test
	mantel.randtest()	Mantelův test pro porovnání dvou distančních matic
	coineria()	koinerční analýza
FactoMineR 1.41 (Lê et al. 2008)	PCA()	PCA, procentuální příspěvky k vysvětlené variabilitě
klaR 0.6.14 (Weihs et al. 2005)	greedy.wilks()	výpočet Wilkova lambda
sp 1.3.1 (Pebesma & Bivand 2005)	SpatialPoints()	vytváří na základě souřadnic objekt třídy SpatialPoints
raster 2.6.7 (Hijmans 2017)	getData()	stahuje data WorldClim
	extract()	k bodům (SpatialPoints) doplní proměnné prostředí (WorldClim)

Tab. V Seznam použitých balíčků (včetně citací) a funkcí statistického programu R použitých v rámci této práce.



## 8.7. Výstupy PCA a CDA

znaky (401 vzorků)	PCA1 zahrada	PCA2 zahrada	PCA1 terén	PCA2 terén	PCA1 vše	PCA2 vše	CDA zahrada	CDA terén	CDA vše
delhor	-	-	<b>6.72</b>	1.17	6.69	1.73	-	<b>0.78</b>	<b>0.81</b>
dzl	4.43	1.80	<b>6.70</b>	0.52	<b>6.78</b>	1.17	<b>0.77</b>	<b>0.78</b>	<b>0.80</b>
delsipro	1.89	<b>8.25</b>	0.99	6.48	0.97	<b>7.76</b>	<b>0.79</b>	<b>0.74</b>	<b>0.75</b>
delsihor	3.04	<b>8.23</b>	1.06	<b>9.95</b>	1.98	<b>9.94</b>	<b>0.72</b>	0.67	0.73
delpro	<b>8.64</b>	1.35	<b>6.75</b>	1.96	<b>7.26</b>	2.08	0.54	0.64	0.66
izahd	0.50	4.02	0.14	6.00	0.03	5.71	0.56	0.61	0.58
olist	0.07	4.12	0.33	7.54	0.03	<b>6.26</b>	-0.51	-0.56	-0.48
terub	5.46	0.20	3.62	1.15	5.19	0.93	0.32	0.52	0.52
<b>bazpro</b>	<b>9.52</b>	0.24	6.08	0.03	<b>6.69</b>	0.15	0.35	0.45	0.47
bazhor	0.58	0.87	5.33	0.28	4.19	0.29	0.35	0.43	0.43
vyska1	6.16	2.20	1.87	6.92	2.19	5.61	0.44	0.41	0.40
izahp	0.96	2.78	0.00	4.54	0.04	4.45	0.39	0.40	0.38
izahh	0.00	1.01	0.77	2.01	0.88	2.09	0.46	0.35	0.43
delsidol	0.34	6.46	0.10	4.95	0.07	6.05	0.43	0.31	0.30
<b>prumer</b>	<b>7.73</b>	0.26	4.32	1.51	5.69	0.53	-0.03	0.30	0.31
deldol	-	-	4.55	1.10	3.69	0.65	-	0.29	0.19
delostz	0.30	1.68	1.49	2.58	0.84	2.70	0.07	0.29	0.19
sirdzar	-	-	6.58	0.12	5.59	0.04	-	0.29	0.21
delvzsihor	1.22	5.22	0.15	4.43	0.13	5.21	0.38	0.29	0.39
sirhzar2	0.09	4.12	3.76	2.23	3.01	2.23	0.09	0.25	0.25
sirdzar2	6.16	1.05	6.50	0.09	5.46	0.55	0.02	0.19	0.11
pomvys	0.41	1.57	0.39	0.81	0.87	0.47	-0.28	-0.19	-0.08
sirhzar	0.40	4.55	3.64	2.41	2.78	2.39	-0.08	0.18	0.17
delhorst	0.03	3.81	1.17	4.14	0.78	4.35	-0.05	0.18	0.11
delvzsivol	0.68	0.28	0.08	1.63	0.06	1.13	0.31	0.18	0.24
bazdol	6.67	0.01	2.81	0.13	2.24	0.07	-0.02	0.18	0.07
delvzsipro	2.73	1.08	0.49	3.58	0.54	2.98	0.21	0.15	0.17
delprost	4.27	1.95	1.73	4.20	2.13	3.78	0.08	0.13	0.12
listy	6.51	0.18	1.44	0.06	2.45	0.03	-0.14	-0.12	-0.04
sirpro2	2.14	5.50	4.05	2.04	4.59	2.29	-0.39	-0.11	-0.07
<b>sirhor2</b>	0.02	<b>7.29</b>	2.66	5.10	1.99	5.40	-0.10	0.11	0.08
delpkr	0.25	1.69	0.01	2.54	0.02	2.38	-0.05	-0.07	-0.03
sumub	5.11	1.41	2.45	0.06	3.89	0.00	-0.47	-0.06	0.02
sirpzar2	1.79	3.10	4.12	0.34	4.40	0.57	-0.24	0.06	0.06
delhkr	1.16	0.46	0.05	1.88	0.00	1.61	0.00	0.02	0.01
deldost	5.05	1.12	0.87	3.01	0.86	2.78	0.13	-0.02	-0.01
sirdol2	5.44	2.76	6.10	1.38	4.96	2.42	-0.20	-0.01	-0.09
deldkr	0.00	0.58	0.11	1.11	0.05	1.20	-0.04	0.01	-0.01

Tab. VI Procentuální příspěvky k vysvětlené variabilitě prvních dvou os PCA a celkové strukturní koeficienty CDA pro soubor 401 jedinců a vybrané znaky. Zvláště pro rostliny pěstované a měřené v experimentální zahradě („zahrada“) a pro rostliny z terénu („terén“) a pro soubor sestávající z obou předchozích („vše“). Seřazeno podle nejvyšších hodnot strukturních koeficientů CDA. Tři nejvyšší hodnoty z každého sloupce vyznačeny tučně. Znaky, pro které nejsou ve sloupcích uvedeny hodnoty, byly vyřazeny z analýz datových souborů, z důvodu vysoké korelace s jinými znaky.

znaky (197 vzorků)	PCA1 zahrada	PCA2 zahrada	PCA1 terén	PCA2 terén	PCA1 vše	PCA2 vše	CDA zahrada	CDA terén	CDA vše
sirhor	4.13	1.32	-	-	0.42	<b>10.17</b>	-0.22	-	-0.02
dzl	2.37	2.15	<b>6.35</b>	0.00	<b>6.24</b>	0.00	<b>0.76</b>	<b>0.65</b>	<b>0.74</b>
delhor	-	-	<b>6.53</b>	0.03	<b>6.29</b>	0.01	-	<b>0.62</b>	<b>0.74</b>
delpro	<b>5.54</b>	0.30	5.27	1.49	<b>6.40</b>	0.93	0.57	<b>0.56</b>	<b>0.62</b>
terub	3.34	0.00	2.05	1.76	4.45	0.86	0.33	0.51	0.48
bazpro	4.80	0.00	5.26	0.01	5.76	0.02	0.44	0.49	0.51
olist	1.36	<b>6.72</b>	0.09	4.61	0.10	3.30	<b>0.67</b>	0.46	0.42
numosth	0.44	4.86	1.64	0.56	1.31	0.09	0.62	0.45	0.55
numostp	1.26	0.95	3.16	0.01	3.71	0.00	0.55	0.44	0.52
delsipro	4.07	3.96	0.72	2.63	1.29	3.23	<b>0.70</b>	0.43	0.60
prumer	3.22	2.15	2.34	5.56	4.94	1.75	0.07	0.42	0.37
sirdzar	0.44	3.65	5.31	0.70	3.99	0.26	-0.07	0.42	0.17
vyska1	<b>6.54</b>	0.73	0.61	<b>11.02</b>	1.75	6.46	0.34	0.41	0.36
sirdzar2	1.00	4.78	<b>6.07</b>	0.14	4.24	0.58	-0.16	0.39	0.11
deldol	<b>5.69</b>	0.38	3.98	0.82	3.08	0.32	0.13	0.39	0.20
proh1	4.73	0.00	0.00	4.47	1.06	3.98	-0.11	-0.37	-0.31
numostd	1.83	0.08	1.80	0.05	2.11	0.04	0.19	0.37	0.30
numostz	1.72	0.02	1.24	1.32	0.70	0.59	-0.01	0.35	0.15
sirdol2	0.88	<b>7.68</b>	6.01	0.01	4.23	0.99	-0.25	0.34	0.06
delsihor	4.22	4.00	0.46	<b>8.00</b>	2.22	<b>6.79</b>	0.73	0.33	0.60
bazhor	0.49	0.29	4.58	1.24	1.93	3.29	0.37	0.32	0.30
sirhzar2	2.10	0.01	2.93	3.67	1.25	6.64	0.22	0.30	0.26
vnz	0.10	0.26	1.44	0.74	0.52	0.99	-0.08	0.30	0.09
sumub	2.39	2.86	2.26	3.93	4.14	1.31	-0.35	0.28	0.18
sirhzar	2.81	0.05	3.18	3.74	1.40	6.29	0.10	0.27	0.22
delprost	0.01	4.76	1.43	1.85	1.35	2.34	-0.08	0.27	0.11
bazdol	-	-	2.88	0.02	1.96	0.04	-	0.26	0.08
sirpro2	0.11	3.95	4.62	0.08	4.63	0.57	-0.19	0.26	0.15
uhel	0.63	3.69	0.35	1.12	1.98	0.52	-0.41	-0.25	-0.06
sirprozar2	0.20	0.59	3.69	0.06	3.93	0.09	0.02	0.24	0.25
<b>sirhor2</b>	2.14	0.88	2.44	5.84	0.91	<b>8.77</b>	-0.04	0.23	0.09
izahd	0.21	6.43	0.10	0.54	0.00	0.38	0.43	0.19	0.34
izahh	0.24	4.46	0.17	1.22	0.23	0.26	0.58	0.18	0.42
listy	3.24	2.28	0.99	4.87	2.29	1.97	-0.27	0.16	0.03
sirdol	-	-	-	-	3.99	0.58	-	-	0.04
delhorst	0.89	1.18	1.37	3.81	0.62	4.77	-0.07	0.16	0.05
deldost	0.45	2.80	0.34	0.43	0.23	0.54	0.04	0.14	0.04
pomvys	0.01	2.60	0.59	0.25	1.67	0.22	-0.30	0.13	0.09
delvzsihor	4.16	0.58	0.47	3.47	0.53	4.84	0.29	0.13	0.34
proh2	3.06	1.25	0.40	5.37	0.01	4.08	0.27	-0.11	0.08
delvzsidol	0.81	1.93	0.06	2.34	0.01	0.53	0.39	0.10	0.28
proh3	0.52	1.68	1.19	0.38	0.09	0.73	0.28	0.10	0.10
deldkr	0.34	0.59	0.00	0.02	0.22	0.02	0.17	0.09	0.20
ostdol	1.30	0.19	0.03	0.70	0.00	0.29	0.10	0.09	0.14
delvzsipro	0.10	1.29	0.89	3.30	0.28	2.43	0.11	-0.07	0.04
stz	0.63	0.00	0.02	0.09	0.25	0.00	0.00	0.07	0.08
osthor	0.27	1.46	0.17	0.92	0.73	0.17	0.04	0.05	-0.01
ostpro	1.94	0.18	0.02	1.80	0.07	1.13	-0.05	-0.04	-0.02
delsidol	3.12	2.90	0.05	0.93	0.00	2.49	0.37	0.02	0.19
delhkr	1.57	0.54	0.04	1.56	0.11	1.28	0.33	-0.02	0.21

znaky (197 vzorků)	PCA1 zahrada	PCA2 zahrada	PCA1 terén	PCA2 terén	PCA1 vše	PCA2 vše	CDA zahrada	CDA terén	CDA vše
izahp	0.01	<b>6.49</b>	0.05	2.11	0.00	1.49	0.51	-0.01	0.32
delpkr	1.69	0.00	0.02	0.37	0.36	0.53	0.25	0.00	0.26

Tab. VII Procentuální příspěvky k vysvětlené variabilitě prvních dvou os PCA a celkové strukturní koeficienty CDA pro soubor 197 jedinců a všechny měřené znaky. Zvlášť pro rostliny pěstované a měřené v experimentální zahradě („zahrada“) a pro rostliny z terénu („terén“) a pro soubor sestávající z obou předchozích („vše“). Seřazeno podle nejvyšších hodnot strukturních koeficientů CDA. Tři nejvyšší hodnoty z každého sloupce vyznačeny tučně. Znaky, pro které nejsou ve sloupcích uvedeny hodnoty, byly vyřazeny z analýz daných datových souborů, z důvodu vysoké korelace s jinými znaky.