

Doktori (PhD) értekezés tézisei

**Molekuláris genetikai módszerek a
botanikai kutatásban**

Malkócs Tamás

Témavezető
Dr. Sramkó Gábor



DEBRECENI EGYETEM

Juhász-Nagy Pál Doktori Iskola

Debrecen, 2022

Bevezetés

Az emberiség növények irányába mutatott érdeklődését kezdetekben azoknak gyógyászati vagy élelemként való felhasználása hajtotta. Ebből az érdeklődésből született meg a 18. században, Linné által létrehozott kettős nevezéktannal a modern botanika tudománya. A növények osztályozását ekkor még kizárólag a szabad szemmel jól látható morfológiai karakterek alapján végezték, majd a mikroszkópia fejlődésével egyre apróbb jelek vizsgálata, és az ezek alapján való elkülönítés is lehetővé vált (Robertson, 1905).

Az 1800-as évek második felében Gregor Mendel borsón kísérletezve dolgozta ki az öröklődés elméletét, amellyel megalapozta a genetika tudományát (Kun, 2017). A kromoszómák, illetve annak felfedezésével, hogy a kromoszómák hordozzák az öröklődés egységeit, a géneket (Carlson, 2004), majd pedig a DNS örökítőanyagként való azonosításával megnyílt az út a molekuláris genetika létrejöttéhez (Avery és mtsai., 1944).

A molekuláris genetika eszköztárának bővülésével a mindössze néhány lókuszon található kis számú allélok vizsgálatától napjainkra eljutottunk az újgenerációs szekvenálásból származó, több ezer egynukleotid-

polimorfizmus (SNP) vizsgálatáig (Kreitman, 1983; Andrews és mtsai., 2016), illetve akár a teljes genomok összehasonlításáig (Henz és mtsai., 2005). A generált adatok mennyiségének nagymértékű növekedése a statisztikai keretrendszerek és a számítógépek számítási kapacitásának fejlődésével vált kezelhetővé.

Általánosan elmondható, hogy a legkorszerűbb molekuláris genetikai módszerek baktérium-, illetve állatmodellekben kerülnek kifejlesztésre, és csak ezt követően kerülnek növényekben is alkalmazásra. A magyar természetvédelmi botanikai kutatás a molekuláris genetikai vonatkozásokat tekintve különösen lemaradásban van, ezt mutatja az is, hogy az első ilyen témájú eredmények csak a 2000-es évek közepén kerültek közlésre (Gulyás és mtsai., 2005).

Dolgozatomban, mintegy követve a molekuláris genetika eszköztárának fejlődését, négy esettanulmányon keresztül mutatom be három módszer, a Sanger-szekvenálással előállított genetikai markerek, a mikroszatellit markerek, illetve a restriktációs helyhez kapcsolt DNS szekvenálás (RADseq) botanikai kutatásban való felhasználásának lehetőségeit, filogenetikai, illetve populációgenetikai megközelítéseken keresztül.

Célkitűzés

1. esettanulmány: Mikroszatellit markerek fejlesztése a fokozottan védett, tetraploid mocsári kardvirágban (*Gladiolus palustris* Gaud.)

Az esettanulmány célja volt (1) a mocsári kardvirág (*G. palustris*) populációgenetikai vizsgálatára alkalmas mikroszatellit markerek azonosítása; (2) ezen markerek populációs mintákon való tesztelése és populációgenetikai mérőszámokkal való jellemzése; (3) a markerek amplifikációjának tesztelése közeli rokon fajokban.

2. esettanulmány: Új SSR markerek fejlesztése a fokozottan védett, endemikus magyarföldi husángban (*Ferula sadleriana* Ledeb.)

Az esettanulmány célja volt (1) a *F. sadleriana* természetvédelmi kezelésében is használható mikroszatellit (SSR) markerek azonosítása; (2) ezen markerek populációs mintákon való tesztelése és jellemzése, valamint (3) amplifikációjuk közeli rokon fajokban való tesztelése.

3. esettanulmány: A veszélyeztetett, endemikus magyarföldi husáng (*Ferula sadleriana*) konzervációgenetikai vizsgálata

Az esettanulmány célja volt a (1) *Ferula* nemzetség filogenetikai törzsfájának rekonstrukciója a sejtmagi

ribosomális ITS (nrITS) szekvencia segítségével, a *F. sadleriana* nemzetségen belül elfoglalt helyének meghatározása érdekében; (2) a *F. sadleriana* populációk genetikai diverzitásának, populációi szerkezetének, és evolúciós történetének elemzése SSR markerek segítségével; (3) genetikailag megalapozott ajánlások megfogalmazása a faj megőrzése érdekében.

4. esettanulmány: A pontuszi sztyeppe eddig ismeretlen, kriptikus zsálya (*Salvia* L.) fajának azonosítása és jellemzése morfometriai és genomi módszerekkel

Az esettanulmány célja volt az általunk azonosított taxon (1) elterjedése földrajzi határainak megállapítása; (2) az általunk azonosított taxon és a *Salvia austriaca* Jacq. közötti, virágmorfológiában megfigyelhető különbség morfometriai vizsgálata; (3) az általunk azonosított taxon és a *S. austriaca* közötti genetikai viszony vizsgálata genomi (RADseq) módszerrel.

Módszerek

A terepen gyűjtött növényi szövetmintákat minden esetben szilikagélben szárítottuk és tároltuk, a genomi DNS kivonását pedig vagy egy módosított CTAB protokoll szerint, vagy a E.Z.N.A. Plant DNA DS Mini Kit (Omega Bio-tek, Inc., GA, USA) segítségével vontuk ki.

1. esettanulmány: Mikroszatellit markerek fejlesztése a fokozottan védett, tetraploid mocsári kardvirágban (*Gladiolus palustris* Gaud.)

A mikroszatellit markerek fejlesztését egy 454 piroszekvenáláson alapuló protokoll szerint végeztük. A több egyed DNS-ét tartalmazó könyvárat mikroszatellitekre dúsítottuk, majd szekvenáltattuk.

2. esettanulmány: Új SSR markerek fejlesztése a fokozottan védett, endemikus magyarföldi husángban (*Ferula sadleriana* Ledeb.)

A mikroszatellit markerek fejlesztéséhez a szekvenáló könyvtárat az Illumina Nextera Kit felhasználásával végeztük, és az Illumina MiSeq berendezésen szekvenáltuk.

1. és 2. esettanulmány:

A szekvenálást követően az első két esettanulmányban azonos módszertant használtunk. Az olvasásokból különböző

szűrésekkel mikroszatelliteket tartalmazó szekvenciákat azonosítottunk és ezekre primereket terveztünk. A primereket több lépcsőben, PCR és elektroforetikus módszerek segítségével teszteltük. A végül kiválasztott markereket szoftveresen, széles körben használt populációgenetikai mérőszámok segítségével jellemeztük.

3. esettanulmány: A veszélyeztetett, endemikus magyarföldi husáng (*Ferula sadleriana*) konzervációgenetikai vizsgálata

A filogenetikai elemzéshez nyolc *F. sadleriana* egyed, két *Ferula songarica* Pall. ex Schult. egyed, és egy-egy *Ferula tatarica* Fisch. ex Spreng. és *Ferula heuffelii* Griseb. ex Heuff. egyed nrITS szekvenciáját PCR-el felszaporítottuk, majd Sanger-szekvenálással szekvenáltattuk. Ezt kiegészítettük a GenBank-ben található *Ferula* nrITS szekvenciákkal, és az így kapott adatsorból ML (Maximum Likelihood) módszerrel filogenetikai rekonstrukciót végeztünk.

A populációgenetikai elemzéshez 13 mikroszatellit markert szaporítottunk fel összesen 109 *F. sadleriana* egyedben, PCR segítségével. Egy 11 lókuszt tartalmazó redukált adatsoron klaszteranalízist végeztünk. Végül, a faj evolúciós történetét ugyanezen adatsor segítségével, az ABC

(Approximate Bayesian Computation) keretrendszerben tártuk fel.

4. esettanulmány: A pontuszi sztyeppe eddig ismeretlen, kriptikus zsálya (*Salvia* L.) fajának azonosítása és jellemzése morfológiai és genomi módszerekkel

A taxon elterjedésének vizsgálatához internetes fotós adatbázisokból olyan *S. austriaca* előfordulásokat gyűjtöttünk, ahol jól azonosítható a virág morfológiája, majd térképen ábrázoltuk a két alak elterjedését.

A morfológiai vizsgálatához a két *Salvia* taxonban különböző vegetatív- és virágjellegeket mértünk, melyek eredményeit statisztikai elemzésekkel hasonlítottuk össze.

A genomi elemzéshez a RADseq módszerrel szekvenáló könyvtárat készítettünk a kérdéses *Salvia* taxon, a *S. austriaca*, és külsoportként három, közelrokon és szimpatikus elterjedésű faj a *Salvia nemorosa* L., *Salvia nutans* L. és *Salvia pratensis* L. egyedeiből. A könyvtárat az Illumina NextSeq platformján, 150 bp hosszúságú, egyirányú olvasásokat adó módszerrel szekvenáltuk. Az olvasásokat egy bioinformatikai folyamatrendszerben dolgoztuk fel, melynek eredményeképp egy egyedenként 2911 független SNP-et tartalmazó mátrixot kaptunk.

Ebből az adatsorból ML filogenetikai törzsfát rekonstruáltunk, többváltozós elemzéssel vizsgáltuk az egyedek közötti genetikai struktúrát a Főkomponensek Diszkriminancia-analízise (DAPC) segítségével, valamint Bayes-faktor alapú faji lehatárolást (BFD) és koaleszcencia alapú datált fajfa-rekonstrukciót végeztünk.

Eredmények

1. esettanulmány: Mikroszatellit markerek fejlesztése a fokozottan védett, tetraploid mocsári kardvirágban (*Gladiolus palustris* Gaud.)

- A *G. palustris* esetében 15 mikroszatellit markert azonosítottunk, melyek közül 11 bizonyult polimorfnak.
- A genetikai diverzitás értékek a két nagyobb populáció esetében normális értékeket mutattak, viszont a kis egyedszámú mátrai populáció rendkívül alacsony genetikai diverzitással rendelkezik.
- A fajok közötti amplifikáció a két közelrokon faj, a *G. imbricatus* és a *G. tenuis* esetében is négy lókuszban volt sikeres.

2. esettanulmány: Új SSR markerek fejlesztése a fokozottan védett, endemikus magyarföldi husángban (*Ferula sadleriana* Ledeb.)

- A *F. sadleriana* esetében 22 mikroszatellit (SSR) markert azonosítottunk, melyek közül 13 bizonyult polimorfnak.
- A genetikai diverzitás értékek alapján néhány lókuszban a Hardy-Weinberg egyensúlytól való eltérést tapasztaltunk, amely a null-allélok jelenlétére vezethető vissza, valamint három lókuszpár esetében genetikai kapcsoltságot figyeltünk meg.
- A 22 lókuszból a fajok közötti amplifikáció a két közelrokon faj, a *F. songarica* és a *F. tatarica* esetében 20 és 19 lókuszban volt sikeres.

3. esettanulmány: A veszélyeztetett, endemikus magyarföldi husáng (*Ferula sadleriana*) konzervációgenetikai vizsgálata

- Az nrITS szekvencia-mátrix filogenetikai vizsgálata alapján a *F. sadleriana* legközelebbi rokonai a *F. tatarica* és a *F. songarica*.
- A *F. sadleriana* kisebb egyedszámú populációi is jelentős genetikai változatosságot rejtenek.

- Nem figyeltünk meg szignifikáns távolság általi izolációt (IBD) a *F. sadlerianában*.
- Genetikai bizonyítékokat találtunk a populációk közötti áttelepítésekre, valamint kimutattuk, hogy a Kis-Kevély populáció forrása minden kétséget kizáróan a Pilis-tetői populáció volt.
- Bár nem találtunk egyértelmű bizonyítékokat a Kis-Kevély és Nagymaros populációk telepített eredetére, eredményeink nem is zárják ki ezeket. Bizonyítottuk viszont az utóbbi populáció kevert eredetét.

4. esettanulmány: A pontuszi sztyeppe eddig ismeretlen, kriptikus zsálya (*Salvia* L.) fajának azonosítása és jellemzése morfológiai és genomi módszerekkel

- Megállapítottuk, hogy az általunk vizsgált *Salvia* taxon elterjedése valószínűleg a Pontuszi sztyepp valódi sztyepp zónájára korlátozódik, míg a *S. austriaca* s.s. a Kárpát-medence, a Pontuszi sztyepp és a Balkán erdős sztyepp zónájában, valamint a Krím-félsziget déli részén található meg.
- Az általunk vizsgált *Salvia* taxon és a *S. austriaca* a vizsgált hatból öt virág-, és egy vegetatív jellegben is szignifikáns eltérést mutattak.

- A filogenomikai elemzés teljes bizonyossággal elkülönítette az általunk vizsgált taxonokat, melyek a koaleszcencia alapú datálás alapján ~197 ezer éve válhattak szét. A Bayes-faktor alapú faji lehatárolás is nagy bizonyossággal a két taxon faji szintű különállását támogatta.
- Bizonyítottuk, hogy a RADseq módszer teljes bizonyossággal képes volt elkülöníteni közelrokon *Salvia* fajokat.

Diszkusszió

Az általunk a *G. palustris*-ban fejlesztett 11 polimorf mikroszatellit marker a közölt mérőszámok alapján alkalmas a faj populációgenetikai vizsgálatára, valamint előrelépést jelentenek a faj vizsgálatára alkalmas markerek tekintetében, ugyanis a korábban használt, domináns AFLP markerek nem alkalmasak a teljes genetikai diverzitás feltérképezésére, valamint különböző vizsgálatok között is problémás az összehasonlításuk (Arnold & Emms, 1998). Az ezen markerekkel végzett vizsgálatok alkalmasak lehetnek a fokozottan védett faj természetvédelmi kezelésének genetikai megalapozására. Eredményeink valószínűsítik, hogy az

általunk azonosított markerek más, közeli rokon fajokban is alkalmazhatók lehetnek.

A *F. sadlerianában* fejlesztett 13 polimorf mikroszatellit marker alkalmasnak bizonyult a faj populációgenetika viszonyainak feltárására. Bár a null allélok jelenléte eltéréseket okoz a Hardy-Weinberg egyensúlytól, ez szoftveresen korrigálható, megbízhatóvá téve a további eredményeket (Chapuis & Estoup, 2007). Ezek az első mikroszatellit markerek, melyek bizonyítottan alkalmasak populációgenetikai vizsgálatok elvégzésére ezen fokozottan védett, endemikus fajban. Ugyanakkor, a nagyarányú fajok közötti amplifikáció valószínűsíti, hogy a markerek szélesebb körben is alkalmazhatóak a *Ferula* nemzetség körében.

Filogenetikai pozícióját illető eredményeink alapján a *F. sadleriana* a nemzetség eurázsiai sztyepp övet meghódító leszármazási vonalának legnyugatibb képviselője. A faj esetében megfigyelt magas genetikai diverzitás a faj hosszú élettartamának tulajdonítható, és egy erősen exogám szaporodási rendszert valószínűsít, melyek így kiegyenlítik a kis populációméret negatív hatását (Hamrick és mtsai., 1991; Karron, 1991; Riley és mtsai., 2010). A szignifikáns "Isolation by Distance" hiánya arra utal, hogy bár az egymáshoz

földrajzilag közel található populációk genetikailag kevésbé differenciáltak, ezt a mintázatot valószínűsíthetően a hosszú ideje tartó izoláció hatására bekövetkező genetikai sodródás elmosta. A Kis-Kevély populáció egyértelműen a Pilis-tetői populációból származik, ugyanis míg köztük rendkívül kicsi, a többi populáció és a Kis-Kevély között nagy F_{st} érték figyelhető meg, mely a Kis-Kevély populációt az alapító hatás klasszikus példájává teszi. Eredményeink genetikai bizonyítékokkal támasztják alá néhány, anekdotikus beszámolót a populációk közti természetvédelmi célú telepítésekről (Somlyay, 2007; Lendvay & Kalapos, 2009), valamint felveti korábban nem ismert telepítések eshetőségét is. A faj evolúciós történetének rekonstrukciója alapján nem bizonyítható, de nem is zárható ki a Kis-Kevély és a Nagymaros populációk telepített eredete, viszont valószínűsíti azt, hogy a faj populációi egy ősi, nagyméretű populációból eredeztethetők, mely néhány ezer–tízezer évvel ezelőtt darabolódhatott fel.

Az általunk azonosított *Salvia* taxon és a *S. austriaca* s.s. feltehetően vikariáns elterjedése arra utal, hogy a két taxon ökológiai igényei nagy valószínűséggel eltérőek. A morfometriai elemzések alapján míg a legtöbb

vegetatív jellegben nem találtunk eltérést, a legtöbb virágjelleg nagyfokú és konzisztens eltérést mutat, ami a zsályáknál a fajképződés fő mozgatórugója lehet (Claßen-Bockhoff és mtsai., 2004). Az összes filogenomikai eredményünk egybehangzóan támogatta a két taxon faji szintű elkülönülését, és az elválásuk idejét korábbra becsültük, mint két általánosan elfogadott faji szintű taxon, a *S. nemorosa* és a *S. pratensis* elválását. A RADseq módszer teljesítményéből arra következtetünk, hogy az alkalmas lehet akár az egész *Salvia* nemzetségen belül a fajok leszármazási viszonyainak feltárására.

A négy esettanulmány megmutatja, hogy különböző, a botanikai kutatásban felmerülő kérdésekre különböző genetikai módszerek lehetnek a legalkalmasabbak. Az nrITS szekvencia használata alacsony költségekkel, és kevés munkaórával kivitelezhető, viszont leginkább csak mély filogenetikai elválások rekonstrukciójára alkalmas. A mikroszatellit markereknek főleg a fejlesztése rendkívül költség- és időigényes, ezért (amennyiben nem állnak rendelkezésre már azonosított és tesztelt markerek), használatuk leginkább nagy egyedszámú, vagy hosszú távú populációgenetikai vizsgálatokban indokolt. A RADseq módszer a magas

költségek mellett egy könyvtárban limitált egyedszám vizsgálatát teszi lehetővé, viszont viszonylag kis munkával az előbbi két módszernél nagyságrendekkel több adatot kapunk. Hátrányai például, hogy kevésbé alkalmas a mély filogenetikai elválások vizsgálatára, valamint az adatokat feldolgozása során használt beállítások befolyásolják az eredményeket, ezért két vizsgálat közötti eredmények összehasonlítása kevésbé megbízható.

Köszönetnyilvánítás

Köszönettel tartozom témavezetőmnek, Sramkó Gábornak a kutatásaim során nyújtott segítségével és tanácsaiért. Köszönöm Laczkó Leventének az elmúlt évek során adott tanácsokat és segítőkészségét. Köszönöm Vincze Orsolyának a munkám során nyújtott támogatást. Köszönöm minden társszerzőmnek, akik szakértelmükkel és tanácsaikkal hozzájárultak a munkám színvonalának emeléséhez, Alexandru Sabin Badaraunak, Bereczki Juditnak, Cservenka Juditnak, Ivan Moysiyenkonak, Kuhn Thomasnak, Máti Attilának, Meglécz Emesének, Molnár V. Attilának, Shyryn Almerekovának, Szabó Annának, Szövényi Péternek. Köszönöm Fekete Rékának, Jordán Sándornak, Süveges Kristófnak és Takács Attilának a terepen nyújtott segítséget.



Nyilvántartási szám: DEENK/172/2022.PL
Tárgy: PhD Publikációs Lista

Jelölt: Malkócs Tamás
Doktori Iskola: Juhász-Nagy Pál Doktori Iskola
MTMT azonosító: 10057910

A PhD értekezés alapjául szolgáló közlemények

Idegen nyelvű tudományos közlemények külföldi folyóiratban (2)

1. **Malkócs, T.**, Laczkó, L., Bereczki, J., Meglécz, E., Szővényi, P., Sramkó, G.: Development and characterization of novel markers in the endangered endemic species *Ferula sadleriana*.
Appl. Plant Sci. 8 (2), 1-4, 2020. ISSN: 2168-0450.
DOI: <http://dx.doi.org/10.1002/aps3.11321>
IF: 1.936
2. **Malkócs, T.**, Almerekova, S., Bereczki, J., Cservenka, J., Meglécz, E., Sramkó, G.: Isolation and characterization of 15 SSR loci for the endangered European tetraploid species *Gladiolus palustris* (Iridaceae).
Appl. Plant Sci. 7 (5), 1-4, 2019. ISSN: 2168-0450.
DOI: <http://dx.doi.org/10.1002/aps3.1245>
IF: 1.591

További közlemények

Idegen nyelvű tudományos közlemények hazai folyóiratban (1)

3. Nagy, T., Nótári, K., Takács, A., **Malkócs, T.**, Tökölly, J., Molnár, V. A.: Precipitation and timing of flowering in ghost orchids (*Epipogium aphyllum* Sw.).
Acta bot. Hung. 60 (1-2), 223-230, 2018. ISSN: 0236-6495.
DOI: <http://dx.doi.org/10.1556/034.60.2018.1-2.10>

Idegen nyelvű tudományos közlemények külföldi folyóiratban (3)

4. **Malkócs, T.**, Viricel, A., Becquet, V., Evín, L., Dubillot, E., Pante, E.: Complex mitogenomic rearrangements within the Pectinidae (Mollusca: Bivalvia).
BMC Ecol. Evo. 22 (1), 1-21, 2022. EISSN: 2730-7182.
DOI: <http://dx.doi.org/10.1186/s12862-022-01976-0>





5. Fekete, R., Bódis, J., Fülöp, B., Süveges, K., Urgyán, R., **Malkócs, T.**, Vincze, O., Silva, L., Molnár, V. A.: Roadsides provide refuge for orchids: characteristic of the surrounding landscape.
Ecol. Evol. 10, 13236-1324, 2020. ISSN: 2045-7758.
IF: 2.912
6. Jamieson, A. J., **Malkócs, T.**, Piertney, S. B., Fujii, T., Zhang, Z.: Bioaccumulation of persistent organic pollutants in the deepest ocean fauna.
Nat. Ecol. Evol. 1 (3), 1-4, 2017. EISSN: 2397-334X.
DOI: <http://dx.doi.org/10.1038/s41559-016-0051>

Magyar nyelvű absztrakt kiadványok (1)

7. **Malkócs, T.**, Almerekova, S., Laczkó, L., Meglécz, E., Cservanka, J., Bereczki, J., Sramkó, G.: A *Gladiolus palustris* populációgenetikai vizsgálata a Kárpát-medencében = The population genetic structure of *Gladiolus palustris* in the Carpathian Basin.
In: XII. Aktuális flóra- és vegetációkutatás a Kárpát-medencében : nemzetközi konferencia : program és összefoglalók : Debrecen, 2018. február 23-25. = 12th International Conference "Advances in research on the flora and vegetation of the Carpato-Pannonian region" : programme and abstracts : Debrecen, 23-25 February 2018. Szerk.: Molnár V. Attila, Sonkoly Judit, Takács Attila, DE TTK Növénytani Tanszék, Debrecen, 38, 2018. ISBN: 9789634739265

Ismeretterjesztő, népszerűsítő cikkek (1)

8. Bak, H., Fekete, R., **Malkócs, T.**, Süveges, K.: Olajligetek és orchideák.
Egy. Élet 58 (9), 26-27, 2020. ISSN: 0230-7731.

A közlő folyóiratok összesített impact faktora: 6,439

A közlő folyóiratok összesített impact faktora (az értekezés alapjául szolgáló közleményekre): 3,527

A DEENK a Jelölt által az iDEa Tudóstérbe feltöltött adatok bibliográfiai és tudománymetriai ellenőrzését a tudományos adatbázisok és a Journal Citation Reports Impact Factor lista alapján elvégezte.

Debrecen, 2022.04.11.



**Short theses for the degree of doctor of philosophy
(PhD)**

**Molecular genetic methods in botanical
research**

Tamás Malkócs

Supervisor

Dr. Gábor Sramkó



UNIVERSITY OF DEBRECEN

Juhász-Nagy Pál Doctoral School

Debrecen, 2022

Introduction

Humanity's interest in plants was initially driven by their medicinal or food uses. From this interest, the science of modern botany was born in the 18th century, starting with the double nomenclature created by Carl Linnaeus. At that time, plants were classified solely on the basis of morphological characters visible to the naked eye, but with the development of microscopy, it became possible to examine and differentiate plants on the basis of characters of ever finer detail (Robertson, 1905).

In the second half of the 1800s, Gregor Mendel developed the theory of heredity by experimenting on peas, which laid the foundations for the science of genetics (Kun, 2017). The discovery of chromosomes, together with the fact that the chromosomes carry the units of heredity, the genes (Carlson, 2004), and the identification of DNA as the hereditary material, opened the way to the establishment of molecular genetics (Avery et al., 1944).

With the expansion of the toolkit of molecular genetics, we have moved from the study of a small number of alleles at just a few loci to the study of thousands of single nucleotide polymorphisms (SNPs) from next-generation sequencing

(Kreitman, 1983; Andrews et al., 2016) and even the comparison of whole genomes (Henz et al., 2005). The increasingly large amount of data generated has become manageable with the constant development of statistical frameworks and the computational power of computers.

Most often, state-of-the-art molecular genetic methods are developed in bacterial and animal models and only then are they transferred to plants. Hungarian conservation and botanical research are particularly lagging behind in the use of molecular genetics, as shown by the fact that the first results on this topic were published only in the mid-2000s (Gulyás et al., 2005).

In my thesis—in a way, following the development of the tools of molecular genetics—I present four case studies to illustrate the potential of three methods, Sanger sequencing, microsatellite markers and restriction site-associated DNA sequencing (RADseq), in botanical research, using phylogenetic and population genetic approaches.

Objectives

Case Study 1: Isolation and characterization of 15 SSR loci for the endangered European tetraploid species *Gladiolus palustris* (Iridaceae).

The objectives of this case study were (1) to identify microsatellite markers suitable for population genetic research in *Gladiolus palustris* Gaud.; (2) to test these markers on population samples and characterize them using population genetics metrics; (3) to test the amplification of markers in closely related species.

Case Study 2: Development and characterization of novel SSR markers in the endangered endemic species *Ferula sadleriana*.

The objectives of this case study were (1) to identify microsatellite (SSR) markers that can be used in the conservation management of *Ferula sadleriana* Ledeb.; (2) to test and characterize these markers in population samples and (3) to test their amplification in closely related species.

Case Study 3: Conservation genetics of the endemic endangered *Ferula sadleriana*.

The objectives of the case study were (1) to reconstruct the phylogenetic tree of the genus *Ferula* using the nuclear

ribosomal ITS (nrITS) sequence to determine the position of *F. sadleriana* within the genus; (2) to analyze the genetic diversity, population structure and evolutionary history of *F. sadleriana* populations using SSR markers; (3) to make genetically grounded recommendations for the conservation of the species.

Case Study 4: Identification and characterisation of a previously unknown species of cryptic sage (*Salvia* L.) from the Pontic steppe using morphometric and genomic methods.

The objectives of the case study were to (1) establish the geographical limits of the distribution of the taxon we identified; (2) morphometrically investigate the difference in floral morphology between the taxon we identified and *Salvia austriaca* Jacq.; (3) investigate the genetic relationship between the taxon we identified and *S. austriaca* using genomic (RADseq) methods.

Methods

In every case study, plant tissue samples collected in the field were dried and stored in silica gel, and genomic DNA was extracted either according to a modified CTAB protocol or using the E.Z.N.A. Plant DNA DS Mini Kit (Omega Bio-tek, Inc., GA, USA).

Case Study 1: Isolation and characterization of 15 SSR loci for the endangered European tetraploid species *Gladiolus palustris* (Iridaceae).

The development of microsatellite markers was carried out according to a protocol based on 454 pyrosequencing. A library containing DNA from several individuals was enriched for microsatellites and sequenced.

Case Study 2: Development and characterization of novel SSR markers in the endangered endemic species *Ferula sadleriana*.

For the development of microsatellite markers, the sequencing library was constructed using the Illumina Nextera Kit and sequenced on the Illumina MiSeq instrument.

Case studies 1 and 2:

The same methodology was used in the first two case studies following sequencing. Sequences containing microsatellites were identified from the reads using different screening methods and primers were designed for these sequences. The primers were tested in several steps using PCR and electrophoretic methods. The finally selected markers were characterized by software using widely used population genetic metrics.

Case Study 3: Conservation genetics of the endemic endangered *Ferula sadleriana*.

For the phylogenetic analysis, the nrITS sequences of eight *F. sadleriana* individuals, two individuals of *Ferula songarica* Pall. ex Schult., and one individual of *Ferula tatarica* Fisch. ex Spreng. and *Ferula heuffelii* Griseb. ex Heuff. each were PCR amplified and sequenced by Sanger sequencing. This was complemented with *Ferula* nrITS sequences downloaded from GenBank, and the resulting data set was used for phylogenetic reconstruction using ML (Maximum Likelihood).

For the population genetic analysis, 13 microsatellite markers were amplified in a total of 109 *F. sadleriana* individuals using PCR. Clustering analysis was performed on a

reduced data set containing 11 loci. Finally, the evolutionary history of the species was investigated using the same dataset in the ABC (Approximate Bayesian Computation) framework.

Case Study 4: Identification and characterisation of a previously unknown species of cryptic sage (*Salvia* L.) from the Pontic steppe using morphometric and genomic methods.

To investigate the distribution of the taxon, we collected occurrences of *S. austriaca* from online photographic databases where the flower morphology could be identified on the photo, and then mapped the distribution of the two different floral forms.

For the morphometric study, several vegetative and floral characters were measured in the two *Salvia* taxa, and the results were compared by statistical analysis.

For the genomic analysis, a sequencing library was constructed using the RADseq method from individuals of the *Salvia* taxon of interest, *S. austriaca*, and three closely related species with a sympatric distribution, *Salvia nemorosa* L., *Salvia nutans* L. and *Salvia pratensis* L., as outgroups. The library was sequenced on Illumina's NextSeq platform using a 150 bp length method with single-end reads. Reads were

processed in a bioinformatic pipeline, resulting in a matrix of 2911 independent SNPs per individual.

From this dataset, we reconstructed a ML phylogenetic tree, performed multivariate analysis of genetic structure between individuals using Discriminant Analysis of Principal Components (DAPC), Bayes-factor Species Delimitation (BFD) and coalescent-based dated species tree reconstruction.

Results

Case Study 1: Isolation and characterization of 15 SSR loci for the endangered European tetraploid species *Gladiolus palustris* (Iridaceae).

- Fifteen microsatellite markers were identified in *G. palustris*, 11 of which proved to be polymorphic.
- Genetic diversity values were as expected in the two larger populations, whereas the small population in the Mátra mountains showed extremely low genetic diversity.
- Interspecific amplification was also successful at four loci in the two closely related species, *G. imbricatus* and *G. tenuis*.

Case Study 2: Development and characterization of novel SSR markers in the endangered endemic species *Ferula sadleriana*.

- In *F. sadleriana*, 22 microsatellite (SSR) markers were identified, of which 13 were found to be polymorphic.
- Based on genetic diversity values, deviations from the Hardy-Weinberg equilibrium were observed in some loci due to the presence of null alleles, and genetic linkage was observed in three pairs of loci.
- Of the 22 loci, interspecific amplification was successful in 20 and 19 loci in the two closely related species, *F. songarica* and *F. tatarica*.

Case Study 3: Conservation genetics of the endemic endangered *Ferula sadleriana*.

- Based on phylogenetic analysis of the nrITS sequence matrix, the closest relatives of *F. sadleriana* are *F. tatarica* and *F. songarica*.
- Smaller populations of *F. sadleriana* also harbour considerable genetic diversity.
- No significant isolation by distance (IBD) was observed in *F. sadleriana*.

- We found genetic evidence for translocations between populations and showed that the source of the Kis-Kevély population was without doubt the population of Pilis Hill.
- Although we did not find clear evidence for the introduced origin of the Kis-Kevély and Nagymaros populations, our results do not exclude these scenarios either. However, we provide evidence for a mixed origin of the latter population.

Case Study 4: Identification and characterisation of a previously unknown species of cryptic sage (*Salvia* L.) from the Pontic steppe using morphometric and genomic methods.

- We found that the distribution of the *Salvia* taxon we studied is seemingly restricted to the true steppe zone of the Pontic steppe, while *S. austriaca* s.s. is found in the forest-steppe zone of the Carpathian Basin, the Pontic steppe and the Balkans, and in the southern part of the Crimean Peninsula.
- The *Salvia* taxon we studied and *S. austriaca* showed significant differences in five out of the six floral, and one out of four vegetative characters examined.

- The phylogenomic analysis separated the taxa we studied with complete certainty, which have diverged ~197,000 years ago based on coalescence dating. The Bayes-factor species delimitation also supported the species-level separation of the two taxa with high confidence.
- We demonstrated that the RADseq method is able to separate closely related *Salvia* species with complete confidence.

Discussion

The 11 polymorphic microsatellite markers we have developed in *G. palustris* are suitable for population genetics of the species based on the reported metrics, and represent an improvement in the markers suitable for the study of the species, as the previously used dominant AFLP markers are not able to reveal the full genetic diversity, and comparison between different studies is problematic (Arnold & Emms, 1998). Studies using the markers we have developed may be useful to provide a genetic foundation for conservation management of this highly protected species. Our results make

it likely that the markers we have identified can be applicable to other, closely related species.

The 13 polymorphic microsatellite markers developed in *F. sadleriana* have been shown to be suitable for elucidating the population genetic structure of this species. Although the presence of null alleles causes deviations from the Hardy-Weinberg equilibrium, this can be corrected computationally, making further results reliable (Chapuis & Estoup, 2007). These are the first microsatellite markers that are suitable for population genetic studies in this highly protected, endemic species. At the same time, the high degree of interspecific amplification makes it likely that the markers could be used more widely in the genus *Ferula*.

Based on our results on the phylogenetic position of *F. sadleriana*, it is the westernmost representative of the lineage of the genus that has colonised the Eurasian steppe belt. The high genetic diversity observed in this species can be attributed to its long lifespan and suggests a highly outcrossing reproductive system, which thus offsets the negative effect of small population size (Hamrick et al., 1991; Karron, 1991; Riley et al., 2010). The lack of significant Isolation by Distance suggests that although geographically close populations are less

differentiated genetically, this pattern is likely to have been obscured by genetic drift resulting from long periods of isolation. The Kis-Kevély population clearly descends from the Pilis Hill population, because while the F_{st} value between the two is extremely small, the F_{st} value between the Kis-Kevély and every other population is quite large, making the Kis-Kevély population a classic example of the founder effect. Our results provide genetic evidence to support some of the anecdotal reports of inter-population introductions for conservation purposes (Somlyay, 2007; Lendvay & Kalapos, 2009), and also raise the possibility of previously unknown introductions. Reconstructions of the evolutionary history of the species do not prove or exclude an introduced origin of the Kis-Kevély and Nagymaros populations, but suggest that the current populations of the species may have originated from an ancient, large population that may have been fragmented some thousands to tens of thousands of years ago.

The seemingly vicariant distributions of the *Salvia* taxon we identified and *S. austriaca* s.s. suggest that the ecological requirements of the two taxa are likely to be very different. Morphometric analyses suggest that while no differences were found in most vegetative characters, most

floral characters showed a high and consistent divergence, which can be a major driver of speciation in sages (Claßen-Bockhoff et al., 2004). All of our phylogenomic results were in agreement, supporting the species-level separation of the two taxa, and we estimated the time of their separation to be earlier than the separation of two generally accepted species-level taxa, *S. nemorosa* and *S. pratensis*. We conclude from the performance of the RADseq method that it may be suitable for exploring the phylogenetic relationships within the entire *Salvia* genus.

The four case studies show that different genetic methods may be best suited to address different questions in botanical research. The use of nrITS sequencing is feasible at low cost and with few person-hours, but most often it is only suitable for reconstructing deep phylogenetic separations. The development of microsatellite markers in particular is extremely costly and time-consuming, and therefore (where no previously developed and tested markers are available) their use is most appropriate when large numbers of individuals are to be studied, or in long-term population genetic studies. The RADseq method allows for the analysis of a limited number of individuals in a library at high cost, but with relatively little

work, and with it, orders of magnitude more data can be obtained than the first two methods. Disadvantages include its reduced suitability for studying deep phylogenetic separations and that the settings used for data processing can affect the results, making comparisons between different studies less reliable.

Acknowledgements

I would like to thank my supervisor Gábor Sramkó for his continuous help and support during my research. I thank Levente Laczkó for his advices and helpfulness over the years. I also thank Orsolya Vincze for her support during my work. I am grateful to all my co-authors who contributed to the quality of my work with their expertise and advice: Alexandru Sabin Badarau, Judit Bereczki, Judit Cservenka, Ivan Moysiyyenko, Thomas Kuhn, Attila Mátis, Emese Meglécz, Attila Molnár V., Shyryn Almerekova, Anna Szabó, Péter Szövényi. I appreciated the help of Réka Fekete, Sándor Jordán, Kristóf Süveges and Attila Takács during field-trips.



Registry number: DEENK/172/2022.PL
Subject: PhD Publication List

Candidate: Tamás Malkócs

Doctoral School: Pál Juhász-Nagy Doctoral School of Biology and Environmental Sciences

MTMT ID: 10057910

List of publications related to the dissertation

Foreign language scientific articles in international journals (2)

1. **Malkócs, T.**, Laczkó, L., Bereczki, J., Meglécz, E., Szövényi, P., Sramkó, G.: Development and characterization of novel markers in the endangered endemic species *Ferula sadleriana*.
Appl Plant Sci. 8 (2), 1-4, 2020. ISSN: 2168-0450.
DOI: <http://dx.doi.org/10.1002/aps3.11321>
IF: 1.936
2. **Malkócs, T.**, Almerekova, S., Bereczki, J., Cservenka, J., Meglécz, E., Sramkó, G.: Isolation and characterization of 15 SSR loci for the endangered European tetraploid species *Gladiolus palustris* (Iridaceae).
Appl. Plant Sci. 7 (5), 1-4, 2019. ISSN: 2168-0450.
DOI: <http://dx.doi.org/10.1002/aps3.1245>
IF: 1.591

List of other publications

Foreign language scientific articles in Hungarian journals (1)

3. Nagy, T., Nótári, K., Takács, A., **Malkócs, T.**, Tökölly, J., Molnár, V. A.: Precipitation and timing of flowering in ghost orchids (*Epipogium aphyllum* Sw.).
Acta bot. Hung. 60 (1-2), 223-230, 2018. ISSN: 0236-6495.
DOI: <http://dx.doi.org/10.1556/034.60.2018.1-2.10>

Foreign language scientific articles in international journals (3)

4. **Malkócs, T.**, Viricel, A., Becquet, V., Evin, L., Dubillot, E., Pante, E.: Complex mitogenomic rearrangements within the Pectinidae (Mollusca: Bivalvia).
BMC Ecol Evo. 22 (1), 1-21, 2022. EISSN: 2730-7182.
DOI: <http://dx.doi.org/10.1186/s12862-022-01976-0>





5. Fekete, R., Bódis, J., Fülöp, B., Süveges, K., Urgyán, R., **Malkócs, T.**, Vincze, O., Silva, L., Molnár, V. A.: Roadsides provide refuge for orchids: characteristic of the surrounding landscape.
Ecol. Evol. 10, 13236-1324, 2020. ISSN: 2045-7758.
IF: 2.912
6. Jamieson, A. J., **Malkócs, T.**, Pieltney, S. B., Fujii, T., Zhang, Z.: Bioaccumulation of persistent organic pollutants in the deepest ocean fauna.
Nat. Ecol. Evol. 1 (3), 1-4, 2017. EISSN: 2397-334X.
DOI: <http://dx.doi.org/10.1038/s41559-016-0051>

Hungarian abstracts (1)

7. **Malkócs, T.**, Almerekova, S., Laczkó, L., Meglécz, E., Cservenka, J., Bereczki, J., Sramkó, G.: A *Gladiolus palustris* populációgenetikai vizsgálata a Kárpát-medencében = The population genetic structure of *Gladiolus palustris* in the Carpathian Basin.
In: XII. Aktuális flóra- és vegetációkutatás a Kárpát-medencében : nemzetközi konferencia : program és összefoglalók : Debrecen, 2018. február 23-25. = 12th International Conference "Advances in research on the flora and vegetation of the Carpato-Pannonian region" : programme and abstracts : Debrecen, 23-25 February 2018. Szerk.: Molnár V. Attila, Sonkoly Judit, Takács Attila, DE TTK Növénytani Tanszék, Debrecen, 38, 2018. ISBN: 9789634739265

Informational/educational articles (1)

8. Bak, H., Fekete, R., **Malkócs, T.**, Süveges, K.: Olajligetek és orchideák.
Egy. Élet. 58 (9), 26-27, 2020. ISSN: 0230-7731.

Total IF of journals (all publications): 6,439

Total IF of journals (publications related to the dissertation): 3,527

The Candidate's publication data submitted to the iDEa Tudóstér have been validated by DEENK on the basis of the Journal Citation Report (Impact Factor) database.

11 April, 2022



Irodalomjegyzék – Bibliography:

- Andrews, K.R., Good, J.M., Miller, M.R., Luikart, G., & Hohenlohe, P.A. 2016. Harnessing the power of RADseq for ecological and evolutionary genomics. *Nat. Rev. Genet.* 17: 81–92.
- Arnold, M.L., & Emms, S.K. 1998. Molecular Markers, Gene Flow, and Natural Selection. *Mol. Syst. Plants II* 442–458.
- Avery, O.T., MacLeod, C.M., & McCarty, M. 1944. Studies on the chemical nature of the substance inducing transformation of pneumococcal types induction of transformation by a desoxyribonucleic acid fraction isolated from *Pneumococcus* type III. *J. Exp. Med.* 79: 137–158.
- Carlson, E.A. 2004. *Mendel's Legacy: The Origin of Classical Genetics*, Cold Spring Harbor: Cold Spring Harbor Laboratory Press.
- Chapuis, M.-P., & Estoup, A. 2007. Microsatellite Null Alleles and Estimation of Population Differentiation. *Mol. Biol. Evol.* 24: 621–631.
- Claßen-Bockhoff, R., Speck, T., Tweraser, E., Wester, P., Thimm, S., & Reith, M. 2004. The staminal lever mechanism in *Salvia* L. (Lamiaceae): A key innovation for adaptive radiation? *Org. Divers. Evol.* 4: 189–205.
- Gulyás, G., Sramkó, G., Molnár V., A., Rudnóy, S., Illyés, Z., Balázs, T., Bratek, Z., Molnár, V.A., Rudnóy, S., Illyés, Z., Balázs, T., & Bratek, Z. 2005. Nuclear ribosomal DNA ITS paralogs as evidence of recent interspecific hybridization in the genus *Ophrys* (Orchidaceae). *Acta Biol. Cracoviensia Ser. Bot.* 47: 61–67.
- Hamrick, J.L., Godt, M.J.W., Murawski, D.A., & Loveless, M.D. 1991. „Correlations between species traits and allozyme

- diversity: implications for conservation biology”, in Falk, D. A. & Holsinger, K. E. (szerk.) *Genetics and Conservation of Rare Plants*. 3–30.
- Henz, S.R., Huson, D.H., Auch, A.F., Nieselt-Struwe, K., & Schuster, S.C. 2005. Whole-genome prokaryotic phylogeny. *Bioinformatics* 21: 2329–2335.
- Karron, J. D. (1991) „Breeding Systems in Rare Plant Species”, in Falk, D. A. & Holsinger, K. E. (szerk.) *Genetics and Conservation of Rare Plants*. 87–89.
- Kreitman, M. 1983. Nucleotide polymorphism at the alcohol dehydrogenase locus of *Drosophila melanogaster*. *Nature* 304: 412–417.
- Kun, Á. 2017. *Evolúcióbíológia*. Budapest: Typotex.
- Lendvay, B., & Kalapos, T. 2009. A magyarföldi husáng (*Ferula sadleriana*) populációinak állapotfelmérése 2008-ban. *Természetvédelmi Közlemények* 15: 486–492.
- Riley, L., McGlaughlin, M.E., & Helenurm, K. 2010. Genetic diversity following demographic recovery in the insular endemic plant *Galium catalinense* subspecies *acrispum*. *Conserv. Genet.* 11: 2015–2025. <https://doi.org/10.1007/s10592-010-0094-8>
- Robertson, A. 1905. Cytology and classification. *New Phytol.* 4: 134–139.
- Somlyay, L. 2007. A *Ferula sadleriana* Ledeb. „újabb” hazai lelőhelye. *Kiteibelia* XII.: 106–107.