
EÖTVÖS LORÁND TUDOMÁNYEGYETEM

TERMÉSZETTUDOMÁNYI KAR

BIOLÓGIAI INTÉZET

GENETIKAI TANSZÉK



**Gímszarvasok egyedazonosítása és populációgenetikai vizsgálata
autoszómás tetranukleotid mikroszatellita markerekkel**

DOKTORI ÉRTEKEZÉS TÉZISEI

Készítette:

Szabolcsi Zoltán

BIOLÓGIA DOKTORI ISKOLA

Iskolavezető: Dr. Erdei Anna, MTA rendes tagja

KLASSZIKUS ÉS MOLEKULÁRIS GENETIKA DOKTORI PROGRAM

Programvezető: Dr. Orosz László, MTA rendes tagja

Témavezető: Dr. Orosz László, MTA rendes tagja

Budapest

2013

1. BEVEZETÉS

Az emberi népesség gyarapodása környezetre gyakorolt negatív hatásunk növekedését okozta, ami természeti értékeinket és így saját fennmaradásunkat is veszélyezteti. A vadvilág igazságügyi genetika (az ún. „wildlife forensic genetics”) a természetkárosítással kapcsolatos bűnügyek megoldásán keresztül támogatja a természeti kincseink védelméért és hosszútávú megőrzéséért tett erőfeszítéseket. E tudományág alapvető kérdése annak a megállapítása, hogy a törvényszéki eljárás során bizonyítékként szolgáló biológiai nyom milyen biztonsággal eredeztethető egy adott fajtól, annak egy egyedétől vagy populációjából (*Ogden et al., 2009*). A vadvilág igazságügyi genetika (az ún. „wildlife forensic genetics”) az utóbbi két évtizedben a molekuláris genetika eszköztárának bővülésével párhuzamosan rendkívül gyors fejlődést mutatott. Az első humán multiplex STR kitek 1994-es megjelenése (*Kimpton et al., 1994*) indította el az egyedazonosítás céljából vadon élő állat-, és növényfajra alkalmazható mikroszatellita markerek multiplex PCR rendszerek fejlesztését. Manapság több vadonélő állatfajokra is léteznek már törvényszéki eljárásokban is alkalmazható multiplex STR rendszerek, amelyekkel több orvvadászati esetet is megoldottak már (*Caniglia et al., 2010; Barbanera et al., 2012*).

2. CÉLKITŰZÉSEK

Célunk az volt, hogy Magyarországon elsőként egy olyan több tetranukleotid mikroszatellitából álló PCR rendszert fejlesszünk ki, amely a törvényszéki eljárások (pl. egy orvvadászati eset) során DNS alapú bizonyítékként szolgálhat (*Szabolcsi et al.*). Az alkalmazott STR markereknek megfelelően polimorfnak kell lennie ahhoz, hogy az egyedeket az igazságügyi vizsgálatokhoz szükséges megfelelő statisztikai erővel el tudjuk különíteni. A mikroszatelliták allélszáma és azok gyakorisági eloszlása populációnként változó ezért meg kellett vizsgálni a mikroszatelliták populáción belüli polimorfizmus fokát. Ehhez elengedhetetlen volt egy populációs referencia adatbázis létrehozása, amelynek segítségével a genetikai profil, mint DNS bizonyíték statisztikai ereje megbecsülhető. Hartl és munkatársainak a magyarországi gímszarvasok populációgenetikai vizsgálata során kiderült, hogy bizonyos élőhelyek állományai között szignifikáns genetikai különbség mérhető (*Hartl et al., 1990*). A populációs adatbázis inhomogenitásának mértéke megváltoztathatja a DNS bizonyíték populációstatisztikai erejét, ezért megvizsgáltuk adatbázisunk genetikai szerkezetét.

Az STR markerek végzett populációgenetikai vizsgálatok igen sok információval szolgálhatnak a vadbiológusok és a vadgazdálkodók számára is. A mikroszatellita rendszereinkkel a másik célunk egy olyan genetikai eszköztár létrehozása is volt, amellyel az állományon belül populációk (démek) térbeli elhelyezkedése genetikai módszerrel is kimutathatóvá váljon. A vadgazdálkodási egységek határait a populációk térbeli elhelyezkedéséhez igazítva megvalósíthatóvá válna az ökológiai szemléletű, valós populációdinamikai paramétereken alapuló gímszarvas gazdálkodás (Csányi *et al.*, 2010).

A megvalósítás érdekében a munkám fő céljai a következők voltak:

- I. A mikroszatellita markerek polimorfizmusainak felméréséhez populációs minták gyűjtése, a populációs felmérések kiszélesítése olyan állományokra is, amelyeket korábban más szerzők még nem vizsgáltak.
- II. A csoportunk által kifejlesztett „zoo-kloning” eljárás alkalmazása a szakirodalomban meglévő - öszvérszarvasból és vapatiból izolált - 22 tetranukleotid mikroszatellita lokusz gímszarvas ortológjának DNS szekvencia vizsgálatához.
- III. STR markerek polimorfizmus alapú szelekciója a PhastSystem technológia alkalmazásával.
- IV. Több STR lokusz paralel amplifikálását végző multiplex PCR rendszerek létrehozása.
- V. A multiplex PCR rendszerekkel végzett genotipizálás kidolgozása. DS-33 fluoreszcens festékekkel jelölt STR markerek PCR fragmenseinek elválasztása ABI 3130 kapilláris-gélelektroforézis készülékkel. A fragmensek méretmeghatározása a GeneScan 500 LIZ standard és a GeneMapper v3.2 szoftver használatával.
- VI. Az STR lokuszok polimorfizmusának felmérése és a detektált allélok szekvencia vizsgálata. A nemzetközileg elfogadott nevezéktani szabályok alkalmazása az alléltípusok meghatározásában.
- VII. Az STR multiplex identifikáló erejének felmérése érdekében meg kellett vizsgálnunk a lokuszon belüli allélok random párosodását és a mikroszatellita lokuszok független öröklődését.

- VIII. A DNS-vizsgálatok törvényszéki bizonyító erejének megbecsülése érdekében meg kellett vizsgálni a populációk genetikai szerkezetét.
- IX. A populációs minták igazságügyi adatbázisként való alkalmazásának vizsgálata. Az STR lokuszok együttes identifikáló erejének becslése a populációs adatbázisokon belül és azok között mért genetikai variancia korrigálásával.
- X. A genetikai profil alapján végzett populációs eredet meghatározás lehetőségének vizsgálata.

3. ANYAGOK ÉS MÓDSZEREK

A populációs mintákból a DNS kinyerése standard protokollok szerint történt. A DNS koncentráció meghatározáshoz spektrofotometriás eljárást alkalmaztunk. A két szarvasfajból - az öszvérszarvasból (*Odocoileus hemionus*) valamint amerikai vapitiből (*Cervus canadensis*) – származó tetranukleotid mikroszatelliták gímszarvas ortológjának azonosítását a csoportunk által kifejlesztett „zoo-kloning módszerrel végeztük (Gyurján et al., 2007; Borsy et al., 2009). A gímszarvas mikroszatelliták – PhastSystem rendszeren végzett - polimorfizmus szelekciója során a kevesebb, mint három alléllal rendelkező mikroszatellitákat kizártuk a multiplex rendszer fejlesztésének további lépéseiből. A forward primerek 5' végét a DS-33 (Applied Biosystems) standard mátrix öt különböző mátrix fluorofórával jelöltük, a kapilláris elektroforézist ABI 3130xl készüléken végeztük. A detektált STR allélok ismétlődő régiójának összetételét, a specifikus primer oligokkal végzett direkt DNS szekvenálással határoztuk meg.

A allélfrekvenciák, a megfigyelt és várt heterozigóta gyakoriságok kalkulációját a Genepop v4.0 szoftverrel végeztük el. HW egyensúly tesztelésére az Arlequin v3.1 (Excoffier et al., 2006) és a Genepop v4.0 (Rousset, 2008) szoftvereket használtuk, amelyek tesztstatisztikának a Fisher-féle exakt teszt Guo és Thompson által kifejlesztett permutációs változatát alkalmazzák (Guo and Thompson, 1992). A lokuszok kapcsoltsági viszonyait kétféle módszerrel vizsgáltuk. A kapcsoltsági egyensúly teszteléséhez (*LD teszt*) az Arlequin v3.1 szoftver valószínűségi hányados (*likelihood ratio*) permutációs tesztjét használtuk, amelynek szignifikancia szintjét a Bonferroni eljárással korrigáltuk (Excoffier et al., 2006; Reiczigel, 2010). Az STR markerek kromoszómás elhelyezkedését a szarvasmarha – gímszarvas genom szekvenciájának hasonlóságára alapozott térképezéssel is vizsgáltuk.

A populációk genetikai összetételének vizsgálatát molekuláris varianciaanalízissel (AMOVA) becsült Wright-féle F-statisztikai paraméterekkel kezdtük, amelyhez ugyancsak az Arlequin v3.1 szoftvert használtuk. Az identifikáló erő kifejezésére kétféle számítást alkalmaztunk. Az egyik a teoretikusan „leggyakoribb” genetikai profil *egyediségének* (un. „*uniqueness*”) becslése (NRC, 1996), a másik az *átlagos profil egyezési valószínűség* (*probability of match*, $pM \sim \text{probability of identity}$, PI_{ave}) kalkulációja, amit az API-CALC v1.0 szoftver segítségével a szignifikáns F-statisztikai paraméterekkel korrigáltuk (Ayres and Overall, 2004). A populációk genetikai homogenitását az egyedek „populációs hűségének” kalkulációjával is elvégeztük, amelyet az ún. *valószínűségi hányados* (LBM- „*likelihood based method*”) módszerrel fejeztünk ki. A profilgyakorisági értékek mintavételi hibából adódó bizonytalanságát a Chakraborty-féle konfidencia intervallum becslésével végeztük (Chakraborty et al., 1993).

4. TÉZISEK

4.1. A „zoo-klónozási” eljárással a munkánk kezdetén rendelkezésre álló 22 - öszvérszarvasból (*Odocoileus hemionus*) és amerikai vapitiből (*Cervus canadensis*) izolált tetranukleotid STR markerből 14 gímszarvas ortológ teljes szekvenciáját sikerült meghatározzuk (Jones et al., 2002; Meredith et al., 2004). A 14 lokusz teljes DNS szekvenciáit összehasonlítottuk az ortológ párjával. Az ún. flanking (határoló) régiók illesztése során egyedül a T123 lokusz esetében nem mutattunk ki szekvencia eltérést. A határoló régiók primer kötőhelyeit vizsgálva a C01, T26, T156, T172, T193 és a T507 lokusznál detektáltunk szekvencia eltérést, ami a szekvencia meghatározások nélkül PCR amplifikációs problémákat okozhatott volna. A határoló régió méretéhez viszonyított legtöbb mutációt az öszvérszarvas eredetű C01, C143, C229 valamint a vapitiből izolált T26 STR markerek esetében mutattuk ki. Ludt és Pitra (Ludt et al., 2004; Pitra et al., 2004) filogenetikai eredményei alapján a gímszarvas - öszvérszarvas mikroszatellita ortológok (C01, C143, C229) között detektált nagyszámú szekvencia eltérés nem meglepő, hiszen a két faj evolúciós értelemben távolabb áll egymástól, mint a vapiti a gímszarvastól, ezért a felhalmozódott mutációk száma is magasabb. A T26 marker flanking régiójának nagyfokú szekvencia-, és a későbbiekben tárgyalt allél polimorfizmusa arra enged következtetni, hogy evolúciós tekintetben a vizsgált mikroszatellita gének között a legősibb lehet. A mikroszatellita evolúciósan korábbi keletkezését a szarvasmarha ortológgal történt szekvencia összehasonlítások eredményei is jól mutatják. A szarvasmarha ortológ lokuszok flanking

régióival végzett szekvencia összehasonlítások során ennél a markernél tapasztaltuk a legalacsonyabb fokú szekvencia hasonlóságot, ami 79,37% volt.

4.2. A mintákat a korábbiakban más szerzők által is vizsgált (*Hartl et al., 1990*) északkelet-, és délnyugat-magyarországi gímszarvas állományokból (EK és DNY) gyűjtöttük kiegészítve őket a gödöllői dombsági, cserhádi-, bükki- és mátrai populációkból származó egyedekkel. Ezekből 14 véletlenszerűen kiválasztott egyedden végeztük el a 14 tetranukleotid mikroszatellita polimorfizmus tesztjét, amelynek során 10 polimorf marker bizonyult alkalmasnak az egyedek elkülönítésére alkalmas marker készlet létrehozásához. A teljes populációs felmérés eredményeképpen minden lokusznak meghatároztuk a populáción belüli alléleloszlását. 135 allélt detektáltunk, amelyek között 28 intermedier köztes (ún. intermedier) allélt különítettünk el. 50 allél DNS szekvencia meghatározását végeztük el. A mikroszatellita allélok szekvencia vizsgálatával megteremtettük a lehetőséget a nemzetközileg elfogadott STR allél nevezéktan alkalmazásának és igazoltuk - a T26, T501, T507 és a T156 lokuszok esetében detektált - egy, két vagy három bázis egységgel eltérő intermedier típusú allélok jelenlétét és azok korrekt elkülönítését.

A legkevesebb, 5 alléllal a C229 rendelkezett, a legpolimorfabbnak a T26 és a T501 markerek (21 és 23 allél) bizonyultak. A T501 Jones és mtsai (2002) által elvégzett, 43 vapii egyed alapján végzett felmérésben is a legváltozékonyabb volt, összesen 11 db T501 allélt azonosítottak. A T501 nagy allélszáma az ismétlődő régiójában lévő A₁₀ homopolimer ismétlődéseinek köszönhető. Az ismétlődő régió szerkezete alapján a T108, T156, T172 és a T507 egyszerű, a C229, T123, T193 összetett, míg a C01, T26 és a T501 komplex felépítésű mikroszatellita markerek. A stabil és a polimorf ismétlődő egységek valamint az allélszám alapján a T26 és a T501 hiperpolimorf STR lokuszok.

4.3. Magyarországon elsőként valósítottuk meg tíz polimorf gímszarvas tetranukleotid STR marker két 5-5 lokuszt tartalmazó multiplex PCR rendszerbe (5-Plex) illesztését, amelyet *DeerPlex I-II*-nek neveztünk el. A *DeerPlex I-II* amplifikációja során detektált nem allélikus eredetű termékek (artifaktumok) relatív mennyiségét a multiplex PCR reakciók optimalizált dNTP, Mg²⁺ és primerkoncentrációjának valamint annealási hőmérsékletének eredményeképpen sikeresen csökkentettük. A T172 és a T501 lokuszok allélcsúcsai mellett detektált artifaktumok (double peak, shoulder peak) csökkentését primer módosítással, az ún. „pig-tailing” eljárással valósítottuk meg.

A markerek fluorofórijait a DS-33 (Applied Biosystems) spektrális kalibráló rendszerének festékeiből állítottuk össze. A fragmensméret meghatározáshoz az ABI PRISM 3130 kapilláris-gélelektroforézis készüléket, a GeneScan™ 500 LIZ (Applied Biosystems) belső méretstandardot és a GeneMapper® ID ver3.2 szoftvert alkalmaztuk. A populációs minták analízise során kalkulált allélcúcs méreteket ± 0.5 nt kiterjedésű és maximálisan 0,25 nt szórású tartományokba csoportosultak. A *DeerPlex I-II* –vel végzett genotipizálás (profilozás) manuális módját valósítottuk meg, ami alapvetően két részből áll. Első lépésben az egyedhez tartozó két multiplex PCR reakció során felsokszorozott allélikus eredetű DNS-szakaszok fragmens-méretének meghatározására került sor. Az ismeretlen allélok típusát az határozza meg, hogy az allélcúcs mérete melyik - a populációs vizsgálatok során detektált - allélhoz tartozó, átlaggal és 0,25 nt szórással jellemezhető, maximálisan 1 nt terjedelmű mérettartományba illeszthető be és a kimutatott allél fragmens-mérete hogyan viszonyul a fragmensméret meghatározás során használt standard allélokhoz. Minden genotipizáláshoz standardként egy olyan gímszarvas tehén *DeerPlex I-II* profilját használtuk, amelynek minden allélméret és allélszekvencia adata ismert volt.

4.4. A *DeerPlex I-II* markerek keresztregáló képességét az emberen, és néhány párosujjú patás fajon is elvégeztük. *DeerPlex I-II* -vel végzett PCR reakciók során bebizonyosodott, hogy az ember, a vadsertés, szarvasmarha, muflon és az őz mintákból allélikus eredetű PCR termék nem mutatható ki. A dámszarvasból 9 STR lokusz genotípusa értékelhető volt, vagyis allélikus eredetű ampikonokat detektáltunk. Vizsgálataink alapján ebből a szarvasfajból részleges *DeerPlex I-II* genetikai profil detektálható, egyedül a T26 lokusz nem amplifikálódik. A gímszarvas populációs minták genetikai profiljaiban a T26 marker kiesését nem észleltük, ezért joggal feltételeztük, hogy a minták közé dámszarvas egyed nem került. A vizsgálatok eredményeképpen kijelenthető, hogy genotipizáló rendszerünk gímszarvasra specifikus, egyéb patás vadfajok (őz, muflon, vaddisznó, dámszarvas) és az ember nem kontaminálhatja. A *Cervinae* alcsalád filogenetikai vizsgálata alapján az amerikai vapitivel együtt külön csoportba (Keleti kládba) tartozó szikaszarvas szekvenciálisan igen hasonló lehet, ezért ezen a szarvasfajon végzett fajspecifikussági teszt indokolt lehet (*Ludt et al., 2004; Pitra et al., 2004*). Állati eredetű biológiai anyagmaradványok vizsgálatánál problémát okozhat, hogy egy mintán belül nemcsak egyedek-, hanem közelrokon fajok keveredése is előfordulhat, amely gondot okozhat a kevert profilok kiértékelése során. A biológiai nyom *DeerPlex I-II* –vel végzett analízisét megelőzően indokolt lehet az adott minta fajspektrumának kvalitatív és kvantitatív vizsgálata.

4.5. A Slate-féle (Slate *et al.*, 2002) kapcsoltsági térkép alapján 4 STR marker kromoszómás elhelyezkedése már ismert volt, de a többi 6 STR mikroszatellita kapcsoltsági viszonyait tisztáznunk kellett. Kétféle megközelítést alkalmaztunk. A statisztikai *LD* kapcsoltsági tesztek során nem mutattunk ki egyetlen olyan lokuszpárt sem, amelyik mindhárom populációban a kapcsoltsági egyensúlytalanság állapotában lett volna. A függetlenségi vizsgálatok statisztikai oldalról közelítése bizonyos lokuszpárok esetében (pl. T108-C01, C01-T501) nem tette egyértelművé a markerek egymáshoz való viszonyát.

Az összehasonlító szekvencia hasonlóságon alapuló térképezés eredményeképpen egyedül a T507 lokusz lokalizációja nem volt egyértelműen meghatározható. A T26 és a C01 markerek ortológjai a 16-os szarvasmarha kromoszómán lokalizáltak, amely a Slate-féle összehasonlító térkép szerint a gímszarvas 14-es kapcsoltsági csoportjával homológ, vagyis a nem független öröklődésük feltételezhető. Annak ellenére, hogy a szarvasmarha T26 szekvencia hasonlósága alacsony gímszarvas ortológjához képest, és a repeat struktúra alapján mikroszatellita gén megkérdőjelezhető, a későbbi vizsgálatok során a lehetségesen kapcsolt T26-C01 párt figyelembe vettük. Ennek bizonyítása csak a teljes genom ismeretében oldható meg, ami a jelenlegi munka írásakor még nem állt rendelkezésünkre.

4.6. Magyarországon elsőként kezdtük el a vadon élő gímszarvas állományok genetikai összetételének vizsgálatát mikroszatellita markerekkel. A géndiverzitás *H*, *PIC*, *PD* értékeinek maximumát az északkeleti populációban mutattuk ki, vagyis a két populáció közül az utóbb említett tekinthető a genetikailag változékonyabbnak és egyben az igazságügyi egyedazonosítás hatékonyságának szempontjából a kedvezőbbnek is. A két alpopuláció közötti genetikai elkülönülés mértéke (F_{ST}) szignifikáns volt, de más európai gímszarvas állományokhoz viszonyítva nem volt jelentős mértékű (Kuehn *et al.*, 2003, 2004). A teljes populáción végzett globális AMOVA eredménye alapján a lokuszokra és a populációkra átlagolt beltenyészet paramétere (F_{IS}) 0,0373, a genetikai elkülönülés paraméterére (F_{ST}) 0,034 értéket kaptunk. Az emberi populációk beltenyészetének és a populációk közötti genetikai különbségek mértékét figyelembe véve mindkét paraméter magasnak számít. A humán és a nem-humán igazságügyi genetikában ajánlott a profilgyakorosság (ezzel együtt az *LR*) közös leszármazási együtthatóval (θ -theta) végzett korrigálása indokolt (Linacre *et al.*, 2011), amely a gímszarvas teljes adatbázisának használatakor az F_{IT} -nek felel meg. A teljes populációs adatbázisból kalkulált *LR* számításához az F_{IS} paraméter lokuszokra átlagolt, míg az F_{ST} paraméter lokuszonként becsült értékét használtuk.

Az alpopulációk inhomogenitását az egyedek populációs hűségének *LBM* alapú számításával analizáltuk. Az LR_{LBM} kalkulációk alapján az egyedek 44%-a nem a referencia populációjába volt sorolható, vagyis az EK és DNY populációban további struktúráltságot feltételeztünk. A *P* érték konfidencia intervallumát figyelembe véve a populációs eredet becslése a nem-referencia populációba soroló egyedek viszonylag magas aránya (EK és DNY egyedek 17%-ában), vagyis a genetikai inhomogenitás miatt feltételezhetően nem adna megbízható értéket. Az alpopulációk (démek) kimutatásához a magyarországi állomány szélesebb körű vizsgálatára lenne szükség. A gímszarvas profilok *LR* értékének kalkulációjához a 100 egyedből álló referencia allélgyakorisági adatbázis használatát ajánljuk, az EK és a DNY alkalmazása, a feltételezhető struktúráltság és a kis populáció méret miatt torzított becslést adhat.

4.7. Célunk a gímszarvas egyedek genetikai úton történő azonosítása volt ezért meg kellett vizsgálnunk a polimorf mikroszatellitákból álló két 5-Plex PCR rendszerünk identifikáló erejét. A kapcsoltsági vizsgálatok eredményeképpen a *DeerPlex I-II* gímszarvas egyedek elkülönítésére való alkalmasságának kalkulációját 10 gímszarvas STR markerre és a C01 lokuszt kivéve a multiplex rendszerből, 9 markerre is elvégeztük.

A 100 egyed allélgyakorisági értékeiből megbecsültük a *DeerPlex I-II* genotípusok *átlagos egyezési/azonosság valószínűségét* (pM/PI_{ave}) a teljes populációban. Az API-CALC v1.0 szoftver alkalmazásával a pM/PI_{ave} értéket korrigálni tudtuk az alpopuláción belül mért szignifikáns beltenyészet (F_{IS}) valamint a teljes populációban kimutatott szubdivíziós (F_{ST}) paraméterrel. Az elvégzett számítások alapján annak valószínűsége, hogy a populációból véletlenszerűen kiválasztott két gímszarvasnak megegyező 10/ 9 lokuszos genetikai profilja legyen $2,6236 \times 10^{-15} / 1,0056 \times 10^{-12}$. Rendkívül alacsony értékek, amelyek – amennyiben a laboratóriumi hiba kizárt - a téves egyed meghatározás valószínűségét is kifejezik.

A *DeerPlex I-II* multiplex rendszerünk identifikáló erejének számszerű kifejezésének másik módszere a teoretikusan „leggyakoribb” genetikai profil *egyediségének* (un. „*uniqueness*”) becslése. Az egyediség a populáció nagyságának függvényében annak a valószínűségét adja meg, hogy az adott profil egyedi. Ha nem egy adott profil, hanem a teoretikusan „leggyakoribb” profil egyediségével számolunk, akkor a marker készletünk adott populáción belül értelmezett minimális egyediségének valószínűségét kapjuk meg. A „leggyakoribb” profil számításakor figyelembe vettük az adott EK-DNY populációkon belül létező beltenyészet (F_{IS}) és a teljes populáción belül kimutatható inhomogenitás (F_{ST}) mértékét is. A 10 lokuszos egyediség tartalma $1 - 6,73 \times 10^{-5}$, a 9 lokuszt tartalmazó marker készlet

egyediségének valószínűségére $1 - 7,93 \times 10^{-4}$ kaptunk. A számításokhoz az európai gím-, és szikaszarvas állomány 2 millióra becsült egyedszámával dolgoztunk.

Mindkét számítás azt igazolja, hogy amennyiben a laboratóriumi hiba kizárt, akár a 10 akár a 9 polimorf mikroszatellitát alkalmazó készlet nagymértékben alkalmas vad populációban élő gímszarvas egyedek genetikai alapú azonosítására és elkülönítésére. Identifikáló ereje az Egyesült Államok kaukázusi népességére kalkulált Identifiler multiplex PCR „*uniqueness*” és kombinált PD értékéhez hasonló (Butler, 2011).

ÖSSZEFOGLALÁS

A gímszarvas (*Cervus elaphus hippelaphus*), mint nagytestű herbivor hazai faunánk ökológiai és vadgazdálkodási szempontból is nagy jelentőséggel bíró tagja. Becsült állománya 100.000 feletti egyedszámra tehető, amelyből évente legálisan jelentős mennyiséget, hozzávetőlegesen 40.000 példányt ejtenek el. Az illegális veszteség a szakemberek becslései szerint igen magas lehet, évente akár 10.000 gímszarvas is orvvadászok zsákmánya lehet. Egyik célunk az volt, hogy egy olyan - gímszarvasok egyedazonosítására alkalmas – tetranukleotid (tetramer) mikroszatellitákból (STR) álló multiplex PCR rendszert hozzunk létre, amellyel az elkövetőt a tettenérés szüksége nélkül lehessen összefüggésbe hozni közvetlenül a tett helyszínével. A mikroszatellita lokuszoknak megfelelő polimorfizmus fokkal kell rendelkezniük ahhoz, hogy egy biológiai nyomot egy egyedről származtatni tudjunk.

A lokuszok polimorfizmusának felméréséhez a délnyugati megyékből (DNY - Zala, Somogy, Baranya és Bács-kiskun (Gemenc)) valamint az ország északi területeiről (EK - Börzsöny, gödöllői-dombság, Cserhát, Bükk, Mátra) 100 mintát gyűjtöttünk. A munkánk másik célja a magyarországi gímszarvas állomány, mint metapopuláció genetikai struktúrájának részletesebb megismerése volt.

A szakirodalomban rendelkezésünkre álló amerikai vapatiból (*Cervus canadensis*) és ösvérszarvasból (*Odocoileus hemionus*) izolált 22 tetramer STR ortológból az ún. „zoo-klónozási” eljárással 14 lokusz szekvenciáját sikerült meghatározni. A PhastSystem és a nagy-felbontású PAGE technológia segítségével végzett polimorfizmus szűréssel 4 STR markert szelektáltunk ki. A 10 polimorf mikroszatellita lokuszból két 5 STR markert tartalmazó multiplex PCR rendszert készítettünk (*DeerPlex I-II*). A markerek jelöléséhez a DS-33 (Applied Biosystems) standard mátrix fluorofórijait választottuk. Genotipizáló rendszernek az ABI 3130 kapilláris-gélelektroforézis készüléket és a Genemapper® ID 3.2 szoftvert (Applied Biosystems) alkalmaztuk.

A markerek identifikáló erejének becsléséhez megvizsgáltuk lokuszaink polimorfizmus fokát, amelyet a Nei-féle géndiverzitással, polimorfizmus információtartalommal és a megkülönböztető erővel jellemeztünk. A C229 lokuszt kivéve egyenként magas értékeket kaptunk. A lokuszokon belül az allélok random asszociációját a Hardy-Weinberg egyensúly (HWE) tesztel az Arlequin v3.11 és a Genepop v4.0 szoftver felhasználásával végeztük. Az 5 %-os szignifikancia szinten a T501, C01 és T108 markereknél volt elvethető a HWE. A lokuszok függetlenségi vizsgálatait sem statisztikai (LD teszt) sem szekvencia hasonlóságon alapuló térképezés alapján nem zárták ki a C01 és a T26 genetikai kapcsoltságát, ezért az igazságügyi számításokhoz a C01 markert nem tartalmazó 9 lokuszt használtuk.

DeerPlex I-II egyedazonosító képességének vizsgálatához szükség volt a populációk genetikai struktúrájának ismeretére. Az Arlequin szoftver segítségével molekuláris variancia analízist végeztünk, amelynek eredményeképpen a két alpopuláció között szignifikáns genetikai elkülönülést mutattunk ki ($F_{ST} = 0.034$), az alpopulációkon belüli átlagos teljes beltenyészet paraméterre, $F_{IS} = 0.037$ értéket kaptunk. A két paraméterrel korrigált *átlagos egyezési/azonosság valószínűséggel* (pM/PI_{ave}) valamint a teoretikusan leggyakoribb genetikai profil egyediségével („uniqueness”) fejeztük ki a *DeerPlex I-II* identifikáló erejét. Minimum $1 - 7,93 \times 10^{-4}$ annak a valószínűsége, hogy egy 9 lokuszos *DeerPlex I-II* profil egyedi és $1,0056 \times 10^{-12}$ a valószínűsége annak, hogy a populációból két véletlenszerűen kiválasztott egyed 9 lokuszos genotípusa megegyezik. A vizsgálatok eredményeképpen igen magas identifikáló erőt kaptunk, megközelítőleg olyan erős, mint az Egyesült Államok kaukázusi népességére számolt Identifiler igazságügyben használt kit egyedazonosító ereje.

A marker készlet igazságügyi alkalmazhatóságát jól példázza az általunk megoldott orvvadászati eset is, amelyben a helyszíni nyomokból kimutatott *DeerPlex I-II* genetikai profilokat és a gyanúsított birtokán lefoglalt gímszarvasbika trófeáját megfelelő statisztikai erővel tudtuk egy egyedtől származtatni.

A *DeerPlex I-II* rendszereinket felhasználtuk az alpopulációk genetikai struktúrájának vizsgálatára is, amit az egyedek populációs eredetének valószínűsítésével – az ún. „assignment tesztel” - végeztünk el. Az EK-DNY adatbázisból kalkulált profilgyakorisági értékpárok alapján mindkét populációt alkotó egyedek 17 %-a nem a referencia populációba sorolt, ebből következően heterogénnek tekinthetők. Igazságügyben DNS alapú bizonyítékok statisztikai interpretációjához az alpopulációk - genetika struktúráltságuk következtében - nem használhatók, helyettük a teljes adatbázist ajánljuk. Az „assignment teszt” alapú populációs eredet meghatározás az egyedek két populáció szerint kalkulált profilgyakorisági értékeinek

nagyarányú (44%) átfedő konfidencia intervalluma alapján torzított lenne, ezért igazságügyi statisztikai interpretációhoz nem alkalmazható.

IRODALOMJEGYZÉK

Az értekezés alapjául szolgáló közlemények

- Szabolcsi Z**, Egyed B, Zenke P, Padar Z, Borsy A, Steger V, Pasztor E, Csanyi S, Buzas Z, Orosz L (2014). Constructing STR multiplexes for individual identification of Hungarian red deer. *J For Sci*.
- Szabolcsi Z**, Egyed B, Zenke P, Borsy A, Padar Z, Zoldag L et al. (2008). Genetic identification of red deer using autosomal STR markers. *Forensic Sci Int Genet*, 1 Suppl: 623-624
- Borsy A**, Podani J, Stéger V, Balla B, Horváth A, Kósa JP, Gyurján IJr, Molnár A, **Szabolcsi Z**, Szabó L, Jakó E, Zomborszky Z, Nagy J, Semsey S, Vellai T, Lakatos P, Orosz L (2009). Identifying novel genes involved in both deer physiological and human pathological osteoporosis. *Mol Genet Genomics*, 281:301-313

Tézishez tartozó publikációk

- Ayres KL**, **Overall ADJ** (2004). API-CALC 1.0: a computer program for calculating the average probability of identity allowing for substructure, inbreeding and the presence of close relatives. *Mol Ecol Notes*, 4(2): 315-318.
- Barbanera F**, Guerrini M, Beccani C, Forcina G, Anayiotos P, Panayides P (2012). Conservation of endemic and threatened wildlife: Molecular forensic DNA against poaching of the Cypriot mouflon (*Ovis orientalis ophion*, Bovidae). *Forensic Sci Int Genet*, 6(5): 671-675.
- Butler JM** (2011). *Advanced topics in forensic DNA typing: methodology*. San Diego: Academic Press.
- Caniglia R**, Fabbri E, Greco C, Galaverni M, Randi E (2010). Forensic DNA against wildlife poaching: Identification of a serial wolf killing in Italy. *Forensic Sci Int Genet*, 4(5): 334-338.
- Chakraborty R**, Srinivasan MR, Daiger SP (1993). Evaluation of standard error and confidence interval of estimated multilocus genotype probabilities, and their implications in DNA forensics. *Am J Hum Genet*, 52(1): 60-70.
- Csányi S**, Lehoczki R (2010). Ungulates and their management in Hungary. pp. 291-318 in: Apollonio, M., Andersen, R. és Putman, R. (eds.) *Ungulate management in Europe in the XXI century*. Cambridge University Press, Cambridge.
- Excoffier L**, Laval G, Schneider S (2006). Arlequin ver 3.1: An Integrated Software Package for Population Genetics Data Analysis. Computational and Molecular Population Genetics Lab (CMPG), Institute of Zoology, University of Berne.
- Gyurján IJr**, Molnár A, Borsy A, Stéger V, Hackler LJr, Zomborszky Z, Papp P, Duda E, Deák F, Lakatos P, Puskás LG, Orosz L (2007). Gene expression dynamics in deer antler: mesenchymal differentiation toward chondrogenesis. *Mol Genet Genomics*, 277(3): 221-235.
- Guo SW**, **Thompson EA** (1992). Performing the exact test of Hardy-Weinberg proportion for multiple alleles. *Biometrics*, 48: 361-372.

- Hartl** GB, Willing R, Lang G, Klein F, Köller J (1990). Genetic variability and differentiation in red deer (*Cervus elaphus* L) of Central Europe, *Genetics Selection Evolution*, 22: 289-306.
- Hartl** DL, Clark AG (2007). *Principles of Population Genetics*, Fourth Edition, Sinauer Associates, Sunderland, Massachusetts.
- Jones** KC, Levine KF, Banks JD (2002). Characterization of 11 polymorphic tetranucleotide microsatellites for forensic applications in California elk (*Cervus elaphus canadensis*). *Mol Ecol Notes*, 2: 425-427.
- Kimpton** CP, Fisher D, Watson S, Adams M, Urquhart A, Lygo J, Gill P (1994). Evaluation of an automated DNA profiling system employing multiplex amplification of four tetrameric STR loci. *Int J Legal Med*, 106:302-311.
- Kuehn** R, Schroeder W, Pirchner F, Rottman O (2003). Genetic diversity, gene flow and drift in Bavarian red deer populations (*Cervus elaphus*). *Conserv Genet*, 4(2): 157-166.
- Kuehn** R, Haller H, Schroeder W, Rottmann O (2004). Genetic roots of the red deer (*Cervus elaphus*) population in Eastern Switzerland. *J Hered*, 95(2):136-43.
- Linacre** A, Gusmao L, Hecht W, Hellmann AP, Mayr WR, Parson W, Princz M, Schneider PM, Morling N (2011). ISFG: Recommendations regarding the use of non-human (animal) DNA in forensic genetic investigations. *Forensic Sci Int Genet*, 5(5): 501-505.
- Ludt** CJ, Schroeder W, Rottmann O, Kuehn R (2004). Mitochondrial DNA phylogeography of red deer (*Cervus elaphus*). *Mol Phylogenet Evol*, 31(3): 1064-1083.
- Meredith** EP, Rodzen JA, Levine KF, Banks JD (2004). Characterization of an additional 14 microsatellite loci in California Elk (*Cervus elaphus*) for use in forensic and population applications. *Conservation Genetics*, 6: 151-153.
- NRC** - National Research Council (1996). *The evaluation of forensic DNA evidence*. National Academy Press Washington, D.C.
- Ogden** R, Dawnay N, McEwing R (2009). Wildlife DNA forensics - bridging the gap between conservation genetics and law enforcement. *Endangered Species Research*, 9: 179-195.
- Pitra** C, Fickel J, Meijaard E, Groves PC (2004). Evolution and phylogeny of old world deer. *Mol Phylogenetic Evol*, 33(3): 880-895.
- Reiczigel** J, Harnos A, Solymosi N (2010). *Biostatistika nem statisztikusoknak*. Javított utánnomás, Pars Kft., Budapest.
- Rousset** F (2008). Genepop'007: a complete reimplement of the Genepop software for Windows and Linux. *Mol Ecol Resour*, 8(1): 103-106.
- Slate** J, Van Stijn TC, Anderson RM, McEwan KM, Maqbool NJ, Mathias HC, Bixley MJ, Stevens DR, Molenaar AJ, Beever JE, Galloway SM, Tate ML (2002). A Deer (Subfamily Cervinae) Genetic Linkage Map and the Evolution of Ruminant Genomes. *Genetics*, 160: 1587-1597.