

## ANALIZA MOLEKULARNE VARIJANCE (AMOVA) I TESTIRANJE HOMOGENOSTI MOLEKULARNE VARIJANCE (HOMOVA) NA TEMELJU RAPD BILJEGA

Z. ŠATOVIĆ<sup>1</sup>, Belén ROMAN<sup>2</sup>, Ana María TORRES<sup>2</sup>,  
I. KOLAK<sup>1</sup> i J. I. CUBERO<sup>3</sup>

<sup>1</sup>Agronomski fakultet Sveučilišta u Zagrebu  
Zavod za sjemenarstvo  
Faculty of Agronomy, University of Zagreb  
Department of Seed Science and Technology

<sup>2</sup>CIFA, Cordoba, ŠPANJOLSKA  
CIFA, Cordoba, SPAIN

<sup>3</sup>E.T.S.I.A.M., Corboda, ŠPANJOLSKA  
E.T.S.I.A.M., Corboda, SPAIN

### SAŽETAK

U radu je raspravljena analiza molekularne varijance (AMOVA) i neparametrijski test homogenosti molekularne varijance (HOMOVA). Tijek provedbe analize je prikazan na primjeru RAPD podataka populacija volovoda (*Orobanche crenata* Forsk.).

Ključne riječi: AMOVA, HOMOVA, *Orobanche crenata* Forsk.

### UVOD

Procjena genetske varijabilnosti, genetske udaljenosti i genetske strukture unutar i između populacija određene biljne vrste omogućava nam pomno praćenje određenih genetskih pojava u svrhu bolje organizacije i provođenja oplemenjivačkih i sjemenarskih programa kao i programa očuvanja biljnih genetskih izvora. Klasična se genetika uglavnom temeljila na praćenju određenih morfoloških i fizioloških svojstava, a skorašnjim se razvitkom tehnika molekularne genetike varijabilnost može uočiti izravno na razini DNA.

RAPD (Random Amplified Polymorphic DNA - Slučajno amplificirana polimorfna DNA) biljezi (Williams et al., 1990; Welsh i McClelland, 1990; Caetano-Annoles, 1994) se temelje na lančanoj reakciji polimerazom (Polymerase Chain Reaction - PCR) (Saiki et al., 1985; Mullis i Faloona, 1987). Zbog visoke razine otkrivene polimorfnosti, te relativne jednostavnosti i niske cijene postupka RAPD se biljezi sve više koriste kako u proučavanju postojeće genetske raznolikosti tako i u genetsko-oplemenjivačkim istraživanjima koja imaju za svrhu izradu genetske karte i upotrebu biljega u odabiru

pomoću biljega (Marker-assisted selection - MAS). Pri proučavanju postojeće genetske raznolikosti upotreba vrijednost RAPD biljega potvrđena je u brojnim istraživanjima na području očuvanja i upotrebe biljnih genetskih izvora, sistematike i populacijske genetike, identifikacije roditelja za buduća križanja kao i identifikacije čistih linija i kultivara (Šatović, 1999).

Analiza molekularne varijance (Analysis of Molecular Variance - AMOVA) relativno je nova statistička metoda kojom je moguće raspodijeliti genotipsku varijancu između i unutar populacija (odnosno bilo kojih drugih pretpostavljenih razina strukture) kao i testirati signifikantnost svakog izvora varijabilnosti (Excoffier et al., 1992). Iako je bila osmišljena za analizu podataka RFLP biljega AMOVA se je vrlo fleksibilna statistička metoda te se može provesti na temelju podataka različitih genskih biljega (odnosno genskih sustava) koji se međusobno uvelike razlikuju po svojim svojstvima i pretpostavkama o tijeku njihove evolucije. Michalakis i Excoffier (1996) su tako pokazali mogućnosti provedbe AMOVA-e na temelju mikrosatelita (Simple Sequence Repeats /SSR/ - Ponavljajuće jednostavne sekvence). No, potrebno je istaći da RAPD biljezi imaju dominantnu ekspresiju za razliku od kodominantne ekspresije izoenzimskih biljega, RFLP biljega i dr. No, kako su pokazali Huff et al. (1993) AMOVA analiza se uspješno može koristiti i na temelju RAPD biljega ako se trake na gelu smatraju RAPD fenotipovima tako da se radi o analizi *fenotipske*, a ne *genotipske* raznolikosti kao u slučaju kodominantnih biljega.

AMOVA analiza na temelju RAPD biljega vrlo je brzo našla svoju primjenu u proučavanju postojeće genetske raznolikosti naročito kod prirodnih populacija različitih vrsta. Ciljevi istraživanja često su bili procjena razine i zemljopisne raspodjele genetske varijabilnosti, utvrđivanje srodnosti populacija iz različitih područja i testiranje korelacije između zemljopisne i genetske udaljenosti. Svrha takovih istraživanja je optimiziranje strategija uzorkovanja prilikom budućih prikupljačkih ekspedicija, te pomažu učinkovitijem očuvanju biljnih genetskih izvora u bankama gena. U svrhu boljeg uvida u populacijsku dinamiku i genetsku varijabilnost AMOVA na temelju RAPD biljega koristila se na prirodnim populacijama različitog krmnog bilja kao što je lucerna (Crochemore et al., 1996; Jenczewski et al., 1999), bijela djetelina (Gustine et al., 1999), višegodišnji ljulj (Huff, 1997), livadna vlasulja (Kolliker et al., 1998) i *Buchloe dactyloides* (Peakall et al., 1995). Predmet istraživanja često su bile prirodne populacije različitih drvenastih kultura (*Abies alba* i *Abies nebrodensis*, Vicario et al., 1995), ljekovite biljke (*Digitalis obscura* L., Nebauer et al., 1999), rijetke i ugrožene vrste (*Grevillea barklyana*, Hogbin et al., 1998), te divlji srodnici kulturnih biljnih vrsta (*Oryza rufipogon*, Ge Song et al., 1999). Uspoređivali su se također i kultivari (grah, Duarte et al., 1999), te primke prirodnih populacija kao i primitivnih varijeteta (grah lima, Fofana et al., 1997).

## BILJNI MATERIJAL

AMOVA analiza bit će objašnjena na primjeru šest populacija volovoda (*Orobanche crenata* Forsk.), po tri iz južne Španolske (Cordoba, Mengibar, Pinos Puente) i iz Izraela (Alonin, Latroon, Masuot). Iz svake je populacije uzorkovano osam biljaka. Stoga je hijerarhijska analiza raznolikosti provedena pomoću dvosmjerne AMOVA-e u svrhu rastavljanja fenotipske varijance na njezine sastavnice: između regija (Izrael i Španjolska), između populacija unutar regija kao i unutar populacija.

Volodod (*Orobanche crenata* Forsk.) je parazitski korov koji napada mnoge kulturne mahunarke kao što su bob, leća, grašak, slanutak i grahorica kao i mnoge divlje vrste mahunarki na području Sredozemlja (Cubero i Moreno, 1979). Stoga poznavanje populacijske strukture volovoda može olakšati izbor mjera suzbijanja, pridonijeti učinkovitijem testiranju otpornosti na volodod kod kulturnih vrsta mahunarki, te usmjeriti potrabu za novim izvorima otpornosti na volodod unutar kolekcija biljnih genetskih izvora mahunarki (Cubero, 1983; Cubero, 1991).

DNA je izolirana iz cvjetnih pupoljaka po metodi opisanoj u Fulton et al. (1995). Jedinke su analizirane pomoću ISSR i RAPD sustava biljega. ISSR (Inter-Simple Sequence Repeats) sustav biljega temelji se na lančanoj reakciji polimerazom i postupak je istovjetan onom za dobivanje RAPD biljega, no kao klice za amplifikaciju ISSR biljega koriste se oligonukleotidi jednostavnih ponavljajućih sekvenci (Gupta et al., 1994). Nakon pregleda 30 ISSR klica 5 je klica upotrebjeno za dobivanje ISSR biljega. Također je korištena i jedna RAPD klica. ISSR biljezi su dominantnih kao i RAPD biljezi tako da se njihova analiza provodi na isti način.

Uspješnost amplifikacije fragmenata DNA omeđenih sekvencama koje su komplementarne upotrebjenim ISSR odnosno RAPD klicama praćena je na gelu i označena kao nazočnost (1) ili odsutnost (0) trake u svrhu izrade binarne matrice različitih ISSR i RAPD fenotipova. Nakon isključivanja biljega koji su bili monomorfni za cijeli set podataka 115 polimorfni biljega je uključeno u daljnju analizu. U Tablici 1. prikazani su RAPD fenotipovi analiziranih jedinki volovoda (*Orobanche crenata* Forsk.).

## ANALIZA MOLEKULARNE VARIJANCE (AMOVA)

Prva faza statističke analize je izračunavanje euklidskih udaljenosti između parova jedinki. Euklidska se udaljenost često koristi u populacijskoj genetici kao

mjera različitosti između jedinki odnosno populacija (Nei i Tajima, 1981). Euklidska udaljenost (E) se izračunava prema formuli:

$$\delta_{jk}^2 = \sum_{s=1}^S (p_{sj} - p_{sk})^2 \quad [1]$$

gdje je,

$\delta_{jk}^2$  - Euklidska udaljenost između jedinki  $j$  i  $k$ ;

$S$  - ukupan broj polimorfnih RAPD biljega u cjelokupnom setu podataka;

$p_{sj}$  - stanje polimorfnog biljega  $s$  u jedinci  $j$  ( $p = 1$  ukoliko je traka nazočna, a  $p = 0$  ukoliko je odsutna).

Stoga je euklidska udaljenost jednaka broju polimorfnih biljega između dviju jedinki. U Tablici 2. prikazana je matrica euklidski udaljenosti analiziranih jedinki volovoda.

Dobivena matrica euklidskih udaljenosti koristi se za provedbu AMOVA-e. AMOVA se temelji na linearnom modelu:

$$p_{jig} = p + a_g + b_{ig} + c_{jig} \quad [2]$$

gdje je;

$p_{jig}$  - učinak  $j$ -te jedinice,  $i$ -te populacije iz  $g$ -te skupine populacija (regije);

$p$  - nepoznata očekivana srednja vrijednost učinka svih jedinki;

$a$  - učinak skupine populacija (regije);

$b$  - učinak populacije unutar skupine populacija (regije);

$c$  - učinak jedinice unutar populacije.

Pretpostavka je da su navedeni učinci ( $a$ ,  $b$  i  $c$ ) aditivni, slučajni i nekorelirani, te da su im pridružene sastavnice varijance  $\sigma_a^2$  (između regije),  $\sigma_b^2$  (između populacija / unutar regija),  $\sigma_c^2$  (unutar populacija).

Za razliku od uobičajene ANOVA analize pri kojoj izračunavamo sume kvadrata pojedinih izvora varijabilnosti (SS) kod AMOVA analize koristimo sume kvadratnih odstupanja (Sums of Squared Deviations - SSD). Sume kvadratnih odstupanja su definirane kao odstupanja od centroida u multidimenzionalnom prostoru jer kao ishodišne podatke u analizi koristimo euklidske udaljenosti između jedinki, a ne vrijednosti samih jedinki. Stoga je Ukupna suma kvadratnih odstupanja (SSD(Ukupno)) jednaka:

$$SSD(Ukupno) = \frac{1}{2N} \sum_{j=1}^N \sum_{k=1}^N \delta_{jk}^2 \quad [3]$$

gdje je,

$N$  - Ukupni broj jedinki u pokusu;

$\delta_{jk}^2$  - Euklidska udaljenost između jedinki  $j$  i  $k$ ,

Sukladno dekompoziciji suma kvadrata, sume kvadratnih odstupanja također se mogu razdijeliti između i unutar pretpostavljenih razina strukture tako da je:

$$SSD(Ukupno) = SSD(Između) + SSD(Unutar) [4]$$

U navedenom primjeru postoje dvije razine strukture: regije i populacije unutar regija, te se stoga radi o dvosmjernom AMOVA analizi. U Tablici 3. prikazan je opći obrazac dvosmjerne hijerarhijske analize molekularne varijance.

Sume kvadratnih odstupanja (SSD) pojedinih izvora varijabiliteta dobivaju se sljedećim izrazima:

$$SSD(a) = \left( \frac{\sum_{j=1}^N \sum_{k=1}^N \delta_{jk}^2}{2N} - \sum_{g=1}^G \frac{\sum_{i=1}^{I_g} \sum_{j=1}^{N_{ig}} \sum_{i'=1}^{I_g} \sum_{k=1}^{N_{i'g}} \delta_{jk}^2}{\sum_{i=1}^{I_g} 2N_{ig}} \right) [5]$$

$$SSD(b) = \sum_{g=1}^G \left( \frac{\sum_{i=1}^{I_g} \sum_{j=1}^{N_{ig}} \sum_{i'=1}^{I_g} \sum_{k=1}^{N_{i'g}} \delta_{jk}^2}{\sum_{i=1}^{I_g} 2N_{ig}} - \sum_{i=1}^{I_g} \frac{\sum_{j=1}^{N_{ig}} \sum_{k=1}^{N_{ig}} \delta_{jk}^2}{2N_{ig}} \right) [6]$$

$$SSD(c) = \sum_{g=1}^G \sum_{i=1}^{I_g} \frac{\sum_{j=1}^{N_{ig}} \sum_{k=1}^{N_{ig}} \delta_{jk}^2}{2N_{ig}} [7]$$

gdje je,

$\delta_{jk}^2$  - Euklidska udaljenost između jedinki  $j$  i  $k$ ;

$G$  - broj skupina populacija (regija);

$I$  - broj populacija;

$N$  - broj jedinki.





Tablica 2. Matrica euklidskih udajenosti analiziranih jedinki volovoda (*Orobanchae cretana* Forsk.)

	Latroon	Masuot	Alonin	Mengibar	Pinos Puente	Cordoba
0						
23 0						
28 27 0						
21 25 35 0						
30 35 33 28 0						
27 25 32 30 24 0						
25 25 30 28 27 30 0						
27 29 35 30 29 30 24 0						
26 25 37 27 37 24 36 28 0						
27 22 27 32 34 30 29 23 27 0						
27 26 30 33 32 25 43 27 33 21 0						
18 17 26 24 30 25 23 21 25 22 26 0						
24 25 29 30 28 15 28 28 24 27 27 24 0						
24 27 33 29 33 16 30 24 18 22 23 21 20 0						
20 21 29 32 33 24 28 24 28 22 19 21 22 22 0						
27 18 28 29 34 25 31 23 21 24 26 20 21 23 19 0						
30 41 42 38 26 29 41 33 37 41 32 30 38 31 33 40 0						
30 30 44 37 27 26 32 30 40 36 31 27 27 26 26 29 37 0						
35 39 37 35 34 33 36 22 38 32 33 32 32 28 24 35 36 31 0						
23 26 30 27 23 30 21 24 32 31 33 23 31 25 20 29 21 24 25 0						
27 28 41 28 27 26 30 24 26 31 28 23 29 15 25 31 32 17 24 17 0						
30 30 37 28 17 26 29 25 38 28 31 31 28 32 24 28 32 27 28 21 24 0						
34 45 44 42 24 31 41 37 40 43 35 34 39 35 36 44 4 39 38 25 34 34 0						
36 31 33 39 30 29 30 24 28 30 32 30 31 28 27 27 31 32 29 26 29 25 35 0						



29 34 39 37 35 43 39 32 39 33 31 32 33 40 26 29 44 36 40 34 40 36 47 34 0  
35 38 40 37 31 32 30 22 34 32 35 25 33 30 28 27 39 40 33 30 34 32 38 32 31 0  
33 42 42 39 33 40 40 30 38 34 35 27 37 36 28 33 43 40 33 34 36 38 42 39 29 18 0  
30 40 48 29 34 37 41 27 33 36 36 29 38 37 31 28 35 37 37 36 36 34 36 31 28 18 21 0  
33 41 41 38 31 37 37 24 37 32 31 30 35 31 27 35 41 37 23 32 30 33 39 33 27 15 12 21 0  
32 32 41 32 28 31 30 13 26 31 29 23 29 26 27 23 34 33 27 28 29 31 33 23 26 13 18 17 14 0  
32 37 46 45 38 40 36 32 38 34 35 34 39 40 24 33 42 42 40 35 41 34 41 35 27 25 26 31 21 18 0  
36 40 46 41 37 35 37 25 35 39 38 33 31 35 33 32 43 36 37 36 38 34 44 27 32 20 24 25 19 17 24 0  
28 30 43 35 33 32 29 26 24 30 34 24 35 28 25 27 36 38 34 25 28 32 35 33 25 14 24 25 19 14 23 26 0  
29 35 36 36 32 34 26 26 36 32 32 32 30 32 19 25 39 36 31 26 33 27 38 29 24 11 21 21 17 17 16 21 19 0  
36 43 50 39 31 32 35 29 37 35 39 28 36 35 30 28 40 40 39 36 37 34 39 38 34 17 16 26 17 15 21 21 23 22 0  
35 37 45 40 32 38 30 24 38 30 39 26 30 32 28 27 46 39 38 33 39 31 45 34 30 19 17 24 19 13 25 25 25 20 17 0  
30 39 43 40 39 38 34 22 34 30 35 28 37 34 32 33 38 46 38 37 39 35 37 38 33 23 24 27 22 17 26 24 19 26 22 23 0  
30 39 43 36 29 39 33 29 35 39 36 30 33 33 27 32 36 36 36 23 31 33 35 33 24 17 20 23 22 21 28 24 19 16 27 21 28 0  
26 42 44 39 34 43 39 27 33 39 34 31 36 35 29 34 31 35 37 28 36 34 34 28 18 28 27 22 23 21 29 25 25 25 31 24 25 19 0  
32 38 46 45 42 43 37 31 33 35 42 33 38 37 29 26 49 45 45 36 40 36 52 40 24 26 31 34 28 20 25 33 18 25 27 22 25 31 24 0  
24 30 38 33 31 35 35 25 35 28 28 23 27 31 17 26 39 32 31 26 32 28 42 31 17 22 20 23 16 20 18 22 22 19 23 19 26 20 21 25 0  
39 49 54 48 46 52 46 42 41 46 47 38 52 45 38 39 52 51 47 42 42 47 51 47 24 25 23 27 25 23 30 39 20 26 31 34 32 26 29 19 33 0  
36 36 48 35 47 41 43 29 35 37 38 27 39 39 31 28 45 48 43 40 42 38 48 37 25 24 24 24 27 17 25 24 22 27 23 23 23 31 26 24 24 27 0  
22 35 41 28 36 37 43 33 33 35 32 28 36 33 23 28 36 41 40 29 33 33 39 38 24 23 25 17 28 28 30 29 21 20 31 28 30 18 25 25 19 30 21 0  
20 35 41 38 36 38 34 22 30 32 33 28 34 32 20 31 32 37 34 25 31 31 35 29 26 25 23 23 19 20 20 25 21 20 27 22 18 20 15 19 17 28 25 20 0  
27 31 40 36 30 37 33 26 31 28 33 29 32 28 24 29 33 31 31 19 27 29 37 30 27 18 20 21 22 22 24 25 21 15 28 21 30 13 19 27 18 31 32 17 17 0  
35 45 41 40 34 40 38 34 46 44 41 38 40 42 30 35 42 43 34 33 42 30 41 39 30 19 19 23 14 21 24 23 26 15 19 24 26 18 25 33 21 32 31 26 26 23 0  
42 47 49 48 44 55 49 43 51 43 42 38 48 49 37 46 54 51 46 43 51 49 53 49 32 27 21 30 19 33 27 29 33 28 35 28 35 21 34 40 25 34 37 26 28 23 24 0

Tablica 3. Opći obrazac dvosmjernje hijerarhijske analize molekularne varijance (AMOVA)

Izvori varijabilnosti	Stupnjevi slobode	Suma kvadratnih odstupanja (SSD)	Prosječno kvadratno odstupanje (MSD)	Očekivane vrijednosti prosječnog kvadratnog odstupanja E(MSD)
Između regija	$G - 1$	SSD(a)	MSD(a)	$\sigma_c^2 + n'\sigma_b^2 + n''\sigma_a^2$
Između populacija / Unutar regija	$\sum_{g=1}^G I_g - G$	SSD(b)	MSD(b)	$\sigma_c^2 + n\sigma_b^2$
Unutar populacija	$N - \sum_{g=1}^G I_g$	SSD(c)	MSD(c)	$\sigma_c^2$
Ukupno	$N - 1$	SSD(T)		

Odgovarajuća prosječna kvadratna odstupanja (Mean Squared Deviations - MSD) izračunavaju se dijeljenjem suma kvadratnih odstupanja s odgovarajućim stupnjevima slobode kao što je to slučaj i kod ANOVA-e. Sastavnice varijance ( $\sigma_a^2$ ,  $\sigma_b^2$  i  $\sigma_c^2$ ) moguće je izračunati poznavanjem prosječnih kvadratnih odstupanja i njihovih očekivanih vrijednosti pri čemu moramo poznavati  $n$  koeficijente koji omogućavaju analizu podataka iz uzoraka nejednakih veličina:

$$n = \frac{\sum_{g=1}^G \sum_{i=1}^{I_g} N_{ig} - \sum_{g=1}^G \left( \frac{\sum_{i=1}^{I_g} N_{ig}^2}{\sum_{i=1}^{I_g} N_{ig}} \right)}{\sum_{g=1}^G I_g} \quad [8]$$

$$n' = \frac{\sum_{g=1}^G \left( \frac{\sum_{j=1}^{I_g} N_{ig}^2}{\sum_{i=1}^{I_g} N_{ig}} \right) - \frac{\sum_{g=1}^G \sum_{j=1}^{I_g} N_{ig}^2}{\sum_{g=1}^G \sum_{i=1}^{I_g} N_{ig}}}{G - 1} \quad [9]$$

$$n'' = \frac{\sum_{g=1}^G \sum_{i=1}^{I_g} N_{ig} - \frac{\sum_{g=1}^G \left( \sum_{i=1}^{I_g} N_{ig} \right)^2}{\sum_{g=1}^G \sum_{i=1}^{I_g} N_{ig}}}{G - 1} \quad [10]$$

Tablica 4. Analiza molekularne varijance (AMOVA) populacija volovoda (*Orobanche crenata* Forsk.): sume kvadratnih odstupanja (SSD) i prosječna kvadratna odstupanja (MSD)

Izvori varijabilnosti	Stupnjevi slobode	Suma kvadratnih odstupanja (SSD)	Prosječno kvadratno odstupanje (MSD)
Između regija	1	122,44	122,44
Između populacija / Unutar regija	4	78,29	19,57
Unutar populacija	42	524,63	12,49
<b>Ukupno</b>	<b>47</b>	<b>725,35</b>	
Izrael: Između populacija	2	54,54	27,27
Izrael: Unutar populacija	21	277,13	13,20
<b>Izrael: Ukupno</b>	<b>23</b>	<b>331,67</b>	
Španjolska: Između populacija	2	23,75	11,88
Španjolska: Unutar populacija	21	247,50	11,79
<b>Španjolska: Ukupno</b>	<b>23</b>	<b>271,25</b>	

U Tablici 4. prikazane su sume kvadratnih odstupanja (SSD) i prosječna kvadratna odstupanja (MSD) prilikom provedbe AMOVA-e na populacijama volovoda. Poznavajući ukupnu varijancu ( $\sigma^2 = \sigma_a^2 + \sigma_b^2 + \sigma_c^2$ ) kao i njene sastavnice možemo izračunati udio varijabilnosti pojedinog izvora ( $[\sigma_a^2/\sigma^2] \times 100$  itd.). U svrhu testiranja signifikantnosti pretpostavljene strukture izračunavaju se  $\Phi$  vrijednosti koje su analogne  $F$  vrijednostima pri provedbi ANOVA-e:

$$\Phi_{ST} = \frac{\sigma_a^2 + \sigma_b^2}{\sigma^2} \quad [11]$$

$$\Phi_{CT} = \frac{\sigma_a^2}{\sigma^2} \quad [12]$$

$$\Phi_{SC} = \frac{\sigma_b^2}{\sigma_b^2 + \sigma_c^2} \quad [13]$$

$\Phi$  vrijednosti se pritom definiraju na sljedeći način:  $\Phi_{ST}$  je korelacija udaljenosti između dviju slučajno odabranih jedinki iz iste populacije i udaljenosti između dviju slučajno odabranih jedinki iz cjelokupnog pokusa;  $\Phi_{CT}$  je korelacija udaljenosti između dviju slučajno odabranih jedinki iz iste regije i

udaljenosti između dviju slučajno odabranih jedinki cjelokupnog pokusa, dok je  $\Phi_{SC}$  je korelacija udaljenosti između dviju slučajno odabranih jedinki iste populacije i udaljenosti između dviju slučajno odabranih jedinki iz iste regije.

U Tablici 5. navedene su sastavnice varijance, udio varijabilnosti i  $\Phi$  vrijednosti za volovod. Testiranje signifikantnosti  $\Phi$  vrijednosti nepoznate nulte distribucije provodi se neparametrijski pomoću permutacija. Permutiranjem svih podataka iz matrice euklidskih udaljenosti (Mantel, 1967) i izračunavanjem odgovarajućih sastavnica varijance kao i  $\Phi$  vrijednosti možemo dobiti nultu distribuciju parametra  $\Phi_{ST}$  kao i  $\sigma_c^2$ . Nultu distribuciju parametra  $\Phi_{SC}$  i  $\sigma_b^2$  možemo dobiti permutiranjem jedinki unutar regija bez obzira kojoj populaciji pripadali, dok se nulta distribucija parametra  $\Phi_{CT}$  i  $\sigma_a^2$  dobiva permutiranjem cjelokupnih populacija između regija.

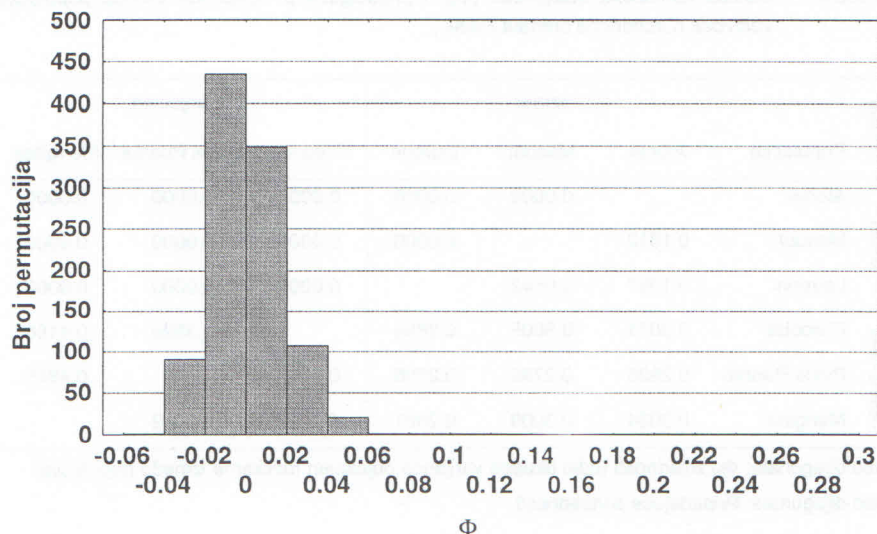
Tablica 5. Analiza molekularne varijance (AMOVA) populacija volovoda (*Orobancha crenata* Forsk.): sastavnice varijance, udio varijabilnosti i  $\Phi$  vrijednosti

Izvori varijabilnosti	Stupnjevi slobode	Sastavnice varijanci	Udio varijabilnosti (%)	$\Phi$ vrijednosti
Između regija (Izrael i Španjolska)	1	$\sigma_a^2 = 4,286$	24,27	$\Phi_{CT} = 0,243$
Između populacija / Unutar regija	4	$\sigma_b^2 = 0,885$	5,01	$\Phi_{SC} = 0,066$
Unutar populacija	42	$\sigma_c^2 = 12,491$	70,72	$\Phi_{ST} = 0,293$
Ukupno	47	$\sigma^2 = 17,6621$		
Izrael: Između populacija	2	$\sigma_a^2 = 1,759$	11,76	$\Phi_{ST} = 0,118$
Izrael: Unutar populacija	21	$\sigma_b^2 = 13,196$	88,24	
Izrael: Ukupno	23	$\sigma^2 = 14,956$		
Španjolska: Između populacija	2	$\sigma_a^2 = 0,011$	0,09	$\Phi_{ST} = 0,001$
Španjolska: Unutar populacija	21	$\sigma_b^2 = 11,786$	99,991	
Španjolska: Ukupno	23	$\sigma^2 = 11,797$		

Kao što je vidljivo iz Grafikona 1. na kojem je prikazana raspodjela učestalosti pojavljivanja određenih  $\Phi_{ST}$  vrijednosti, nakon 1000 permutacija niti jednom permutacijom nije dobivena veća  $\Phi_{ST}$  vrijednost od procijenjene iz

stvarnih podataka (0,293). Isti tako, procijenjene  $\Phi_{SC}$  i  $\Phi_{CT}$  vrijednosti su vrlo signifikantne ( $p < 0,001$ ), te možemo zaključiti da je pretpostavljena struktura varijabilnosti i stvarna. No, bilo kako bilo, većina raznolikosti uzrokovana je razlikama između jedinki (70,72 %). Ostala raznolikost bila je nejednako raspoređena na razlike između regija (24,27 %) i razlike između populacija (5,01 %).

Graf. 1. Raspodjela učestalosti  $\Phi_{ST}$  vrijednosti nakon 1000 permutacija



Budući da je analiza obuhvaćala dvije razine strukture (populacije i regije) bilo je moguće provesti jednosmjerne AMOVA-e za svaku regiju (Izrael i Španjolska) posebno da bismo dobili bolji uvid u populacijsku strukturu unutar regija. Za razliku od Izraela gdje je populacijska struktura potvrđena signifikantnom  $\Phi_{ST}$  vrijednosti, unutar Španjolske raznolikost je uzrokovana isključivo razlikama između jedinki, a ne i njihovom pripadnosti različitim populacijama. Udio varijabilnosti između populacija u Španjolskoj iznosio je samo 0,09 %.

#### MATRICA $\Phi_{ST}$ VRIJEDNOSTI IZMEĐU POPULACIJA I ANALIZA SKUPINA

Nadalje, AMOVA analiza može se provesti posebno za svaki par populacija i utvrditi pripadajuće  $\Phi_{ST}$  vrijednosti koje su u ovom slučaju ekvivalentne udjelu

ukupne varijance koja je raspodijeljena između ispitivanih populacija (Gustine i Huff, 1999). Stoga se  $\Phi_{ST}$  vrijednosti između svakog para populacija mogu smatrati prosječnim međupopulacijskim udaljenostima, a istodobno je moguće testirati njihovu signifikantnost pomoću permutacija.

U Tablici 6. prikazana je matrica fenotipskih udaljenosti (odnosno  $\Phi_{ST}$  vrijednosti između svih parova populacija) kao i pripadajuće  $p$  vrijednosti između populacija volovoda.

Tablica 6. Matrica fenotipska udaljenosti ( $\Phi_{st}$ ) i pripadajuće  $p$  vrijednosti između populacija volovoda (*Orobanche crenata* Forsk.)

Regija	Populacija	Izrael			Španjolska		
		Alonin	Masuot	Latroon	Cordoba	Pinos Puente	Mengibar
Izrael	Alonin		0.0000	0.0000	0.0000	0.0000	0.0000
	Masuot	0.1812		0.0000	0.0000	0.0000	0.0000
	Latroon	0.1097	0.0542		0.0000	0.0000	0.0000
Španjolska	Cordoba	0.3071	0.3005	0.2898		0.3896	0.4166
	Pinos Puente	0.2935	0.2792	0.2708	0.0047		0.4815
	Mengibar	0.3034	0.3009	0.2881	-0.0022	0.0002	

Ispod dijagonale:  $\Phi_{st}$  vrijednosti (udio ukupne varijance objašnjen razlikama između populacija)

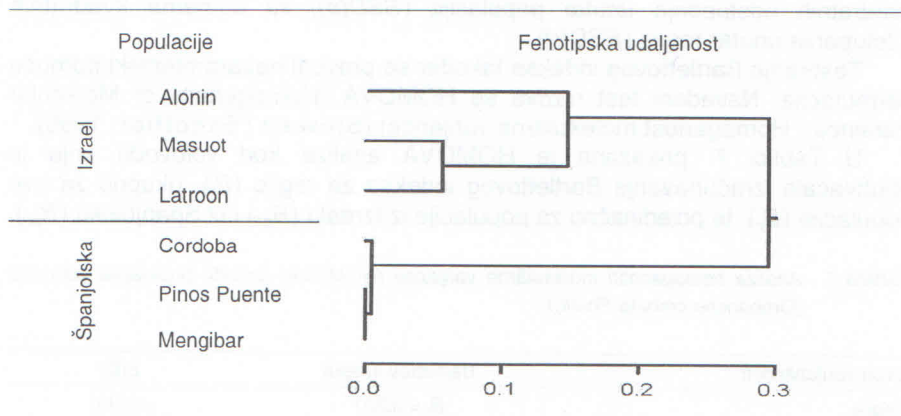
Iznad dijagonale: Pripadajuće  $p$  vrijednosti

Na temelju matrice  $\Phi_{ST}$  vrijednosti između svih parova populacija moguće je provesti analizu skupina (cluster analysis) u svrhu grafičkog predovačavanja odnosa između populacija. Pritom se često koristi metoda UPGMA (Unweighted Pair-Group Method of Arithmetic Averages) (Huff et al., 1997). Međupopulacijske udaljenosti dobivene nakon analize skupina mogu se usporediti s ishodišnom matricom  $\Phi_{ST}$  vrijednosti pomoću kofenetičkog koeficijenta čija se signifikantnost ispituje neparametrijski, Mantelovim testom (Mantel, 1967) pomoću permutacija matrica.

Grafikon 2. prikazuje dendrogram na temelju fenotipskih udaljenosti između populacija volovoda nakon provedene analize skupina (cluster analysis; metoda UPGMA).

Kofenetički koeficijent između matrice međupopulacijskih udaljenosti nakon provedene analize skupina i ishodišne matrice bio je signifikantan ( $r = 0,9917$ ;  $p = 0,012$ ).

Graf. 2. Dendrogram na temelju fenotipskih udaljenosti između populacija volovoda nakon provedene analize skupina (cluster analysis; metoda UPGMA)



#### ANALIZA HOMOGENOSTI MOLEKULARNE VARIJANCE (HOMOVA)

Homogenost varijanci između analiziranih izvora populacija nužna je pretpostavka u testiranju parametara ANOVA-e. No, budući da se pri AMOVA-i  $\Phi$  vrijednosti kao i sastavnice varijance testiraju neparametrijski pomoću permutacija homogenost varijanci nije potrebna. Bilo kako bilo, u mnogim je istraživanjima bioraznolikosti poznavanje načina raspodjele varijabilnosti između različitih populacija samo po sebi bitno. Bartlettov test homogenosti varijance (Bartlett, 1937) se stoga može preinačiti tako da odražava odstupanja varijanci pojedinih populacija (ili regija) od srednje ukupne varijance istraživanja. Bartlettov indeks ( $B_p$ ) heteroscedastičnosti (homogenosti varijanci) između ispitivanih populacija može se izraziti u sumama kvadratnih odstupanja na sljedeći način:

$$B_p = \frac{(N - I) \ln \left( \frac{SSD(T)}{N - I} \right) - \sum_{i=1}^I (N_i - 1) \ln \left( \frac{SSD(p)_i}{N_i - 1} \right)}{1 + \frac{1}{3(I - 1)} \left( \sum_{i=1}^I \frac{1}{N_i - 1} - \frac{1}{N - 1} \right)} \quad [14]$$

gdje je,

$I$  - broj populacija;

$N$  - broj jedinki.

$SSD(T)$  - suma kvadratnih odstupanja svih jedinki;

$SSD(p)_i$  - suma kvadratnih odstupanja unutar  $i$ -te populacije.

Na isti način možemo izračunati Bartlettov indeks homogenosti varijanci između regija ( $B_r$ ) tako da zamjenimo  $I$  (broj populacija) s  $G$  (broj regija) i sume kvadratnih odstupanja unutar populacija ( $SSD(p)$ ) sa sumama kvadratnih odstupanja unutar regija ( $SSD(r)$ ).

Testiranje Bartlettovog indeksa također se provodi neparametrijski pomoću permutacija. Navedeni test naziva se HOMOVA (Homogeneity of Molecular Variance - Homogenost molekularne varijance) (Stewart i Excoffier, 1996).

U Tablici 7. prikazana je HOMOVA analiza kod volovoda koja je obuhvaćala izračunavanje Bartlettovog indeksa za regije ( $B_r$ ), ukupno za sve populacije ( $B_p$ ), te pojedinačno za populacije iz Izraela ( $B_{p1}$ ) i iz Španjolske ( $B_{p2}$ ).

Tablica 7. Analiza homogenosti molekularne varijance (HOMOVA) između populacija volovoda (*Orobanche crenata* Forsk.)

Izvori varijabilnosti	Bartlettov indeks	$p(B)$
Regije	$B_r = 0,227$	< 0,001
Populacije	$B_p = 0,181$	0,198
Izrael: Populacije	$B_{p1} = 0,079$	0,105
Španjolska: Populacije	$B_{p2} = 0,037$	0,449

$p(B)$  - signifikantnost Bartlettovog indeksa dobivena nakon 1000 permutacija

Analiza homogenosti molekularne varijance otkiva da su molekularne varijance signifikantno heterogene između regija, no ne i između populacija. Isto tako potvrđena je homogenost varijanci između izraelskih populacija kao i između španjolskih populacija. Bez obzira na nesignifikantnost heterogenosti varijanci izraelskih populacija one se ipak međusobno signifikantno razlikuju (signifikantan  $\Phi_{ST}$ ). No, većina varijabilnosti je svakako uzrokovana razlikama između jedinki (udio varijabilnosti između populacija je 11,76 %), a ne između populacija.

#### NAPOMENA O SOFTWARE-U:

Pojedine faze AMOVA i HOMOVA analize mogu provesti upotrebom standardnih statističkih programa. U SAS-u (SAS, 1990) je moguće izračunati euklidske udaljenosti između jedinki upotrebom makroa (Kuo, 1997) koji daje mnogobrojne mogućnosti za izračunavanje različitih udaljenosti (kao npr. euklidska udaljenosti, Jaccardov koeficijent, Diceov koeficijent, Gowerov koeficijent itd.). NTSYS-pc (Rohlf, 1990) je program specijaliziran za numeričku taksonomiju i multivarijatnu analizu, te je pogodan za izračunavanje kofenetičkih koeficijenata kao i za provedbu Mantelovog testa. Velik broj raznolikih *shareware* programa za upotrebu u filogenetskih istraživanjima može se naći na: <http://evolution.genetics.washington.edu/phylip/software.html>.



Bilo kako bilo, Arlequin ver. 2.000 (Schneider et al., 2000) specijaliziranih je program za analizu podataka iz populacijske genetike. Program kao i uputstva za njegovu instalaciju i upotrebu nalaze se na: <http://anthro.unige.ch/arlequin>. Dosad je stavljeno na raspolaganje više verzija programa Arlequin (Schneider et al., 1997), a prije izlaska Arlequina postojali su manji programi istih autora pod nazivom AMOVA i WinAMOVA. No, HOMOVA analiza nije uključena u program Arlequin.

## THE ANALYSIS OF MOLECULAR VARIANCE (AMOVA) AND THE TEST FOR THE HOMOGENEITY OF MOLECULAR VARIANCE (HOMOVA) WITH RAPD MARKERS

### SUMMARY

The analysis of molecular variance (AMOVA) and the nonparametric test for the homogeneity of molecular variance (HOMOVA) have been reviewed. The calculation procedure was illustrated with broomrape (*Orobanche crenata* Forsk.) RAPD population data.

Key words: AMOVA, HOMOVA, *Orobanche crenata* Forsk.

### LITERATURA - REFERENCES

1. Bartlett, M.S. 1937. Some examples of statistical methods of research on agriculture and applied biology. *Journal of the Royal Statistical Society Supplement* 4: 137-170
2. Caetano-Anolles, G. 1994. MAAP - A versatile and universal tool for genome analysis. *Pl. Molec. Biol.* 25: 1011-1026
3. Crochemore, M.L., Huyghe, C., Kerlan, M.C., Durand, F. i Julier, B. 1996. Partitioning and distribution of RAPD variation in a set of populations of the *Medicago sativa* complex. *Agronomie* 16(7): 421-432
4. Cubero, J.I. i Moreno, M.T., 1979. Agronomical control and sources of resistance in *Vicia faba* to *Orobanche crenata* Forsk. U: Bond, D.A., Scarascia-Mugnozza, G. T. i Poulsen, M. H. (ur.) Some current research on *Vicia faba* in Western Europe. Commission of the European Communities, Luksemburg
5. Cubero, J.I. 1983. Parasitic diseases in *Vicia faba* L. With special reference to broomrape (*Orobanche crenata* Forsk.). U: Hebblethwaite, P.B. (ur.) *The Faba Bean*. Butterworths, London, Velika Britanija
6. Cubero, J.I., 1991. Breeding for resistance to *Orobanche* species: A review. U: Wegmann, K. i Musselman, L.J. (ur.) *Progress in Orobanche Research*. Eberhard-Karls-Universität, Tübingen, Njemačka
7. Duarte, J.M., dos Santos, J.B. i Melo, L.C. 1999. Genetic divergence among common bean cultivars from different races based on RAPD markers. *Genetics and Mol. Biol.* 22(3): 419-426
8. Excoffier, L., Smouse, P.E. i Quattro, J.M. 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction sites. *Genetics* 131: 479-491

9. Fofana, B., Vekemans, X., du Jardin, P. i Baudoin, J.P. 1997. Genetic diversity in lima bean (*Phaseolus lunatus* L.) as revealed by RAPD markers. *Euphytica* 95(2): 157-165
10. Fulton, T.M., Chunwongse, J., Tanksley, S.D. 1995. Microprep protocol for extraction of DNA from tomato and other herbaceous plants. *Plant Mol. Biol. Rep.* 13: 207-209
11. Ge Song, Oliveira, G.C.X., Schaal, B.A., Gao Li Zhi, Hong De Yuan, Ge, S., Gao, L.Z. i Hong, D.Y. 1999. RAPD variation within and between natural populations of the wild rice *Oryza rufipogon* from China and Brazil. *Heredity* 82(6): 638-644
12. Gupta, M., Chyi, Y. S., Romero-Severson, J. i Owen, J. L. 1994. Amplification of DNA markers from evolutionarily diverse genomes using singel primers of simple-sequence repeats. *Theor. Appl. Genet.* 89: 998-1006
13. Gustine, D.L. i Huff, D.R. 1999. Genetic variation within and among white clover populations from managed permanent pastures of the northeastern USA. *Crop Sci.* 39: 524-530
14. Hogbin, P.M., Ayre, D.J. i Whelan, R.J. 1998. Genetic variation and reproductive success of road verge populations of the rare shrub *Grevillea barklyana* (Proteaceae). *Heredity* 80(2): 180-186
15. Huff, D.R., Peakall, R. and Smouse, P.E. 1993. RAPD variation within and among natural populations of outcrossing buffalograss [*Buchloë dactyloides* (Nutt.) Engelm.]. *Theor. Appl. Genet.* 86: 927-934
16. Huff, D.R. 1997. RAPD characterization of heterogeneous perennial ryegrass cultivars. *Crop Sci.* 37(2): 557-564
17. Jenczewski, E., Prosperi, J.M. i Ronfort, J. 1999. Differentiation between natural and cultivated populations of *Medicago sativa* (Leguminosae) from Spain: analysis with random amplified polymorphic DNA (RAPD) markers and comparison to allozymes. *Mol. Ecol.* 8(8): 1317-1330
18. Kolliker, R., Stadelmann, F.J., Reidy, B. i Nosberger, J. 1998. Fertilization and defoliation frequency affect genetic diversity of *Festuca pratensis* Huds. in permanent grasslands. *Mol. Ecol.* 7(11): 1557-1567
19. Kuo, A. 1997. The DISTANCE Macro: Preliminary Documentation, 2nd ed. Multivariate & Numerical R & D, Application Division, SAS Institute Inc. Cary, NC, SAD
20. Mantel, N. 1967. The detection of disease clustering and a generalized regression approach. *Cancer Res.* 27: 209-220
21. Michalakis, Y. i Excoffier, L. 1996. A generic estimation of population subdivision using distances between alleles with special reference for microsatellite loci. *Genetics* 142: 1061-1064
22. Mullis, K. B. i Faloona, F. A. 1987. Specific synthesis of DNA in vitro via a polymerase-catalyzed chain reaction. *Methods in Enzymology* 155: 335-350
23. Nebauer, S.G., del Castillo Agudo, L. i Segura, J. 1999. RAPD variation within and among natural populations of outcrossing willow leaved foxglove (*Digitalis obscura* L.). *Theor. Appl. Genet.* 98(6-7): 985-994
24. Nei, M. i Tajima, F. 1981. DNA polymorphism detectable by restriction endonucleases. *Genetics* 97: 145-163
25. Peakall, R., Smouse, P.E. i Huff, D.R. 1995. Evolutionary implications of allozyme and RAPD variation in diploid populations of dioecious buffalograss *Buchloë dactyloides*. *Mol. Ecol.* 4(2): 135-147
26. Rohlf, F.J. 1990. NTSYS: Numerical Taxonomy and Multivariate Analysis System Version 1.60. Exeter Publ. Ltd., Setauket, NY, SAD
27. Saiki, R.K., Scharf, S., Faloona, F., Mullis, K.B., Horn, G.T., Erlich, H.A. i Arnheim, N. 1985. Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. *Science* 230(4732): 1350-1354
28. SAS 1990. SAS/STAT user's guide, ver. 6, 4th ed. vol. 2, SAS Institute Inc., Cary, NC, SAD
29. Šatović, Z. 1999. Genetski biljezi i njihova uporaba u biljnoj genetici, oplemenjivanju bilja i sjemenarstvu. *Sjemenarstvo* 1-2: 73-95
30. Schneider, S., Kueffer, J.M., Roessly, D. i Excoffier, L. 1997. Arlequin ver. 1.1: A software for population genetic data analysis. Genetics and Biometry Laboratory, University of Geneva, Švicarska

31. Schneider, S., Roessli, D. i Excoffier, L. 2000. Arlequin ver. 2.000: A software for population genetics data analysis. Genetics and Biometry Laboratory, University of Geneva, Švicarska
32. Stewart, C. N. and Excoffier, L. 1996. Assessing populaton structure and variability with RAPD data: Application to *Vaccinium macrocarpon* (American cranberry). J. Evol. Biol. 9: 153-171
33. Vicario, F., Vendramin, G.G., Rossi, P., Lio, P. i Giannini, R. 1995. Allozyme, chloroplast DNA and RAPD markers for determining genetic relationships between *Abies alba* and the relic population of *Abies nebrodensis*. Theor. Appl. Genet. 90(7-8): 1012-1018
34. Welsh, J. i McClelland, M. 1990. Fingerprinting genomes using PCR with arbitrary primers. Nucleic Acids Res. 18: 7213-7218.
35. Williams, J.G.K., Kubelik, A.R., Livak, K.J., Rafalski, J.A. and Tingey, S.V. 1990. DNA polymorphism amplified by arbitrary primers are useful as genetic markers. Nucleic Acids Res. 18:6531-6535.

**Adresa autora – Authors' addresses:**

Doc. dr. sc. Zlatko Šatović  
Prof. dr. sc. Ivan Kolak  
Agronomski fakultet Sveučilišta u Zagrebu  
Zavod za sjemenarstvo  
Svetošimunska 25  
HR-10000 Zagreb

Dr. sc. Belén Román del Castillo  
Dr. sc. Ana María Torres Romero  
CIFA "Alameda de obispo"  
Departamento de Mejora y Agronomía  
Apartado 3092  
14080 Córdoba  
Španjolska

Prof. dr. sc. José Ignacio Cubero Salmerón  
Universidad de Córdoba  
Escuela Técnica Superior de Ingenieros Agrónomos y Montes (E.T.S.I.A.M.)  
Departamento de Genética  
Apartado 3048  
14080 Córdoba  
Španjolska

**Primljeno - Received:**

12. 12. 2000.