



EFOP-3.4.3-16-2016-00014



Pénzes Zsolt, Markó Bálint

Molekuláris ökológia: populációk szerkezete

Jelen tananyag a Szegedi Tudományegyetemen
készült az Európai Unió támogatásával.

Projekt azonosító: EFOP-3.4.3-16-2016-00014

Szegedi Tudományegyetem
Cím: 6720 Szeged, Dugonics tér 13.
www.u-szeged.hu
www.szechenyi2020.hu





Európai Unió
Európai Szociális
Alap



MAGYARORSZÁG
KORMÁNYA

BEFEKTETÉS A JÖVŐBE

A molekuláris ökológia előadások célja a molekuláris módszerek néhány alkalmazási lehetőségének bemutatása ökológiai és evolúcióbio-
lógiai problémák megfogalmazásában/megválaszolásában. Kérdése-
ink populációkra, fajokra vonatkoznak – például populációk izoláci-
ójának mértéke, egy invazív faj eredete, leszármazási kapcsolatok.
A válasz keresése során a molekuláris módszerek eszközökként szol-
gálnak.

Az előadáson a populáció szerkezetének tradicionális jellemzését és
a populációk közötti migráció, génáramlás koncepcionális modelljeit
tárgyaljuk, továbbá kiegészítjük a standard koaleszcencia modellt
migrációval.

Ajánlott irodalom: Nielsen R, Slatkin M (2013) An Introduction to
Population Genetics. Sinauer. (4. fejezet)

Feltételezett előismeretek: molekuláris markerek, a genetikai válto-
zatosság mérőszámai, ideális populáció, neutrális változás, koalesz-
cens modell.

Az olvasólecke átnézése kb. 20 percet vesz igénybe.

Populációk szerkezete

- Szerkezet: nem random kombinálódás – földrajzi tényezők
- Természetvédelmi vonatkozások – genetikai egységek
- Genetikai változatosság – modellek
 - a változatosság jellemzése
 - változás – folyamatok, tényezők hatása
 - génáramlás (migráció, izoláció)
 - populációk divergenciája
- Példa: hosszúszárnyú bálna (*Megaptera novaeangliae*)
 - sok vizsgálat – pl. mikroszatellit elemzések
 - nagy mozgáskörzet – populációk?
 - az atlanti és pacifikus populációk jelentős differenciálódása (nagy populációk közötti különbség) – kvantitatív kifejezése?

Az allél kombinálódás egy kiválasztott lokuszon gyakran nem véletlenszerű, ennek számos oka lehet, így például a populációk földrajzi tagolódása. Például egy sziget populáció egy nőstény egyede gyakrabban szaporodik egy a szigeten élő hímmel, mint egy a kontinensen élővel. **Populáció szerkezet** (struktúra) alatt ezt a nem véletlenszerű kombinálódást, annak következményeit értjük. A struktúra és annak változása földrajzi léptékű (geográfiai, térbeli) és történeti (időbeli) mintázatok meghatározó tényezője, így hatását jellemoznünk és értelmeznünk kell. A szerkezet jellemzése a molekuláris markerek alkalmazásának egyik alapvető célja, lehetővé teszi például a szaporodási egységek elkülönítését egy nagy elterjedési területen belül.

A szerkezet következményeit a genetikai változatosságra **modellek** segítségével becsüljük. Célunk egyrészt a populációk közötti eltérés mértékének kvantitatív jellemzése erre alkalmas mérőszámok segítségével („leírás”), másrészt betekintést nyerni a populációk közötti különbséget befolyásoló – generáló, fenntartó, csökkentő – evolúciós tényezőkbe, végső soron a szerkezet hatását kifejezni az evolúciós változás sebességére és irányára.

Populációk szerkezete

- Modellek – alapeset:

- két azonos méretű populáció (N), mindkettőben Hardy-Weinberg egyensúly (HWE), egy lokusz – két allél (A , a) modell
- átlagos allél gyakoriság és populációk közötti különbség:

$$f_A = \frac{f_{A1} + f_{A2}}{2} \quad \text{és} \quad \delta = |f_{A1} - f_{A2}|$$

- Heterozigóta arány

- H_S (subdivided): szerkezet figyelembe véve, populációk átlaga

$$H_S = \frac{2f_{A1}(1 - f_{A1}) + 2f_{A2}(1 - f_{A2})}{2} = f_{A1}(1 - f_{A1}) + f_{A2}(1 - f_{A2})$$

- H_T (total): egyesített populáció, A allél gyakorisága f_A

$$H_T = 2f_A(1 - f_A) = f_{A1}(1 - f_{A1}) + f_{A2}(1 - f_{A2}) + \delta^2/2$$

- azonos allél gyakoriság ($\delta = 0$): $H_T = H_S$ – egyesített populációban is HWE, nincs kimutatható szerkezet

Legegyszerűbb esetben két allél (A és a) szegregálódik egy lokuszon két diploid populációban (1 és 2), melyek mérete N_1 és N_2 és az allél gyakoriságok $f_{A1} = 1 - f_{a1}$, illetve $f_{A2} = 1 - f_{a2}$. Az A allél gyakorisága a két populációban együttesen (egyesített populációban, az **átlagos allél gyakoriság**) ekkor

$$f_A = \frac{2N_1f_{A1} + 2N_2f_{A2}}{2N_1 + 2N_2} \quad \text{és} \quad f_a = 1 - f_A.$$

A továbbiakban az egyszerűség kedvéért tegyük fel, hogy a populációk mérete azonos ($N_1 = N_2 = N$), ez esetben az átlagos allél gyakoriság $f_A = (f_{A1} + f_{A2})/2$. Tegyük fel továbbá, hogy mindkét populáció Hardy-Weinberg egyensúlyban (HWE) van. Ez esetben a genotípus gyakoriságokat is ismerjük, a heterozigóták aránya (**heterozigotitás**, H) ekkor $H_1 = 2f_{A1}(1 - f_{A1})$ és $H_2 = 2f_{A2}(1 - f_{A2})$. Ez utóbbi esetben a populációk átlagos heterozigotitása $H_S = (H_1 + H_2)/2$, ezzel a mérőszámmal a populáció tagolódást is figyelembe vettük. A két populációt egyesítve, mintha egy tagolódás nélküli populáció lenne, a heterozigotitás $H_T = 2f_A(1 - f_A) = H_S + \delta^2/2$, ahol δ a populációk közötti allél gyakoriság különbség és f_A az átlagos allélgyakoriság. Azonos allél gyakoriság ($\delta = 0$) esetén $H_T = H_S$, az egyesített populáció is Hardy-Weinberg egyensúlyban van.

Populációk szerkezetének jellemzése

- Szerkezet ($\delta > 0 \Rightarrow H_T > H_S$) következménye a vártnál kevesebb heterozigóta (több homozigóta) – Wahlund hatás
- Szerkezet: allél gyakoriság különbség a populációkban
 - jellemzése: F_{ST}
 - számos változat, különböző definíciók – pl. H_T -vel standardizált eltérés:

$$F_{ST} = \frac{H_T - H_S}{H_T}$$

- Több populáció – a szerkezet jellemzése
 - azonos allél gyakoriság ($\delta = 0$): $H_T = H_S$, így $F_{ST} = 0$
 - maximális allél gyakoriság különbség ha eltérő allél fixálódása (pl. $f_{A1} = 0$ és $f_{A2} = 1$): $H_T > H_S$ és $H_S = 0$, így $F_{ST} = 1$
 - $0 \leq F_{ST} \leq 1$ – empirikus szabályok (Wright), de tesztelhető...
 $F_{ST} < 0,05$ nincs, vagy kis mértékű differenciálódás
 $F_{ST} > 0,25$ jelentős differenciálódás
- F -statisztikák (F beltenyésztési együtttható, F_{ST}) a Hardy-Weinberg egyensúlytól való eltérés magyarázatára

Eltérő allél gyakoriság esetén a populációk átlagos heterozigóta aránya mindig kisebb (vagy homozigotizáció nagyobb) a vártnál ($H_T > H_S$), annál az esetnél, ha egy tagolódás nélküli populációknak lenne. Ezt **Wahlund-hatásnak** nevezzük, amely általánosabb körülmények között is igaz (eltérő populáció méretek, több populáció, több allél/lokusz).

A populáció szerkezetet definiáljuk a továbbiakban populációk/alpopulációk közötti allél gyakoriság különbségként. Jellemzésének általánosan használt mérőszáma az F_{ST} érték (Wright F_{ST}). Több különböző koncepciója létezik, egyik általánosan használt az egyesített populációra (vagyis a várt heterozigotizációra) vonatkoztatott, standardizált eltérés. Azonos allél gyakoriság esetén $H_T = H_S$, így $F_{ST} = 0$, míg maximális allél gyakoriság különbség esetén, vagyis ha eltérő allél fixálódik a populációkban (pl. $f_{A1} = 0$ és $f_{A2} = 1$), $H_T > H_S = 0$, így $F_{ST} = 1$. Wright ugyan empirikus szabályokat is megfogalmazott az F_{ST} értelmezésére, azonban ez csak tájékoztató céllal alkalmazható a gyakorlatban.

A hosszúszárnnyú bálna mikroszatellit elemzés során azt találták, hogy az észak-atlanti mintavételi helyek közötti eltérés $F_{ST} \approx 0,04$, míg az atlanti és pacifikus populációk közötti differenciálódásra $F_{ST} > 0,4$. Ez utóbbi jelentős differenciálódásra utal.

Populációk szerkezetének jellemzése

- F_{ST} becslésére különböző módszerek – információ kombinálása több lokusz alapján
- Több lokusz – különbségek genom léptékben
 - mutáció+drift eltérő hatása – de genom léptékben átlagolt F_{ST} a genetikai differenciálódás mértékének jó mérőszáma
 - egyes lokuszokra kiugró értékek egyéb allél gyakoriságra ható folyamatokra utalhatnak – pl. szelekció...
- Átlagos F_{ST}
 - emlősök: fajok közötti különbségekre $F_{ST} > 0,8$ általában, allopatrikus populációkra $0,1 < F_{ST} < 0,8$
 - ember: nagyobb elterjedés, de kisebb különbségek, pl. Európában $F_{ST} < 0,025$
kontinentális csoportok között $0,05 < F_{ST} < 0,2$

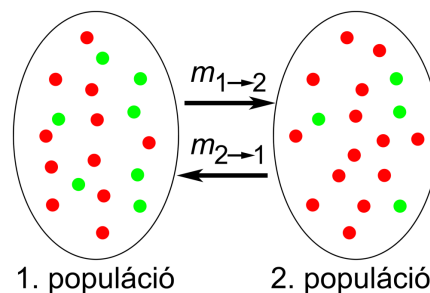
Az F_{ST} becslésére különböző módszerek állnak rendelkezésre. A gyakorlatban általában több különböző (független) lokuszra kapott információt kombinálva becsüljük. Mivel az allél gyakoriságok a genom egyes lokuszain különbözhetnek, ezért a lokuszonkénti F_{ST} értékek is eltérhetnek. Ennek egyik oka az, hogy neutrális esetben a populációk változatosságát mutációk és genetikai sodródás formálják, együttes hatásuk eltérő módon realizálódhat a különböző genom régiókra. Általános nézet, hogy a genom szinten **átlagolt** F_{ST} a genetikai differenciálódás megbízható jellemzője.

Azonban egyes lokuszok kiugró F_{ST} értékeket mutathatnak a populációk közötti jelentős allél gyakoriság különbség következtében. Ez utalhat például természetes szelekció hatására, eltérő allél lehet előnyös a különböző populációkban. Ilyen módon egy genom régióra ható szelekció hatására következtetünk, ez egyike a genom szinten kimutatható szelekcióra utaló jelnek.

Fajok, illetve egy adott faj populációi között jelentős különbségeket tapasztalhatunk a gyakorlatban. Ez alól az ember genetikai változatossága szembetűnő kivétel, különösen ha figyelembe vesszük nagy földrajzi elterjedését. A tapasztalt eltérés hátterében különböző folyamatok, tényezők állhatnak – magyarázhatja a populációk múltbeli divergenciája vagy a recens génáramlás egyaránt.

Migráció

- Folyamatok hatása – migráció (populációk közötti génáramlás)
- Alapeset: a Wright-Fisher (WF) modell kiterjesztése, két populáció
 - populációk a WF modell szerint evolválódnak
 - alkalmanként helyettesítés a másik populáció egyedével – valószínűsége $m_{1 \rightarrow 2}$ és $m_{2 \rightarrow 1}$
 - m migrációs ráta/egyed/generáció
 - allél gyakoriságok idővel azonossá válnak a két populációban...



A populációk allél gyakoriságának eltérő mértékű módosulásával az F_{ST} is változik, melynek hátterében különböző tényezők állhatnak. Belátható, hogy például a populációk közötti folyamatos **génáramlás** következtében a populációk allél gyakoriságai azonossá válnak a két populációban, így F_{ST} a 0 értékhez tart az idő függvényében ($F_{ST} \rightarrow 0$). Génáramlás alatt az egyedek populációk közötti **migrációját** és a bevándorlóknak a célpopuláció rezidens egyedeivel történő szaporodását értjük. Mérőszáma a generációra és egyedre vonatkoztatott **migrációs ráta** (m). Intenzív génáramlás a populációk közötti hasonlóságot nagyobb mértékben növeli, hiányában pedig a különbségek növekedését várjuk mutáció és genetikai sodródás következtében egy neutrális lokuszon. Ez utóbbi magyarázza a hosszúsárnyú bálna populációk két óceán közötti jelentős differenciálódását, mivel az óceánok közötti migráció ritka esemény. Az egyszerűség kedvéért két olyan populációból indulunk ki, amelyek a Wright-Fisher (WF) modell szerint evolválódnak. Alkalmanként génáramlás történik közöttük, ennek valószínűsége generációnként m egy egyedre vonatkoztatva (pontosabban egyedek helyettesítése történik, a WF modellben a populáció mérete konstans).

Migráció

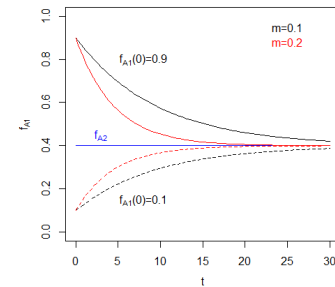
- A következő generáció várható allél gyakoriságai az adott lokuszon:

$$E[f_{A1}(t+1)] = (1 - m_{2 \rightarrow 1})f_{A1}(t) + m_{2 \rightarrow 1}f_{A2}$$

$$E[f_{A2}(t+1)] = (1 - m_{1 \rightarrow 2})f_{A2}(t) + m_{1 \rightarrow 2}f_{A1}$$

- Speciális eset – a kontinens-sziget modell

- 1. sziget populáció, $f_{A1}(t)$ – WF populáció és csak migráció a kontinensről ($m_{2 \rightarrow 1}$)
- 2. kontinens, $m_{1 \rightarrow 2} = 0$ és f_{A2} konstans
- egyensúlyi allél gyakoriság a szigeten: $f_{A1} = f_{A2}$
– szigeten a kontinenssel azonos allél gyakoriság
- a változás sebessége kizárólag m függvénye



- Általában jelentős migrációval hasonló allélgyakoriság a populációkban
– de természetes populációkban mutáció, drift, szelekció is ...

Ha $f_{A1}(t)$ és $f_{A2}(t)$ az A allél gyakorisága a t generációban a két populációban, a következő generáció várható allél gyakoriságai a modellben $E[f_{A1}(t+1)]$ és $E[f_{A2}(t+1)]$ (a bevándorló egyedek hozzájárulása m , a rezidenseké $(1 - m)$ arányú generációnként).

Ennek speciális esete a **kontinens-sziget modell**, ahol a migráció egyirányú ($m_{1 \rightarrow 2} = 0$ és $f_{A2}(t) = f_{A2}$ konstans). Az A allél gyakoriságának változása a szigeten kifejezhető az idő függvényében:

$$f_{A1}(t) = f_{A2} + (f_{A1}(0) - f_{A2})(1 - m_{2 \rightarrow 1})^t$$

Az egyensúlyi allél gyakoriság a szigeten (vagyis ha $t \rightarrow \infty$, $E[f_{A1}(t+1)] = f_{A1}(t) = f_{A1}$): $f_{A1} = f_{A2}$ – nem meglepő módon a szigeten egyensúlyban az allél gyakoriság a kontinensével azonos. Továbbá a konvergencia sebessége kizárólag a migrációs ráta függvénye, minél nagyobb $m_{2 \rightarrow 1}$ annál gyorsabb a változás. Az ábra az allél gyakoriság változást szemlélteti az idő függvényében két különböző kezdeti allél gyakorisággal ($t = 0$, $f_{A1}(0) = 0,1$ és $0,9$) két különböző migrációs rátával ($m_{2 \rightarrow 1} = 0,1$ és $m_{2 \rightarrow 1} = 0,2$), illetve $f_{A2} = 0,4$.

Természetes populációkban számolnunk kell egyéb folyamatokkal is, amelyek hatással lehetnek az allél gyakoriságára a kérdéses lokuszon.

Koaleszcencia migrációval

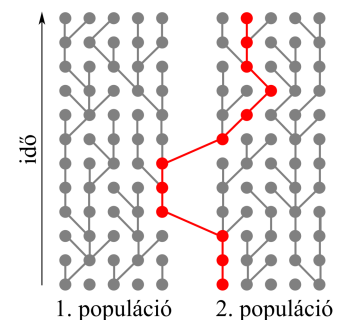
- Koaleszcencia kiterjesztése – ős másik populációból is származhat
 - két populáció, WF modell + migráció
 - tfh. szimmetrikus migráció, $m_{1 \rightarrow 2} = m_{2 \rightarrow 1} = m$, és azonos populáció méret ($N_1 = N_2 = N$)
 - 1. populáció egy egyede a 2. populáció egy előző generációbeli egyedétől származik m valószínűséggel. . .
 - az 1. populációbeli minta őse az 1. populációban r generáción keresztül: $(1 - m)^r$

- Időt $2N$ generációban mérve, $r = 2Nt$, $M = 2Nm$

- ősök ua. populációban t időn keresztül ($N \rightarrow \infty$):

$$(1 - M/(2N))^{2Nt} \rightarrow e^{-Mt}$$

- M a populáció mérettel átskálázott migrációs ráta – a generációnkénti bevándorló egyedszám



A **koaleszcencia folyamatot** migrációval kiegészítve feltételezzük, hogy az ős egy másik populációból is származhat. Az egyszerűség kedvéért feltételezzük, hogy a két WF populáció mérete azonos és a migráció szimmetrikus. Bármely egyed (**minta**) előző generációbeli őse a másik populációból származik m , azonos populációból származik $1 - m$ valószínűséggel. A már megismert módon származtatva, az időt $2N$ generációban mérve és $N \rightarrow \infty$ feltétellel az első migrációig eltelt idő eloszlása $1/M$ várható értékű exponenciális, ahol $M = 2Nm$ ($m \rightarrow 0$ tetszőleges M esetén, vagyis a modell feltételei: nagy N és kis m). Mivel a populáció mérete $2N$ és egy egyed m valószínűséggel bevándorló, $M = 2Nm$ a **generációnkénti bevándorló egyedszám**. A gyakorlatban M értékét becsüljük, m becsléséhez $2N$ ismerete is szükséges.

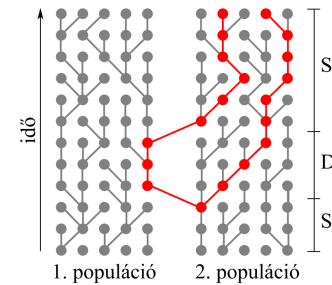
Tételezzük fel, hogy mintánk két azonos populációból származó egyedből áll ($n = 2$). A kiterjesztett koaleszcencia folyamatot koaleszcencia és migráció események sorozata alkotja, két különböző múltbéli eseménnyel kell számolnunk: a két leszármazási sor koaleszcenciájával (rátája 1 az időt $2N$ generációban mérve) vagy migrációval (rátája $2M$, mivel 2 leszármazási sor van és szimmetrikus a migráció a két azonos méretű populáció között).

Koaleszcencia migrációval

- A várható koaleszcencia idők két populációra, $n = 2$ eset

- $N_1 = N_2$, szimmetrikus migráció
- leszármazási sor egy (S) vagy két (D) populációban

$$E_S[t] = 2 \text{ és } E_D[t] = \frac{1}{2M} + 2$$



- Koaleszcencia idők és M
 - kis M esetén nagy $E_D[t]$ – hosszú idő míg egy populációba kerülnek
 - nagy M esetén $E_D[t] \approx E_S[t]$ – mintha egy populáció lenne
 - $E_S[t]$ nem függ M -től – értéke $4N$ természetes időskálán
- Kiterjesztése több populációra – a sziget modell
 - d darab $2N$ méretű populáció (dém)
 - szimmetrikus migráció, rátája $m (= M/(2N))$ bármely párra

$$E_S[t] = d \text{ és } E_D[t] = \frac{1}{2M} + d$$

A modell feltételei alapján nem következhet be koaleszcencia és migráció egyazon generációban. A modell alapján a várható koaleszcencia idő ($E[t]$) becsülhető, de szét kell választanunk két esetet: a kételemű minta származhat egy (S konfiguráció, „same”) vagy két (D konfiguráció, „different”) populációból. Hiszen a koaleszcencia feltétele az, hogy a két leszármazási sor egy populációban legyen. Ha a minta két különböző populációból származik ($E_D[t]$) és M kicsi (kevés migráns egyed), nagy a várható koaleszcencia idő, mivel nagy a várakozási idő az egy populációba kerülésig. Nagy M esetén $E_D[t] \approx E_S[t]$, vagyis a koaleszcencia idő hasonló ahhoz, amit egy populációból származó mintákra várunk – mintha egy populációnk lenne. Továbbá ha a minta egy populációból származik ($E_S[t]$), a várható koaleszcencia idő a migrációs rátától független, természetes időskálán (generációk számában) értéke $2 * 2N = 4N$.

A koaleszcencia modell kiterjeszthető további populáció szerkezet modellekre. A **sziget modell** például feltételezi, hogy a populációk („sziget”, dém, al-populáció) között tetszőleges irányú génáramlás lehetséges. A továbbiakban feltételezzük, hogy d azonos méretű ($2N$) populációnk van és bármely kettő között a génáramlás szimmetrikus, a migrációs ráta m . A koaleszcencia idő várható értéke a két populációs esethez hasonlóan értelmezhető.

Ellenőrző kérdések

- ❶ Mit értünk populáció szerkezet alatt?
- ❷ Mit értünk Wahlund hatás alatt?
- ❸ Értelmezze az F_{ST} jelentését!
- ❹ Mekkora az F_{ST} érték, ha a populációkban az allél gyakoriságok azonosak?
- ❺ Mi a kontinens-sziget modell, milyen feltételekkel definiálhatjuk?
- ❻ Mi a sziget modell, milyen feltételekkel definiálhatjuk?
- ❼ Mit értünk migrációs ráta alatt?
- ❽ Milyen feltételei vannak a migrációval kiterjesztett standard koaleszcencia modell alkalmazhatóságának?
- ❾ Miért utalhat természetes szelekcióra egy lokuszon a két populáció közötti magas F_{ST} érték?
- ❿ Miért különbözhet két neutrális lokuszra a két populáció közötti F_{ST} érték?



EFOP-3.4.3-16-2016-00014



Jelen tananyag a Szegedi Tudományegyetemen
készült az Európai Unió támogatásával. Projekt
azonosító: EFOP-3.4.3-16-2016-00014

Szegedi Tudományegyetem
Cím: 6720 Szeged, Dugonics tér 13.
www.u-szeged.hu
www.szechenyi2020.hu



MAGYARORSZÁG
KORMÁNYA

Európai Unió
Európai Szociális
Alap



BEFEKTETÉS A JÖVŐBE