

Pràctica amb ordinador 2. Selecció i deriva

Principals Transicions Evolutives

2023-04-19

Preparació de l'ordinador

En aquesta pràctica utilitzarem el paquet `learnPopGen` de R (Revell 2019). Per tant, ens hem d'assegurar que està instal·lat (`install.packages('learnPopGen')`) i l'hem de carregar en la sessió mitjançant el comandament:

```
library('learnPopGen')
```

Trobaràs útil escriure i guardar en un arxiu de text (un *script* de R) els comandaments que necessites executar. En RStudio, pots executar línies concretes del teu *script* o l'*script* sencer amb els botons **Run** i **Source**, respectivament.

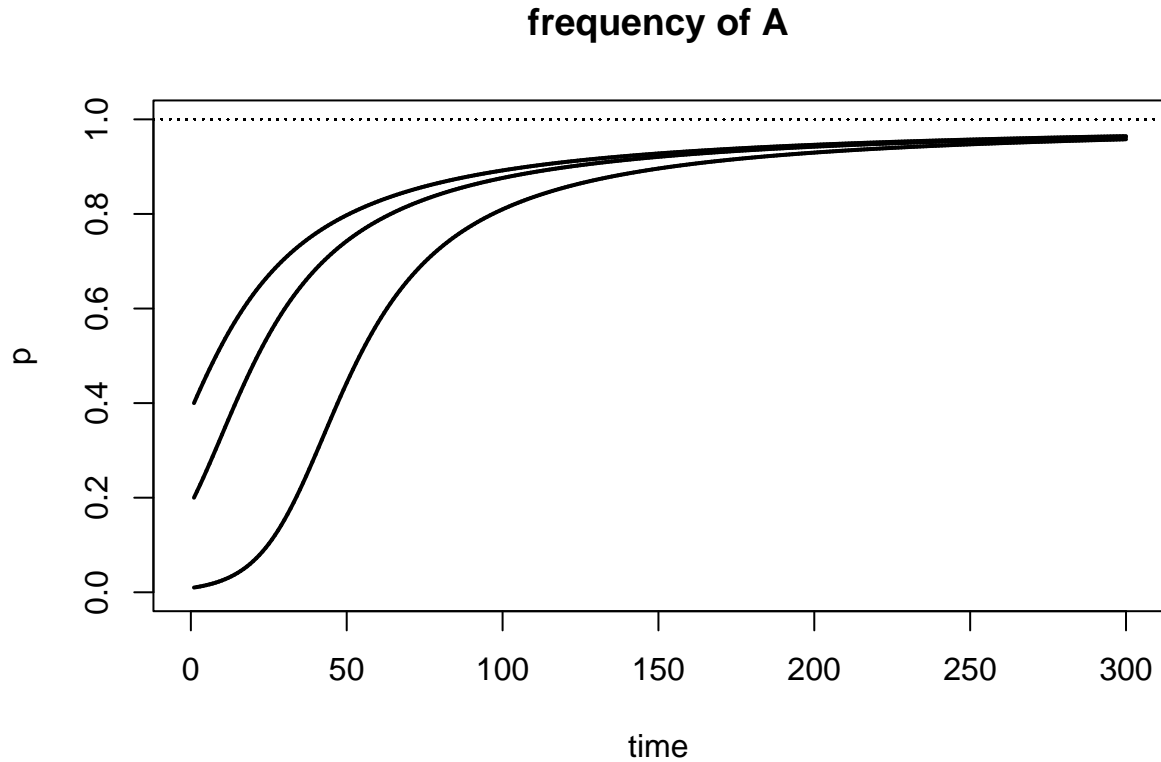
Selecció natural

La funció `selection()` permet representar el canvi de freqüència al·lèlica en un locus amb dos al·lels sota selecció natural. El paràmetre `p0` determina la freqüència inicial de l'al·lel **A**. El paràmetre `w` és el vector de les tres eficàcies biològiques, corresponents als genotips **AA**, **Aa** i **aa**. Podem fixar també el nombre de generacions durant les quals ha d'evolucionar la població (`time = 400`, per exemple) i si volem que al gràfic es mostri la freqüència d'equilibri (`equil = TRUE`, recomanable). A més, podem afegir una segona simulació a un gràfic anterior si utilitzem el paràmetre `add = TRUE`. Per exemple:

```
selection(p0 = 0.01, w = c(1, 1, 0.9), time = 300, pause = 0, show = 'p', equil = TRUE)
selection(p0 = 0.20, w = c(1, 1, 0.9), time = 300, pause = 0, show = 'p', add = TRUE)
selection(p0 = 0.40, w = c(1, 1, 0.9), time = 300, pause = 0, show = 'p', add = TRUE)
```

Taula 1: Tres models de selecció a simular.

Model	Eficàcia d'AA	Eficàcia d'Aa	Eficàcia d'aa
Selecció a favor del dominant	1.0	1.0	0.9
Selecció a favor del recessiu	1.0	0.8	0.8
Selecció a favor de l'heterozigot	0.8	1.0	0.9



Comprova que pots utilitzar la funció `selection()` i aleshores observa què passa en cada una de les situacions descrites a la taula 1. Observa que les distribucions d'eficàcia biològica proposades per a cada model fan que l'al·lel **A** (la freqüència del qual és la representada) siga dominant en un cas, recessiu en un altre, i parcialment recessiu en el tercer. Assegura't d'utilitzar diferents valors de freqüència inicial.

Contesta les preguntes següents:

1. Com influeix la manera d'herència (recessiva o dominant) sobre la fixació d'un al·lel afavorit per la selecció natural?
2. Pots explicar el patró observat?
3. Quina peculiaritat caracteritza la selecció a favor de l'heterozigot?
4. Els equilibris absoluts, són estables o inestables?

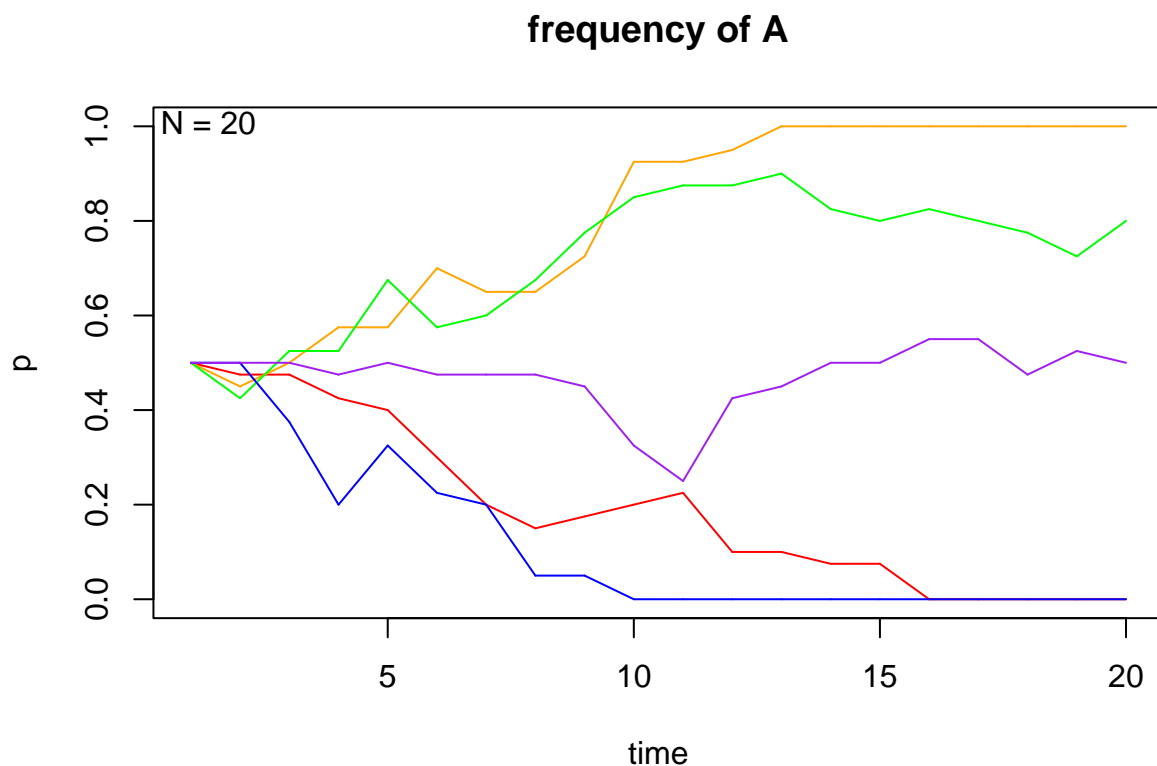
Deriva genètica

Per estudiar l'efecte de la deriva genètica sobre una població de mida finita, utilitzarem la funció `genetic.drift()`, que accepta els paràmetres següents:

- `p0`: Freqüència inicial de l'al·lel **A** (0.5 per defecte).
- `Ne`: Mida poblacional efectiva (20 per defecte).
- `nrep`: Nombre de rèpliques a simular (10 per defecte).
- `time`: Nombre de generacions a simular.

- **show:** Opcions de representació gràfica:
 - 'p': la freqüència de l'al·lel **A** respecte del temps (per defecte).
 - 'genotypes': histograma dinàmic de freqüències genotípiques.
 - 'fixed': proporció de poblacions on s'ha fixat cada al·lel.
 - 'heterozygosity': l'heterozigositat observada i la teòrica respecte del temps.
- **pause:** Temps (segons) de pausa entre generacions, per suavitzar les animacions. Per defecte, 0.1.
- **colors:** Vector dels colors amb què es vol representar cada rèplica.

```
sim01 <- genetic.drift(p0 = 0.5,
                      Ne = 20,
                      nrep = 5,
                      time = 20,
                      pause = 0,
                      colors = c('red', 'orange', 'green', 'blue', 'purple'),
                      show = 'p')
```



Familiaritza't amb els diferents paràmetres i gràfics. Després realitza i guarda en variables diferents (`sim01`, `sim02`, ... `sim12`) les simulacions amb les 12 combinacions de paràmetres indicades en la taula 2.

A continuació, utilitza la funció `plot()` per observar com varia amb el temps la proporció de rèpliques de cada simulació en què es fixa l'al·lel **A**, o l'**a**, o cap dels dos. Per exemple: `plot(sim01, show = 'p')` o `plot(sim01, show = 'fixed')`.

Intenta contestar les preguntes següents:

1. Si en una simulació fas 100 rèpliques (`nrep = 100`), cada rèplica representa una població. És possible predir els canvis de freqüència al·lèlica en una població concreta?
2. Hauràs observat que si es deixa passar prou de temps en qualsevol població només queda un dels dos al·lells. Quin factor influeix més en el *temps* que tarda en desaparèixer el polimorfisme?
3. Quin efecte té la freqüència al·lèlica inicial, p_0 ?
4. Sabent que en última instància es fixarà un al·lel o l'altre, sabries dir quina és la probabilitat de que acabe fixant-se cada al·lel?

Taula 2: Combinacions de paràmetres per la simulació de la deriva genètica

p0	Ne	nrep	time
0.1	10	100	300
0.2	10	100	300
0.4	10	100	300
0.5	10	100	300
0.1	50	100	300
0.2	50	100	300
0.4	50	100	300
0.5	50	100	300
0.1	100	100	300
0.2	100	100	300
0.4	100	100	300
0.5	100	100	300

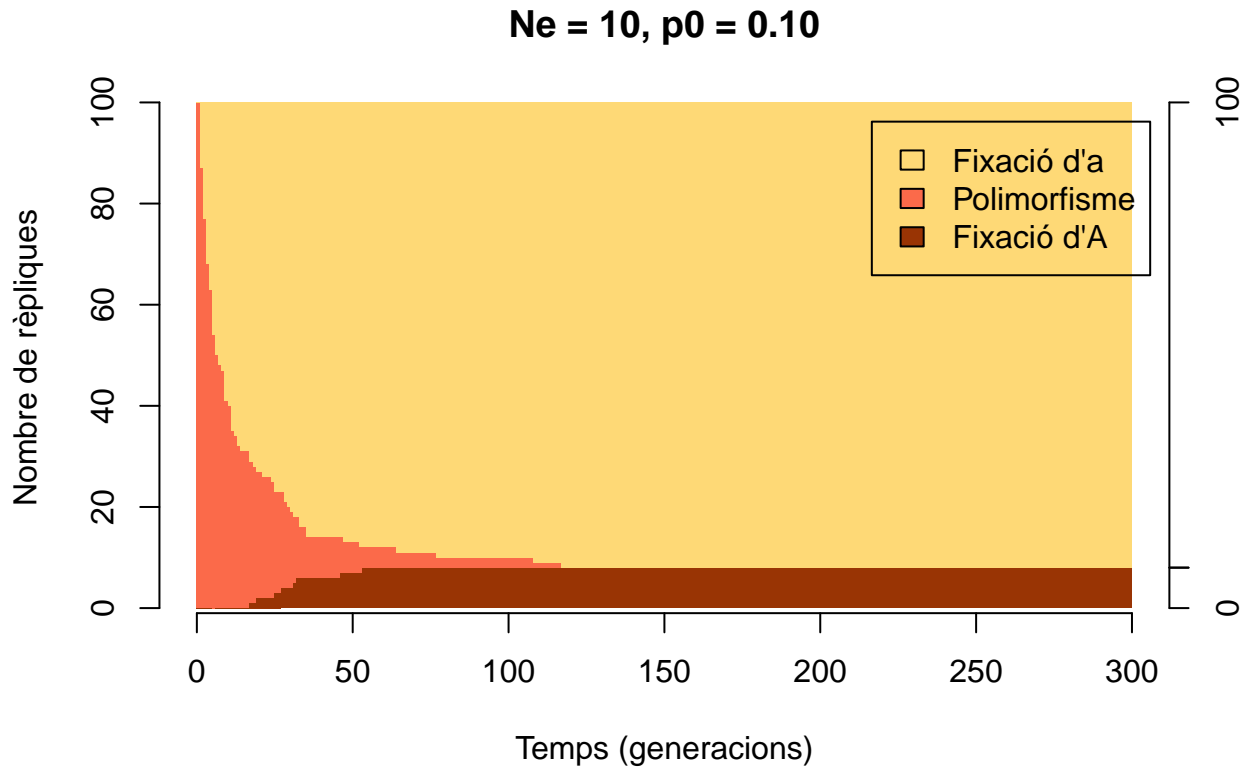
Els comandaments següents defineixen una nova funció que t'ajudarà a visualitzar les proporcions de rèpliques de cada simulació en què es fixa un al·lel o l'altre al llarg del temps. Per poder utilitzar la funció `fixacions()`, copia les línies següents, pega-les en el terminal i prem **enter**.

```
fixacions <- function(x){
  if (class(x) == 'drift.selection') {
    x <- matrix(unlist(x), ncol = length(x))
  } else {
    stopifnot(class(x) == 'genetic.drift')
  }
  fixades <- cbind(
    A_fixat = apply(x, 1, function(z) sum(z == 1)),
    polimorfic = apply(x, 1, function(z) sum(z * (1 - z) > 0)),
    a_fixat = apply(x, 1, function(z) sum(z == 0))
  )
  fixades <- t(fixades)
  barplot(fixades,
    border = NA,
    space = 0,
    xlab = 'Temps (generacions)',
    ylab = 'Nombre de rèpliques',
    col = c('#993404', '#FB6A4A', '#FED976'),
    legend.text = c("Fixació d'A", "Polimorfisme", "Fixació d'a"),
    main = sprintf('Ne = %.0f, p0 = %.2f', attr(x, 'Ne'), attr(x, 'p0'))
  )
  axis(1, at = seq(from = 0, to = dim(fixades)[2], by = dim(x)[1] %/% 6))
  axis(4, at = c(0, fixades[1, dim(x)[1]], sum(fixades[1:2, dim(x)[1]]), dim(x)[2]))
}
```

A partir d'aleshores, podràs aplicar la funció a cada una de les simulacions que has realitzat. Per exemple:

```
sim01 <- genetic.drift(p0 = 0.1, Ne = 10, nrep = 100, time = 300, pause = 0)
```

```
fixacions(sim01)
```



Potser aquestes gràfiques ajuden a confirmar les teues respostes a les preguntes anteriors.

Selecció i deriva

Quan un locus polimòrfic afecta l'eficàcia biològica en una població finita, els canvis en les freqüències al·lèliques d'eixe locus estan determinades conjuntament per la selecció natural i per la deriva. Per simular poblacions en què actuen alhora la selecció natural i la deriva utilitzarem la funció `drift.selection()`, que admet els paràmetres següents:

- **p0**: freqüència inicial de l'al·lel **A**.
- **Ne**: Mida poblacional efectiva.
- **w**: vector d'eficàcies biològiques dels genotips **AA**, **Aa** i **aa**.
- **nrep**: Nombre de rèpliques a simular.
- **ngen**: Nombre de generacions total.
- **colors**: Colors per a les rèpliques.

Per estudiar l'efecte de la freqüència inicial (**p0**), la mida poblacional (**Ne**), i el mode de selecció natural (**w**) en els canvis de freqüències al·lèliques, cal realitzar (i guardar en variables diferents) les 36 simulacions descrites a la taula 3. Per facilitar l'anàlisi, la taula 3 inclou el nom recomanat per la variable on es guardarà cada simulació, on els prefixos *Dom*, *Rec* i *Het* indiquen si la selecció és a favor del *Dominant*, del *Recessiu*, o de l'*Heterozigot*. Suggerim que totes les simulacions tinguin almenys 300 generacions i no més 30 rèpliques. Com que són moltes simulacions, trobaràs útil escriure primer els comandaments en un **script de R**, on és fàcil copiar, pegar i editar els comandaments tots alhora.

Amb la funció `plot()` podràs recuperar el gràfic de les freqüències al·lèliques de cada simulació. El paràmetre `lwd = 1` en esta funció farà les línies més fines. També pots utilitzar la funció `fixacions()`, definida anteriorment, amb aquestes simulacions per observar les proporcions de rèpliques (poblacions) en què s'ha fixat un al·lel o l'altre, o cap dels dos. Després de realitzar i observar els resultats d'estes simulacions, compara-les amb les simulacions anteriors, on només actuava la deriva o només la selecció natural, i respon les preguntes següents:

Taula 3: Combinacions de paràmetres per simular poblacions amb selecció natural i deriva genètica.

Nom	w	Ne	p0	ngen	nrep
DomNe10p0.1	c(1.0, 1.0, 0.9)	10	0.1	300	30
DomNe10p0.2	c(1.0, 1.0, 0.9)	10	0.2	300	30
DomNe10p0.4	c(1.0, 1.0, 0.9)	10	0.4	300	30
DomNe10p0.5	c(1.0, 1.0, 0.9)	10	0.5	300	30
DomNe50p0.1	c(1.0, 1.0, 0.9)	50	0.1	300	30
DomNe50p0.2	c(1.0, 1.0, 0.9)	50	0.2	300	30
DomNe50p0.4	c(1.0, 1.0, 0.9)	50	0.4	300	30
DomNe50p0.5	c(1.0, 1.0, 0.9)	50	0.5	300	30
DomNe100p0.1	c(1.0, 1.0, 0.9)	100	0.1	300	30
DomNe100p0.2	c(1.0, 1.0, 0.9)	100	0.2	300	30
DomNe100p0.4	c(1.0, 1.0, 0.9)	100	0.4	300	30
DomNe100p0.5	c(1.0, 1.0, 0.9)	100	0.5	300	30
RecNe10p0.1	c(1.0, 0.8, 0.8)	10	0.1	300	30
RecNe10p0.2	c(1.0, 0.8, 0.8)	10	0.2	300	30
RecNe10p0.4	c(1.0, 0.8, 0.8)	10	0.4	300	30
RecNe10p0.5	c(1.0, 0.8, 0.8)	10	0.5	300	30
RecNe50p0.1	c(1.0, 0.8, 0.8)	50	0.1	300	30
RecNe50p0.2	c(1.0, 0.8, 0.8)	50	0.2	300	30
RecNe50p0.4	c(1.0, 0.8, 0.8)	50	0.4	300	30
RecNe50p0.5	c(1.0, 0.8, 0.8)	50	0.5	300	30
RecNe100p0.1	c(1.0, 0.8, 0.8)	100	0.1	300	30
RecNe100p0.2	c(1.0, 0.8, 0.8)	100	0.2	300	30
RecNe100p0.4	c(1.0, 0.8, 0.8)	100	0.4	300	30
RecNe100p0.5	c(1.0, 0.8, 0.8)	100	0.5	300	30
HetNe10p0.1	c(0.8, 1.0, 0.9)	10	0.1	300	30
HetNe10p0.2	c(0.8, 1.0, 0.9)	10	0.2	300	30
HetNe10p0.4	c(0.8, 1.0, 0.9)	10	0.4	300	30
HetNe10p0.5	c(0.8, 1.0, 0.9)	10	0.5	300	30
HetNe50p0.1	c(0.8, 1.0, 0.9)	50	0.1	300	30
HetNe50p0.2	c(0.8, 1.0, 0.9)	50	0.2	300	30
HetNe50p0.4	c(0.8, 1.0, 0.9)	50	0.4	300	30
HetNe50p0.5	c(0.8, 1.0, 0.9)	50	0.5	300	30
HetNe100p0.1	c(0.8, 1.0, 0.9)	100	0.1	300	30
HetNe100p0.2	c(0.8, 1.0, 0.9)	100	0.2	300	30
HetNe100p0.4	c(0.8, 1.0, 0.9)	100	0.4	300	30
HetNe100p0.5	c(0.8, 1.0, 0.9)	100	0.5	300	30

1. Quan un al·lel acaba fixant-se, pots determinar si ha sigut per causa de la deriva o per causa de la selecció natural?
2. Quan apareix un al·lel per mutació, la seua freqüència és $1/2N$, si la població és diploide. Si aquest al·lel és beneficiós, en quina població creus que tindrà major probabilitat d'arribar a fixar-se, en una gran o en una petita? Per què?
3. La deriva dificulta la fixació de l'al·lel beneficiós. Ho fa de la mateixa manera quan l'al·lel és recessiu o dominant?
4. La selecció a favor de l'heterozigot tendeix a mantenir el polimorfisme. En les poblacions que has simulat, creus que el polimorfisme pot mantenir-se per sempre? Per què? Com ho comprovaries?

Bibliography

Revell, Liam J. 2019. "learnPopGen: An R package for population genetic simulation and numerical analysis." *Ecology and Evolution* 9 (14): 7896–7902. <https://doi.org/https://doi.org/10.1002/ece3.5412>.