Facultat de
Biociències
UAB

Olimpíada de Biologia
de Catalunya

# XVI Olimpíada de Biologia de Catalunya

Prova pràctica

12 de febrer de 2026

## PROVA PRÀCTICA 2: ADAPTACIÓ LOCAL, CANVI CLIMÀTIC I GENÒMICA DELS TEOSINTS

Autora: Dra. Sònia Casillas. Institut de Biotecnologia i Biomedicina “Vicent Villar Palasí” i Departament de Genètica i de Microbiologia de la Universitat Autònoma de Barcelona. 08193 Bellaterra, Barcelona.

**Enllaç a la prova:** https://t.ly/LDRiQ

---

NOM I COGNOMS: ____________________________________________

DNI: __________________________

Marca la teva resposta amb una creu: ☒

| | A | B | C | D |
|---|---|---|---|---|
| **1** | ☐ | ☐ | ☐ | ☒ |
| **2** | ☐ | ☐ | ☒ | ☐ |
| **3** | ☐ | ☒ | ☐ | ☐ |
| **4** | ☐ | ☐ | ☒ | ☐ |
| **5** | ☐ | ☐ | ☒ | ☐ |
| **6** | ☐ | ☒ | ☐ | ☐ |
| **7** | ☐ | ☒ | ☐ | ☐ |
| **8** | ☐ | ☒ | ☐ | ☐ |
| **9** | ☐ | ☐ | ☒ | ☐ |
| **10** | ☐ | ☒ | ☐ | ☐ |

### ACLARACIONS PRÈVIES

- Aquesta prova pràctica s’avalua mitjançant 10 preguntes
- Totes les preguntes tenen el mateix pes a la nota final
- Totes les preguntes tenen quatre opcions de resposta, de les quals només una és correcta
- Els errors penalitzen (tres errors resten un encert)

### Instruccions generals

Aquesta prova pràctica es realitza mitjançant un *Jupyter Notebook*, que combina blocs de text i blocs de codi.

Els **blocs de text** contenen les explicacions teòriques, les instruccions de cada apartat i les preguntes que hauràs de respondre.

Els **blocs de codi** tenen el fons gris i contenen instruccions de programació que permeten analitzar les dades i generar gràfiques i simulacions.

Per executar un bloc de codi, cal:

1. Fer clic **dins** del bloc de codi.
2. Fer clic al **botó** ▶ que apareix al menú superior (o prémer *Shift+Enter*).
3. Pots executar qualsevol bloc de codi tantes vegades com vulguis, fent clic repetidament al botó ▶ (o prement repetidament *Shift+Enter*), per repetir una simulació.

En general, no caldrà modificar el codi; només caldrà executar-lo per obtenir els resultats. Als apartats de simulació, però, les instruccions indicaran explícitament que cal modificar algun paràmetre del codi i tornar a executar el bloc per observar com canvien els resultats.

## Adaptació local, canvi climàtic i genòmica dels teosints

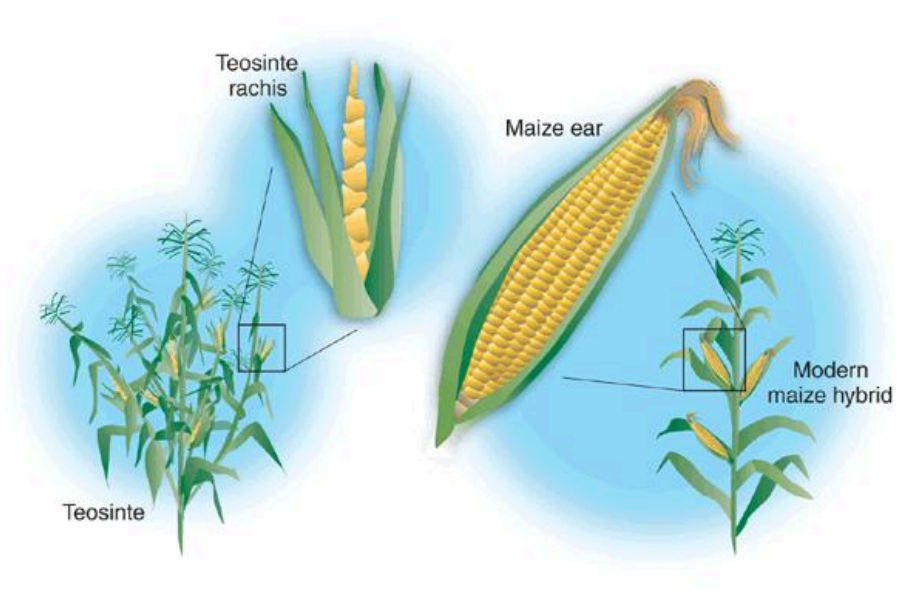
El **canvi climàtic** està transformant ràpidament les condicions ambientals del planeta. L'augment de les temperatures, els canvis en la disponibilitat d'aigua i les alteracions en l'estacionalitat afecten totes les espècies, que es veuen obligades a respondre d'alguna manera. Aquestes respostes poden incloure:

1. **Migració cap a zones més favorables**, desplaçant-se geogràficament per mantenir-se dins del seu rang òptim de condicions ambientals.
2. **Plasticitat fenotípica**, és a dir, la capacitat d'un mateix genotip de produir fenotips diferents en funció de l'ambient, sense que hi hagi canvis genètics. Aquesta resposta pot ser ràpida i permetre la supervivència a curt termini davant canvis ambientals sobtats.
3. **Adaptació genètica**, és a dir, la selecció natural d'aquelles variants genètiques que confereixen una major aptitud en les noves condicions ambientals, augmentant la seva freqüència al llarg de les generacions.

L'estudi de milers de **variants genètiques** distribuïdes al llarg del genoma (SNPs, *Single Nucleotide Polymorphisms*) ens permet detectar patrons d'adaptació, identificar poblacions vulnerables i predir com els organismes podrien respondre als escenaris climàtics futurs.

## El model: els teosints, ancestres silvestres del blat de moro

Els teosints són els ancestres directes del blat de moro modern i, per tant, constitueixen un model excel·lent per estudiar com les poblacions vegetals responen a gradients ambientals i com poden adaptar-se al canvi climàtic.



En aquesta prova pràctica analitzarem dades reals de **dues subespècies de teosints**:

1. *Zea mays ssp. parviglumis* – habita zones baixes, més càlides i humides
2. *Zea mays ssp. mexicana* – es troba en zones més altes i fredes

L'estudi de referència del qual parteixen les dades analitza **33.464 SNPs** en **646 individus** de **47 poblacions** de Centre-Amèrica.

## La hipòtesi de l'abundància al centre de la distribució

La **hipòtesi de l'abundància al centre de la distribució** (*central-abundance hypothesis*) proposa que les poblacions d'una espècie no són iguals a tot el seu rang de distribució. Segons aquesta hipòtesi, les característiques genètiques d'una població depenen de com de lluny es troba del centre de la distribució de l'espècie, tant des del punt de vista geogràfic com ambiental.

En general:

1. Les **poblacions centrals** viuen en condicions ambientals més favorables. Això fa que tinguin més individus, una diversitat genètica elevada i una pressió selectiva relativament baixa.
2. Les **poblacions perifèriques** es troben als límits de la distribució de l'espècie, on les condicions ambientals són més extremes. Aquestes poblacions solen tenir menys diversitat genètica, però estan sotmeses a una selecció natural més intensa, que pot afavorir l'aparició d'adaptacions locals.

Aquesta hipòtesi ens ajuda a entendre per què les poblacions que viuen en ambients més difícils poden mostrar empremptes genètiques d'adaptació, especialment rellevants en un context de canvi climàtic.

## Exploració del genoma de *Zea mays* a NCBI Datasets

Abans d'analitzar dades genòmiques i de variants genètiques, és important conèixer algunes característiques bàsiques del genoma de l'espècie estudiada.

En aquest apartat explorarem el genoma de referència del blat de moro (*Zea mays*) utilitzant la plataforma NCBI Datasets: 🔗 https://www.ncbi.nlm.nih.gov/datasets. En concret, treballarem amb el genoma de referència: **Zm-B73-REFERENCE-NAM-5.0** 🔗 https://www.ncbi.nlm.nih.gov/datasets/genome/GCF_902167145.1/.

### Instruccions

1. Accedeix a l'enllaç del genoma de referència.
2. Identifica la informació més rellevant del genoma (mida del genoma i número de gens).
3. Respon les preguntes següents.

### Preguntes

**1. Quina és la mida del genoma de *Zea mays*?**

A. 220 nucleòtids
B. ~2.200 nucleòtids
C. ~2.200.000 nuceòtids
D. ~2.200.000.000 nucleòtids

*Nota: Kb = ×10$^3$ nucleòtids; Mb = ×10$^6$ nucleòtids; Gb = ×10$^9$ nucleòtids.*

**2. Per quants gens codifica el genoma de *Zea mays*?**

A. < 1.000 gens
B. 1.000 - 20.000 gens
C. 20.000 - 50.000 gens
D. > 50.000 gens

## Càrrega i exploració inicial de les dades genòmiques

El fitxer **teosints_poblacions.csv** conté informació genètica, geogràfica i ambiental de 47 poblacions de teosints.

El fitxer està en format CSV (valors separats per punt i coma ;) i cada fila correspon a una població. Inclou:

1. Població
2. Subespècie (`parviglumis` o `mexicana`)
3. Coordenades geogràfiques de longitud (`X`) i latitud (`Y`)
4. Distància al centre geogràfic (`Dist_geogràfica`) i ambiental (`Dist_ambiental`) de l'espècie
5. Diversitat genètica (`Hs`)
6. Grau de consanguinitat (`Fis`)
7. Número de SNPs adaptatius (`Num_SNPs`)

### Instruccions

1. Executa el codi de sota per carregar i visualitzar el fitxer:

In [1]:

```python
import pandas as pd

# Carregar el fitxer CSV (separador ;)
df = pd.read_csv("https://raw.githubusercontent.com/soniacasillas/teosints/refs/heads/main/teosints_poblacions.csv", sep=";")

# Mostrar totes les files del fitxer (per mostrar només les primeres files, escriu `df.head()` enlloc de `df`)
df
```

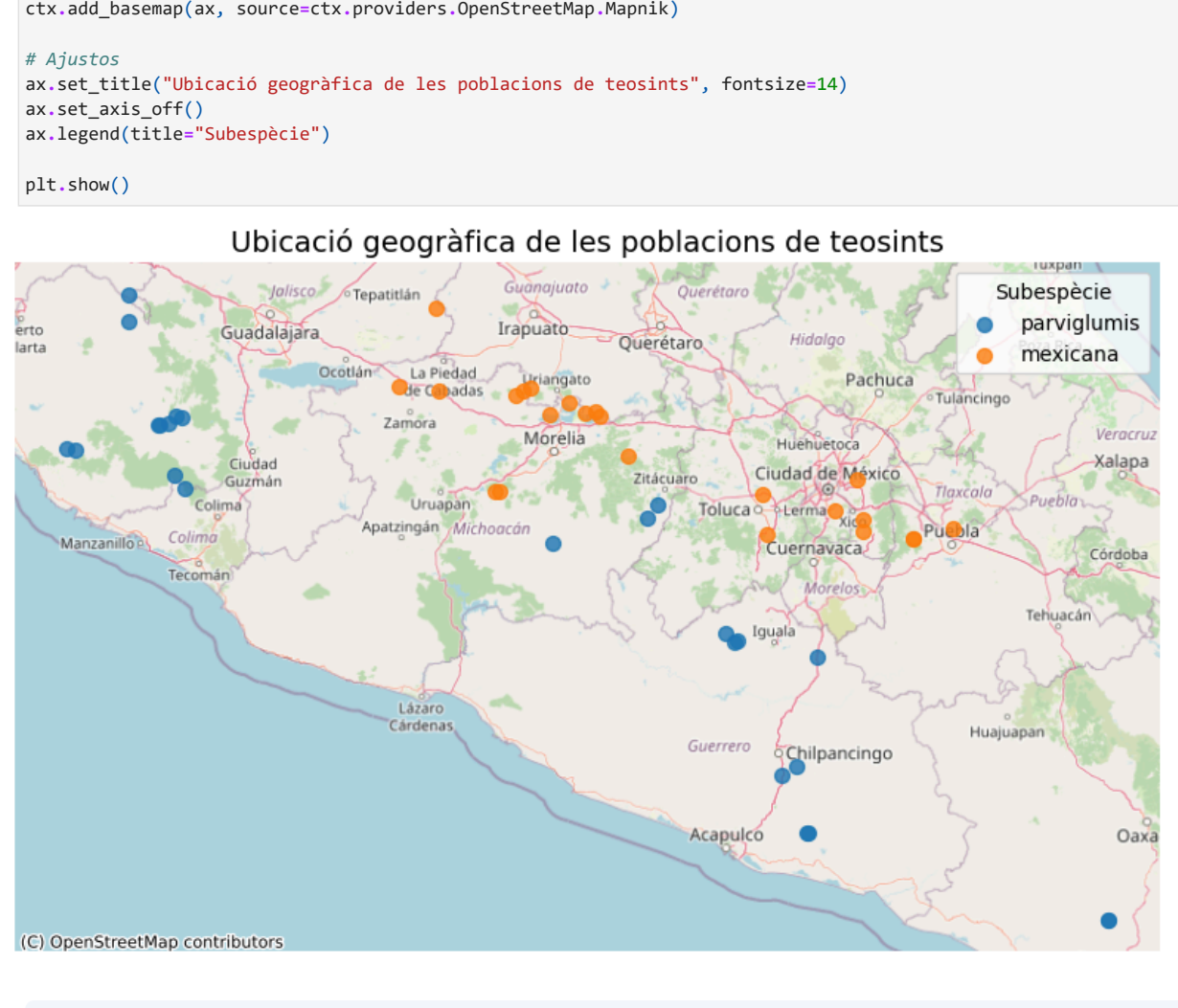
Out[1]:

| | Població | Subespècie | X | Y | Dist_geogràfica | Dist_ambiental | Hs | Fis | Num_SNPs |
|---|---|---|---|---|---|---|---|---|---|
| **0** | VillaSeca_Otzolotepec | mexicana | -99.627056 | 19.407472 | 0.808897 | 4.418383 | 0.24 | 0.01 | 30 |
| **1** | Churintzio_Churintzio | mexicana | -102.068444 | 20.139389 | 1.794106 | 4.080644 | 0.22 | 0.06 | 24 |
| **2** | SMH571_Acambaro | mexicana | -100.884889 | 19.991722 | 0.603063 | 1.883979 | 0.23 | 0.04 | 16 |
| **3** | SMH572_Acambaro | mexicana | -100.849667 | 19.958194 | 0.562040 | 2.474122 | 0.25 | 0.03 | 20 |
| **4** | SMH573_Acambaro | mexicana | -100.959917 | 19.982806 | 0.674893 | 2.879088 | 0.26 | 0.10 | 26 |
| **5** | SMH575_SantaAnaMaya | mexicana | -101.088111 | 20.053194 | 0.814904 | 3.137147 | 0.24 | 0.06 | 29 |
| **6** | SMH576_Yuriria | mexicana | -101.373083 | 20.160722 | 1.117767 | 3.863060 | 0.22 | 0.10 | 21 |
| **7** | SMHMGCH579_Puruandiro | mexicana | -101.434167 | 20.133806 | 1.170218 | 3.366603 | 0.25 | 0.21 | 27 |
| **8** | SMHMGCH580_Huandacareo | mexicana | -101.235417 | 19.964611 | 0.945116 | 2.826503 | 0.26 | 0.03 | 22 |
| **9** | SMHMGCH582_JesusMaria | mexicana | -102.089833 | 20.719556 | 1.988210 | 2.452162 | 0.21 | 0.07 | 14 |
| **10** | Cocotitlan_Cocotitlan | mexicana | -98.866139 | 19.226250 | 1.564877 | 2.897349 | 0.23 | 0.13 | 29 |
| **11** | SanNicolas_SnNicolasBuenosAires | mexicana | -98.185417 | 19.159472 | 2.224434 | 2.567262 | 0.23 | 0.04 | 25 |
| **12** | Tenancingo_Tenancingo | mexicana | -98.863722 | 19.139778 | 1.604141 | 2.883401 | 0.23 | 0.08 | 25 |
| **13** | Calpan_Calpan | mexicana | -98.487444 | 19.083889 | 1.968884 | 4.075388 | 0.21 | 0.03 | 28 |
| **14** | Texcoco_Texcoco | mexicana | -98.914611 | 19.502528 | 1.427359 | 4.778473 | 0.21 | 0.03 | 31 |
| **15** | ElPorvenir | mexicana | -100.640000 | 19.680000 | 0.389560 | 3.734284 | 0.23 | 0.06 | 12 |
| **16** | Ixtlan | mexicana | -102.370000 | 20.170000 | 2.096837 | 5.044943 | 0.21 | 0.11 | 29 |
| **17** | Opopeo | mexicana | -101.610000 | 19.420000 | 1.386665 | 4.667299 | 0.22 | 0.03 | 25 |
| **18** | Puruandiro | mexicana | -101.490000 | 20.110000 | 1.219545 | 3.746131 | 0.26 | 0.05 | 11 |
| **19** | SanPedro | mexicana | -98.490000 | 19.090000 | 1.964126 | 4.075388 | 0.21 | 0.02 | 27 |
| **20** | SantaClara | mexicana | -101.640000 | 19.420000 | 1.415129 | 4.688319 | 0.19 | 0.05 | 23 |
| **21** | TenangodelAire | mexicana | -99.590000 | 19.120000 | 1.024454 | 4.213120 | 0.21 | 0.06 | 23 |
| **22** | Xochimilco | mexicana | -99.080000 | 19.290000 | 1.343878 | 2.633053 | 0.16 | 0.02 | 14 |
| **23** | Teloloapan_Teloloapan | parviglumis | -99.841083 | 18.349833 | 0.928659 | 2.491114 | 0.28 | 0.15 | 12 |
| **24** | Alcholoa_Teloloapan | parviglumis | -99.908417 | 18.410000 | 0.845023 | 1.991818 | 0.29 | 0.06 | 15 |
| **25** | Teconoapan_Teconoapa | parviglumis | -99.285472 | 16.981139 | 2.211668 | 6.758742 | 0.22 | 0.07 | 27 |
| **26** | Chilpancingo_Chilpancingo | parviglumis | -99.477639 | 17.391750 | 1.777902 | 2.654884 | 0.26 | 0.04 | 19 |
| **27** | Mochitlan_Mochitlan | parviglumis | -99.368306 | 17.459917 | 1.809193 | 2.208371 | 0.25 | 0.08 | 19 |
| **28** | PasoMorelos_Huitzuco | parviglumis | -99.210000 | 18.237694 | 1.556001 | 2.795941 | 0.23 | 0.19 | 16 |
| **29** | Guachinango_Guauchinango | parviglumis | -104.407889 | 20.627306 | 4.181349 | 4.605261 | 0.28 | 0.08 | 16 |
| **30** | Telpitita_VillaPurificacion | parviglumis | -104.804889 | 19.715444 | 4.223607 | 4.654015 | 0.14 | 0.03 | 24 |
| **31** | SMH565_Ejutla | parviglumis | -104.176750 | 19.896139 | 3.674561 | 3.329858 | 0.24 | 0.08 | 14 |
| **32** | SMH577_VilladePurificacion | parviglumis | -104.871861 | 19.732472 | 4.292710 | 4.519874 | 0.10 | 0.06 | 24 |
| **33** | SMH578_Toliman | parviglumis | -104.058250 | 19.535361 | 3.455641 | 3.103771 | 0.24 | 0.11 | 14 |
| **34** | SMHMGCH581_Zitacuaro | parviglumis | -100.421444 | 19.327389 | 0.718177 | 2.802745 | 0.24 | 0.06 | 13 |
| **35** | SanCristobalHonduras_SanJeronimoCoatlan | parviglumis | -97.020000 | 16.350000 | 4.361287 | 1.942905 | 0.19 | 0.06 | 22 |
| **36** | Ahuacatitlan | parviglumis | -99.810000 | 18.360000 | 0.954528 | 2.192164 | 0.27 | 0.13 | 20 |
| **37** | AmatlandeCasas | parviglumis | -104.410000 | 20.820000 | 4.276730 | 4.159920 | 0.23 | 0.11 | 19 |
| **38** | CruceroLagunitas | parviglumis | -99.280000 | 16.980000 | 2.216072 | 6.765714 | 0.22 | 0.12 | 26 |
| **39** | EjutlaA | parviglumis | -104.180000 | 19.900000 | 3.678914 | 3.329858 | 0.25 | 0.04 | 15 |
| **40** | EjutlaB | parviglumis | -104.110000 | 19.910000 | 3.616426 | 2.437917 | 0.25 | 0.06 | 11 |
| **41** | ElRodeo | parviglumis | -97.020000 | 16.350000 | 4.361287 | 1.942905 | 0.19 | 0.07 | 16 |
| **42** | ElSauz | parviglumis | -103.980000 | 19.440000 | 3.356810 | 2.248043 | 0.18 | 0.07 | 14 |
| **43** | LaCadena | parviglumis | -101.210000 | 19.060000 | 0.631630 | 4.249338 | 0.26 | 0.10 | 14 |
| **44** | LaMesa | parviglumis | -104.050000 | 19.960000 | 3.577829 | 3.659975 | 0.20 | 0.06 | 17 |
| **45** | LosGuajes | parviglumis | -100.490000 | 19.230000 | 0.601809 | 1.412015 | 0.28 | 0.03 | 8 |
| **46** | SanLorenzo | parviglumis | -104.010000 | 19.950000 | 3.527595 | 3.051800 | 0.14 | 0.02 | 20 |

2. Executa el codi de sota per representar la localització de les 47 poblacions de teosints sobre un mapa:

In [2]:

```python
import pandas as pd
import matplotlib.pyplot as plt
import geopandas as gpd
import contextily as ctx

# Crear GeoDataFrame (X = longitud, Y = latitud)
gdf = gpd.GeoDataFrame(
    df,
    geometry=gpd.points_from_xy(df["X"], df["Y"]),
    crs="EPSG:4326"  # WGS84 (lat/long)
)

# Reprojectar a Web Mercator per al mapa base
gdf = gdf.to_crs(epsg=3857)

# Separar subespècies
gdf_parv = gdf[gdf["Subespècie"] == "parviglumis"]
gdf_mex = gdf[gdf["Subespècie"] == "mexicana"]

# Crear figura
fig, ax = plt.subplots(figsize=(10, 8))

# Dibuixar punts
gdf_parv.plot(ax=ax, color="tab:blue", markersize=50, label="parviglumis", alpha=0.8)
gdf_mex.plot(ax=ax, color="tab:orange", markersize=50, label="mexicana", alpha=0.8)

# Afegir mapa cartogràfic de fons
ctx.add_basemap(ax, source=ctx.providers.OpenStreetMap.Mapnik)

# Ajustos
ax.set_title("Ubicació geogràfica de les poblacions de teosints", fontsize=14)
ax.set_axis_off()
ax.legend(title="Subespècie")

plt.show()
```



### Preguntes

**3. A partir del mapa de la distribució geogràfica de les poblacions de teosints, quina de les afirmacions següents és correcta?**

A. Les poblacions de *Zea mays ssp. parviglumis* es concentren a l'interior del territori i presenten una distribució limitada en comparació amb les de *mexicana*.
B. Les poblacions de *Zea mays ssp. parviglumis* es distribueixen principalment a zones costaneres i ocupen una àrea geogràfica més àmplia que les de *mexicana*.
C. Les poblacions de *Zea mays ssp. parviglumis* es concentren al nord del territori i les de *mexicana* es concentren al sud.
D. Ambdues subespècies presenten una distribució geogràfica similar.

## Diversitat genètica i distància al centre de la distribució

Una de les prediccions principals de la hipòtesi de l'abundància al centre de la distribució és que la diversitat genètica d'una població disminueix a mesura que ens allunyem del centre de la distribució geogràfica i ambiental de l'espècie.

En aquest apartat analitzarem si aquesta predicció es compleix en els teosints, estudiant la relació entre la **diversitat genètica** de cada població (`Hs`), i:

- la seva **distància al centre geogràfic** (`Dist_geogràfica`),
- la seva **distància al centre ambiental** (`Dist_ambiental`).

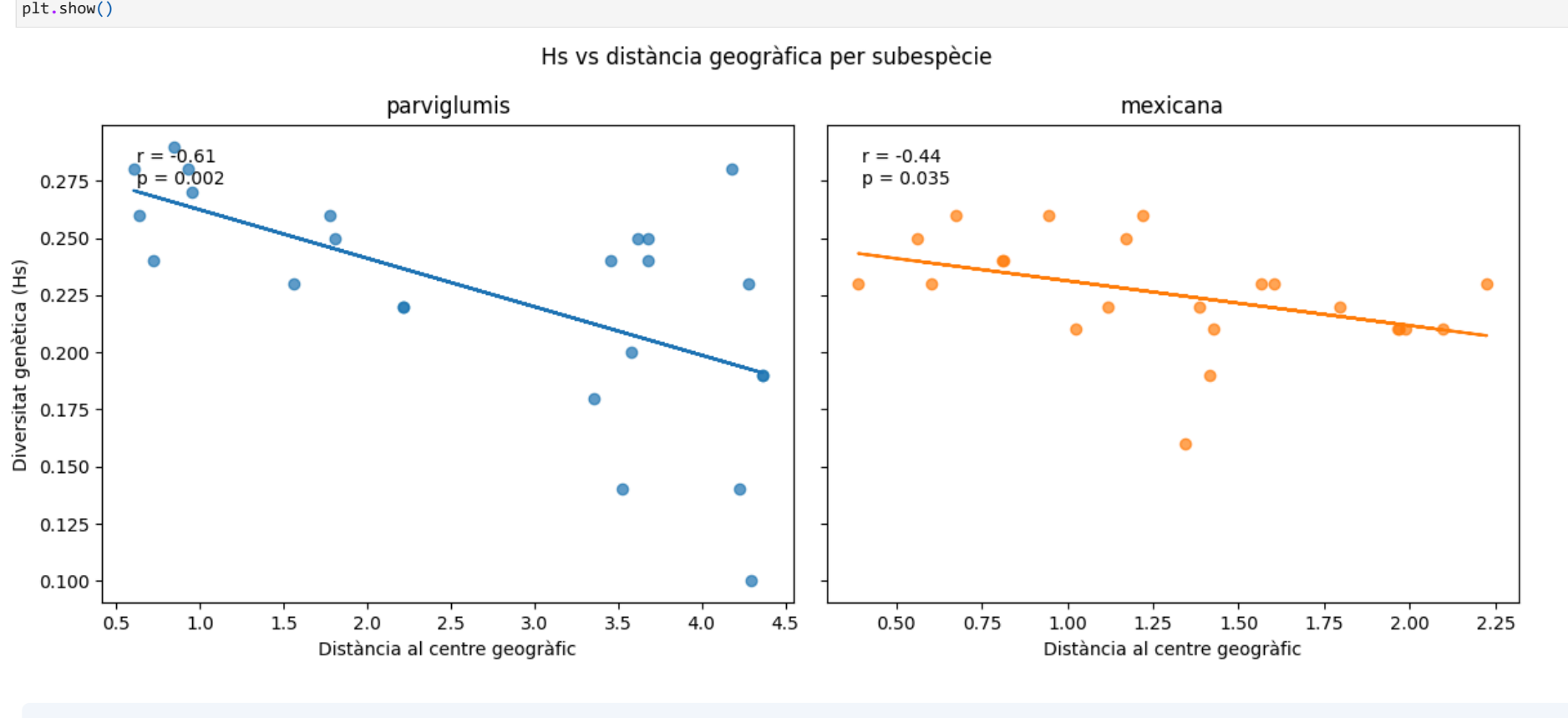
Per fer-ho, representarem gràficament aquestes relacions i calcularem el coeficient de correlació (`r`), que ens permet quantificar la força i el sentit de l'associació entre dues variables, i la significació estadística (`p`) de la relació.

### Instruccions

1. Executa el codi de sota per representar la relació entre la diversitat genètica i la distància geogràfica:

In [3]:

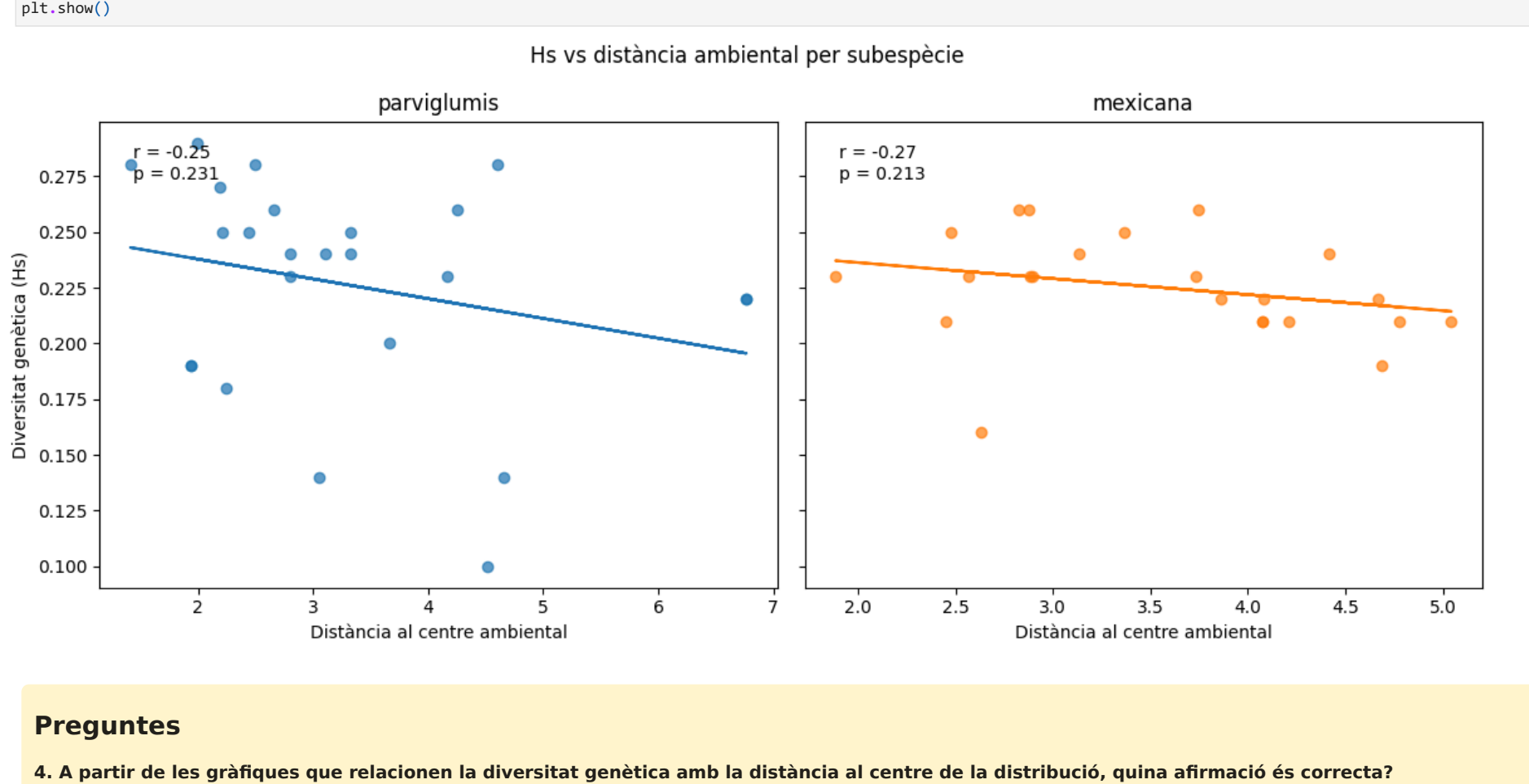
```python
import numpy as np
import matplotlib.pyplot as plt
from scipy.stats import pearsonr

# Separar subespècies i eliminar NaNs
parv = df[df["Subespècie"] == "parviglumis"][["Hs", "Dist_geogràfica"]].dropna()
mex  = df[df["Subespècie"] == "mexicana"][["Hs", "Dist_geogràfica"]].dropna()

fig, axes = plt.subplots(1, 2, figsize=(12, 5), sharey=True)

# --- parviglumis ---
x = parv["Dist_geogràfica"]
y = parv["Hs"]
r, p = pearsonr(x, y)

axes[0].scatter(x, y, color="tab:blue", alpha=0.7)
m, b = np.polyfit(x, y, 1)
axes[0].plot(x, m*x + b, color="tab:blue")
axes[0].set_title("parviglumis")
axes[0].set_xlabel("Distància al centre geogràfic")
axes[0].set_ylabel("Diversitat genètica (Hs)")
axes[0].text(0.05, 0.95, f"r = {r:.2f}\np = {p:.3f}",
             transform=axes[0].transAxes, va="top")

# --- mexicana ---
x = mex["Dist_geogràfica"]
y = mex["Hs"]
r, p = pearsonr(x, y)

axes[1].scatter(x, y, color="tab:orange", alpha=0.7)
m, b = np.polyfit(x, y, 1)
axes[1].plot(x, m*x + b, color="tab:orange")
axes[1].set_title("mexicana")
axes[1].set_xlabel("Distància al centre geogràfic")
axes[1].text(0.05, 0.95, f"r = {r:.2f}\np = {p:.3f}",
             transform=axes[1].transAxes, va="top")

plt.suptitle("Hs vs distància geogràfica per subespècie")
plt.tight_layout()
plt.show()
```



2. Executa el codi de sota per representar la relació entre la diversitat genètica i la distància ambiental:

In [4]:

```python
# Separar subespècies i eliminar NaNs
parv = df[df["Subespècie"] == "parviglumis"][["Hs", "Dist_ambiental"]].dropna()
mex  = df[df["Subespècie"] == "mexicana"][["Hs", "Dist_ambiental"]].dropna()

fig, axes = plt.subplots(1, 2, figsize=(12, 5), sharey=True)

# --- parviglumis ---
x = parv["Dist_ambiental"]
y = parv["Hs"]
r, p = pearsonr(x, y)

axes[0].scatter(x, y, color="tab:blue", alpha=0.7)
m, b = np.polyfit(x, y, 1)
axes[0].plot(x, m*x + b, color="tab:blue")
axes[0].set_title("parviglumis")
axes[0].set_xlabel("Distància al centre ambiental")
axes[0].set_ylabel("Diversitat genètica (Hs)")
axes[0].text(0.05, 0.95, f"r = {r:.2f}\np = {p:.3f}",
             transform=axes[0].transAxes, va="top")

# --- mexicana ---
x = mex["Dist_ambiental"]
y = mex["Hs"]
r, p = pearsonr(x, y)

axes[1].scatter(x, y, color="tab:orange", alpha=0.7)
m, b = np.polyfit(x, y, 1)
axes[1].plot(x, m*x + b, color="tab:orange")
axes[1].set_title("mexicana")
axes[1].set_xlabel("Distància al centre ambiental")
axes[1].text(0.05, 0.95, f"r = {r:.2f}\np = {p:.3f}",
             transform=axes[1].transAxes, va="top")

plt.suptitle("Hs vs distància ambiental per subespècie")
plt.tight_layout()
plt.show()
```



### Preguntes

**4. A partir de les gràfiques que relacionen la diversitat genètica amb la distància al centre de la distribució, quina afirmació és correcta?**

A. En totes dues subespècies, la diversitat genètica augmenta amb la distància al centre geogràfic i ambiental.
B. Només *Zea mays ssp. mexicana* mostra una relació entre la diversitat genètica i la distància al centre de la distribució.
C. En totes dues subespècies, la diversitat genètica disminueix amb la distància al centre geogràfic i ambiental.
D. No s'observa cap relació entre la diversitat genètica i la distància al centre de la distribució en els teosints.

**5. Tenint en compte els resultats obtinguts i la hipòtesi de l'abundància al centre de la distribució (*central-abundance hypothesis*), quina interpretació és correcta?**

A. La diversitat genètica depèn només de factors ambientals i no de la posició geogràfica de les poblacions.
B. La relació més forta observada en *parviglumis* indica que la seva diversitat genètica varia menys amb la distància al centre.
C. *Zea mays ssp. parviglumis* ocupa una àrea geogràfica més extensa, on les poblacions perifèriques estan més aïllades i presenten una menor diversitat genètica.
D. Els resultats indiquen que la hipòtesi de l'abundància al centre de la distribució no és aplicable als teosints.

## Número de SNPs adaptatius i distància al centre de la distribució

La diversitat genètica informa sobre la variabilitat global d'una població, però no totes les variants genètiques tenen el mateix paper en l'adaptació. Algunes variants poden estar associades a la resposta a condicions ambientals concretes i mostrar senyals d'adaptació local.

En aquest apartat analitzarem la relació entre el nombre de SNPs adaptatius identificats a cada població (`Num_SNPs`), i:

- la distància al centre geogràfic (`Dist_geogràfica`),
- la distància al centre ambiental (`Dist_ambiental`).

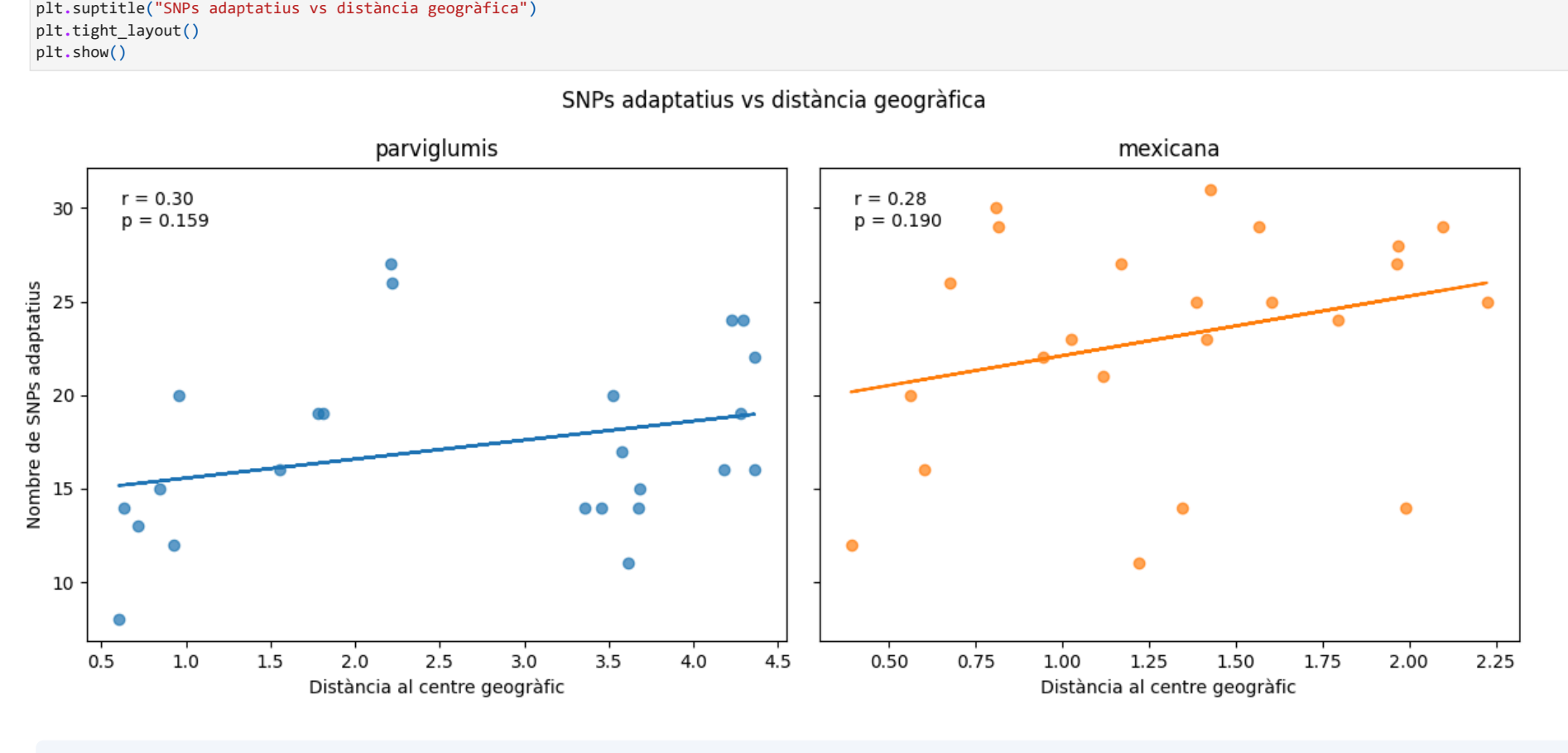
Segons la hipòtesi de l'abundància al centre de la distribució (*central-abundance hypothesis*), s'espera que les poblacions situades als marges del rang de distribució, tot i tenir una menor diversitat genètica, puguin mostrar un enriquiment en SNPs adaptatius com a resultat d'una selecció natural més intensa.

### Instruccions

1. Executa el codi de sota per representar la relació entre el número de SNPs adaptatius i la distància geogràfica:

In [5]:

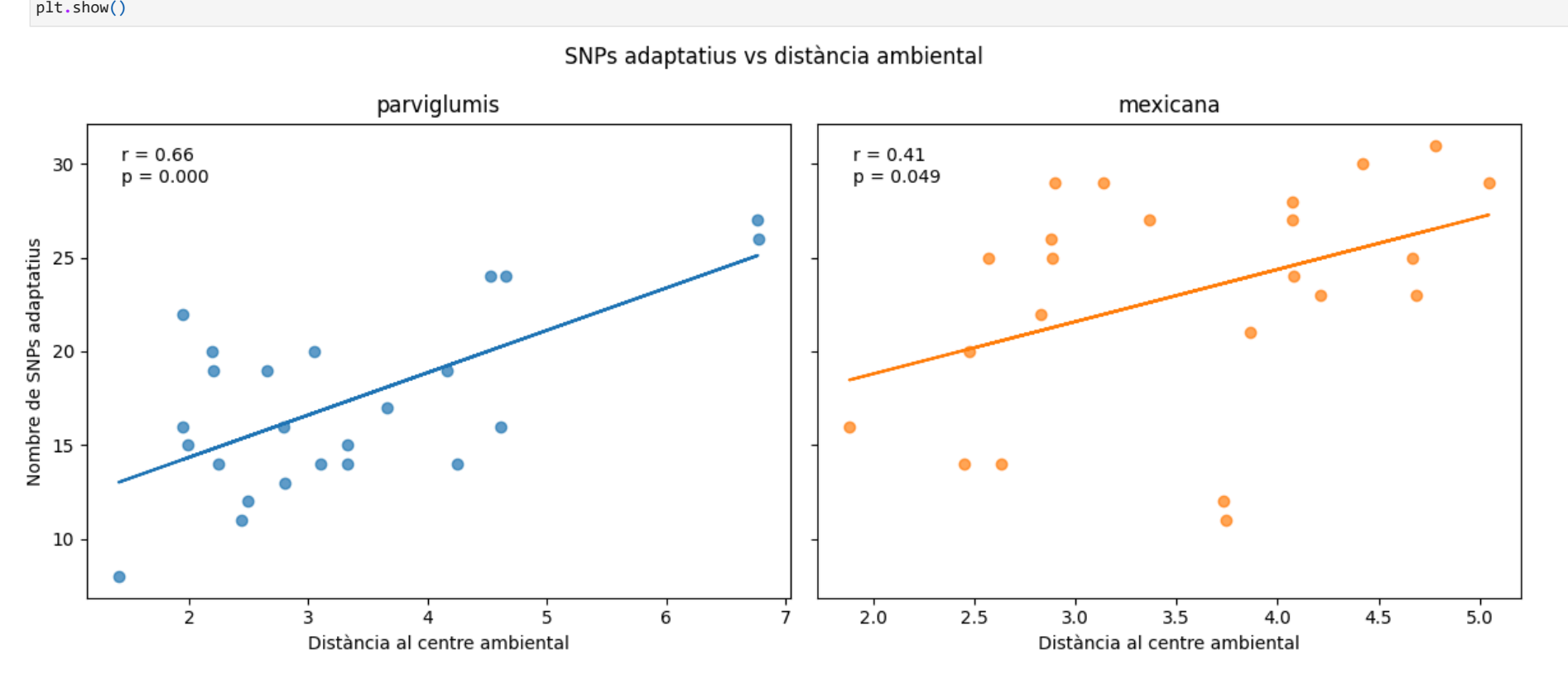
```python
import numpy as np
import matplotlib.pyplot as plt
from scipy.stats import pearsonr

# Preparar dades (eliminar NaNs)
parv = df[df["Subespècie"] == "parviglumis"][["Num_SNPs", "Dist_geogràfica"]].dropna()
mex  = df[df["Subespècie"] == "mexicana"][["Num_SNPs", "Dist_geogràfica"]].dropna()

fig, axes = plt.subplots(1, 2, figsize=(12, 5), sharey=True)

# --- parviglumis ---
x = parv["Dist_geogràfica"]
y = parv["Num_SNPs"]
r, p = pearsonr(x, y)

axes[0].scatter(x, y, color="tab:blue", alpha=0.7)
m, b = np.polyfit(x, y, 1)
axes[0].plot(x, m*x + b, color="tab:blue")
axes[0].set_title("parviglumis")
axes[0].set_xlabel("Distància al centre geogràfic")
axes[0].set_ylabel("Nombre de SNPs adaptatius")
axes[0].text(0.05, 0.95, f"r = {r:.2f}\np = {p:.3f}",
             transform=axes[0].transAxes, va="top")

# --- mexicana ---
x = mex["Dist_geogràfica"]
y = mex["Num_SNPs"]
r, p = pearsonr(x, y)

axes[1].scatter(x, y, color="tab:orange", alpha=0.7)
m, b = np.polyfit(x, y, 1)
axes[1].plot(x, m*x + b, color="tab:orange")
axes[1].set_title("mexicana")
axes[1].set_xlabel("Distància al centre geogràfic")
axes[1].text(0.05, 0.95, f"r = {r:.2f}\np = {p:.3f}",
             transform=axes[1].transAxes, va="top")

plt.suptitle("SNPs adaptatius vs distància geogràfica")
plt.tight_layout()
plt.show()
```



2. Executa el codi de sota per representar la relació entre el número de SNPs adaptatius i la distància ambiental:

In [6]:

```python
# Preparar dades (eliminar NaNs)
parv = df[df["Subespècie"] == "parviglumis"][["Num_SNPs", "Dist_ambiental"]].dropna()
mex  = df[df["Subespècie"] == "mexicana"][["Num_SNPs", "Dist_ambiental"]].dropna()

fig, axes = plt.subplots(1, 2, figsize=(12, 5), sharey=True)

# --- parviglumis ---
x = parv["Dist_ambiental"]
y = parv["Num_SNPs"]
r, p = pearsonr(x, y)

axes[0].scatter(x, y, color="tab:blue", alpha=0.7)
m, b = np.polyfit(x, y, 1)
axes[0].plot(x, m*x + b, color="tab:blue")
axes[0].set_title("parviglumis")
axes[0].set_xlabel("Distància al centre ambiental")
axes[0].set_ylabel("Nombre de SNPs adaptatius")
axes[0].text(0.05, 0.95, f"r = {r:.2f}\np = {p:.3f}",
             transform=axes[0].transAxes, va="top")

# --- mexicana ---
x = mex["Dist_ambiental"]
y = mex["Num_SNPs"]
r, p = pearsonr(x, y)

axes[1].scatter(x, y, color="tab:orange", alpha=0.7)
m, b = np.polyfit(x, y, 1)
axes[1].plot(x, m*x + b, color="tab:orange")
axes[1].set_title("mexicana")
axes[1].set_xlabel("Distància al centre ambiental")
axes[1].text(0.05, 0.95, f"r = {r:.2f}\np = {p:.3f}",
             transform=axes[1].transAxes, va="top")

plt.suptitle("SNPs adaptatius vs distància ambiental")
plt.tight_layout()
plt.show()
```



### Preguntes

**6. A partir de les gràfiques que relacionen el nombre de SNPs adaptatius amb la distància al centre de la distribució, quina afirmació és correcta?**

A. En totes dues subespècies, el nombre de SNPs adaptatius disminueix amb la distància al centre geogràfic i ambiental.
B. El nombre de SNPs adaptatius augmenta amb la distància al centre, especialment amb la distància ambiental, on la correlació és més forta i estadísticament significativa.
C. Només *Zea mays ssp. mexicana* mostra una correlació entre el nombre de SNPs adaptatius i la distància al centre.
D. No s'observa cap relació entre el nombre de SNPs adaptatius i la distància al centre de la distribució.

**7. Tenint en compte aquests resultats, quina interpretació és la més adequada en el marc de la hipòtesi de l'abundància al centre de la distribució (*central-abundance hypothesis*)?**

A. Les poblacions centrals acumulen més SNPs adaptatius perquè tenen una diversitat genètica més elevada.
B. Les poblacions perifèriques mostren més SNPs adaptatius perquè estan sotmeses a una pressió selectiva més forta, especialment associada a condicions ambientals diferents de les òptimes.
C. El nombre de SNPs adaptatius és independent tant de la posició geogràfica com de l'ambient.
D. Les diferències observades indiquen que l'adaptació local només es produeix en *Zea mays ssp. mexicana*.

## Simulació de l'evolució d'un al·lel adaptatiu en condicions de canvi climàtic

Fins ara hem analitzat dades genètiques reals de diverses poblacions de teosints i les hem relacionat amb dades geogràfiques i ambientals. Però, què pot passar en el futur si les condicions ambientals continuen canviant?

En aquest apartat simularem l'evolució de la freqüència d'un al·lel adaptatiu al llarg del temps, assumint que aquest al·lel confereix un avantatge selectiu en un ambient més càlid o més sec, com el que es preveu amb el canvi climàtic.

La simulació ens permetrà explorar com influeixen factors com:

- la intensitat de la selecció natural (`s`),
- i la mida poblacional (`N`),

en la probabilitat que un al·lel adaptatiu augmenti de freqüència o es perdi.

### Model de simulació

Assumirem que:

- hi ha dos al·lels: **A** i **a**
- l'al·lel **A** té un avantatge selectiu (`s > 0`),
- la població té una mida constant al llarg del temps (`N`),
- la reproducció és aleatòria (model *Wright–Fisher*).

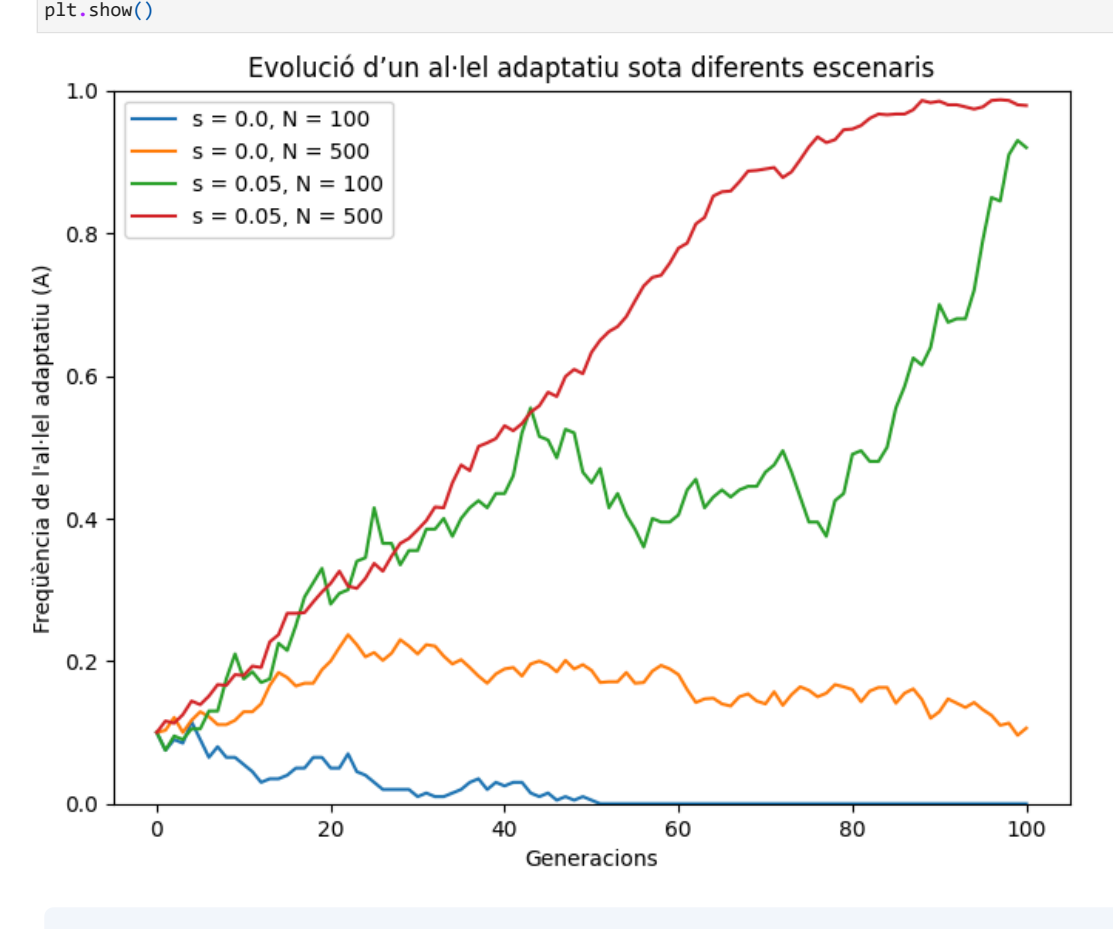
### Instruccions

1. El codi que trobaràs a continuació conté dues línies amb els paràmetres que hauràs de modificar (`valors_s` i `valors_N`). Segueix les instruccions de cada apartat a continuació, modifica els valors `valors_s` i `valors_N` del bloc de codi, i torna a executar-lo per observar com canvien els resultats.

2. La simulació inclou una component d'atzar (deriva genètica). Això fa que els resultats no siguin exactament iguals cada vegada que s'executa el codi, fins i tot mantenint els mateixos paràmetres. Executa la simulació diverses vegades amb els mateixos valors per identificar el comportament general o la tendència més probable.

In [7]:

```python
import numpy as np
import matplotlib.pyplot as plt

def simular_al_lel_adaptatiu(p0, s, N, generacions):
    p = p0
    freq = [p]
    for _ in range(generacions):
        # Selecció
        p_sel = (p * (1 + s)) / (p * (1 + s) + (1 - p))
        # Deriva genètica
        p = np.random.binomial(2*N, p_sel) / (2*N)
        freq.append(p)
    return freq

# --- PARÀMETRES FIXOS ---
p0 = 0.1
generacions = 100

# --- PARÀMETRES A MODIFICAR ---
valors_s = [0.0, 0.05]        # intensitat de la selecció
valors_N = [100, 500]        # mida poblacional

# --- SIMULACIÓ ---
plt.figure(figsize=(8, 6))

for s in valors_s:
    for N in valors_N:
        freq = simular_al_lel_adaptatiu(p0, s, N, generacions)
        plt.plot(freq, label=f"s = {s}, N = {N}")

# --- GRÀFIC ---
plt.xlabel("Generacions")
plt.ylabel("Freqüència de l'al·lel adaptatiu (A)")
plt.title("Evolució d'un al·lel adaptatiu sota diferents escenaris")
plt.ylim(0, 1)
plt.legend()
plt.show()
```



### Efecte de la intensitat de la selecció

1. Al bloc de codi de dalt, fixa la mida poblacional a `N` a 500:

`valors_N = [500]`

2. Prova diversos valors de selecció `s` entre 0 i 0.05, per exemple:

`valors_s = [0.0, 0.01, 0.03, 0.05]`

3. Executa el bloc de codi i observa com canvia la freqüència de l'al·lel adaptatiu **A** al llarg de les generacions. Recorda executar el codi diverses vegades per identificar el comportament general.

### Preguntes

**8. Què passa a mesura que augmenta el valor de `s` (intensitat de la selecció)?**

A. L'al·lel adaptatiu A disminueix de freqüència perquè la selecció elimina la variació genètica.
B. L'al·lel adaptatiu A augmenta de freqüència més ràpidament com més gran és el valor de \`s\`.
C. El valor de \`s\` no té cap efecte observable sobre l'evolució de la freqüència de l'al·lel adaptatiu A.
D. L'al·lel adaptatiu A només augmenta de freqüència quan \`s\` = 0.

### Efecte de la mida poblacional

1. Al bloc de codi de dalt, fixa la intensitat de la selecció a `s` = 0.05.

`valors_s = [0.05]`

2. Prova diferents mides poblacionals `N` entre 100 i 10000, per exemple:

`valors_N = [100, 1000, 2000, 10000]`

3. Executa el bloc de codi i observa com canvia la freqüència de l'al·lel adaptatiu **A** al llarg de les generacions. Recorda executar el codi diverses vegades per identificar el comportament general.

### Preguntes

**9. Quin efecte té la mida poblacional `N` sobre l'evolució de l'al·lel adaptatiu segons la simulació?**

A. En poblacions petites, l'al·lel adaptatiu **A** s'imposa més ràpidament.
B. En poblacions grans, la deriva genètica és més forta i genera més variabilitat entre simulacions.
C. En poblacions petites, la freqüència de l'al·lel **A** fluctua més i pot perdre's, tot i ser adaptatiu.
D. La mida poblacional no influeix en l'evolució de la freqüència de l'al·lel adaptatiu **A**.

## Conclusió

Al llarg d'aquesta prova pràctica hem analitzat dades genòmiques reals de poblacions de teosints per entendre com responen les espècies als gradients geogràfics i ambientals.

Després, les simulacions evolutives han permès explorar com la selecció natural i la mida poblacional influeixen en l'evolució futura d'al·lels adaptatius.

### Preguntes

**10. Tenint en compte els resultats de les anàlisis genòmiques i les simulacions realitzades en aquesta prova pràctica, quina afirmació resumeix millor la resposta dels teosints al canvi climàtic?**

A. Les poblacions centrals presenten menys diversitat genètica però acumulen més SNPs adaptatius, cosa que facilita la seva adaptació al canvi climàtic.
B. Les poblacions perifèriques tenen menys diversitat genètica però mostren més senyals d'adaptació local, i les simulacions indiquen que la seva resposta futura dependrà fortament de la selecció natural i de la mida poblacional.
C. Les simulacions mostren que tots els al·lels adaptatius s'imposen ràpidament, independentment de la mida poblacional i de la intensitat de la selecció.
D. Els resultats indiquen que la genòmica no permet predir la resposta de les espècies al canvi climàtic.

## Referències

Aguirre-Liguori, J. A., Tenaillon, M. I., Vázquez-Lobo, A., Gaut, B. S., Jaramillo-Correa, J. P., Montes-Hernandez, S., Souza, V., & Eguiarte, L. E. (2017). Connecting genomic patterns of local adaptation and niche suitability in teosintes. *Molecular ecology*, 26(16), 4226–4240. https://doi.org/10.1111/mec.14203

Bernatchez, L., Ferchaud, A. L., Berger, C. S., Venney, C. J., & Xuereb, A. (2024). Genomics for monitoring and understanding species responses to global climate change. *Nature reviews. Genetics*, 25(3), 165–183. https://doi.org/10.1038/s41576-023-00657-y