

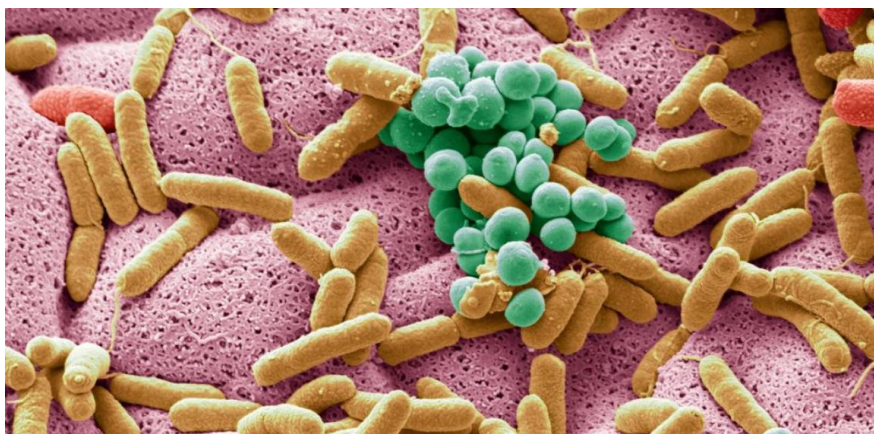


UNIVERSITAT  
ROVIRA I VIRGILI

## ANÀLISI DE LA COMPOSICIÓ DE LA MICROBIOTA INTESTINAL D'INDIVIDUS SALVATGES I CAPTIUS EN ESPÈCIES DE VERTEBRAT

Arey Ferrero Ramos

TREBALL FINAL DE GRAU BIOTECNOLOGIA



Tutor acadèmic i supervisor:

Nom: Marta Sales Pardo.

Departament: Departament Enginyeria  
Química, ETSEQ, URV.

Correu: [marta.sales@urv.cat](mailto:marta.sales@urv.cat)

En cooperació amb:

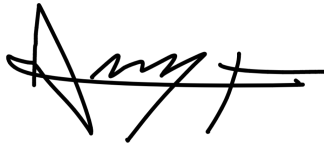
SEES:lab

Tarragona, setembre 2023



Jo, Arey Ferrero Ramos, amb DNI 41588224S, soc coneixedor de la guia de prevenció del plagi a la URV Prevenció, detecció i tractament del plagi en la docència: guia per a estudiants (aprovaada el juliol 2017) (<http://www.urv.cat/ca/vidacampus/serveis/crai/que-us-oferim/formacio-competencies-nuclears/plagi/>) i afirmo que aquest TFG no constitueixen cap de les conductes considerades com a plagi per la URV.

Tarragona, 1 de setembre de 2023

A handwritten signature in black ink, consisting of a stylized 'A' followed by a series of loops and a horizontal line extending to the right.

# Índex

|  |    |
|--|----|
| Resum.....   | 3  |
| Introducció .....  | 4  |
| Hipòtesis del treball i objectius .....  | 6  |
| Hipòtesis del treball .....  | 6  |
| Objectius .....  | 6  |
| Metodologies .....   | 7  |
| Obtenció de les dades de microbiotes intestinals d'animals.....                                | 7  |
| Matrius d'abundàncies relatives .....  | 7  |
| Descripció preliminar de les dades .....   | 8  |
| Distribució d'abundàncies relatives .....  | 9  |
| Selecció de factors dietètics i ambientals .....   | 11 |
| Classe.....  | 11 |
| Habitat i bioma.....   | 11 |
| Dieta .....  | 12 |
| Estat de conservació.....  | 12 |
| Càlcul de l' $\alpha$ -diversity .....   | 13 |
| Test <i>t</i> de Student .....   | 13 |
| Càlcul de la distància euclidiana.....   | 14 |
| Càlcul del coeficient de correlació de Pearson .....   | 14 |
| Diagrama de caixes o <i>boxplot</i> .....  | 15 |
| Discretització de les matrius d'abundàncies relatives .....                                    | 15 |
| Càlcul del valor de <i>nestedness</i> .....  | 15 |
| Obtenció del p-valor en un estudi de <i>nestedness</i> .....                                   | 16 |
| Resultats i discussió.....   | 17 |
| Estudi de diversitats .....  | 17 |
| Resultats del càlcul de les diversitats.....   | 17 |
| Influència dels factors dietètics i ambientals en els resultats del càlcul de les diversitats. | 18 |
| Estudis de similituds i de correlació entre similituds i diversitats .....                     | 26 |
| Estudi de similituds .....   | 26 |
| Estudi de correlació entre similituds i diversitats .....                                      | 29 |
| Estudi de <i>nestedness</i> .....  | 31 |
| Selecció dels llindars .....   | 31 |
| Valors de <i>nestedness</i> .....  | 32 |

|   |    |
|---|----|
| Generació dels <i>clustermaps</i> ..... | 32 |
| Obtenció del p-valor.....               | 33 |
| Conclusions .....                       | 35 |
| Referències.....                        | 36 |
| Autoavaluació.....                      | 43 |

## Resum

La microbiota intestinal té una importància capital en moltes funcions biològiques dels éssers humans. Per aquest motiu, es tracta de la comunitat microbiana més estudiada del nostre organisme. Per altra banda, sobre la composició i funcionament de la microbiota intestinal dels animals es tenen molts menys coneixements. Aquest estudi aprofundeix en les diferències en la composició de la microbiota intestinal de vint-i-quatre espècies de vertebrat en funció de si aquestes pertanyen a individus que es troben en llibertat o en captiveri. Les dades per desenvolupar aquest projecte s'han obtingut de l'estudi (Alberdi et al., 2021), i s'han organitzat en matrius d'abundàncies relatives de microbis en individus o grups d'individus (salvatges o captius). Això ha permès realitzar un estudi de les diferències entre les diversitats d'aquests dos grups d'individus per cada espècie de vertebrat i de quins factors dietètics o ambientals les poden explicar. També ha permès realitzar un estudi similar respecte a les similituds i trobar una correlació entre el comportament d'aquestes i el de les diversitats per totes les espècies de vertebrat analitzades. Finalment, s'ha calculat el valor de *nestedness* de totes les matrius d'abundàncies relatives de la mateixa manera que en l'estudi (Cobo-López et al., 2022). El projecte no ha quedat tancat, ja que caldria realitzar l'últim pas de l'estudi de *nestedness*, es a dir, comparar els valors calculats amb els de la hipòtesi nul·la per saber si aquests són o no són estadísticament significatius.

**Paraules clau:** microbiota, captiveri, diversitat, patrons ecològics (*nestedness*).

## Introducció

La microbiota intestinal és una comunitat microbiana allotjada al tracte gastrointestinal, que s'encarrega de modular molts processos fisiològics imprescindibles per garantir la salut dels éssers humans (Singh et al., 2017). Tot i que es tracta de la comunitat microbiana més estudiada, encara no s'ha aconseguit establir quina és la composició característica d'una microbiota saludable (Cobo-López et al., 2022). Això es degut a la gran variabilitat que existeix entre les composicions de les microbiotes dels individus sans (Huttenhower et al., 2012). Per altra banda, sí que s'ha detectat una certa correspondència entre una elevada diversitat en la microbiota intestinal i una millor salut de l'hoste (Gilbert et al., 2019).

Tot i la variabilitat observada en la composició de les microbiotes, s'ha detectat que les microbiotes dels individus sans realitzen les mateixes funcions essencials, cosa que suggereix l'existència de patrons comuns en la composició de les microbiotes saludables (Costea et al., 2018). Els enterotips són un mecanisme de classificació de la microbiota en base a aquests patrons que ha demostrat certes consistències (Costea et al., 2018). No obstant, la utilitat de la classificació basada en enterotips és motiu de controvèrsia, ja que no és prou robusta respecte a poblacions geogràficament diferenciades (Costea et al., 2018).

L'estudi de les microbiotes intestinals humanes ha mostrat que si aquestes s'organitzen en matrius binàries d'abundàncies relatives d'espècies bacterianes (o agrupacions d'aquestes espècies bacterianes) en un conjunt d'individus (o agrupacions d'aquests individus), aquestes segueixen un determinat patró jeràrquic anomenat *nested* (Cobo-López et al., 2022). Aquest patró es va observar primerament en les faunes d'habitats fragmentats o arxipèlags (Patterson et al., 1986) i més endavant en altres sistemes ecològics com comunitats d'endoparàsits d'animals (Poulin et al., 2001) (Warburton et al., 2018) i organitzacions mutualistes (Bastolla et al., 2009) (Mariani et al., 2019). En les microbiotes d'individus de l'espècie humana, un patró *nested* indica, per una banda, l'existència de microbis generalistes que estan presents en la major part dels individus (o agrupacions d'aquests individus) i microbis especialistes que

estan presents tan sols en una part d'aquests, i, per l'altra banda, l'existència d'hostes generalistes que presenten la major part dels microbis (o agrupacions d'aquests microbis) i d'hostes especialistes que presenten únicament una part d'aquests (Cobo-López et al., 2022). La presència d'aquest patró en les microbiotes és un factor que cal tenir en consideració en la investigació en tractaments amb medicina personalitzada (Cobo-López et al., 2022).

La microbiota intestinal també es troba al tracte gastrointestinal de molts animals (McFall-Ngai et al., 2013) a on també desenvolupa processos fisiològics importants (Lynch et al., 2019). Aquest processos han estat molt menys estudiats que aquells realitzats per la microbiota humana (Levin et al., 2021). Es per aquesta raó que el descobriment d'un patró *nested* en les microbiotes d'altres espècies animals podria tenir interès científic i, fins i tot, aplicacions igual de rellevants a l'esmentada en éssers humans.

La microbiota intestinal es veu afectada des del naixement per diversos factors dietètics i ambientals, que porten a l'observació d'una gran variabilitat entre individus de la mateixa espècie, tant en humans (Cobo-López et al., 2022) com en altres espècies d'animals (Alberdi et al., 2021). Alguns d'aquests factors dietètics i ambientals poden tenir una implicació en les diferències entre la microbiota dels animals en estat salvatge i la dels animals en condicions de captiveri, tot i que encara no s'ha pogut determinar gairebé cap patró comú en totes les espècies d'animals (Alberdi et al., 2021). L'única característica comuna en totes les microbiotes d'animals en captiveri és un augment de microorganismes propis de la microbiota dels éssers humans respecte al seu estat salvatge (Alberdi et al., 2021). Per altra banda, existeixen evidències tant de que la diversitat de la microbiota intestinal disminueix com de que augmenta en condicions de captiveri (McKenzie et al., 2017).

La investigació sobre com les condicions de captiveri afecten a la microbiota intestinal dels animals és rellevant, ja que les alteracions observades poden tenir diversos efectes, com desequilibris nutricionals (Gibson et al., 2019) o majors dificultats en la reintroducció d'espècies al seu habitat natural (Cheng et al., 2015).



## Hipòtesis del treball i objectius

### Hipòtesis del treball

Existeixen un o més factors dietètics o ambientals que poden explicar les diferències entre la diversitat del conjunt de les microbiotes intestinals d'individus salvatges i la del conjunt de les microbiotes intestinals d'individus captius per a totes les espècies de vertebrat utilitzades en l'estudi (Alberdi et al., 2021).

Pot haver-hi una correlació entre la similitud de la microbiota de dos individus de la mateixa espècie segons les condicions en que s'han desenvolupat i la diversitat típica de la microbiota de l'espècie per totes les espècies de vertebrat utilitzades en l'estudi (Alberdi et al., 2021).

La microbiota intestinal dels vertebrats pot presentar el mateix patró *nested* que s'ha observat en la microbiota intestinal humana.

### Objectius

- Processar les dades de l'estudi (Alberdi et al., 2021) per obtenir les matrius d'abundàncies relatives de gèneres bacterians.
- Realitzar un estudi comparatiu de les diversitats de les microbiotes intestinals entre individus salvatges i individus captius de cada espècie de vertebrat i obtenir informació sobre els factors dietètics i ambientals que puguin explicar els resultats.
- Realitzar un estudi comparatiu de les distàncies entre individus salvatges, entre individus captius i entre individus salvatges i individus captius de cada espècie de vertebrat i estudiar si existeix una correlació entre aquests resultats i els de l'estudi de diversitat previ.
- Estimar el valor de *nestedness* de les matrius d'abundàncies relatives de gèneres bacterians.

## Metodologies

### Obtenció de les dades de microbiotes intestinals d'animals

Inicialment, les dades de microbiotes per portar a terme aquest projecte es volien obtenir de l'estudi (Levin et al., 2021). Malauradament, no es trobaven disponibles en el format necessari. En conseqüència, s'han hagut de replantejar alguns objectius, i les dades necessàries s'han obtingut del repositori de GitHub associat a l'estudi (Alberdi et al., 2021). En concret, se n'han extret tres fitxers. Primer, el fitxer 'count\_Genus\_all.tsv', que és un fitxer estructurat en forma matricial que conté el número d'espècies (abundàncies absolutes) de cada gènere bacterià en cada individu del qual s'han recollit mostres. Segon, el fitxer 'metadata.csv', que conté tota la informació de cada individu. D'aquest fitxer només ens interessarà saber el codi de l'espècie de vertebrat a la que pertany l'individu i el tipus d'individu (salvatge o captiu). Per últim, el fitxer 'sp\_code.txt', que és un fitxer de text que conté la correspondència entre el codi que se li ha assignat a una espècie de vertebrat i el seu nom científic. Aquest últim només serveix per a identificar en els gràfics les espècies de vertebrat pel nom científic, en lloc de pel codi.

### Matrius d'abundàncies relatives

Per poder portar a terme els estudis detallats posteriorment, les dades de partida s'han organitzat en matrius d'abundàncies relatives. Una matriu d'abundàncies relatives és una matriu  $M$  en la que cada element  $M_{ij}$  representa l'abundància relativa d'un gènere bacterià  $i$  en un individu o grup d'individus  $j$ , de manera que es compleix que  $\sum_i M_{ij} = 1$ .

$$M_{ij} = \frac{\text{Número de bacteris del gènere bacterià } i \text{ en l'individu/grup } j}{\text{Número de bacteris totals en l'individu/grup } j}$$

L'individu o grup d'individus j fa referència a un individu d'una espècie de vertebrat o a la totalitat d'individus d'un sol tipus (salvatge o captiu) d'una espècie de vertebrat. Això dona lloc a diferents tipus de matrius d'abundàncies relatives en funció de l'estudi concret que s'estigui realitzant.

### Descripció preliminar de les dades

L'estudi (Alberdi et al., 2021) proporciona mostres de microbiotes de vint-i-quatre espècies de vertebrat, una de les quals ha estat tractada com a dues espècies separades degut a que les mostres s'han obtingut de bases de dades diferents. El número de mostres obtingut es molt diferent depenent de l'espècie de vertebrat, sent onze mostres el menor número de mostres que s'ha obtingut i seixanta-tres el major número de mostres que s'ha obtingut, per una sola espècie de vertebrat. Per altra banda, en cada espècie de vertebrat, el número de mostres d'individus salvatges i el número de mostres d'individus captius és molt semblant. Aquestes dades es detallen de manera més específica en la taula 1.

| <b>Espècie</b>                 | <b>Núm. individus salvatges</b> | <b>Núm. individus captius</b> | <b>Núm. individus totals</b> |
|--------------------------------|---------------------------------|-------------------------------|------------------------------|
| <i>Ailuropoda melanoleuca</i>  | 30                              | 31                            | 61                           |
| <i>Aldabrachelys gigantea</i>  | 6                               | 6                             | 12                           |
| <i>Aphanius iberus</i>         | 11                              | 12                            | 23                           |
| <i>Bos gaurus</i>              | 9                               | 10                            | 19                           |
| <i>Cervus nippon</i>           | 7                               | 12                            | 19                           |
| <i>Chelonia mydas</i>          | 8                               | 7                             | 15                           |
| <i>Elaphurus davidianus</i>    | 6                               | 6                             | 12                           |
| <i>Equus kiang</i>             | 18                              | 21                            | 39                           |
| <i>Gorilla gorilla</i>         | 6                               | 6                             | 12                           |
| <i>Lasiorhinus latifrons</i>   | 19                              | 14                            | 34                           |
| <i>Moschus chrysogaster</i>    | 14                              | 13                            | 27                           |
| <i>Myrmecophaga trydactyla</i> | 6                               | 5                             | 11                           |
| <i>Papio anubis</i>            | 9                               | 9                             | 18                           |
| <i>Panthera tigris</i>         | 13                              | 13                            | 26                           |

| <b>Espècie</b>                  | <b>Núm. individus salvatges</b> | <b>Núm. individus captius</b> | <b>Núm. individus totals</b> |
|---------------------------------|---------------------------------|-------------------------------|------------------------------|
| <i>Pan troglodytes</i>          | 5                               | 6                             | 11                           |
| <i>Peromyscus leucopus</i>      | 17                              | 18                            | 35                           |
| <i>Peromyscus maniculatus</i>   | 32                              | 31                            | 63                           |
| <i>Pygathrix nemaeus</i>        | 27                              | 27                            | 54                           |
| <i>Rana dybowskii</i>           | 12                              | 12                            | 24                           |
| <i>Rhinopithecus brelichi</i>   | 9                               | 10                            | 19                           |
| <i>Sarcophilus harrisii</i> 1   | 10                              | 11                            | 21                           |
| <i>Sarcophilus harrisii</i> 2   | 6                               | 6                             | 12                           |
| <i>Shinisaurus crocodilurus</i> | 16                              | 11                            | 27                           |
| <i>Tursiops truncatus</i>       | 10                              | 12                            | 22                           |
| <i>Valencia hispanica</i>       | 16                              | 13                            | 29                           |

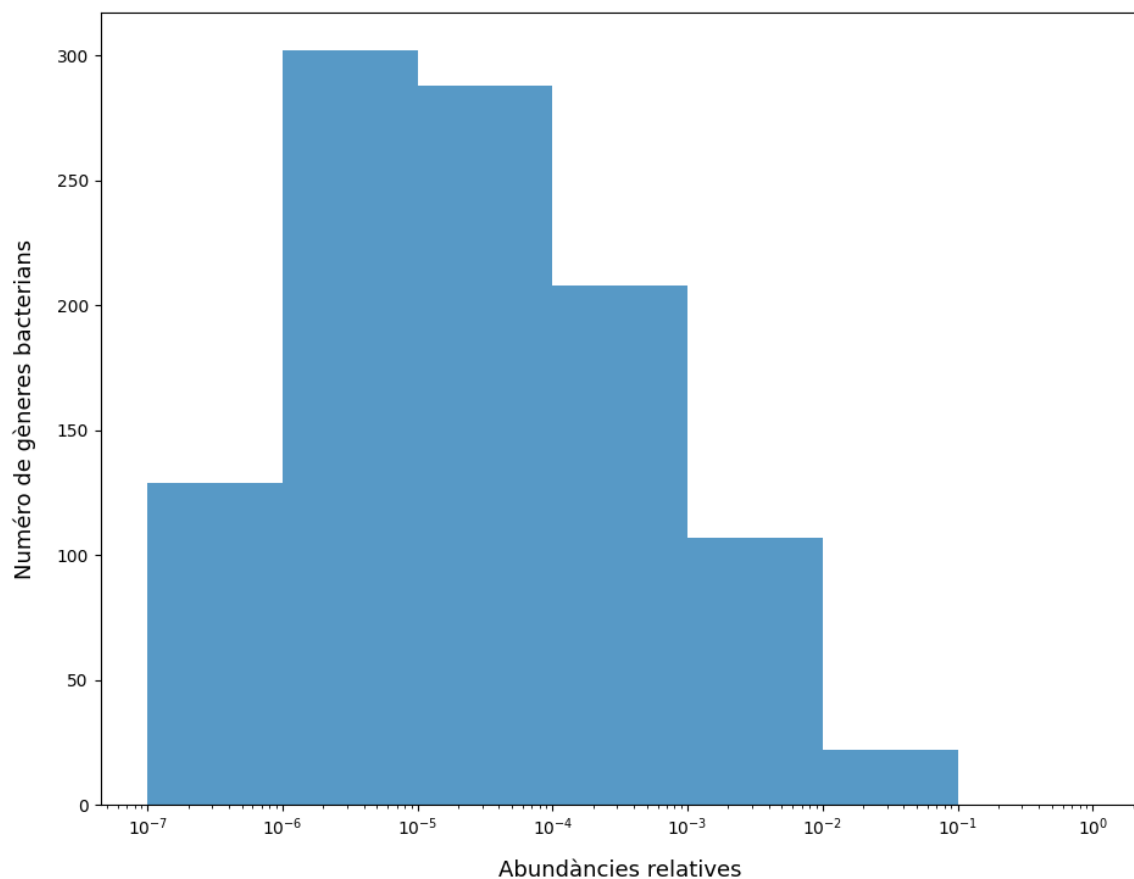
Taula 1. Número d'individus salvatges, número d'individus captius i número d'individus totals de cada espècie de vertebrat de l'estudi.

En el conjunt de totes les mostres, s'han detectat un total de mil cinquanta-sis gèneres bacterians diferents.

#### Distribució d'abundàncies relatives

Amb les dades obtingudes i estructurades, s'ha volgut caracteritzar el seu contingut. Per fer-ho, cal utilitzar la matriu d'abundàncies relatives de gèneres bacterians en tots els individus. Aquesta matriu, que té 1056 gèneres bacterians (files) i 625 individus (columnes), és la que s'ha obtingut dividint l'abundància absoluta de cada gènere bacterià d'un individu entre la suma d'abundàncies absolutes d'aquell individu. Posteriorment, hem discretitzat el contingut d'aquesta matriu en base al conjunt de valors  $\{10^{-1}, 10^{-2}, 10^{-3}, 10^{-4}, 10^{-5}, 10^{-6}, 10^{-7}\}$ . Tots els valors d'abundàncies relatives de gèneres bacterians que es troben per sota de  $10^{-7}$  s'ha considerat que son zero i se n'ha comptat la quantitat. Així, el percentatge d'abundàncies relatives amb valor zero és del 95.81%. Es tracta d'un valor esperat. La major part dels bacteris presents en la microbiota humana no es troben en la microbiota de la major part dels individus utilitzats en l'estudi

(Cobo-López et al., 2022). Per tant, hagués estat encara menys probable que la major part dels bacteris presents en la microbiota de les vint-i-quatre espècies de vertebrat es trobessin presents en la microbiota de la major part dels individus que pertanyen a aquestes espècies dels que s'han recopilat mostres per l'estudi (Alberdi et al., 2021). Totes les abundàncies relatives amb un valor diferent de zero s'han representat en un histograma a escala logarítmica (Figura 1).



*Fig 1. Freqüències relatives de gèneres bacterians. Histograma en el que es representen les freqüències relatives de gèneres bacterians per tots els individus, és a dir, el nombre de gèneres bacterians que tenen cada un dels valors d'abundàncies relatives discretitzades per tots els individus de l'estudi. La freqüència d'abundàncies relatives amb valor igual a 0 és del 95.81%.*

La informació mostrada en aquest gràfic tampoc és diferent del que s'hauria d'esperar. Els gèneres bacterians comencen a tenir representació amb un valor d'abundància relativa de  $10^{-1}$  i, per tant, no hi ha cap microbiota que tingui només bacteris d'un únic gènere bacterià. A partir d'aquí, el nombre de gèneres bacterians augmenta amb cada nou valor llindar d'abundàncies relatives fins

arribar al seu màxim quan les abundàncies relatives son  $10^{-4}$  i  $10^{-5}$ . Després torna a disminuir fins al límit inferior de  $10^{-7}$ .

D'aquest histograma també es pot extreure informació que permeti intuir quin pot ser el valor llindar en el pas de binarització de l'estudi de *nestedness* posterior. Així, el valor que sembla més probable és  $10^{-4}$ , ja que és el valor que divideix l'histograma en dos grups de manera més equitativa.

### Selecció de factors dietètics i ambientals

Els factors dietètics i ambientals seleccionats, explicats a continuació, son la categoria taxonòmica de classe, l'habitat, el bioma, la dieta i l'estat de conservació.

#### Classe

La classe és un taxó o categoria taxonòmica més específic que el fílum i més genèric que l'ordre (Sadava et al., 2009) (IUCN Red List, 2022). Els vertebrats (*vertebrata*), un subfílum del fílum *chordata* que es caracteritza per una columna vertebral articulada que ha substituït a la notocorda com a estructura de suport en les primeres fases del desenvolupament, estan dividits en nou classes: *Myxini* (mixines), *Hyperoartia* (llamprees), *Chondrichthyes* (peixos cartilaginosos), *Actinopterygii* (peixos ossis amb aletes radiades), *Sarcopterygii* (peixos ossis amb aletes lobulades), *Amphibia* (amfibis), *Reptilia* (rèptils), *Aves* (aus) i *Mammalia* (mamífers) (Sadava et al., 2009).

#### Habitat i bioma

Pot existir certa confusió entre els conceptes d'habitat i bioma. En el nostre estudi, l'habitat fa referència a la regió geogràfica on aquests vertebrats es localitzen en estat salvatge (IUCN Red List, 2022) mentre que el bioma fa referència a un ambient terrestre definit pel creixement de les formes de les

seves plantes (Sadava et al., 2009). Degut a que els límits entre els biomes son molt difusos (Sadava et al., 2009), la bibliografia científica pot arribar a reconèixer molts biomes diferents. En conseqüència, per reduir la complexitat d'aquest estudi, s'ha considerat una de les divisions més senzilles: bosc tropical o selva, bosc temperat o bosc, prat tropical o sabana, prat temperat o estepa, matoll, tundra i desert (Sadava et al., 2009) (IUCN Red List, 2022). Encara que l'aiguamoll correspondria, més aviat, al límit entre dos biomes (Sadava et al., 2009), degut a la idiosincràsia d'algunes de les espècies avaluades en el nostre estudi, se l'ha considerat un bioma més (IUCN Red List, 2022).

### Dieta

Els vertebrats tenen nutrició heteròtrofa, es a dir, la seva energia depèn del consum d'altres organismes (Sadava et al., 2009). Segons la seva dieta, els vertebrats es divideixen en tres grans grups: Herbívors (consumeixen plantes), carnívors (consumeixen altres animals) i omnívors (consumeixen plantes i altres animals) (Sadava et al., 2009). Per a reduir la complexitat de l'estudi, aquells animals que, tot i no seguir una dieta estricta, s'alimenten principalment de plantes o de carn, s'han classificat com a herbívors o carnívors, respectivament; sense fer cap distinció d'aquells animals que sí que segueixen una dieta estricta.

### Estat de conservació

L'estat de conservació d'una espècie fa referència al nivell de risc d'extingir-se que aquesta espècie té en estat salvatge, cosa que és determinada per nombrosos factors com la mida de la població, la distribució i les amenaces per a la seva supervivència (IUCN Red List, 2022). Sense tenir en compte aquells cassos en els que no s'ha fet una avaluació o en els que les dades que se n'han extret son deficientes, l'estat de conservació es divideix en vuit grups de menor a major risc: Preocupació menor (LC), gairebé amenaçat (NT), vulnerable (VU), en perill (EN), en perill crític (CR), extint en certes regions (RE), extint en estat salvatge (EW) i extint (EX) (IUCN Red List, 2022).

### Càlcul de l' $\alpha$ -diversity

L' $\alpha$ -diversity (Whittaker, 1960) (Whittaker, 1972) mesura la heterogeneïtat de perfils d'abundàncies relatives d'un individu. Aquesta mesura de diversitat pot calcular-se emprant diferents índexs (Lande, 1996). Un dels més utilitzats és l'índex de Shannon (Lande, 1996) (Grice et al., 2009).

$$\alpha - diversity_j = - \sum_i M_{ij} \ln M_{ij}$$

En el nostre estudi, aquesta mètrica ens permet obtenir la diversitat de gèneres bacterians en la microbiota d'un individu.

$$M_{ij} = \frac{\text{Número de bacteris del gènere bacterià } i \text{ en l'individu } j}{\text{Número de bacteris totals en l'individu } j}$$

### Test *t* de Student

El test *t* de Student és un test estadístic que s'utilitza per comprovar si la similitud entre dues poblacions és estadísticament significativa (Teplow, 2020). Aquest test s'utilitza quan les distribucions son normals i, en el cas de treballar amb diversitats, quan aquestes s'han calculat mitjançant l'índex de Shannon (Teplow, 2020). El format de correspondència entre els lindars dels p-valors escollits i la significança estadística es mostra a la taula 2 (*New England Journal of Medicine*).

En el llenguatge de programació Python aquesta prova es pot portar a terme mitjançant la funció `stats.ttest_ind()` de la llibreria `scipy`, que retorna un objecte que conté el t-estadístic i el p-valor (`scipy.org`, 2008-2023).



| <b>p-valor</b>                     | <b>Significança</b>     |
|------------------------------------|-------------------------|
| $p\text{-valor} \leq 0.001$        | Significatiu (***)      |
| $0.001 < p\text{-valor} \leq 0.01$ | Significatiu (**)       |
| $0.01 < p\text{-valor} \leq 0.05$  | Significatiu (*)        |
| $p\text{-valor} > 0.05$            | No significatiu (n. s.) |

*Taula 2. Correspondència entre el llinar dels p-valors i el format de significança estadística (New England Journal of Medicine)*

### Càlcul de la distància euclidiana

La distància euclidiana s'utilitza com a mesura de la similitud entre les abundàncies relatives de gèneres bacterians dels individus salvatges, entre les dels individus captius i entre les d'individus salvatges i individus captius de cada espècie de vertebrat. Es calcula mitjançant la següent equació:

$$d(M_{j_1}, M_{j_2}) = \sqrt{\sum_i (M_{j_{1i}} - M_{j_{2i}})^2}$$

En el llenguatge de programació Python, per obtenir la distància euclidiana normalitzada, s'ha d'utilitzar la funció `linalg.norm()` de la llibreria `numpy`, que rebrà com a paràmetre el valor resultant de l'equació anterior ([numpy.org](https://numpy.org), 2008-2022).

### Càlcul del coeficient de correlació de Pearson

En el llenguatge de programació Python, el coeficient de correlació de Pearson es calcula amb la funció `stats.pearsonr()` de la llibreria `scipy` ([scipy.org](https://scipy.org), 2008-2023).

### Diagrama de caixes o *boxplot*

Un diagrama de caixes o *boxplot* és un tipus de gràfic que proporciona un representació visual de la mediana, els quartils, la desviació estàndard i els valors anormals o *outliers* d'un conjunt de dades (Dutoit, 2012).

En el llenguatge de programació Python, els diagrames de caixes es poden generar amb la funció `boxplot()` de la llibreria `seaborn` ([seaborn.pydata.org](https://seaborn.pydata.org), 2012-2022).

### Discretització de les matrius d'abundàncies relatives

Per a poder fer un estudi de *nestedness* sobre una matriu, cal que aquesta sigui una matriu binària. Per tant, hem de discretitzar la matriu d'abundàncies relatives  $M$  per donar lloc a una matriu binària  $Q$ . Es a dir, cal escollir un valor llindar  $T$  per sota del qual es pot considerar que les abundàncies relatives  $M_{ij}$  son negligibles.

$$Q = \begin{cases} q_{ij} = 1 & \text{si } M_{ij} \geq T \\ q_{ij} = 0 & \text{si } M_{ij} < T \end{cases}$$

El llindar adient es pot intuir a partir de l'histograma d'abundàncies relatives de tots els gèneres bacterians en tots els individus i dels *clustermaps* resultants de discretitzar la matriu d'abundàncies relatives per cada un dels llindars possibles. Quan l'estudi de *nestedness* s'ha completat, el llindar adient per cada matriu d'abundàncies relatives ha de ser aquell que maximitzi el valor de *nestedness* de la matriu respecte al de les seves randomitzacions.

### Càlcul del valor de *nestedness*

Per dissenyar l'algorisme per calcular el valor de *nestedness* d'una matriu binària es parteix de la següent equació (Cobo-López et al., 2022):

$$n = \frac{\sum_{a < b} q_{ab}^{(i)} + \sum_{a < b} q_{ab}^{(j)}}{\sum_{a < b} \min(q_a^{(i)}, q_b^{(i)}) + \sum_{a < b} \min(q_a^{(j)}, q_b^{(j)})}$$

On **i** representa el gènere bacterià (files); **j** representa l'individu o espècie de vertebrat (columnes); **q\_a** (**q\_b**) representa el número d'interaccions (elements diferents de zero) en una fila (columna); **q\_ab** representa el número d'interaccions compartides entre dues files o dues columnes.

#### Obtenció del p-valor en un estudi de *nestedness*

Per si sol, el valor de *nestedness* d'una matriu binària no ens proporciona prou informació per determinar si aquesta té o no té un patró *nested*. Es necessari comparar aquest valor de *nestedness* amb el de la hipòtesi nul·la. La hipòtesi nul·la proposa que el valor de *nestedness* calculat és només una conseqüència de la densitat de la matriu d'abundàncies discretitzades Q.

Per portar a terme aquest test, generem mil randomitzacions. Una randomització és una matriu binària que es genera redistribuint aleatòriament aquelles posicions de la matriu Q que tenen un valor igual a 1. Per tant, les randomitzacions tenen el mateix nombre de files, el mateix nombre de columnes i el mateix nombre de valors iguals a 1 que la matriu Q, però es diferencien en les posicions que aquests valors iguals a 1 ocupen. Els valors de *nestedness* calculats per la matriu Q i per les mil randomitzacions s'ordenen i se'n calcula el p-valor, que és la fracció de randomitzacions que tenen un valor de *nestedness* més gran que el de la matriu Q. El valor de *nestedness* es considerarà significatiu si el p-valor és inferior o igual a 0.05.

## Resultats i discussió

### Estudi de diversitats

#### Resultats del càlcul de les diversitats

L'objectiu d'aquest apartat ha estat fer un estudi comparatiu entre les diversitats de les microbiotes dels individus salvatges i les diversitats de les microbiotes dels individus captius per cada espècie de vertebrat. Primer de tot, s'ha calculat la diversitat de gèneres bacterians en cada individu  $j$  a partir de la matriu d'abundàncies relatives de gèneres bacterians (1056) en individus (625). La mètrica utilitzada ha estat l' $\alpha$ -diversity. A continuació, les diversitats calculades s'han organitzat segons el tipus d'individu (salvatge i captiu) i segons l'espècie de vertebrat a la que pertanyen i s'han representat en un diagrama de caixes o *boxplot* (Figura 2).

Ara bé, la representació de les diversitats de les microbiotes dels individus salvatges al costat de les de les microbiotes dels individus captius en un gràfic no proporciona la suficient informació per comparar aquestes dues poblacions. Això es degut a que el nombre d'elements de les poblacions influeix en la percepció que es té respecte a com de diferents són realment les dues poblacions. Es a dir, dues poblacions que tinguin un gran nombre d'individus poden estar molt separades en un gràfic sense que la diferència real entre aquestes sigui molt gran mentre que en una població que tingui pocs individus una diferència real considerable es pot correspondre amb una petita diferència en la seva representació gràfica. Per tant, per determinar si les diferències observades són significatives o no, es necessari fer un test  $t$  de *Student* als perfils d'abundàncies relatives de cada espècie de vertebrat. El conjunt de les diversitats dels individus salvatges correspondrà a la primera població mentre que el conjunt de les diversitats dels individus captius a la segona. El format de significança estadística que correspon al p-valor (Taula 2) i el nombre d'individus de cada tipus en cada espècie de vertebrat també es representen en el diagrama de caixes (Figura 2).

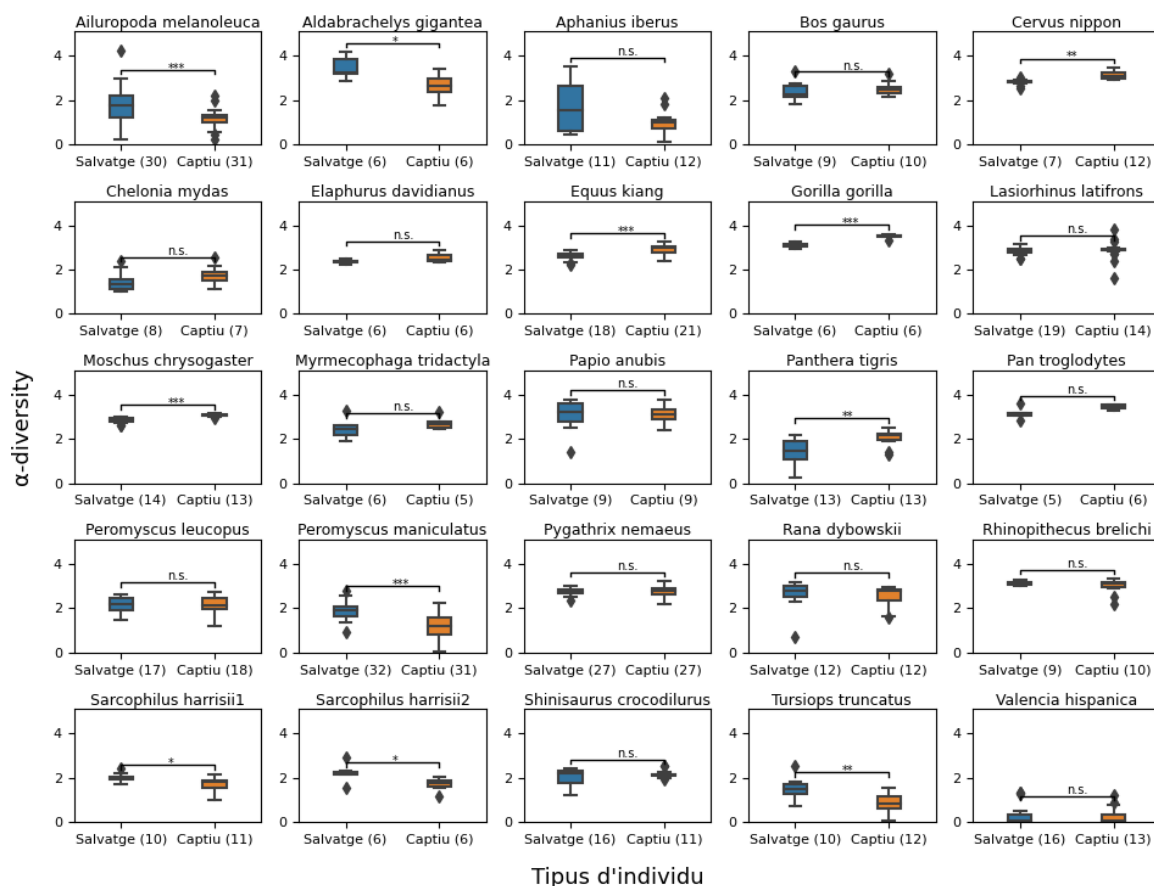


Fig 2. Diversitats de gèneres bacterians en espècies de vertebrat. Diagrames de caixes amb la comparativa entre les diversitats dels individus salvatges i les diversitats dels individus captius en cada espècie de vertebrat. P-valors (n.s.: p-valor > 0.05; \*: 0.05 ≥ p-valor > 0.01; \*\*: 0.01 ≥ p-valor > 0.001; \*\*\*: p-valor ≤ 0.001).

Un cop fets el càlcul de les diversitats, l'anàlisi estadístic i la representació gràfica dels resultats, es pot observar que, de les vint-i-cinc espècies de vertebrat de l'estudi (Alberdi et al., 2021), les diferències entre la diversitat en la microbiota dels individus salvatges i la diversitat en la microbiota dels individus captius és significativa en onze de les espècies i no significativa en les catorze espècies restants.

#### Influència dels factors dietètics i ambientals en els resultats del càlcul de les diversitats

S'ha recopilat i organitzat tota la informació necessària per discutir els resultats de l'estudi de diversitat (Taula 3). La primera columna de la taula correspon al nom científic de cada espècie de vertebrat de l'estudi i pot utilitzar-se com a clau per obtenir la informació de cada una d'aquestes. La segona columna indica el

nom comú de l'espècie i la seva funció es merament informativa, ja que possibilita que un revisor conegui l'espècie que s'està avaluant sense haver-la de cercar.

La tercera, quarta, cinquena, sisena i setena columna proporcionen les característiques d'aquestes espècies que podrien proporcionar una explicació sobre perquè es produeixen les variacions en la diversitat de la microbiota d'individus salvatges i individus captius. Finalment, la vuitena columna especifica el valor de significança resultant del test *t* de *Student* realitzat sobre els valors d' $\alpha$ -diversity del conjunt d'individus salvatges i captius de cada espècie de vertebrat.

Cal recordar que, tot i que sembla que les mostres utilitzades en l'estudi provinguin de vint-i-cinc espècies de vertebrats diferents, en realitat es tracta de vint-i-quatre espècies, ja que *Sarcophilus harrisii* 1 i *Sarcophilus harrisii* 2 en realitat no pertanyen a espècies diferents, sinó que són mostres d'individus de la mateixa espècie obtingudes de bases de dades d'estudis diferents (Alberdi et al, 2021).

### Classe

El peixet de sequiol (*Aphanius iberus*) i l'espècie *Valencia hispànica* pertanyen a la classe *Actinopterygii* i totes dues presenten diferències entre les diversitats dels seus individus salvatges i captius que no són significatives. *Rana dybowskii* és la única espècie que pertany a la classe *Amphibia* i també presenta una diferència entre les diversitats dels seus grups d'individus no significativa. La tortuga gegant d'Aldabra (*Aldabrachelys gigantea*), la tortuga verda (*Chelonia mydas*) i el llargardaix cocodril xinès (*Shinisaurus crocodilurus*) pertanyen a la classe *Reptilia*. En el cas de la primera d'aquestes tres espècies la diferència entre les diversitats és lleugerament significativa, mentre que en les altres dos les diferències són no significatives. La resta d'espècies d'aquest estudi, les quals en conjunt abasten tots els possibles intervals de significança que s'han utilitzat en aquest estudi, pertanyen a la classe *Mammalia*.

| Espècie                  | Nom comú                       | Classe         | Habitat                           | Bioma                                       | Dieta       | Conservació | Significança | Referència               |
|--------------------------|--------------------------------|----------------|-----------------------------------|---|-------------|-------------|--------------|--------------------------|
| Ailuropoda melanoleuca   | Panda gegant                   | Mammalia       | Xina                              | Bosc  | Herbívor VU |             | ***          | (Swaigood et al., 2016)  |
| Aldabrachelys gigantea   | Tortuga gegant d'Aldabra       | Reptilia       | Aldabra (Oceà Índic)              | Illa d'escull de corall                     | Herbívor VU |             | *            | (Grubb, 1971)            |
| Aphanis iberus           | Peixet de sequiol              | Actinopterygii | Espanya                           | Costa mediterrània                          | Omnívor EN  |             | n.s.         | (Crivelli, 2006)         |
| Bos gaurus               | Gaur (bisont de la Índia)      | Mammalia       | Àsia sud i sudoriental            | Bosc, sabana, estepa, matoll                | Herbívor VU |             | n.s.         | (Duckworth et al., 2016) |
| Cervus nippon            | Cérvol sika                    | Mammalia       | Àsia oriental                     | Bosc, estepa, matoll                        | Herbívor LC |             | **           | (Harris, 2015)           |
| Chelonia mydas           | Tortuga verda                  | Reptilia       | Oceà                              | Oceà tropical                               | Herbívor EN |             | n.s.         | (Seminoff, 2004)         |
| Elaphurus davidianus     | Cérvol del pare David          | Mammalia       | Xina                              | Estepa, aiguamoll                           | Herbívor EW |             | n.s.         | (Jiang et al., 2016)     |
| Equus kiang              | Kiang                          | Mammalia       | Altiplà tibetà                    | Estepa, desert, matoll                      | Herbívor LC |             | ***          | (Shah et al., 2015)      |
| Gorilla gorilla          | Gorila occidental              | Mammalia       | Àfrica occidental                 | Selva                                       | Herbívor CR |             | ***          | (Maisels et al., 2018)   |
| Lasiorhinus latifrons    | Uombat de musell pelut del sud | Mammalia       | Austràlia sud i occidental        | Estepa, bosc, sabana                        | Herbívor NT |             | n.s.         | (Woinarski et al., 2016) |
| Moschus chrysogaster     | Cérvol mesquer alpi            | Mammalia       | Himalaya                          | Bosc, estepa, matoll                        | Herbívor EN |             | ***          | (Harris, 2016)           |
| Myrmecophaga tridactyla  | Ós formiguer gegant            | Mammalia       | Amèrica del sud i central         | Selva, sabana, matoll, estepa               | Carnívor VU |             | n.s.         | (Miranda et al., 2014)   |
| Papio anubis             | Babuí d'oliva                  | Mammalia       | Àfrica subshariana                | Bosc, sabana, matoll, estepa                | Omnívor LC  |             | n.s.         | (Wallis, 2020)           |
| Panthera tigris          | Tigre                          | Mammalia       | Àsia                              | Selva, bosc, sabana, matoll, estepa         | Carnívor RE |             | **           | (Breitenmoser, 2010)     |
| Pan troglodytes          | Ximpanzé                       | Mammalia       | Àfrica central i occidental       | Selva, sabana                               | Omnívor EN  |             | n.s.         | (Humle et al., 2016)     |
| Peromyscus leucopus      | -                              | Mammalia       | Nord-Amèrica central i occidental | Bosc, matoll, estepa, desert, àrees urbanes | Omnívor LC  |             | n.s.         | (Cassola, 2016)          |
| Peromyscus maniculatus   | Ratolí cérvol nord-americà     | Mammalia       | Nord-Amèrica                      | Bosc, matoll, estepa, aiguamoll, desert     | Omnívor LC  |             | ***          | (Cassola, 2016)          |
| Pygathrix nemaeus        | Douc langur de tija vermella   | Mammalia       | Àsia sudoriental                  | Selva                                       | Herbívor EN |             | n.s.         | (Coudrat et al., 2020)   |
| Rana dybowskii           | -                              | Amphibia       | Riu Amur (Àsia occidental)        | Aiguamoll, bosc, matoll                     | Carnívor LC |             | n.s.         | (Kuzmin et al., 2004)    |
| Rhinopithecus brelichi   | Langur gris de nas xat         | Mammalia       | Gizhou (Xina)                     | Bosc  | Herbívor CR |             | n.s.         | (Long et al., 2022)      |
| Sarcophilus harrisi      | Diable de Tasmània             | Mammalia       | Tasmània (Austràlia)              | Bosc, matoll, estepa                        | Carnívor EN |             | *            | (Hawkins et al., 2008)   |
| Shinisaurus crocodilurus | Llangardaix cocodríl xinès     | Reptilia       | Xina, Vietnam                     | Aiguamoll, selva, bosc                      | Carnívor EN |             | n.s.         | (Nguyen et al., 2022)    |
| Tursiops truncatus       | Dofí nas d'ampolla comú        | Mammalia       | Oceà                              | Oceà tropical, oceà temperat                | Carnívor LC |             | **           | (Wells et al., 2019)     |
| Valencia hispanica       | -                              | Actinopterygii | Espanya                           | Costa mediterrània                          | Carnívor CR |             | n.s.         | (Crivelli, 2006)         |

Taula 3. Factors dietètics i ambientals de cada espècie de vertebrat.

Per tant, d'entre les espècies que no son mamífers tan sols n'hi ha una en la que la diferència sigui significativa i és tracta de l'interval més baix. En condicions de captiveri els animals entren en major contacte amb l'espècie humana, i per tant, existeix un traspàs de bacteris de la microbiota humana a la d'aquests animals (Alberdi et al., 2021)., Així, una possible explicació al fenomen observat podria ser que, com l'ésser humà pertany a la classe *Mammalia* (Sadava, D. et al., 2009), els gèneres bacterians presents a la microbiota humana, un cop traspassats, tenen més facilitat per adaptar-se a les condicions de la microbiota d'altres espècies quan aquestes pertanyen a aquesta classe que a d'altres de diferents. Per tant, la diferència entre les diversitats de les microbiotes d'una espècie en condicions de llibertat de les d'aquesta espècie en condicions de captiveri podria ser més gran si aquesta pertany a la classe *Mammalia*, ja que els bacteris de la microbiota humana tindran més facilitat per adaptar-se a les condicions de la seva microbiota.

Per desgracia, amb un nombre tan reduït d'exemplars de les altres classes no es pot arribar a cap conclusió sòlida sobre com varia la diversitat de la microbiota de condicions de llibertat a condicions de captiveri tenint en compte aquesta característica.

#### *Habitat*

Per aquest factor es poden plantejar diverses hipòtesis. La primera és que la diversitat de la microbiota d'aquelles espècies de vertebrat que habiten en regions de gran amplitud tindrà més variacions en condicions de captiveri que la de la microbiota de les espècies que habiten en regions molt reduïdes, degut a que les espècies de vertebrat que han colonitzat regions de gran amplitud tenen més facilitat per incorporar nous bacteris i a que, d'aquesta manera, la composició de la seva microbiota variï amb més facilitat. Una altra proposta vàlida podria ser la contrària. Aquelles espècies que en llibertat es troben només en regions molt concretes tindran variacions més elevades en la diversitat de les seves microbiotes en condicions de captiveri que les espècies que en llibertat han colonitzat amplies extensions territorials. Aquesta proposta es podria



justificar en que aquelles espècies de vertebrat que han colonitzat grans extensions territorials tindran una microbiota molt estable en la que pocs bacteris no autòctons s'hi podran adaptar mentre que aquelles espècies que s'ubiquen en regions molt concretes tindran una microbiota molt més concreta que pot variar fàcilment amb un canvi d'habitat, en aquest cas al captiveri.

Al estudiar els resultats, hem comprovat que cap de les dues propostes es compleix. Entre les microbiotes que tenen variacions de diversitat molt significatives entre els grup d'individus salvatges i el grup d'individus captius, n'hi ha que pertanyen a espècies que habiten en regions molt àmplies com el kiang (*Equus kiang*) o el ratolí cérvol nord-americà (*Peromyscus maniculatus*) i n'hi ha que es troben en d'altres molt més específiques com és el cas del panda gegant (*Ailuropoda melanoleuca*) o, en menor mesura, del goril·la occidental (*Gorilla gorilla*). Igualment, d'entre les microbiotes que no tenen variacions significatives entre aquests dos grups d'individus, algunes pertanyen a espècies localitzades en zones molt concretes com el peixet de sequiol (*Aphanius iberus*) o el llangardaix cocodril xinès (*Shinisaurus crocodilurus*) i d'altres a espècies que han colonitzat grans extensions territorials com la tortuga verda (*Chelonia mydas*) o *Peromyscus leucopus*.

Per tant, en aquest cas tampoc s'han pogut extreure conclusions sobre com la regió geogràfica en la que habita una espècie pot explicar com variarà la composició de la seva microbiota en condicions de captiveri.

### Bioma

Per aquest factor ambiental es poden plantejar també diverses hipòtesis. La primera té una forta equivalència amb el que s'ha plantejat pel factor habitat. La diferència és que, en lloc de parlar d'espècies situades en regions geogràfiques molt concretes, es parla d'espècies situades en un únic bioma i, en lloc de parlar d'espècies que han colonitzat grans extensions territorials, es parla d'espècies que s'han adaptat a diversos biomes. De nou, els resultats mostren que aquesta hipòtesi no es compleix. El kiang (*Equus kiang*) o el ratolí cérvol nord-americà (*Peromyscus maniculatus*) es poden localitzar en diversos biomes mentre que el

goril·la occidental (*Gorilla gorilla*) o el panda gegant (*Ailuropoda melanoleuca*) es troben en un de sol i tots quatre tenen diferències entre les diversitats de les seves microbiotes molt significatives. Igual que amb l'habitat, aquesta hipòtesis es podria plantejar en el sentit contrari, però tampoc queda avalada pels resultats. El douc langur de tija vermella (*Pygathrix nemaeus*) o el langur gris de nas xat (*Rhinopithecus brelichi*) s'ubiquen en un únic bioma mentre que el babuí d'oliva (*Papio anubis*) o l'os formiguer gegant (*Myrmecophaga tridactyla*) n'han colonitzat una gran quantitat, sent que, tots quatre tenen diferències entre les diversitats de les espècies en condicions de llibertat i en condicions de captiveri que no son significatives.

Una altra pregunta que ens podem plantejar és si aquelles espècies ubicades en biomes amb gran biodiversitat tindran microbiotes amb molta més diversitat en condicions de llibertat que en condicions de captiveri, mentre que aquelles espècies que estan ubicades en biomes amb baixa biodiversitat tindran microbiotes en les que la diversitat augmenta en condicions de captiveri. Els resultats també semblen indicar que la hipòtesi no es compleix. Algunes espècies característiques de la selva com el ximpanzé (*Pan troglodytes*) o el duc langur de tija vermella (*Pygathrix nemaeus*) no tenen variacions significatives en quant a la diversitat de la seva microbiota quan passen de condicions de llibertat a condicions de captiveri. En el cas del goril·la occidental (*Gorilla gorilla*), la diversitat de la seva microbiota fins i tot augmenta de manera molt significativa en condicions de captiveri. Ara bé, la hipòtesi si que semblaria complir-se en part ja que espècies com el cérvol sika (*Cervus nippon*) o el kiang (*Equus kiang*), localitzats en biomes amb més baixa biodiversitat com boscos d'alta muntanya, estepes i, fins i tot, deserts, presenten un augment significatiu de la seva biodiversitat en condicions de captiveri. No obstant, tenint en compte la manca d'espècies localitzades en biomes amb molt baixa biodiversitat i sabent que la hipòtesi ja no es compleix per aquelles espècies que es localitzen en biomes amb biodiversitats molt elevades, no podem considerar-la com a correcta.

El cas que val més la pena recalcar en aquesta secció és el de les dues espècies de ratolí que pertanyen al gènere *Peromyscus*. *Peromyscus leucopus* i *Peromyscus maniculatus* pertanyen a la mateixa classe, habiten en regions

geogràfiques gairebé idèntiques, tenen el mateix estat de conservació i, a grans trets, la seva dieta és igual. Ara bé, la variació en la diversitat de la microbiota entre el grup d'individus salvatges i el grup d'individus captius és no significativa en la primera espècie i molt significativa en el cas de la segona. Una possible explicació seria que *Peromyscus leucopus* no es troba en el bioma d'aiguamoll mentre que *Peromyscus maniculatus* sí, però no sembla la més probable. Una de millor és que *Peromyscus leucopus* s'ha adaptat a la vida en àrees urbanes. Això comporta que, en condicions de llibertat, hagi tingut una exposició important a la dieta dels humans de manera que els bacteris de la microbiota humana s'hauran traspassat a la seva microbiota. Per tant, es raonable que en condicions de captiveri aquesta espècie no pateixi cap canvi significatiu.

#### *Dieta*

De tots els factors analitzats, la dieta és el que es suggereix que té més influència en la composició de la microbiota d'un individu (Cobo-Lopez et al., 2022). De manera més concreta, s'ha constatat que les espècies que són herbívores tendeixen a tenir microbiotes amb més diversitat que les espècies omnívores, i alhora les microbiotes d'aquestes tendeixen a tenir més diversitat que les de les espècies carnívores (Levin et al., 2022). Per tant, seria de suposar que la diversitat de la microbiota d'una espècie criada en condicions de captiveri variï de manera diferent a com es comporta en condicions de llibertat en funció de si aquesta espècie és herbívora, omnívora o carnívora.

Els resultats obtinguts mostren que les espècies carnívores tendeixen a tenir diferències entre la diversitat de les microbiotes d'individus salvatges i la d'individus captius no significatives i, en cas de que siguin significatives, són poc significatives amb excepció del tigre (*Panthera tigris*) (carnívor estricte), en la que ja és bastant més significativa. En el cas de les espècies omnívores, aquestes presenten diferències entre la diversitat de les microbiotes d'individus salvatges i la d'individus captius que són no significatives amb excepció del ratolí cérvol nord-americà (*Peromyscus maniculatus*), que presenta diferències molt significatives. Finalment, en el cas de les espècies herbívores, la microbiota es

comporta de totes les maneres possibles (fins i tot en herbívors estrictes) havent-hi espècies com el kiang (*Equus kiang*) o el cérvol mesquer alpi (*Moschus chrysogaster*) que tenen diferències entre les diversitats dels individus salvatges i captius molt significatives mentre que en d'altres com el cérvol del pare David (*Elaphurus davidianus*) o el uombat de musell pelut del sud (*Lasiorhinus latifrons*) les diferències són no significatives. A la llum d'aquests resultats, no es pot extreure cap conclusió sobre si la dieta que té una espècie en condicions de llibertat serà un indicatiu del comportament de la diversitat de la seva microbiota en condicions de captiveri i de quina manera.

#### *Estat de conservació*

En el cas de l'estat de conservació, la hipòtesi que ens hem plantejat és que aquelles espècies que estan en estadis més propers a l'extinció són les que menys s'han adaptat a la convivència amb els éssers humans i, per tant, en les que la seva microbiota pateix més variacions en condicions de captiveri. Igual que amb la resta de factors avaluats, aquesta hipòtesi es compleix per algunes espècies, com és el cas del goril·la occidental (*Gorilla gorilla*), que està en un estadi molt proper a l'extinció i les diferències entre les diversitats de les seves microbiotes de condicions de llibertat a condicions de captiveri són molt significatives o del babuí d'oliva (*Papio anubis*) i *Rana dybowskii*, que estan en estadis molt allunyats de l'extinció i tenen diferències entre les diversitats de les seves microbiotes segons el grup d'individus no significatives.

Ara bé, com a passat amb la resta de factors, no es compleix amb la major part de les espècies. D'entre les que es troben en estadis propers a l'extinció, el cérvol del pare David (*Elaphurus davidianus*) o el langur gris de nas xat (*Rhinopithecus brelichi*) tenen diferències no significatives entre les diversitats de les microbiotes de condicions de llibertat a condicions de captiveri mentre que espècies en estadis bastant més allunyats de l'extinció com el kiang (*Equus kiang*) o el ratolí cérvol nord-americà (*Peromyscus maniculatus*) tenen diferències entre les diversitats de les microbiotes dels seus dos respectius grups d'individus molt significatives. Una de les possibles explicacions a perquè aquesta hipòtesi no es

compleix, és que algunes espècies estan en perill de extinció per causes diferents a l'acció humana, com es el cas del diable de Tasmània (*Sarcophilus harrisi*), que es troba en aquesta situació degut a una malaltia de tumors facials molt contagiosa (Bender, 2010).

## Estudis de similituds i de correlació entre similituds i diversitats

### Estudi de similituds

Per obtenir les mesures de similitud, s'han calculat les distàncies entre diferents subconjunts d'individus  $j$  obtinguts a partir de la matriu d'abundàncies relatives de gèneres bacterians (1056) en individus (625). En concret, s'ha calculat la distància entre els individus salvatges, entre els individus captius i entre els individus salvatges i els individus captius per cada espècie de vertebrat. La mètrica de distància seleccionada ha estat la distància euclidiana. Els resultats s'han representat en un diagrama de caixes o *boxplot* (Figura 3) i el valor mitjà d'aquests resultats en un *stemplot* (Figura 4).

En els gràfics obtinguts podem observar uns resultats generals molt semblants als de l'estudi (Alberdi et al, 2021) o als obtinguts en l'estudi de diversitat fet per nosaltres. És a dir, hi ha espècies de vertebrat en les que les distàncies entre individus salvatges son molt més grans que les distàncies entre individus captius, com és el cas del peixet de sequiol (*Aphanius iberus*) o del tigre (*Panthera Tigris*), cosa que suggereix una convergència de la microbiota dels individus en condicions de captiveri. Aquest comportament pot ser explicat per una simplificació de l'ambient en el que es desenvolupen els individus captius (Alberdi et al., 2021). En l'altre extrem, hi ha espècies de vertebrat en les que les distàncies entre individus salvatges son més petites que entre individus captius, com és el cas del douc langur de tija vermella (*Pygathrix nemaeus*) o el diable de Tasmània (*Sarcophilus harrisi*). Aquesta situació, menys esperada a priori, indica una diferenciació de la microbiota d'aquests individus en condicions de captiveri i, per tant, suggereix que aquestes condicions poden ser més diferents les unes de les altres del que ho son les condicions de llibertat. Una de les explicacions que es poden proposar per explicar aquest comportament és que

les espècies en les que es dona una convergència de la microbiota en condicions de captiveri tendeixen a ubicar-se en estat salvatge en regions geogràfiques molt àmplies, como és el cas del tigre que es troba per tota Àsia, mentre que aquelles espècies en les que es dona una diferenciació de la microbiota en condicions de captiveri s'ubiquen en estat salvatge en regions geogràfiques molt petites, com és el cas de les dues espècies mencionades, ubicades en unes poques regions del Sud-est asiàtic o en l'illa de Tasmània, respectivament.

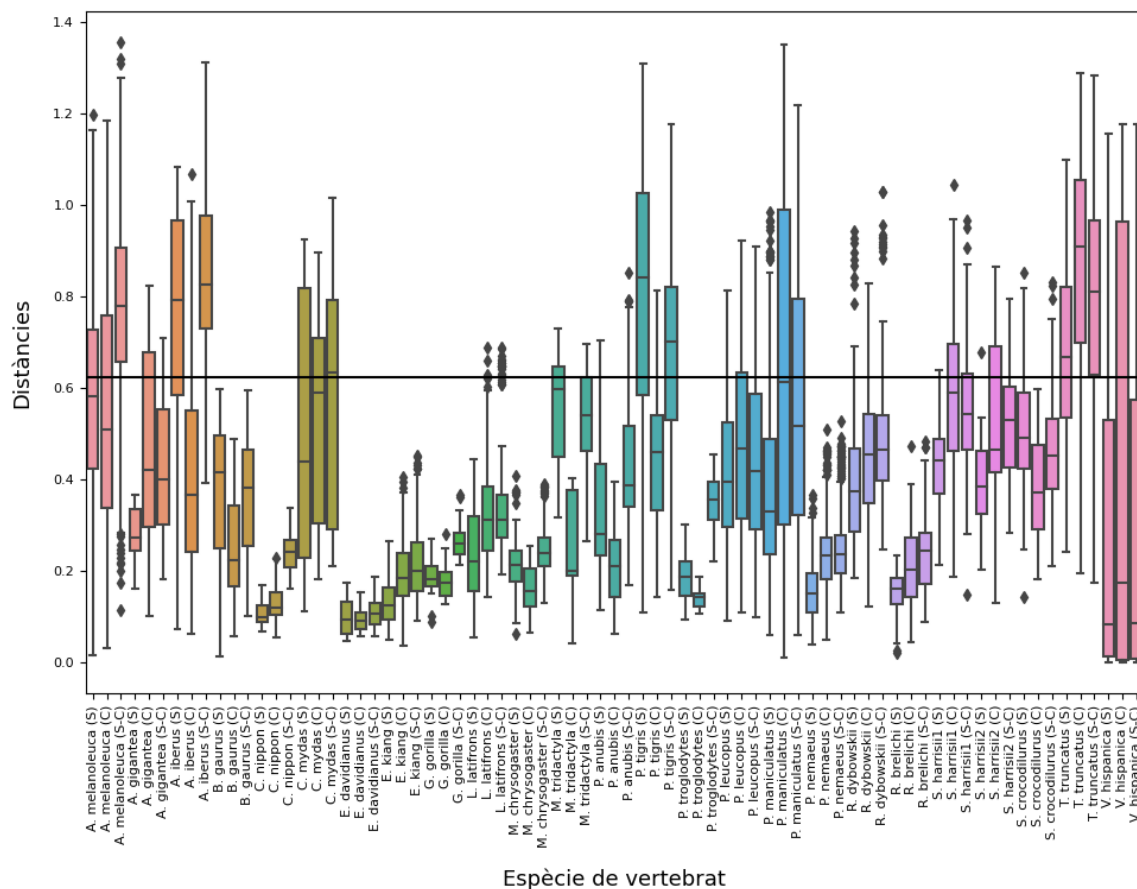


Fig 3. Similitud entre individus (salvatges, captius i salvatges-captius) per espècie de vertebrat. Diagrama de caixes en el que es representen les similituds entre tots els individus salvatges, entre tots els individus captius i entre tots els individus salvatges i captius de cada espècie de vertebrat.

Per altra banda, les distàncies entre individus salvatges i individus captius son, en general, més grans que les distàncies entre individus salvatges i que les distàncies entre individus captius. Aquests resultats son els que s'haurien d'esperar. Tot i així, hi ha alguns cassos en els que alguna d'aquestes dues condicions no es compleix. Així, en el cas de la tortuga gegant d'Aldabra

(*Aldabrachelys gigantea*) o del gaur (*Bos gaurus*), la distància entre individus captius és superior a la distància entre individus salvatges i captius. Això pot suggerir que les condicions de captiveri d'aquests dos animals son més diferents entre sí del que ho són amb les condicions en llibertat. El cas contrari seria, per exemple, el de l'os formiguer gegant (*Myrmecophaga tridactyla*) o el del llangardaix cocodril xinès (*Shinisaurus crocodilurus*), en els que la distància entre individus salvatges és superior a la distància entre individus salvatges i captius, suggerint que les condicions de vida en llibertat d'aquests dos animals son més diferents entre sí del que ho son amb les condicions de captiveri. No hi ha cap cas en que les distàncies entre individus salvatges i captius sigui inferior alhora a la distància entre individus salvatges i a la distància entre individus captius alhora.

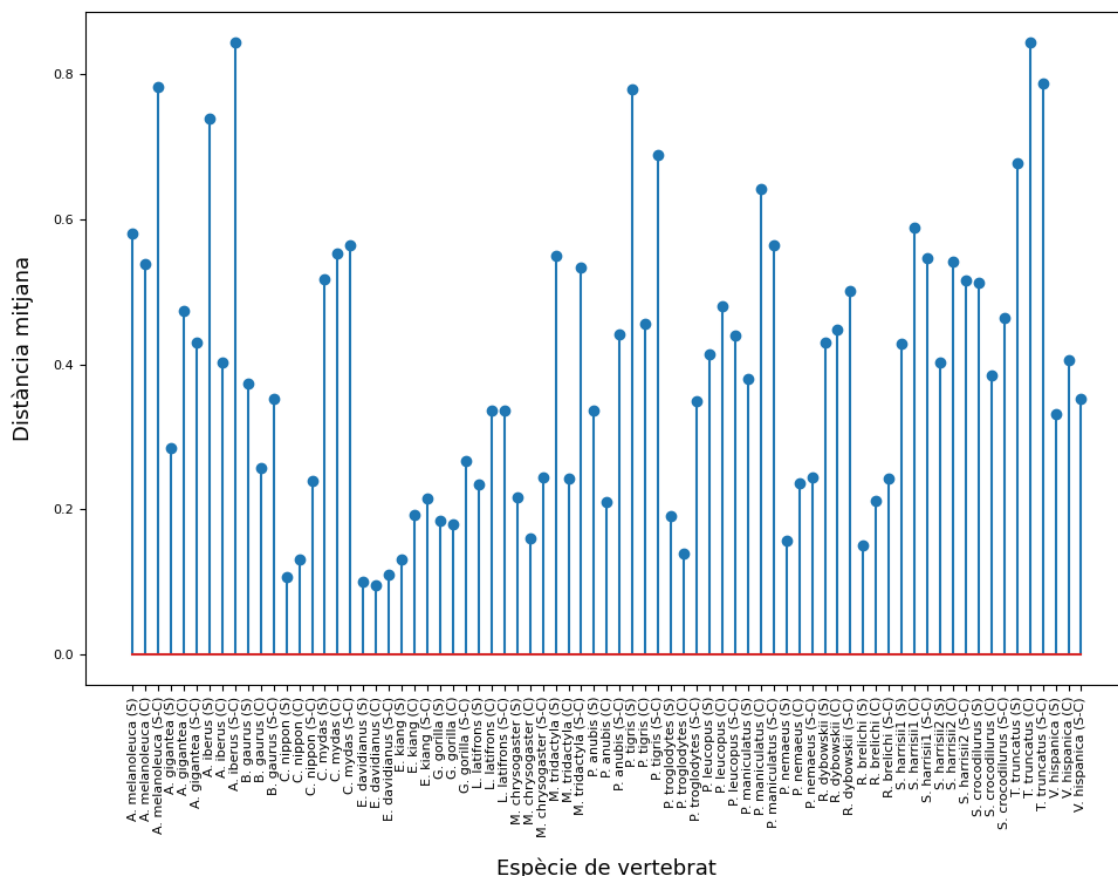


Fig 4. Similitud mitjana entre individus (salvatges, captius i salvatges-captius) per espècie de vertebrat. Stemplot en el que es representen les mitjanes de les similituds calculades entre tots els individus salvatges, tots els individus captius i tots els individus salvatges i captius de cada espècie de vertebrat.

El diagrama de caixes també presenta una línia base o *baseline* (Figura 3). Aquest element representa la mitjana de 10000 valors de distàncies entre individus de diferents espècies de vertebrat seleccionats a l'atzar. Per tant, la línia base permet comparar les diferències entre els individus d'un mateix tipus o espècie amb els individus d'altres espècies de vertebrat. Dit d'una altra manera, en tots aquells grups d'individus en els que el valor de la mediana estigui per sobre de la línia base, es pot establir que la similitud entre les microbiotes d'aquell grup d'individus tindrà un valor inferior al del valor mitjà de les similituds entre microbiotes de totes les espècies de vertebrat.

Per tant, el que s'esperaria és que cap de les medianes hagués tingut un valor superior al de la línia base però, tot i que en la major part dels cassos és així, això no es compleix per a totes les espècies de vertebrat. Per exemple, en el cas del panda gegant (*Ailuropoda melanoleuca*), la distància entre els individus salvatges i la distància entre els individus captius es inferior a la línia base però no així la distància entre els individus salvatges i els individus captius. Això podria suggerir que les condicions de llibertat i les condicions de captiveri pels espècimens del panda gegant son similars entre elles, però tant diferents les unes amb les altres que algunes espècies de vertebrat tenen microbiotes més similars a les d'altres espècies. Un altre cas seria el del dofí nas d'ampolla comú (*Tursiops truncatus*), en el que la distància entre les microbiotes per qualsevol grup d'individus és sempre superior a la línia base. Això suggereix que les microbiotes dels individus de dofí comú son més diferents entre elles del que ho son algunes de les microbiotes de moltes espècies de vertebrat en comparació amb les d'altres espècies de vertebrat.

#### Estudi de correlació entre similituds i diversitats

L'objectiu d'aquesta secció ha estat el de comprovar si existeix una correlació entre la similitud de la microbiota de dos individus de la mateixa espècie i la diversitat típica de la microbiota de l'espècie. La mètrica de diversitat utilitzada ha estat l' $\alpha$ -diversity i la mètrica de distància ha estat la distancia euclidiana, totes dues calculades en estudis anteriors. Una vegada s'han obtingut les dues



mètriques, es representen les distàncies en funció de les diversitats en un gràfic de dispersió o *scatterplot* (Figura 5).

De nou, representar gràficament aquests resultats és insuficient per saber si existeix una correlació entre distàncies i diversitats. Cal calcular el coeficient de correlació i comprovar si el p-valor resultant és significatiu o no. El coeficient de correlació utilitzat ha estat el de la correlació de Pearson que, junt amb el p-valor, estan representats en el gràfic (Figura 5).

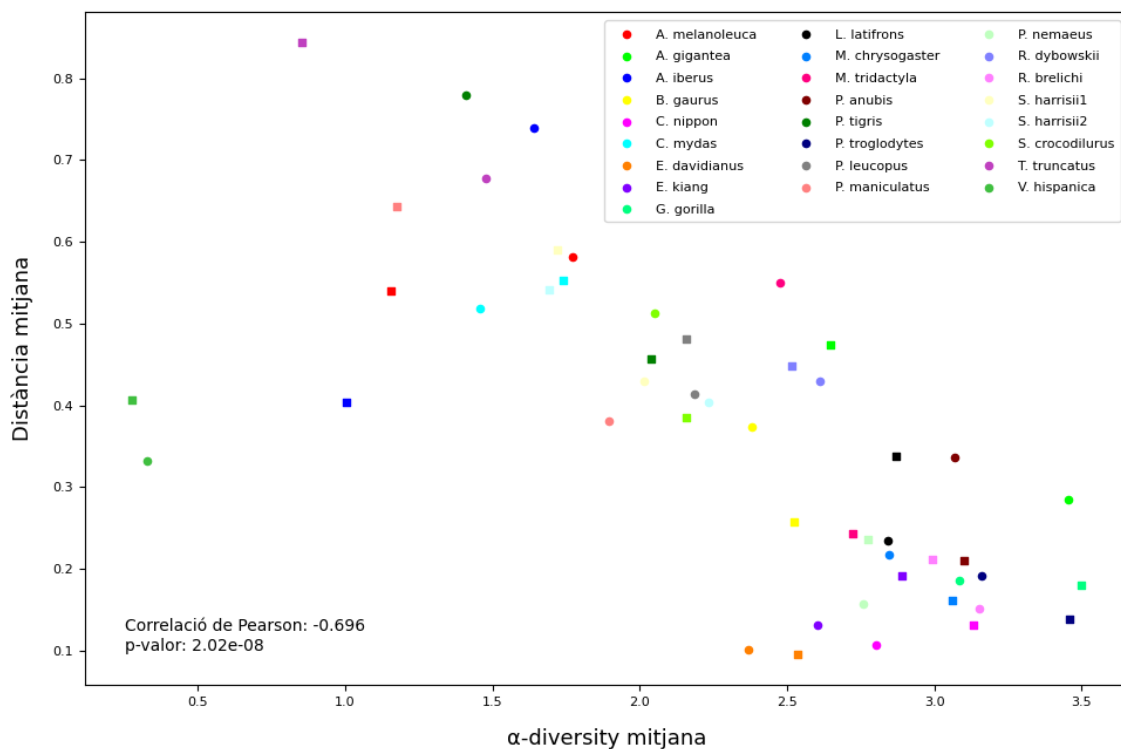


Fig 5. Correlació entre distàncies i diversitats per espècie de vertebrat. Gràfic de dispersió en el que es representa la correlació entre el valor mitjà de les distàncies euclidianes i el valor mitjà de les  $\alpha$ -diversitats del conjunt d'individus salvatges i del conjunt d'individus captius de cada espècie de vertebrat. P-valors (n.s.: p-valor > 0.05; \*:  $0.05 \geq$  p-valor > 0.01; \*\*:  $0.01 \geq$  p-valor > 0.001; \*\*\*: p-valor  $\leq$  0.001).

El p-valor ens indica que la correlació de Pearson és molt significativa. Per tant, existeix una correlació molt forta entre la similitud de dos individus qualsevol d'un tipus (salvatge o captiu) i la diversitat típica de la microbiota d'aquell grup d'individus per a tot el conjunt d'espècies de vertebrat del nostre estudi. El significat d'aquesta correlació és el següent. En aquells grups d'individus salvatges o captius d'una espècie de vertebrat en les que la microbiota tingui una

diversitat molt elevada, les microbiotes d'aquests individus seran molt semblants les unes de les altres. Els dos cassos en els que aquest fenomen és més acurat son les microbiotes dels individus captius del goril·la occidental (*Gorilla gorilla*) i les dels individus captius de ximpanzé (*Pan troglodytes*). És molt probable que el fet de que el goril·la occidental i el ximpanzé pertanyin al mateix infra-ordre que l'ésser humà, amb el que conviuen en condicions de captiveri, estigui relacionat amb el fet de que sigui en aquestes dues espècies en les que aquesta dinàmica es produeix de manera més clara. Per altra banda, en aquelles espècies de vertebrat en les que la microbiota tingui una diversitat molt baixa, els individus d'aquella espècie tindran una microbiota molt diferent els uns dels altres. Aquest és el cas, per exemple, dels individus salvatges del tigre (*Panthera tigris*) o dels individus captius del dofí nas d'ampolla comú (*Tursiops truncatus*). Una possible explicació per aquest fenomen podria ser que, quan més elevada sigui la diversitat de la microbiota, menys susceptible serà aquesta microbiota a una variació de la seva composició degut a canvis dietètics i/o ambientals.

### Estudi de *nestedness*

#### Selecció dels llindars

Com ja s'ha explicat, per realitzar un estudi de *nestedness* és necessari que les matrius utilitzades siguin matrius binàries. Per tant, el primer pas que cal fer és discretitzar les matrius d'abundàncies relatives en base a un valor llindar. Ara bé, el valor llindar idoni és aquell que maximitza el valor de *nestedness* de la matriu binària i, per tant, aquest valor no es pot saber amb certesa des d'un principi. És cert que, a partir de l'observació de representacions gràfiques com els histogrames o els *clustermats*, es pot intuir; però, encara i així, es millor comprovar-ho empíricament. Això significa que s'han de calcular diversos valors de *nestedness* utilitzant un conjunt de llindars. A amb l'excepció del valor  $10^{-1}$ , el conjunt de llindars seleccionats és el mateix que s'ha utilitzat per construir l'histograma de la matriu d'abundàncies relatives de gèneres bacterians en individus (Figura 1).

### Valors de *nestedness*

S'han calculat tots els possibles valors de *nestedness* de totes les matrius d'abundàncies relatives construïdes en estudis anteriors (Taula 4).

| Llindar          | Valor de <i>nestedness</i> |               |                       |              |              |                       |              |              |
|------------------|----------------------------|---------------|-----------------------|--------------|--------------|-----------------------|--------------|--------------|
|                  | Individus                  | Vertebrats    | <i>A. melanoleuca</i> |              |              | <i>P. maniculatus</i> |              |              |
|                  |                            |               | S                     | C            | S&C          | S                     | C            | S&C          |
| 10 <sup>-2</sup> | 0.2161                     | 0.1934        | 0.360                 | <b>0.682</b> | 0.366        | 0.660                 | 0.594        | 0.618        |
| 10 <sup>-3</sup> | 0.2895                     | 0.2679        | 0.456                 | 0.549        | <b>0.383</b> | <b>0.720</b>          | 0.660        | 0.686        |
| 10 <sup>-4</sup> | <b>0.3199</b>              | 0.3299        | 0.462                 | 0.476        | 0.356        | 0.718                 | <b>0.672</b> | <b>0.693</b> |
| 10 <sup>-5</sup> | 0.3197                     | 0.3598        | <b>0.463</b>          | 0.467        | 0.352        | 0.718                 | 0.672        | 0.693        |
| 10 <sup>-6</sup> | 0.3197                     | <b>0.3602</b> | 0.463                 | 0.467        | 0.352        | 0.718                 | 0.672        | 0.693        |
| 10 <sup>-7</sup> | 0.3197                     | 0.3602        | 0.463                 | 0.467        | 0.352        | 0.718                 | 0.672        | 0.693        |

Taula 4. Valors de *nestedness* de totes les matrius binàries resultants de discretitzar cada una de les matrius d'abundàncies relatives per cada un dels valors llindars.

En la taula 4, el valor de *nestedness* més gran que s'ha calculat per cada matriu d'abundàncies està marcat en negreta. D'aquesta manera, a través d'aquesta taula es pot seleccionar el valor llindar idoni per discretitzar cada una de les matrius d'abundàncies relatives de manera que el valor de *nestedness* calculat sigui màxim.

### Generació dels *clustermaps*

Per conèixer l'aspecte que adopta una matriu d'abundàncies relatives després de ser discretitzada amb cada un dels valors llindars i així triar-ne el més adequat, una de les eines utilitzades pot ser el *clustermap* de la matriu. Tot i que s'han generat els *clustermaps* de totes les matrius de les dues primeres columnes de la taula 4 (Ferrero, 2023), en aquesta secció només hem inclòs el *clustermap* de les matrius binàries obtingudes al discretitzar la matriu d'abundàncies relatives de gèneres bacterians en individus amb el valor llindar

$10^{-4}$  i la matriu d'abundàncies relatives de gèneres bacterians en espècies de vertebrat amb el valor llindar  $10^{-6}$ , que són els que permeten obtenir el valor de *nestedness* més gran per aquestes dues matrius (Figures 6 i 7).

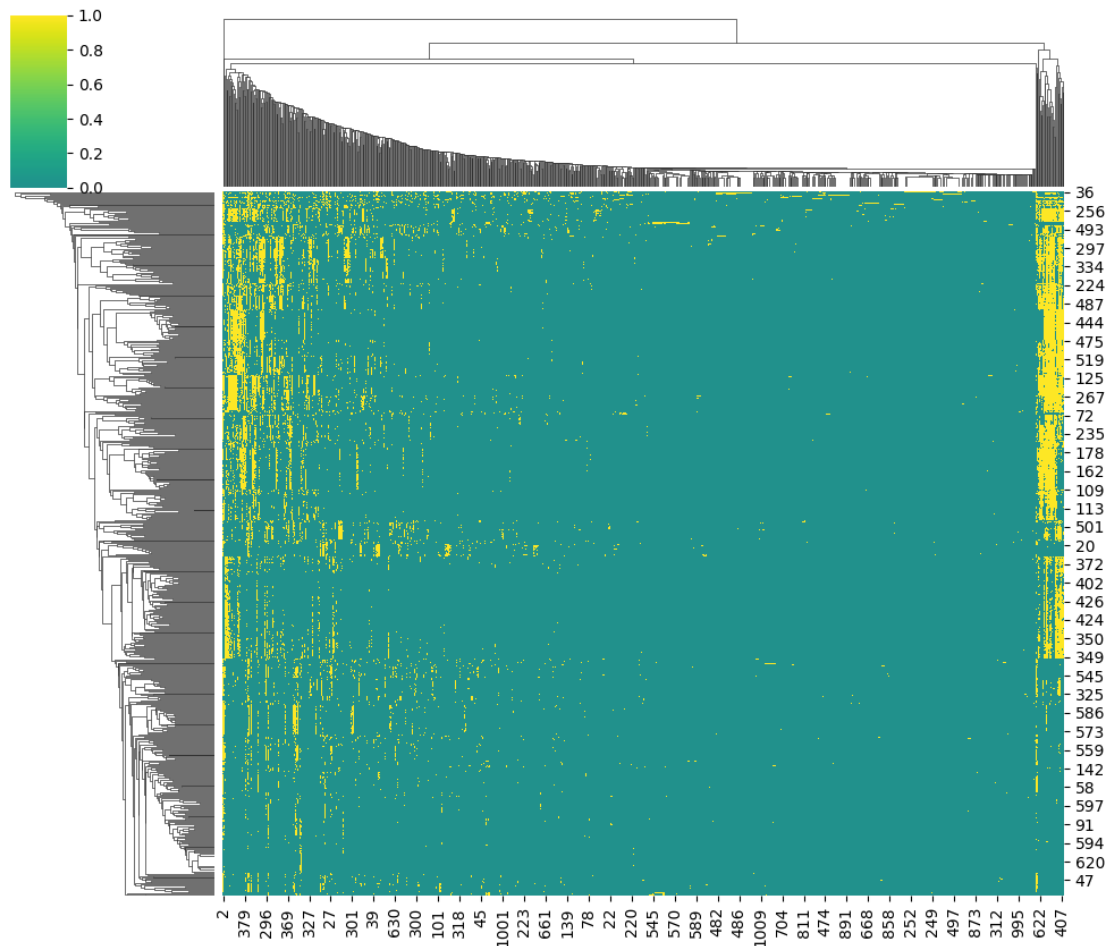


Fig 6. Clustermap matriu binaria d'abundàncies de gèneres bacterians en individus. Clustermap de la matriu binaria resultant de discretitzar la matriu d'abundàncies relatives de gèneres bacterians en individus amb el valor llindar  $10^{-4}$ .

#### Obtenció del p-valor

Per últim, per saber si els valors de *nestedness* seleccionats són o no són significatius, aquests s'han de comparar amb els de les respectives hipòtesis nul·les. Tot i que es va implementar el codi per portar a terme aquest últim pas i es va comprovar que funcionava correctament amb petites matrius de prova, el seu temps d'execució hagués estat extremadament alt per matrius d'aquestes característiques. El codi s'ha optimitzat diverses vegades i se li ha aplicat

paral·lelisme (Ferrero, 2023). Això ha permès reduir el temps d'execució de manera molt significativa. Encara i així, segueix sent de dies per la matriu més gran, massa elevat per a executar-lo en un ordinador portàtil. Caldria executar-lo de manera ininterrompuda en els ordinadors del laboratori de manera indefinida, cosa que, per desgracia, no ha estat possible.

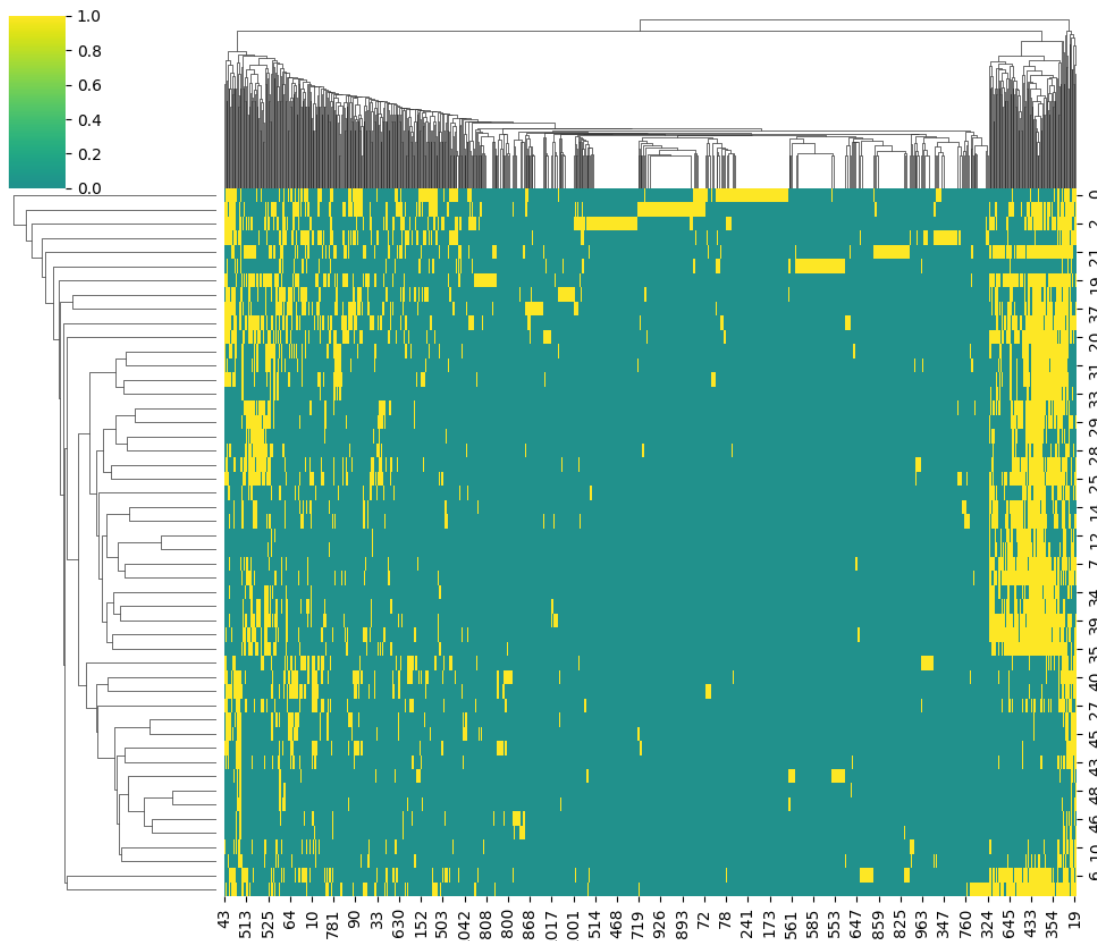


Fig 7. Clustermap matriu binària d'abundàncies de gèneres bacterians en espècies de vertebrat. Clustermap de la matriu binària resultant de discretitzar la matriu d'abundàncies relatives de gèneres bacterians en grups d'individus salvatges i captius de cada espècie de vertebrat amb el valor llindar  $10^{-6}$ .

En conseqüència, no es pot saber amb certesa si les matrius d'abundàncies relatives que hem construït segueixen un patró *nested* o no.

## Conclusions

A grans trets, els resultats de l'estudi han mostrat que la microbiota intestinal d'algunes espècies de vertebrat pateix variacions en la seva composició quan l'espècie es criada en condicions de captiveri.

En el cas de la diversitat, algunes microbiotes presenten diferències entre el grup d'individus salvatges i el grup d'individus captius que poden arribar a ser molt significatives. A més, aquestes diferències es donen en totes dues direccions. Es a dir, hi ha espècies en les que la diversitat de la seva microbiota augmenta en condicions de captiveri i d'altres en les que disminueix. Respecte als motius pels quals aquest fenomen es produeix en algunes espècies i en altres no, cap dels factors dietètics o ambientals avaluats ens ha proporcionat una explicació per a totes o la major part de les espècies. Això ens demostra que encara s'ha de fer molta recerca en el camp de la microbiota intestinal per arribar a comprendre el seu comportament en certes condicions.

En el cas de les similituds, es dona un situació molt semblant. Algunes microbiotes son molt semblants entre qualsevol dels individus de l'espècie, sigui aquest salvatge o captiu, mentre que d'altres son extraordinàriament diferents, fins al punt de que algunes s'assemblen més a les d'individus d'altres espècies que entre elles quan pertanyen a diferents grups. A més, s'ha constatat l'existència d'una correlació entre les similituds i les diversitats, que afecta d'igual manera als grups d'individus salvatges i als grups d'individus captius. Segons els resultats observats, les microbiotes que tinguin una alta diversitat seran molt més semblants entre elles que les que tinguin una baixa diversitat.

Per altra banda, tot i que s'ha pogut calcular el valor de *nestedness* de totes les matrius d'abundàncies relatives, l'estudi en qüestió no s'ha pogut completar. Per arribar a una conclusió sòlida hagués estat necessari comparar els valors calculats amb els de la hipòtesi nul·la i obtenir-ne el p-valor. No obstant, en base als valors de *nestedness* obtinguts per les randomitzacions generades per provar l'algorisme, queda patent que es tracta d'una projecte en el que pot ser interessant continuar investigant, ja que pensem que és probable que aquestes matrius tinguin un patró *nested*.

## Referències

- Alberdi, A., Martin Bideguren, G., & Aizpurua, O. (2021). *Diversity and compositional changes in the gut microbiota of wild and captive vertebrates: a meta-analysis*. Scientific reports, 11(1), 22660. <https://doi.org/10.1038/s41598-021-02015-6>
- Bastolla, U., Fortuna, M. A., Pascual-García, A., Ferrera, A., Luque, B., & Bascompte, J. (2009). *The architecture of mutualistic networks minimizes competition and increases biodiversity*. Nature, 458(7241), 1018–1020. <https://doi.org/10.1038/nature07950>
- Bender, H. S. (2010). *Devil Facial Tumour Disease (DFTD): Using Genetics and Genomics to Investigate Infectious Disease in an Endangered Marsupial*. Marsupial Genetics and Genomics, 499–515. [https://doi.org/10.1007/978-90-481-9023-2\\_23](https://doi.org/10.1007/978-90-481-9023-2_23)
- Breitenmoser, U. (2010). *Panthera tigris (Mediterranean assessment)*. *The IUCN Red List of Threatened Species*. e.T15955A5330476. <https://dx.doi.org/10.2305/IUCN.UK.2010-2.RLTS.T15955A5330476.en>
- Cassola, F. (2016). *Peromyscus leucopus (errata version published in 2017)*. *The IUCN Red List of Threatened Species*. e.T16669A115136270. <https://dx.doi.org/10.2305/IUCN.UK.2016-3.RLTS.T16669A22360579.en>
- Cassola, F. (2016). *Peromyscus maniculatus*. *The IUCN Red List of Threatened Species*. e.T16672A22360898. <https://dx.doi.org/10.2305/IUCN.UK.2016-2.RLTS.T16672A22360898.en>
- Cheng, Y., Fox, S., Pemberton, D., Hogg, C., Papenfuss, A. T., & Belov, K. (2015). *The Tasmanian devil microbiome-implications for conservation and management*. Microbiome, 3, 76. <https://doi.org/10.1186/s40168-015-0143-0>
- Cobo-López, S., Gupta, V. K., Sung, J., Guimerà, R., & Sales-Pardo, M. (2022). *Stochastic block models reveal a robust nested pattern in healthy human gut microbiomes*. PNAS nexus, 1(3), pgac055. <https://doi.org/10.1093/pnasnexus/pgac055>

- Costea, P. I., Hildebrand, F., Arumugam, M., Bäckhed, F., Blaser, M. J., Bushman, F. D., de Vos, W. M., Ehrlich, S. D., Fraser, C. M., Hattori, M., Huttenhower, C., Jeffery, I. B., Knights, D., Lewis, J. D., Ley, R. E., Ochman, H., O'Toole, P. W., Quince, C., Relman, D. A., Shanahan, F., ... Bork, P. (2018). *Enterotypes in the landscape of gut microbial community composition*. *Nature microbiology*, 3(1), 8–16. <https://doi.org/10.1038/s41564-017-0072-8>
- Coudrat, C. N. Z., Quyet, L. K., Duc, H., Phiaphalath, P., Rawson, B. M., Nadler, T., Ulibarri, L. & Duckworth, J. W. (2020). *Pygathrix nemaeus*. *The IUCN Red List of Threatened Species*. e.T39826A17941247. <https://dx.doi.org/10.2305/IUCN.UK.2020-2.RLTS.T39826A17941247.en>
- Crivelli, A. J. (2006). *Aphanius iberus*. *The IUCN Red List of Threatened Species*. e.T1846A8299534. <https://dx.doi.org/10.2305/IUCN.UK.2006.RLTS.T1846A8299534.en>
- Duckworth, J. W., Sankar, K., Williams, A. C., Samba Kumar, N. & Timmins, R. J. (2016). *Bos gaurus*. *The IUCN Red List of Threatened Species*. e.T2891A46363646. <https://dx.doi.org/10.2305/IUCN.UK.2016-2.RLTS.T2891A46363646.en>
- Dutoit, S. H. C. (2012). *Graphical exploratory data analysis*. Springer. ISBN 9781461293712.
- Ferrero, A. (2023). <https://github.com/AreyFerreroRamos/TFGBiotecnologia>
- Gibson, K. M., Nguyen, B. N., Neumann, L. M., Miller, M., Buss, P., Daniels, S., Ahn, M. J., Crandall, K. A., & Pukazhenthi, B. (2019). *Gut microbiome differences between wild and captive black rhinoceros - implications for rhino health*. *Scientific reports*, 9(1), 7570. <https://doi.org/10.1038/s41598-019-43875-3>
- Gilbert, J. A., & Lynch, S. V. (2019). *Community ecology as a framework for human microbiome research*. *Nature medicine*, 25(6), 884–889. <https://doi.org/10.1038/s41591-019-0464-9>
- Grice, E. A., Kong, H. H., Conlan, S., Deming, C. B., Davis, J., Young, A. C., NISC Comparative Sequencing Program, Bouffard, G. G., Blakesley, R. W., Murray, P. R., Green, E. D., Turner, M. L., & Segre, J. A. (2009). *Topographical*



- and temporal diversity of the human skin microbiome. *Science* (New York, N.Y.), 324(5931), 1190–1192. <https://doi.org/10.1126/science.1171700>
- Grubb, P. (1971). *The growth, ecology, and population structure of giant tortoises on Aldabra*. *Philosophical Transactions of the Royal Society of London B*, 260, 327–372. <https://doi.org/10.1098/rstb.1971.0018>
- Harris, R. (2016). *Moschus chrysogaster*. *The IUCN Red List of Threatened Species*. e.T13895A61977139. <https://dx.doi.org/10.2305/IUCN.UK.2016-1.RLTS.T13895A61977139.en>
- Harris, R. B. (2015). *Cervus nippon*. *The IUCN Red List of Threatened Species*. e.T41788A22155877. <https://dx.doi.org/10.2305/IUCN.UK.2015-2.RLTS.T41788A22155877.en>
- Hawkins, C. E., McCallum, H., Mooney, N., Jones, M. & Holdsworth, M. (2008). *Sarcophilus harrisii*. *The IUCN Red List of Threatened Species*. e.T40540A10331066. <https://dx.doi.org/10.2305/IUCN.UK.2008.RLTS.T40540A10331066.en>
- Humle, T., Maisels, F., Oates, J. F., Plumptre, A. & Williamson, E. A. (2016). *Pan troglodytes* (errata version published in 2018). *The IUCN Red List of Threatened Species*. e.T15933A129038584. <https://dx.doi.org/10.2305/IUCN.UK.2016-2.RLTS.T15933A17964454.en>
- Huttenhower, C et al. Human Microbiome Project Consortium. (2012). *Structure, function and diversity of the healthy human microbiome*. *Nature*, 486(7402), 207-214. <https://doi.org/10.1038/nature11234>
- IUCN Red List of threatened species. (2022). <https://www.iucnredlist.org/>
- Jiang, Z. & Harris, R. B. (2016). *Elaphurus davidianus*. *The IUCN Red List of Threatened Species*. e.T7121A22159785. <https://dx.doi.org/10.2305/IUCN.UK.2016-2.RLTS.T7121A22159785.en>
- Kuzmin, S., Ishchenko, V., Maslova, I., Ananjeva, N., Orlov, N., Matsui, M., Feng, X., & Kaneko, Y. (2004). *Rana dybowskii*. *The IUCN Red List of Threatened Species*. e.T58589A11792510. <https://dx.doi.org/10.2305/IUCN.UK.2004.RLTS.T58589A11792510.en>

Lande, R. (1996). *Statistics and Partitioning of Species Diversity, and Similarity among Multiple Communities*. *Oikos*, 76(1), 5–13.

<https://doi.org/10.2307/3545743>

Levin, D., Raab, N., Pinto, Y., Rothschild, D., Zanir, G., Godneva, A., Mellul, N., Futorian, D., Gal, D., Leviatan, S., Zeevi, D., Bachelet, I., & Segal, E. (2021). *Diversity and functional landscapes in the microbiota of animals in the wild*. *Science* (New York, N.Y.), 372(6539), eabb5352.

<https://doi.org/10.1126/science.abb5352>

Long, Y., Li, B., Zhou, J., Ren, B. & Garber, A. (2022). *Rhinopithecus brelichi*. *The IUCN Red List of Threatened Species*. e.T19595A215621115.

<https://dx.doi.org/10.2305/IUCN.UK.2022-1.RLTS.T19595A215621115.en>

Lynch, J. B., & Hsiao, E. Y. (2019). *Microbiomes as sources of emergent host phenotypes*. *Science* (New York, N.Y.), 365(6460), 1405–1409.

<https://doi.org/10.1126/science.aay0240>

Maisels, F., Bergl, R. A. & Williamson, E. A. (2018). *Gorilla gorilla* (amended version of 2016 assessment). *The IUCN Red List of Threatened Species*.

e.T9404A136250858. <https://dx.doi.org/10.2305/IUCN.UK.2018-2.RLTS.T9404A136250858.en>

Mariani, M. S., Ren, Z. M., Bascompte, J., & Tessone, C. J. (2019). *Nestedness in complex networks: observation, emergence, and implications*. *Physics and Society*, 813, 1–90. <https://doi.org/10.1016/j.physrep.2019.04.001>

McFall-Ngai, M., Hadfield, M. G., Bosch, T. C., Carey, H. V., Domazet-Lošo, T., Douglas, A. E., Dubilier, N., Eberl, G., Fukami, T., Gilbert, S. F., Hentschel, U., King, N., Kjelleberg, S., Knoll, A. H., Kremer, N., Mazmanian, S. K., Metcalf, J. L., Nealson, K., Pierce, N. E., Rawls, J. F.,... Wernegreen, J. J. (2013). *Animals in a bacterial world, a new imperative for the life sciences*. *Proceedings of the National Academy of Sciences of the United States of America*, 110(9), 3229–3236. <https://doi.org/10.1073/pnas.1218525110>

McKenzie, V. J., Song, S. J., Delsuc, F., Prest, T. L., Oliverio, A. M., Korpita, T. M., Alexiev, A., Amato, K. R., Metcalf, J. L., Kowalewski, M., Avenant, N. L., Link, A., Di Fiore, A., Seguin-Orlando, A., Feh, C., Orlando, L., Mendelson, J.

R., Sanders, J., & Knight, R. (2017). *The Effects of Captivity on the Mammalian Gut Microbiome*. Integrative and comparative biology, 57(4), 690–704.

<https://doi.org/10.1093/icb/ix090>

Methé, B. A. et al. Human Microbiome Project Consortium. (2012). *A framework for human microbiome research*. Nature, 486(7402), 215–221.

<https://doi.org/10.1038/nature11209>

Miranda, F., Bertassoni, A. & Abba, A. M. (2014). *Myrmecophaga tridactyla*. *The IUCN Red List of Threatened Species*. e.T14224A47441961.

<https://dx.doi.org/10.2305/IUCN.UK.2014-1.RLTS.T14224A47441961.en>

Nguyen, T. Q., Hamilton, P. & Ziegler, T. (2022). *Shinisaurus crocodilurus* (amended version of 2014 assessment). *The IUCN Red List of Threatened Species*. e.T57287221A217811129.

<https://dx.doi.org/10.2305/IUCN.UK.2022-1.RLTS.T57287221A217811129.en>

Patterson, B. D., & Atmar, W. (1986). *Nested subsets and the structure of insular mammalian faunas and archipelagos*. Biological Journal of the Linnean Society, 28(1-2), 65–82. <https://doi.org/10.1111/j.1095-8312.1986.tb01749.x>

Poulin, R., & Valtonen, E. T. (2001). *Nested assemblages resulting from host size variation: the case of endoparasite communities in fish hosts*. International journal for parasitology, 31(11), 1194–1204. [https://doi.org/10.1016/s0020-7519\(01\)00262-4](https://doi.org/10.1016/s0020-7519(01)00262-4)

Sadava, D., Heller, H. C., Orians, G. H., Purves, W. H., & Hillis, D. M. (2009). *Vida. La ciència de la biologia* (8a ed.). Panamericana. ISBN 9789500682695

Seminoff, J. A. (Southwest Fisheries Science Center, U.S.). (2004). *Chelonia mydas*. *The IUCN Red List of Threatened Species*.

e.T4615A11037468. <https://dx.doi.org/10.2305/IUCN.UK.2004.RLTS.T4615A11037468.en>

Shah, N., St. Louis, A. & Qureshi, Q. (2015). *Equus kiang*. *The IUCN Red List of Threatened Species*. e.T7953A45171635.

<https://dx.doi.org/10.2305/IUCN.UK.2015-4.RLTS.T7953A45171635.en>

Singh, R. K., Chang, H. W., Yan, D., Lee, K. M., Ucmak, D., Wong, K., Abrouk, M., Farahnik, B., Nakamura, M., Zhu, T. H., Bhutani, T., & Liao, W. (2017). *Influence of diet on the gut microbiome and implications for human health*. *Journal of translational medicine*, 15(1), 73.

<https://doi.org/10.1186/s12967-017-1175-y>

Swaigood, R., Wang, D. & Wei, F. (2016). *Ailuropoda melanoleuca* (errata version published in 2017). *The IUCN Red List of Threatened Species*. e.

T712A121745669. <https://dx.doi.org/10.2305/IUCN.UK.2016-2.RLTS.T712A45033386.en>

Teplow, D. B. (2020). *The Microbiome in Health and Disease*. Academic Press. p. 397. ISBN 9780128200018.

Wallis, J. (2020). *Papio anubis*. *The IUCN Red List of Threatened Species*.

e.T40647A17953200. <https://dx.doi.org/10.2305/IUCN.UK.2020-2.RLTS.T40647A17953200.en>

Warburton, E. M., Van Der Mescht, L., Khokhlova, I. S., Krasnov, B. R., & Vonhof, M. J. (2018). *Nestedness in assemblages of helminth parasites of bats: a function of geography, environment, or host nestedness?*. *Parasitology research*, 117(5), 1621–1630. <https://doi.org/10.1007/s00436-018-5844-4>

Wells, R. S., Natoli, A. & Braulik, G. (2019). *Tursiops truncatus* (errata version published in 2019). *The IUCN Red List of Threatened Species*.

e.T22563A156932432. <https://dx.doi.org/10.2305/IUCN.UK.2019-1.RLTS.T22563A156932432.en>

Whittaker, R. H. (1960). *Vegetation of the Siskiyou Mountains, Oregon and California*. *Ecological Monographs*, 30, 279–338.

<https://doi.org/10.2307/1943563>

Whittaker, R. H. (1972). *Evolution and Measurement of Species Diversity*.

*Taxon*, 21, 213-251. <https://doi.org/10.2307/1218190>

Woinarski, J. & Burbidge, A. A. (2016). *Lasiorninus latifrons*. *The IUCN Red List of Threatened Species*. e.T40555A21959203.

<https://dx.doi.org/10.2305/IUCN.UK.2016-1.RLTS.T40555A21959203.en>

<https://seaborn.pydata.org/generated/seaborn.boxplot.html>

## Autoavaluació

El motiu pel qual vaig decidir realitzar aquest projecte és perquè la temàtica, emmarcada dins del context de la ciència de dades, semblava combinar de forma molt encertada conceptes biotecnològics amb eines informàtiques. La meua principal expectativa era resoldre els dubtes que tenia sobre si aquest camp és del meu gust o no.

Durant el desenvolupament del treball, he hagut d'aprendre a cercar informació en articles científics. En ocasions, trobar aquesta informació ha estat un repte. A més, la lectura de certes seccions dels articles o d'articles sencers (com és el cas d'aquells en els que he basat el projecte) és complexa i m'ha forçat a una millora del meu nivell d'anglès. El domini en el que probablement hagi adquirit més coneixements ha estat el de les microbiotes intestinals. Inicialment em va sorprendre les poques teories sòlides a les que s'ha arribat tenint en compte la quantitat de recerca que s'ha fet en aquest camp, però després de no poder extraure gairebé cap conclusió concreta dels meus resultats, he entès la dificultat de fer recerca en aquest àmbit i, probablement, en molts altres. Això també demostra com de sacrificat és el món de la ciència.

També he millorat els meus coneixements del llenguatge de programació Python, en el que he implementat tot el codi per processar les dades d'entrada i obtenir els resultats. Altres eines informàtiques en les que he hagut de demostrar un cert coneixement han estat el llenguatge de programació *bash*, l'entorn de desenvolupament integrat PyCharm, el software de control de versions Git o la plataforma GitHub per allotjar projectes.

A grans trets, puc dir que el meu dubte inicial està resolt i que la resposta és afirmativa. Tinc algunes desavinences amb la dificultat de comprendre el contingut d'alguns articles científics i la necessitat de citar informació molt obvia, però, els coneixements que he adquirit durant la realització d'aquest treball han estat molt interessants i les eines que he après a utilitzar molt útils, així que el balanç es positiu. Per tant, el món de la ciència de dades és un món en el que em plantejo seguir treballant si en tinc la oportunitat.