

Problème Pratique de Statistique – 62

Coévolution hôtes-parasites et couplage de phylogénies

Sabine Patot (M2 – EEB – 2005) nous invite à découvrir l'article de Legendre, P., Y. Desdevises, and E. Bazin. 2002. A statistical test for host–parasite coevolution. *Systematic Biology* **51**:217–234.

Son titre est une illustration efficace :



Co-évolution hôte-parasite



1

Sa présentation est claire :

Dans le milieu naturel, les espèces distinctes interagissent continuellement. Certaines interactions privilégiées sont observées dans les générations successives; elles sont qualifiées de relations durables. Il en existe trois types: le commensalisme, le mutualisme et le parasitisme.

*L'équipe à laquelle j'appartiens s'intéresse aux **relations hôte-parasite** avec l'exemple de la bactérie *Wolbachia* (thèse de F. Vavre, 2000; thèse en cours de L. Gavotte).*

L'association hôte-parasite.

Dans ce type d'association étroite qu'est l'association hôte-parasite, les hôtes représentent une partie essentielle de l'environnement du parasite. Aussi, l'évolution de l'hôte peut influencer de façon importante la phylogénie des parasites. C'est le concept de la co-évolution où de nouvelles associations hôte-parasite se forment au cours du temps.

Dans le présent exposé, le terme "co-évolution" est utilisé pour décrire les relations entre la phylogénie des parasites et celle de leurs hôtes. Il s'agit de mesurer à quel point ces deux arbres sont congruents. La congruence parfaite des deux arbres indique une

¹ <http://www.americazoo.com/goto/index/mammals/animals/161y.jpg>

http://www.alyon.org/InfosTechniques/biomedical/biologie/animale/insectes/poux--ordre_des_pthirapteres--pou_de_tete.jpg

co-spéciation des hôtes et de leurs parasites. En revanche, une absence totale de congruence signifie que les associations observées se seraient établies "par hasard" lors de l'évolution.

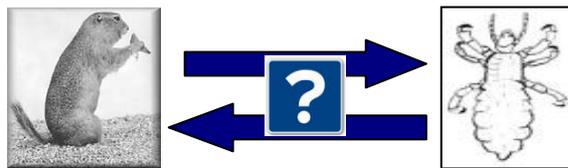
*Différentes méthodes permettent l'étude de la co-évolution pour le système hôte-parasite mais elles ont été souvent jugées trop peu rigoureuses et lourdes à mettre en place. Il a été proposé une **nouvelle méthode de l'étude de la co-évolution hôte-parasite** par Legendre et al. (la méthode Parafit).*

L'hypothèse nulle (H_0) est que l'évolution des hôtes est indépendante de celle des parasites.

Jeux de données.

Le modèle biologique utilisé ici pour l'étude de la co-évolution est constitué des deux partenaires, le poux (le parasite) et le gaufre (l'hôte), rongeur fouisseur du continent nord-américain.

Dans cette association, les poux effectuent l'ensemble de leurs cycles de vie chez les gaufres. De plus, les différentes espèces de gaufres n'interagissent que rarement. Chaque espèce de poux parasite théoriquement une seule espèce de gaufre. Ces caractéristiques font que l'évolution des poux et des gaufres s'est faite de façon similaire.



*On dispose du **jeu de données** formé de trois matrices m_A , m_B et m_C . m_B et m_C sont des matrices de distances phylogénétiques entre hôtes (B) et entre parasites (C) obtenues à partir des phylogénies des organismes étudiés. m_A est une matrice binaire représentant les liens observés dans la nature entre les hôtes et leurs parasites; 1 indique une association observée entre les deux espèces, 0 indique une absence de relation.*

Principe.

Les deux matrices de distances B et C sont transformées en tableaux par une analyse en coordonnées principales. Une combinaison des matrices A , B , C permet d'en obtenir une quatrième $D = C.A'.B$.

*Les **mesures de congruence** entre les matrices des hôtes et des parasites pondérées par les liens hôte-parasite individuels sont alors extraites de cette nouvelle matrice (D). La congruence est ensuite testée par permutation où les parasites sont distribués par hasard sur les hôtes. La valeur de l'association est testée contre la distribution générée aléatoirement par le test de permutation.*

*Il est également possible de tester la **contribution de chaque lien hôte-parasite** à la congruence entre les matrices de distance.*

Il va ainsi être testée l'hypothèse nulle à savoir s'il existe une co-évolution entre le parasite et son hôte.

Les matrices sont dans le manuel d'utilisation du logiciel parafit écrit par P. Legendre pour mettre en œuvre cette méthode, logiciel présenté à :

<http://www.bio.umontreal.ca/casgrain/en/labo/parafit.html>

et diffusé à :

<http://www.bio.umontreal.ca/casgrain/en/telecharger/index.html#ParaFit>

Reprenons le problème posé par les auteurs, en amont de ces choix techniques, à partir de leur figure 8 qui contient l'intégralité de l'information nécessaire :

2002

LEGENDTRE ET AL.—HOST-PARASITE COEVOLUTION

231

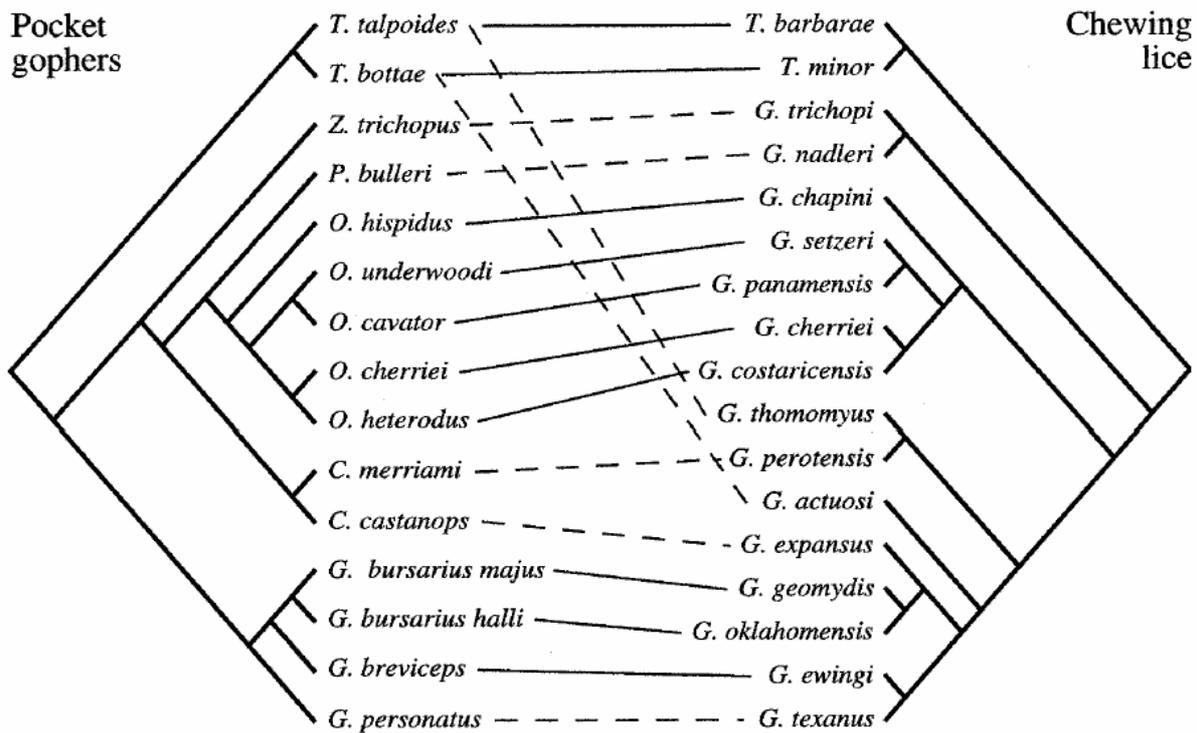


FIGURE 8. Pocket gophers and chewing lice phylogenetic trees and H-P links. Significant H-P links are represented by full lines, nonsignificant links by dashed lines.

A gauche les hôtes (une phylogénie et une liste de noms d'espèces), à droite les parasites (une phylogénie et une liste de noms d'espèces), au centre un lien sous forme de l'indication pour chaque parasite de l'unique espèce d'hôtes qui le supporte. La liste **pps062** contient exactement cette information.

pps062

\$hos.names

H1	H2	H3
"T. talpoides"	"T. bottae"	"Z. trichopus"
H4	H5	H6
"P. bulleri"	"O. hispidus"	"O. underwoodi"
H7	H8	H9
"O. cavator"	"O. cherriei"	"O. heterodus"
H10	H11	H12
"C. merriami"	"C. castanops"	"G. bursarius majus"
H13	H14	H15
"G. bursarius halli"	"G. breviceps"	"G. personatus"

\$hos.tre

```
[1] "((H1:1,H2:1)a:13,((H3:8,((H4:5,(H5:4,((H6:1,H7:1)b:2,(H8:1,H9:1)c:2)h:1)i:1)j:2,"
[2] "(H10:1,H11:1)d:6)k:1)m:4,((H12:1,H13:1)e:1,H14:2)f:1,H15:3)g:9)n:2)root:0;"
```

```

$par.names
      P1          P2          P3          P4
  "T. barbara"    "T. minor"    "G. trichopi"    "G. nadleri"
      P5          P6          P7          P8
  "G. chapini"    "G. setzeri"    "G. panamensis"    "G. cherriei"
      P9          P10         P11         P12
"G. costaricensis"  "G. thomomyus"    "G. perotensis"    "G. actuosi"
      P13         P14         P15         P16
  "G. expansus"    "G. geomydis"    "G. oklahomensis"    "G. ewingi"
      P17
  "G. texanus"

$par.tre
[1] "((P1:1,P2:1)a:13,((P3:1,P4:1)c:11,((P5:4,((P6:1,P7:1)h:2,(P8:1,P9:1)i:2)g:1)e:8,"
[2]
"((P10:1,P11:1)j:5,(P12:4,((P13:2,(P14:1,P15:1)x:1)m:2,(P16:1,P17:1)n:3)l:1)k:2)f:5)d:2)b
:2)root;"

$link
      P1      P2      P3      P4      P5      P6      P7      P8      P9      P10     P11     P12     P13
  "H1"  "H2"  "H3"  "H4"  "H5"  "H6"  "H7"  "H8"  "H9"  "H1"  "H10"  "H2"  "H11"
      P14     P15     P16     P17
  "H12" "H13" "H14" "H15"

```

On peut évidemment transformer les phylogénies en distances et les distances en tableaux, mais peut-on manipuler directement les phylogénies comme classe de données ? Un beau problème de description statistique des structures biologiques.