

Table des matières

Quelques statistiques d'arbres phylogénétiques incongrus 3

 Deux arbres phylogénétiques incongrus d'espèces de crocodiles. 3

 Signification statistiques de 2 arbres phylogéniques incongrus. 4

 Détails Mathématiques 5

Références 6

Titre original : **29+ Evidences for Macroevolution - Some Statistics of Incongruent Phylogenetic Trees**

Auteur : [Douglas Theobald, Ph.D.](#)

Copyright © 1999-2012 - Version 2.89

Permission is granted to copy and print these pages in total for non-profit personal, educational, research, or critical purposes.

La page originale est disponible à l'adresse : <http://www.talkorigins.org/faqs/comdesc/incongruent.html>

Navigation :

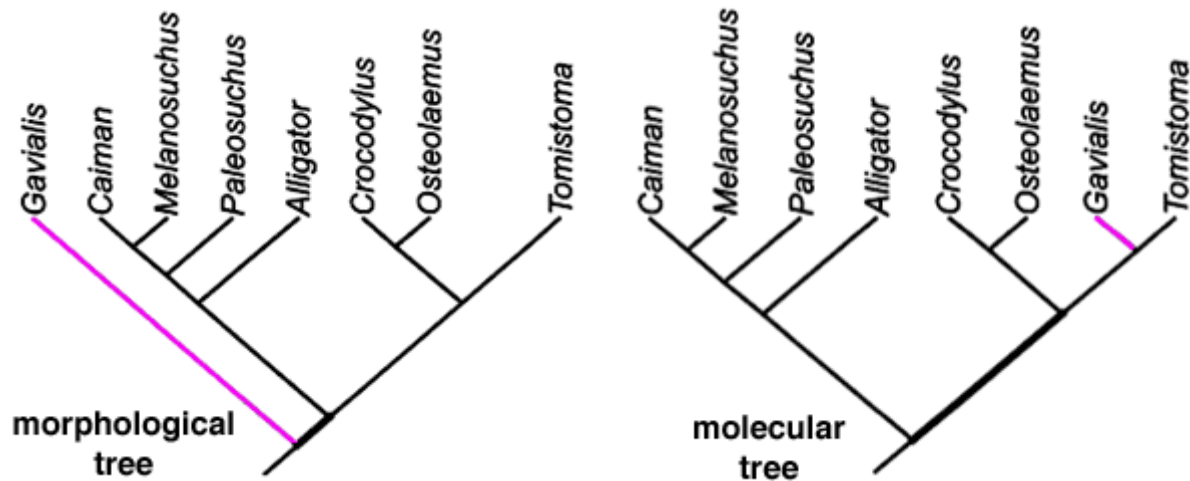
- [Introduction](#)
 - [Preuve scientifique et méthode scientifique](#)
 - [Glossaire](#)
- [Introduction à la phylogénétique](#)
- [Partie 1. Un arbre phylogénétique unique et historique](#)
- [Partie 2. Histoire passée](#)
- [Partie 3. Opportunisme évolutionniste](#)
- [Partie 4. Preuve moléculaire](#)
- [Partie 5. Changement](#)
- [Remarques finales](#)

Quelques statistiques d'arbres phylogénétiques incongrus

Le tableau ci-dessous fournit des valeurs pour la signification statistique d'une correspondance entre deux arbres phylogénétiques incongrus, rapportées en tant que valeurs P (P -values). Ces P -values donnent la probabilité que deux arbres à racines, avec un nombre donné (ou moins) de branches non concordantes, correspondent par hasard.

Le nombre de branches incongrues est déterminé par rapport au sous-arbre d'accord maximum (maximum agreement subtree, MAST) entre deux arbres. Un MAST est le sous-arbre "noyau" commun entre deux arbres. Le nombre de branches incongrues est égal au nombre minimum de branches qui doivent être élaguées de l'un des arbres réels pour obtenir le MAST. Un exemple tiré de l'analyse des espèces de crocodiles par John Harshman est donné dans la figure ci-dessous (Harshman et al. 2003).

Deux arbres phylogénétiques incongrus d'espèces de crocodiles.



Deux phylogénies incongrues de crocodiles. L'arbre à gauche est basé sur des données morphologiques; l'arbre à droite sur la séquence moléculaire du proto-oncogène c-myc (Harshman et al. 2003). Le MAST commun est représenté en noir. Selon la métrique de distance décrite ci-dessus, la distance entre les deux arbres est une branche, en raison de la branche de *Gavialis* égarée indiquée en magenta. La signification de la correspondance entre ces deux phylogénies incongruentes est $P \leq 0,00077$. De plus, Harshman et al. ont effectué une analyse phylogénétique indépendante avec les gènes mitochondriaux, ce qui a donné le même arbre que les données du proto-oncogène c-myc. La signification globale pour ces trois arbres indépendants est $P \leq 7,4 \times 10^{-8}$.

Dans le tableau ci-dessous, les lignes répertorient les valeurs permettant de comparer deux arbres comportant un nombre croissant de taxons. Les colonnes énumèrent la signification pour un nombre donné de différences entre les deux arbres. L'incongruence de "1 adjacent" fait référence au cas où une branche diverge par un seul nœud adjacent (c'est-à-dire que deux branches adjacentes sont échangées par rapport à l'autre arbre). Les colonnes 1 à 10 restantes font référence au cas où x branches ou moins sont mal placées n'importe où dans l'arbre. Une signification statistique élevée ($P < 0,01$ ou une confiance supérieure à 99%) est indiquée en bleu clair. La signification statistique ($p < 0,05$ ou plus de 95% de confiance) est indiquée en rose. Les valeurs équivoques ($0,05 < P < 0,50$) sont indiquées en blanc. Les valeurs très insignifiantes ($P > 0,50$) sont indiquées en rouge et les valeurs impossibles en noir.

Signification statistiques de 2 arbres phylogéniques incongrus.

Statistical Significance of Two Incongruent Phylogenetic Trees

	Maximum <i>P</i> -value for two trees incongruent by given number of branches:											
Number of taxa	exactly congruent	1 adjacent	1	2	3	4	5	6	7	8	9	10
4	0.067	0.20	1.00	1.00								
5	0.0095	0.038	0.28	1.00	1.00							
6	0.0011	0.0052	0.050	0.97	1.00	1.00						
7	9.6 × 10 ⁻⁵	5.8 × 10 ⁻⁴	0.0067	0.20	1.00	1.00	1.00					
8	7.4 × 10 ⁻⁶	5.2 × 10 ⁻⁵	6.8 × 10 ⁻⁴	0.030	0.53	1.00	1.00	1.00				
9	4.9 × 10 ⁻⁷	3.9 × 10 ⁻⁶	6.2 × 10 ⁻⁵	0.0035	0.089	1.00	1.00	1.00	1.00			
10	2.9 × 10 ⁻⁸	2.6 × 10 ⁻⁷	4.6 × 10 ⁻⁶	3.3 × 10 ⁻⁴	0.012	0.22	1.00	1.00	1.00	1.00		
11	1.5 × 10 ⁻⁹	1.5 × 10 ⁻⁸	3.0 × 10 ⁻⁷	2.7 × 10 ⁻⁵	0.0012	0.032	0.49	1.00	1.00	1.00	1.00	
12	7.2 × 10 ⁻¹¹	8.0 × 10 ⁻¹⁰	1.8 × 10 ⁻⁸	1.9 × 10 ⁻⁶	1.1 × 10 ⁻⁴	0.0037	0.076	0.98	1.00	1.00	1.00	1.00
13	3.1 × 10 ⁻¹²	3.8 × 10 ⁻¹¹	9.1 × 10 ⁻¹⁰	1.2 × 10 ⁻⁷	8.3 × 10 ⁻⁶	3.5 × 10 ⁻⁴	0.0095	0.17	1.00	1.00	1.00	1.00
14	1.2 × 10 ⁻¹³	1.6 × 10 ⁻¹²	4.3 × 10 ⁻¹¹	6.6 × 10 ⁻⁹	5.6 × 10 ⁻⁷	2.9 × 10 ⁻⁵	9.9 × 10 ⁻⁴	0.022	0.33	1.00	1.00	1.00
15	4.6 × 10 ⁻¹⁵	6.6 × 10 ⁻¹⁴	1.8 × 10 ⁻¹²	3.3 × 10 ⁻¹⁰	3.3 × 10 ⁻⁸	2.1 × 10 ⁻⁶	8.7 × 10 ⁻⁵	0.0025	0.048	0.62	1.00	1.00
16	1.6 × 10 ⁻¹⁶	2.4 × 10 ⁻¹⁵	5.6 × 10 ⁻¹⁴	1.5 × 10 ⁻¹¹	1.8 × 10 ⁻⁹	1.3 × 10 ⁻⁷	6.7 × 10 ⁻⁶	2.3 × 10 ⁻⁴	0.0056	0.095	1.00	1.00
17	5.2 × 10 ⁻¹⁸	8.3 × 10 ⁻¹⁷	2.1 × 10 ⁻¹⁵	6.4 × 10 ⁻¹³	8.6 × 10 ⁻¹¹	7.5 × 10 ⁻⁹	4.5 × 10 ⁻⁷	1.9 × 10 ⁻⁵	5.6 × 10 ⁻⁴	0.012	0.18	1.00
18	1.5 × 10 ⁻¹⁹	2.7 × 10 ⁻¹⁸	7.4 × 10 ⁻¹⁷	2.5 × 10 ⁻¹⁴	3.8 × 10 ⁻¹²	3.9 × 10 ⁻¹⁰	2.7 × 10 ⁻⁸	1.4 × 10 ⁻⁶	4.9 × 10 ⁻⁵	0.0013	0.024	0.32
19	4.5 × 10 ⁻²¹	8.1 × 10 ⁻²⁰	2.3 × 10 ⁻¹⁸	8.9 × 10 ⁻¹⁶	1.6 × 10 ⁻¹³	1.8 × 10 ⁻¹¹	1.5 × 10 ⁻⁹	8.6 × 10 ⁻⁸	3.7 × 10 ⁻⁶	1.2 × 10 ⁻⁴	0.0027	0.046
20	1.2 × 10 ⁻²²	2.3 × 10 ⁻²¹	7.3 × 10 ⁻²⁰	3.0 × 10 ⁻¹⁷	5.9 × 10 ⁻¹⁵	7.8 × 10 ⁻¹³	7.3 × 10 ⁻¹¹	4.9 × 10 ⁻⁹	2.5 × 10 ⁻⁷	9.2 × 10 ⁻⁶	2.5 × 10 ⁻⁴	0.0054
Number of taxa	exact match	1 adjacent	1	2	3	4	5	6	7	8	9	10

Détails Mathématiques

Pour une correspondance exacte entre deux arbres (pas d'incongruence) :

$$P = (2^{N-2})(N-2)! / (2N-3)!$$

ou

$$P = 1 / (2N-3)!!$$

où "!!" est une **double factorielle** et N = nombre de taxons. Pour une incongruence de "1 adjacent" :

$$P = (2^{N-2})(N-1)! / (2N-3)!$$

Pour une incongruence de I branches, placées n'importe où entre 2 arbres :

$$P \leq (2^{N-I-2})(N-I-2)!N! / (2[N-I]-3)!(N-I)!!$$

ou

$$P \leq (N!/(N-I)!!) / (2[N-I]-3)!!$$

N = nombre de taxons et I = # nombre de branches incongrues.

Ce dernier calcul de la valeur *P* est une limite supérieure. En d'autres termes, cette valeur *P* est une surestimation, car la *P*-value réelle est très probablement inférieure (meilleure). *P* est le ratio entre nombre maximal d'arbres incongrus possibles sur le nombre total d'arbres possibles. Cependant, dans l'équation finale, le nombre maximal calculé d'arbres incongrus comprend les arbres non uniques (c'est-à-dire que certains des arbres incongrus ont la même topologie et sont donc comptés plus d'une fois). Par exemple, pour *N* = 4 et *I* = 1, ce calcul donne $P \leq 1,3333$, alors que *P* exact = 0,73333. Plus *N* et *I* sont grands, plus *P* converge vers la valeur exacte.

Ces équations peuvent être facilement étendues au cas de divergences entre plus de deux arbres, chacun du

même nombre de taxons. La probabilité pour que k arbres enracinés, binaires, de N -taxons aient au plus l branches incongrues est:

$$P \leq (N!/(N-l)!!) / ((2[N-l]-3)!!)^{\{k-1\}}$$

De manière équivalente, il s'agit de la probabilité que deux ou plus arbres de N -taxons partagent le même MAST de taille $N - l$ ou plus. (Une calculatrice Javascript [disponible sur la page originale](#) utilise cette équation pour déterminer ses valeurs P .)

J'apprécierais toute suggestion d'une personne qui a des idées sur la façon de corriger les arbres non uniques. J'ai établi de façon indépendante la plupart de ces équations à l'été 2002. Plus tard, j'ai découvert par correspondance personnelle que Mike Steel avait également établi ces équations et allait bientôt publier le tout dans un livre à venir (Bryant et al. 2002). Il semble que l'équation finale a été établie indépendamment par moi et Mike Steel, et à ma connaissance, elle reste non publiée.

[Introduction à la phylogénétique](#)

Références

- Li, W.-H. (1997). Molecular Evolution. Sunderland, MA, Sinauer Associates. p. 102.
- Bryant, D., MacKenzie, A. and Steel, M. (2002). "The size of a maximum agreement subtree for random binary trees." In: Bioconsensus II. DIMACS Series in Discrete Mathematics and Theoretical Computer Science (American Mathematical Society). ed., M.F. Janowitz.
- Harshman, J., Huddleston, C. J., Bollback, J. P., Parsons, T. J., and Braun, M. J. (2003). "True and false gharials: a nuclear gene phylogeny of crocodylia." Syst Biol. 52: 386-402. [PubMed](#)

From:

<https://evowiki.fr/> - **EvoWiki**

Permanent link:

<https://evowiki.fr/comdesc-incongruent>

Last update: **2019/12/06 18:39**

