

Systematic Palaeontology (Palaeobotany)

Does minimizing homoplasy really maximize homology? MaHo: A method for evaluating homology among most parsimonious trees

Nathanaël Cao^{a,*}, Timothée Le Pechon^b, René Zaragüeta-Bagils^c^a Muséum national d'histoire naturelle, UMR 5143 CNRS, « Paléobiodiversité et paléoenvironnements », département « Histoire de la Terre », laboratoire de « Paléobotanique et Paléoécologie », bâtiment de géologie, CP48, 43, rue Buffon, 75005 Paris, France^b Université Pierre-et-Marie-Curie (Paris-6), UMR5143 CNRS « Paléobiodiversité et paléoenvironnements », laboratoire de « Paléobotanique et paléoécologie », bâtiment de géologie, CP48, 43, rue Buffon, 75005 Paris, France^c Université Pierre-et-Marie-Curie (Paris-6), UMR5143 « Paléobiodiversité et paléoenvironnements », laboratoire « Informatique et systématique », 57, rue Cuvier, CP48, 75005 Paris, France

Received 25 May 2007; accepted after revision 20 December 2007

Available online 15 February 2008

Presented by Philippe Taquet

Abstract

Parsimony analysis aims at finding the tree that best fits hypotheses of homology. However, parsimony does not directly maximize homology, but minimizes homoplasy. When a parsimony analysis results in more than a single most-parsimonious tree (MPT), it is shown that the number of homologous characters may vary significantly. We propose a method called MaHo to identify, among the MPTs, the tree(s) that has (have) the highest number of characters that are homologies. We apply this approach to the phylogenetic relationships of the Dombeyoideae (Malvaceae) of the Mascarene Islands. A parsimony analysis was performed, including 31 representatives of the Dombeyoideae. The search resulted in 29,336 MPTs. MaHo was used in order to improve the resolution of the consensus and to increase the number of unambiguous homologies. The consensus of the 7592 MPTs presenting the highest number of homologies was chosen. This resulted in unravelling five additional synapomorphies and in reducing the number of MPTs. *To cite this article: N. Cao et al., C. R. Palevol 7 (2008).*

© 2008 Académie des sciences. Published by Elsevier Masson SAS. All rights reserved.

Résumé

Minimiser l'homoplasie revient-il à maximiser l'homologie? MaHo : une méthode pour évaluer l'homologie au sein des arbres équiparcimonieux. L'analyse de parcimonie a pour but d'identifier l'arbre s'optimisant le mieux avec les hypothèses d'homologie. Cependant, la méthode de parcimonie ne maximise pas directement l'homologie, mais minimise l'homoplasie. Lorsque l'analyse de parcimonie identifie plusieurs arbres équiparcimonieux, il a été montré que le nombre de caractères homologues varie de façon significative d'un arbre à l'autre. Nous proposons une méthode appelée MaHo, permettant de sélectionner, parmi les arbres équiparcimonieux, le(s) arbre(s) portant le nombre maximum d'homologies. Nous appliquons cette approche aux relations phylogénétiques de 31 représentants de la sous-famille des Dombeyoideae (Malvaceae, ex-Sterculiaceae) originaires de l'archipel des Mascareignes. L'analyse de parcimonie permet de retenir 29 336 arbres équiparcimonieux. MaHo est utilisé afin d'améliorer la résolution et d'augmenter le nombre d'homologies acceptées. Le consensus des 7592 arbres équiparcimonieux présentant le nombre

* Corresponding author.

E-mail address: ncao@snv.jussieu.fr (N. Cao).

maximal d'homologies est sélectionné. Cette sélection « révèle » cinq synapomorphies supplémentaires, réduit de façon importante le nombre des arbres équiparcimonieux et améliore la résolution du consensus strict de ces arbres équiparcimonieux. **Pour citer cet article :** N. Cao et al., C. R. Palevol 7 (2008).

© 2008 Académie des sciences. Published by Elsevier Masson SAS. All rights reserved.

Keywords: MaHo; Parsimony; Homology; Retention index; Dombeya; Mascarene Islands

Mots clés : MaHo ; Parcimonie ; Homologie ; Indice de rétention ; Dombeya ; Mascareignes

1. Introduction

Phylogenetic studies using morphological characters are generally performed through the computerized implementation of standard cladistic analysis, here referred to as parsimony analysis. Cladistic analysis is a method proposed by the German entomologist Willi Hennig. Hennig [22] proposed: “The method of phylogenetic systematics, as that part of biological science whose aim is to investigate the degree of phylogenetic relationship between species and to express this in the system which it has designed, thus has the following basis: that morphological resemblance between species cannot be considered simply as a criterion of phylogenetic relationship, but that this concept should be divided into the concepts of symplesiomorphy, convergence, and synapomorphy, and that only the last-named category of resemblance can be used to establish states of relationship.”

Parsimony analysis identifies the tree that minimizes the Manhattan type of approach [16]. The shortest tree maximizes congruence, and best fits the hypotheses of homology proposed by the systematist as synapomorphies. This shortest tree is interpreted as the one that minimizes the number of transformations between character states. If a character, in the shortest tree, is reconstructed in such a way that each state transforms only once into another state, it is considered as a homology. Each extra step, i.e. extra-interpreted transformation, is regarded as an instance of homoplasy, interpreted as a convergence, parallelism, or reversal [15]. However, the logical justification of parsimony is not to maximize the number of homologous characters, but to minimize homoplasies, i.e. ad hoc hypotheses. Ad hoc hypotheses are put forward to explain the lack of congruence. Thus, parsimony analysis aims to maximize homology because “[if] that point of similarity is in fact a homoplasy, the similarity is irrelevant to evaluating genealogical hypotheses [...]. If the similarity is instead a homology, then only the [...] genealogy can explain it. If there is any chance that the similarity is homologous, the [...] genealogy is to be preferred” [16].

In order to know if similarity can be interpreted as homology, i.e. as synapomorphy, the proposed hypo-

theses of homology or primary homology hypotheses, which are retained as secondary homology hypotheses [14], are the ones that pass the test of congruence [31]. Hypotheses that fail the test are deemed ad hoc hypotheses or homoplasies [28]. It seems clear from the foundations of cladistic and, a fortiori, parsimony analysis that the concepts of homology and homoplasy are qualitative, referring to the capacity of maximizing explanatory power [15,29]. Parsimony analysis, however, aims at proposing hypotheses of relationship among taxa based on homologies through the quantitative minimization of homoplasies.

The parsimony analysis often results in more than one single most-parsimonious tree (MPT). As all MPTs are considered optimal trees, one should expect that all of them have as few ad hoc hypotheses as any other. However, this is not the case. We show here that among MPTs, the number of secondary homologies [14] can vary considerably. This finding addresses the problem of the equivalence between minimizing homoplasies and maximizing homologies. We insist, however, in using this result as a way to discuss a possible choice of the MPTs that bear the highest amount of secondary homologies, i.e. that demands the smallest amount of ad hoc hypotheses, following the rationale of parsimony analysis. We apply this approach to the phylogenetic relationships of a group of plants, the Dombeyoideae (Malvaceae) of the Mascarene Islands.

2. Method

We propose a method to identify the tree or trees that has (have) the highest number (s) of secondary homologies among MPTs. This method is implemented in a program called MaHo, for Maximizing Homology [13]. Choosing the MPT or MPTs with the highest number of secondary homologies is justified in what Hennig [21] may be called his “auxiliary principle”: the presence of apomorphous characters in different species “is always a reason for suspecting kinship (i.e., that the species belong to a monophyletic group), and that their origin by convergence should not be assumed a priori.” The purpose of this study is not to doubt of the relevance of

the principle of parsimony as applied in standard parsimony analysis, but to consider whether the quantity of homology among the MPTs is equivalent. MaHo aims at evaluating the quality of some characters on MPTs previously identified.

When a parsimony analysis identifies several optimal trees, the number of characters interpreted as homologies is generally variable. Indeed, if the aim of the method is to maximize homology, then not all the MPTs, whether they have the same length, are optimal in terms of number of homologies. In order to estimate the best tree, or group of trees, we measure the retention index of the characters. The retention index of a character [3,17] has been proposed as a measure of homoplasy [3]. More precisely, it is a measure of the proportion of hypotheses of primary homology retained as secondary homologies [28]. In the parsimony analysis, it is computed as:

$$RI = \frac{(g - s)}{(g - m)}$$

where g is the maximum number of steps a character may have given the taxonomic sample used, s corresponds to the observed length of the character in the retained cladogram(s), and m is the minimum length of a character, which corresponds to the number of states minus one in the case of unordered characters. As Farris [17] writes it, “on a tree for which $s = m$, $r = 1$, there is no homoplasy, and all similarities between terminals in the character are homologous. On another tree for which $s > m$, some of those similarities are homoplasies. Each additional requirement for a step implies a separate origin for state, and each such new origin reduces the fraction of similarities that can be regarded as homologous.” However, as previously said, a hypothesis of primary homology may be rejected as a secondary homology because of a single instance of homoplasy [14], or because of homology and homoplasy are exclusive concepts, or following Kitching et al. [28], who reported that homoplasy is “any character that is not a synapomorphy (homology).” The exclusivity criterion implies that both homology and homoplasy are qualitative and not quantitative concepts. Yet, parsimony analysis is a quantitative method that counts the number of instances of homoplasy through a set of trees in order to choose those that minimize the number of instances of homoplasy. However, this method has a major drawback. Among the MPTs, some may have instances of homoplasy distributed among different numbers of hypotheses of homology. The parsimony analysis rationale implies that a MPT found from five characters, each being homoplasious because of the presence of a single instance of homoplasy, is as optimal as a MPT that bears four homologies and a character showing five instances of

homoplasy. While parsimony seems to minimize homoplasy, it is problematic to assert that it actually always maximizes homology.

To find MPTs that have the highest number of homologous characters, a parsimony analysis is performed using PAUP* 4b10 [34]. In order to obtain the retention index of every character, once the search is finished, the command *describetrees all/diag* is used. PAUP displays a table with the retention index value, among other information, for each MPT. The displayed log may be saved as a text file using the command *Log output to Disk* from the *File* menu. This file is read using the MaHo computer software [13]. MaHo.pl is a perl script (in order to run it a perl interpreter must first be installed; see <http://www.activestate.com>). MaHo parses PAUP logs to sort the MPTs according to the number of homologies present in the tree. MaHo produces two different output files. The first output contains the number of MPTs for each quantity of homologous characters present in PAUP's log. The second output provides the number given by PAUP to the sorted trees in a nexus format that allows computing different consensus trees with a nexus-based software, e.g., PAUP.

3. A theoretical example

The eleven matrix characters (C1 to C11) scored for 12 taxa, X being considered as the outgroup, are reported in Table 1.

Characters C1 to C9 are mutually congruent. Characters C10 to C11 are also mutually congruent. However, both groups of characters conflict, i.e. they cannot fit a single tree. When the MPTs are searched for, using the option *collapse=no*, PAUP finds 455 MPTs with a length $L=20$ steps and an ensemble retention index $RI=0.7353$. MaHo sorts the trees as a function of the number of characters that pass the congruence test as secondary homologies. As all characters are binary, the length of secondary homologies is $s=1$ and their retention index is $RI=1$. Thus, sorting MPTs by their length (measured through the retention index of characters) is equivalent to sort them by the number of homologies they contain (Fig. 1). MPTs found by PAUP show between two and nine homologies, with a maximum of MPTs showing four homologies among the 11 characters. No MPT is completely devoid of homologous characters and no MPT fits the 11 characters, showing that there is some conflict among the scored hypotheses of primary homology. The maximum number of secondary homologies among the MPTs is 9, with a single tree showing this quantity of synapomorphies (Table 2).

Table 1
The taxon/character matrix used for the simulation
Tableau 1
Matrice taxons/caractères utilisée pour la simulation

	SE1									SE2	
	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11
X	0	0	0	0	0	0	0	0	0	0	0
A	0	0	0	0	0	0	0	0	0	1	0
B	1	0	0	0	0	0	0	0	0	0	1
C	1	1	0	0	0	0	0	0	0	1	0
D	1	1	1	0	0	0	0	0	0	0	1
E	1	1	1	1	0	0	0	0	0	1	0
F	1	1	1	1	1	0	0	0	0	0	1
G	1	1	1	1	1	1	0	0	0	1	0
H	1	1	1	1	1	1	1	0	0	0	1
I	1	1	1	1	1	1	1	1	0	1	0
J	1	1	1	1	1	1	1	1	1	0	1
K	1	1	1	1	1	1	1	1	1	1	0

Characters C1 to C9 are congruent; characters C10 and C11 are congruent, but defend a different topology than the nine first ones. The strict consensus of the 455 MPTs results in a totally unresolved bush.
Les caractères C1 à C9 sont congruents ; les caractères C10 et C11 sont congruents, mais défendent une topologie différente de celle défendue par le premier bloc. Le consensus strict des 455 arbres équiparcimonieux obtenus est entièrement irrésolu.

After the MPTs are sorted, a consensus is calculated for every class (number) of homologies, using the computer program PAUP. Fig. 1 shows that the strict consensus tree for the class of two, four or six homologies is totally unresolved; the class of three, five, and seven homologies shows a single resolved node; the strict consensus of the MPTs having eight homologies shows four resolved nodes; finally, the single tree presenting nine secondary homologies shows nine synapomorphies. This is not a general result; in fact, it is possible that the consensus of the MPTs that show a maximum number of secondary homologies is more or less unresolved.

Table 2
Distribution of trees for each quantity of homologies among the MPTs
Tableau 2
Distribution des arbres en fonction du nombre d’homologies parmi les arbres équiparcimonieux

Number of homologies	Number of MPTs
0	0
1	0
2	66
3	107
4	121
5	87
6	50
7	18
8	5
9	1
10	0
11	0

Are all MPTs equal, concerning the amount of secondary homologies they convey? If trees shown in Fig. 2 are taken as an example, Tree 1 fits only two characters without ad hoc hypotheses of homoplasy. Thus, it has only two secondary homologies. On the other hand, Tree 2 presents nine secondary homologous characters. There seems to be an argument to choose Tree 2. Tree 1 needs seven ad hoc hypotheses to explain observed similarities that have led to the proposition of nine of the 11 primary hypotheses of homology, whereas Tree 2 only needs to propose ad hoc theories for two primary hypotheses of homology. If the original taxon/character matrix is considered, it seems obvious that Tree 2 fits the hypotheses of primary homology much better than Tree 1. This argument has recently been used [33] to choose one among two MPTs found by a parsimony analysis. The reason the authors chose one of the MPTs is because “there are two good characters in support of Eritreum to join Elephantimorpha and only one good character in support of Eritreum to join Elephantida”, “a good character” being defined by Shoshani et al. as one “with consistency index, CI= 1.0”. Note that condition CI= 1 also implies a retention index RI= 1.

4. Biological application

In order to validate MaHo and the theoretical background that supports its proposal, a biological example is proposed. It concerns a phylogenetic analysis of a group of plants belonging to the Malvaceae,

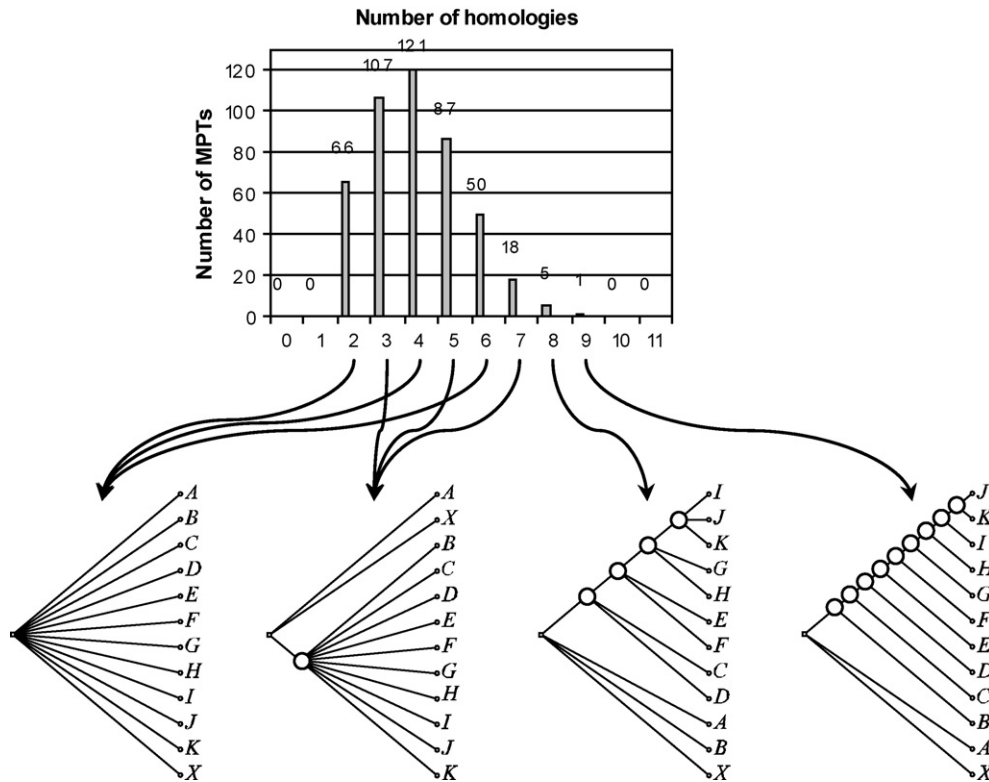


Fig. 1. Distribution of MPTs for each class of homology. Classes are defined by the number of secondary homologies retained by the MPT. The number above the columns indicates the number of MPTs. Below are shown the strict consensus for each of the classes: A, strict consensus of trees showing 2, 4, or 6 homologies; B, strict consensus of trees showing 3, 5, or 7 homologies; C, strict consensus of trees showing 8 homologies; D, strict consensus of trees showing 9 homologies. Each white circle indicates the node at which a secondary homology is placed as a synapomorphy.

Fig. 1. Distribution des arbres équiparcimonieux pour chaque classe d'homologie. Les classes sont définies par le nombre d'homologies secondaires retenues par les arbres équiparcimonieux. Le nombre au-dessus des colonnes indique le nombre d'arbres équiparcimonieux. Au-dessous sont montrés les consensus stricts pour chaque classe : A, consensus strict des arbres présentant 2, 4 ou 6 homologies ; B, consensus strict des arbres présentant 3, 5 ou 7 homologies ; C, consensus strict des arbres présentant 8 homologies ; D, consensus strict des arbres présentant 9 homologies. Chaque cercle blanc indique le nœud où une homologie secondaire est placée en tant que synapomorphie.

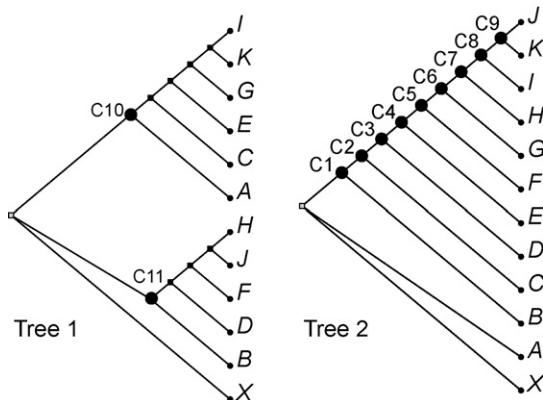


Fig. 2. Distribution of the synapomorphies on two different most parsimonious cladograms.

Fig. 2. Distribution des synapomorphies sur deux cladogrammes équiparcimonieux différents.

Dombeyoideae, based on a set of morpho-anatomical characters.

Relationships within the Malvaceae have recently been clarified and the family now includes nine major sub-families [1,2,10,30]. Among them, the Dombeyoideae is composed of 20 genera, mainly distributed in Madagascar, with more than 250 species present in this island, the Comoros and the Mascarene Archipelago [10]. The latter, located at 600 km east from Madagascar, is composed of three young oceanic islands, less than 8 millions years old: Mauritius, La Réunion, and Rodrigues. These islands are characterized by an exceptionally rich fauna and flora showing a high proportion of endemic taxa [11,12]. Among the Angiosperms, the Dombeyoideae are one of the most diversified groups in the Mascarenes, with 22 species included in four genera, among which 21 are endemic, i.e. inhabit the Mascarenes only, and one is native, i.e. inhabits the Mascarenes and

elsewhere. Phylogenetic relationships within Mascarene Dombeyoideae, but also between the representatives of the Dombeyoideae of the Mascarenes and those of the Indian Ocean, remain uncertain. Here we perform a phylogenetic analysis to validate MaHo.pl software based on the systematics of the Dombeyoideae from the Mascarene archipelago. Thirty-one species were selected to maximize the morphological and ecological diversity of the Mascarene species. Some Continental African and Malagasy taxa were also included as representatives of the global geographical distribution of the

studied group. *Ruizia cordata* Cav. was excluded from the taxonomic sampling because of its high degree of polymorphism.

Recently published systematic studies including some Dombeyoideae and other Malvaceae taxa [7,30], using molecular and morphological characters, showed that the Malagasy genus *Nesogordonia* Baill. represents the sister-group of the remaining members of the sub-family. *Nesogordonia macrophylla* Arènes was therefore selected as an outgroup for rooting the obtained trees.

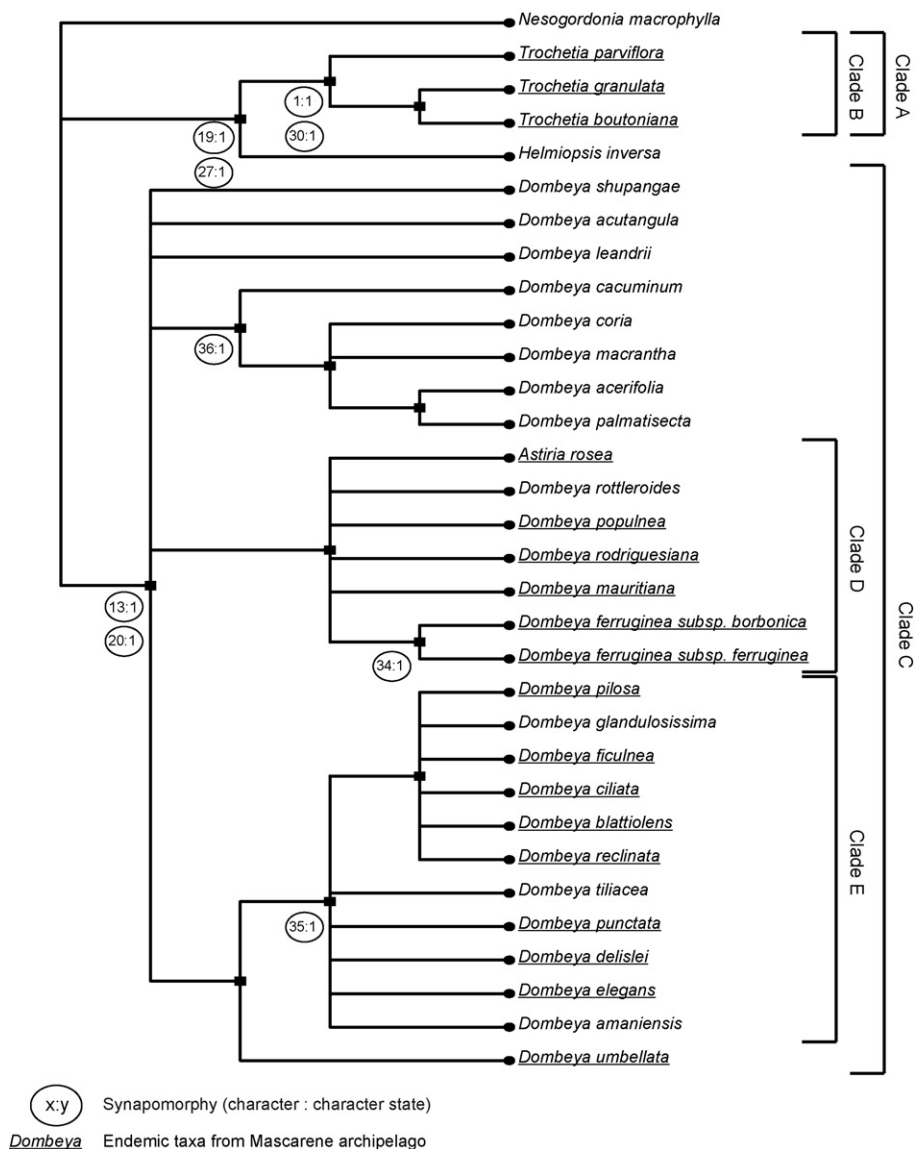


Fig. 3. Strict consensus of the 29,336 MPTs. The unambiguous synapomorphies are detailed at each node of the cladogram. A to E correspond to the five major clade discussed in the text.

Fig. 3. Consensus strict des 29 336 arbres équiparcimonieux. Les synapomorphies non ambiguës sont détaillées à chaque nœud du cladogramme. Les lettres A à E correspondent aux cinq clades majeurs discutés dans le texte.

A ‘taxon/character’ matrix was built based on published literature [4–9,18–20,23–27], field observations, and material from the herbaria of Paris (P). The morphological matrix included 40 characters. The vegetative features, the architecture of the inflorescence, and floral and fruit structures were examined. Missing data were scored in 6.13% of the cells due to lack of information.

A parsimony analysis was performed using PAUP* 4b10 [34], using a branch and bound search option. The search resulted in 29,336 MPTs (tree length $L = 68$, $RI = 0.7986$), and the strict consensus (tree length $L = 85$, $RI = 0.606$) is shown in Fig. 3. This consensus presents nine synapomorphies and six clades that are supported by at least one of unambiguous homology. The *Helmiopsis* H. Perrier and *Trochetia* DC. genera form a monophyletic group (Fig. 3, Clade A) supported by two synapomorphies. The genus *Trochetia*, endemic of the Mascarene archipelago, is monophyletic and supported by two synapomorphies (Fig. 3, Clade B). The *Dombeya* Cav. taxa are grouped in clade C, which also includes the genus *Astiria* Lindl. However, relationships among the members of this clade are totally unresolved.

The Dombeyoideae from the Mascarene archipelago form clades D and E. Clade D is composed of six endemic taxa of the Mascarene archipelago and an endemic taxon of Madagascar. Nevertheless, no synapomorphy unambiguously supports this group and the basal polytomy does not permit to interpret phylogenetic relationships within this clade. Clade E contains eight *Dombeya* taxa from the Mascarenes as well as taxa from Africa and Madagascar. This group is supported by one synapomorphy. Within clade E, a sub-clade is present, but without unambiguous support.

Considering the data matrix, the number of unambiguous synapomorphies appears extremely reduced in the strict consensus. This may partly be due to the high number of MPTs obtained. The lack of support might be due to the quantitative approach for reducing homoplasy of parsimony methods. In fact, MaHo software seems to allow a selection of trees presenting the highest number of homologies.

Within the global population of MPTs, MaHo generated four classes of MPTs (Fig. 4) as a function of their number of homologies.

The 7592 MPTs presenting 14 homologies were chosen in order to study the distribution of synapomorphies among them and their strict consensus was computed (Fig. 5; $L = 75$; $RI = 0.75$).

The selection of trees showing the highest number of homologous characters, selected by MaHo, allowed considerable reduction of the number of MPTs. Consequently, the strict consensus was better resolved

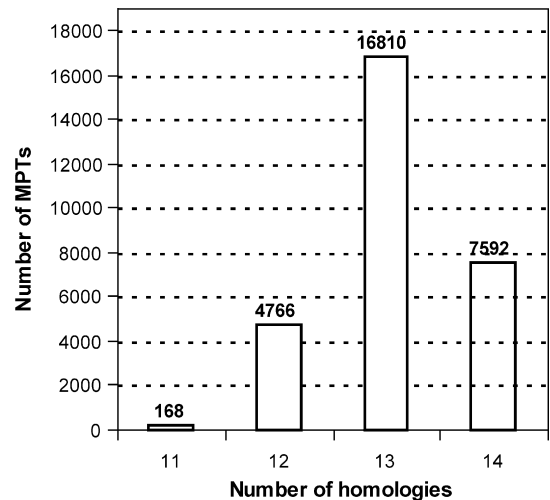


Fig. 4. Distribution of the MPTs as a function of the number of homologies they convey. The 7592 MPTs presenting the maximum number of homologies are included in the hatched class.

Fig. 4. Distribution des arbres équiparcimonieux en fonction de leur nombre d'homologies. Les 7592 arbres équiparcimonieux présentant le nombre maximal d'homologies sont indiqués par la barre hachurée.

ved compared to the strict consensus of the 29,336 MPTs, and the main clades, already supported in the first consensus (Fig. 5), were supported by the same synapomorphies.

Five additional synapomorphies were found after MaHo selection. These synapomorphies unambiguously support several nodes (nodes 1, 2, 3 and 4). As a result, the clade A appears composed by the *Helmiopsis* and *Trochetia* genera, the latter being monophyletic and endemic from the Mascarene Archipelago. The composition of clade C remained unchanged after MaHo treatment, involving the same taxa supported by the same synapomorphies.

The main changes after the application of MaHo were the resolution of clades D and E. Within clade D, a sub-clade noted D', exclusively comprised of endemic taxa from the Mascarene archipelago, can be distinguished. This group is supported by one synapomorphy. In addition, all the internal nodes of clade B' are supported by one homology: this is revealed for the selection performed by MaHo.

Resolution of clade E is improved. A new clade, noted E', is comprised of three taxa from the Mascarene archipelago and one taxa endemic of Madagascar. This node is supported by two synapomorphies that were masked in the global population of MPTs.

From a botanical point of view, the new clades, D' and E', revealed by MaHo, appear morphologically relevant. Clade D' is characterised by several specific

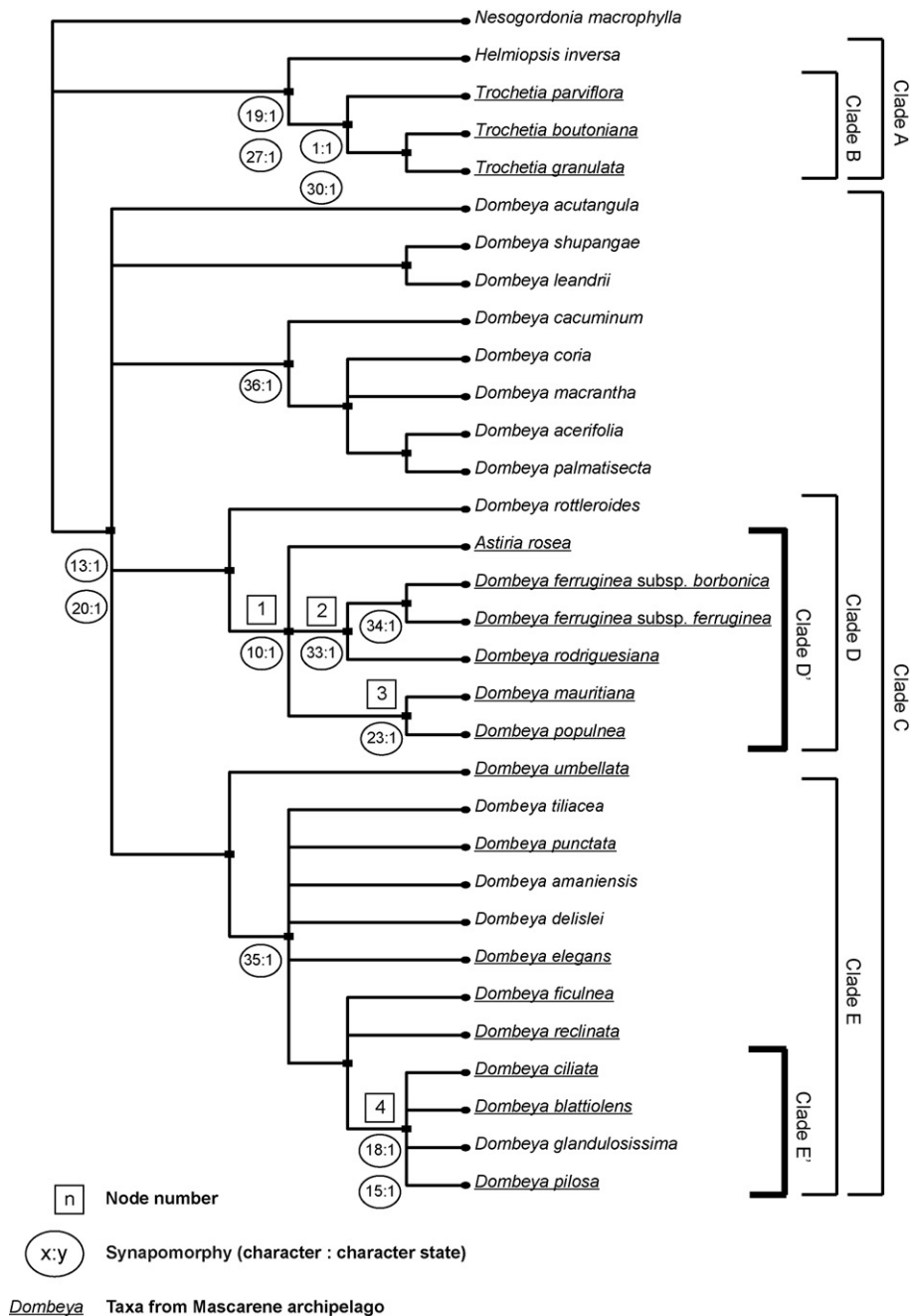


Fig. 5. Strict consensus of the 7592 MPTs ($L=73$, $RI=0.7919$) presenting 14 unambiguous synapomorphies. The unambiguous synapomorphies are detailed for each node of the cladogram. A to E' correspond to the six major clades discussed in the text.

Fig. 5. Le consensus strict des 7592 arbres équiparcimonieux ($L=73$, $RI=0.7919$) présente 14 synapomorphies non ambiguës. Les synapomorphies non ambiguës sont placées sur les nœuds du cladogramme. Les lettres A à E' correspondent aux six clades majeurs discutés dans le texte.

morphological characters. Furthermore, these taxa (*D. populnea* (Cav.) Baker, *D. mauritiana* Friedmann, *D. rodriguesiana* Friedmann and *Astiria rosea* Lindl.) are located in the dry tropical semi-deciduous forest or in areas of mountain ridge with high slopes where desicca-

ting effects of wind are important (*Dombeya ferruginea* Cav.). These two types of habitats can be considered as xeric. This high proportion of endemic species within a single monophyletic taxon might be a signal of a series of differentiation processes that followed the colonisations.

In addition, the supposed diversification of clade D' might have been ecological and related to environmental constraints in xeric habitats. Clade E' is also morphologically homogenous and comprised of taxa living in very moist habitats (precipitations above 2000 mm yr⁻¹).

5. Conclusion

The use of MaHo program on Mascarene Dombeyoideae has allowed the selection of the global population of MPTs that show the maximum number of homologies. This selection reduced the number of trees, improving the resolution of the strict consensus. The new strict consensus of the MPTs presenting the highest number of synapomorphies, revealed by MaHo, allowed retrieving five synapomorphies that were 'masked' among the global population of MPTs. These synapomorphies characterize two unveiled clades (Fig. 5, D' and E'). In this new consensus, relationships of the Dombeyoideae appear more consistent.

From a more general viewpoint, MaHo shows that minimizing homoplasy is not equivalent to maximizing homology. If the dichotomy between synapomorphy and homoplasy, which constitutes the foundations of phylogenetic systematics, is accepted, the parsimony method does minimize homoplasy, but not necessarily always maximizes homology [32]. MaHo may help systematists in selecting, among all MPTs, the ones that present the maximum number of homologies, i.e., that best explain the hypotheses of primary homology. The Dombeyoideae example has shown that the contributions of MaHo may be useful to understand better and interpret more consistently the knowledge conveyed by a real character/taxon matrix, i.e. to find out which trees are supported better if the qualitative distinction between homology and homoplasy is accepted. The use of MaHo may be generally helpful to phylogenetic studies based on morphological or molecular data, resulting in a high number of MPTs, or just in more than a single tree.

Acknowledgments

We are grateful to Estelle Bourdon and Isabella Cattadori for their comments on the manuscript.

References

- [1] W.S. Alverson, K.G. Karol, D.A. Baum, M.W. Chase, S.M. Swensen, R. McCourt, K.J. Sytsma, Circumscription of the Malvales and relationships to other Rosidae: evidence from *rbcL* sequence data, *Am. J. Bot.* 85 (6) (1998) 876–887.
- [2] W.S. Alverson, R. Nyffeler, B.A. Whitlock, D.A. Baum, Phylogenetic analysis of the core Malvales based on sequences of *ndhF*, *Am. J. Bot.* 86 (1999) 1474–1486.
- [3] J.W. Archie, Homoplasy excess ratios: new indices for measuring levels of homoplasy in phylogenetic systematics and a critique of the consistency index, *Syst. Zool.* 38 (1990) 253–269.
- [4] J. Arènes, Les Dombeya des îles des Mascareignes Mémoires de l'institut scientifique de Madagascar 9 (1959) 189–216.
- [5] J. Arènes, 131^e famille Sterculiacées, in: H. Humbert (Ed.), Flore de Madagascar et des Comores, Typ. Firmin-Didot, Paris, 1959.
- [6] L.C. Barnett, An unusual new species of *Helmiopsis* H. Perrier (Sterculiaceae) from Madagascar, *Ann. Mo. Bot. Gard.* 74 (1987) 450–452.
- [7] L.C. Barnett, Tribal realignment of certain paleotropical Sterculiaceae, *Am. J. Bot.* 74 (1987) 724.
- [8] L.C. Barnett, Two new species of *Nesogodonia* (Sterculiaceae) from Madagascar, *Adansonia* 1 (1987) 95–100.
- [9] L.C. Barnett, L.J. Dorr, A new species and variety of *Dombeya* (Sterculiaceae) from Madagascar, *Adansonia* 4 (1986) 365–371.
- [10] C. Bayer, K. Kubitzki, Malvaceae, in: K. Kubitzki (Ed.), The families and Genera of Vascular Plants, Springer, Berlin, 2003, pp. 225–314.
- [11] F. Blanchard, Guide des Milieux Naturels, La Réunion–Maurice–Rodrigues, Ulmer, Paris, 2000.
- [12] T. Cadet, La Végétation de l'île de la Réunion : étude phytocéologique et phytosociologique, Université Aix–Marseille, Marseille, France, 1980.
- [13] N. Cao, T. Le Péchon, G. Sauvenay, Maximum Homology (MaHo). A Perl Script for choosing among MPTs, Published by the author, Paris, 2007.
- [14] M.G.G. de Pinna, Concepts and tests of homology in the cladistic paradigm, *Cladistics* 7 (1991) 367–394.
- [15] P. Darlu, P. Tassy, Reconstruction phylogénétique : Concepts et Méthodes, Masson, Paris, 1993.
- [16] J.S. Farris, The logical basis of phylogenetic analysis, in: N.I. Platnick, V.A. Funk (Eds.), Advances in Cladistics, Proc. First Meeting of the Willi Hennig Society, 1983, pp. 7–36.
- [17] J.S. Farris, The retention index and the rescaled consistency index, *Cladistics* 5 (1989) 417–419.
- [18] F. Friedman, La Réunion, Maurice, Rodrigues, Sterculiacées, in: J. Bosser, T. Cadet, J. Géohy, W. Marais (Eds.), Flore des Mascareignes, MSIRI, Mauritius, ORSTOM, Paris, & Royal Botanical Garden, Kew, UK, 1987.
- [19] L. Gigord, C. Lavigne, J.A. Shykoff, Partial self-incompatibility in an endemic tree species of La Réunion (Mascarene Archipelago, Indian Ocean): consequences for conservation, *Oecologia* 117 (1998) 342–352.
- [20] L. Gigord, F. Picot, J.A. Shykoff, Effects of habitat fragmentation on components of reproductive success in *Dombeya acutangula* Cav. (Sterculiaceae), a partially self-incompatible native tree species of the island of La Réunion (Mascarene Archipelago, Indian Ocean), *Biol. Conserv.* 88 (1999) 43–51.
- [21] W. Hennig, Phylogenetic systematics, *Annu. Rev. Entomol.* 10 (1963) 97–116.
- [22] W. Hennig, Phylogenetic Systematics, University of Illinois Press, Urbana, IL, USA, 1966.
- [23] B.P.G. Hochreutiner, Monographie des *Dombeya* de Madagascar, *Candollea* 3 (1926) 5–120.
- [24] L. Humeau, T. Paillet, J.D. Thompson, Cryptic dioecy and leaky dioecy in endemic species of *Dombeya* (Sterculiaceae) on La Réunion, *Am. J. Bot.* 86 (1999) 1437–1447.

- [25] L. Humeau, T. Pailler, J.D. Thompson, Variation in the breeding system of two sympatric *Dombeya* species on La Réunion island, *Plant Syst. Evol.* 218 (1999) 77–87.
- [26] L. Humeau, T. Pailler, J.D. Thompson, Variation in gender and flower-size dimorphism in the dioecious tree *Dombeya ciliata*, an endemic to La Réunion Island, *Biotropica* 32 (2000) 463–472.
- [27] E. Jacob de Cordemoy, *Flore de l'île de la Réunion*, Paul Klincksieck, Paris, 1895.
- [28] I.J. Kitching, P.L. Forey, C.J. Humphries, D.M. Williams, *Cladistics: the Theory and Practice of Parsimony Analysis*, Oxford University Press, Oxford, UK, 1998.
- [29] A.G. Kluge, Tree-taxon transformation in phylogenetic inference: ambiguity and distortion as regards explanatory power, *Cladistics* 9 (1993) 246–259.
- [30] R. Nyffeler, C. Bayer, W.S. Alverson, A. Yen, B.A. Whitlock, M.W. Chase, D.A. Baum, Phylogenetic analysis of Malvaden-drina clade (Malvaceae s.l.) based on plastid DNA sequences, *Org. Divers. Evol.* 5 (2005) 109–123.
- [31] C. Patterson, Morphological characters and homology, in: K.A. Joysey, A.E. Friday (Eds.), *Problems of phylogenetic reconstruction*, Academic Press, London, 1982, pp. 21–74.
- [32] R.W. Scotland, Taxic homology and three-taxon statement analysis, *Syst. Biol.* 49 (2000) 481–500.
- [33] J. Shoshani, R.C. Walter, M. Abraha, S. Berhe, P. Tassy, W.J. Sanders, G.H. Marchant, Y. Libsekal, T. Ghirmai, D. Zinner, A proboscidean from the Late Oligocene of Eritrea, a 'missing link' between early Elephantiformes and Elephantimorpha, and biogeographic implications, *Proc. Natl Acad. Sci. USA* 103 (2006) 17296–17301.
- [34] D.L. Swofford, *PAUP*: Phylogenetic Analysis Using Parsimony (*and other Methods)*, Sinauer Associates, Sunderland, MA, USA, 1998.