Evidence for a Correlation between Systematics and Bioactivity in New Caledonian Cunoniaceae and Its Implications for Screening and Conservation[1]

Yohan Pillon[2,3] *and Bruno Fogliani*[3]

Abstract: It is generally assumed that there is a good correlation between systematics and the secondary compounds found in plants. However because of the frequent homoplasy of chemical characters this has been difficult to test using statistical methods. Here we applied two nonparametric tests on a published data set, where 50 species of New Caledonian Cunoniaceae were screened for bioactivity against several pathogenic strains. Using Moran's *I* index we showed that in two of nine tests against pathogenic strains there was a significantly higher similarity than expected in bioactivities between species belonging to the same genus and a significantly higher than expected dissimilarity in bioactivity between species belonging to different tribes. When considering the bioactivities against all pathogenic strains with Mantel tests, we also found significant correlation between bioactivity and phylogenetic distance in two of four tests. This has implications in screening and conservation. Searches for new molecules and bioactivity should preferentially be made on species spread across the tree of life. There is also a need to preserve as much phylogenetic diversity as possible to make sure that the widest reservoir of natural compounds remains available for future generations.

Biochemistry has played an important role in plant systematics (see, e.g., Grayer et al. 1999) by providing insights into the relationships between species (Albach et al. 2005, Petrakis et al. 2005), within families such as the Cunoniaceae (Bate-Smith 1977), and between families (Grayer et al. 1999). However, chemical compounds are characters that are often affected by homoplasy because similar compounds can appear independently several times in the course of evolution or might be lost secondarily (Wink and Mohamed 2003, Albach et al. 2005). Some of these problems may also be explained by sampling issues because a single taxon may not always be representative of a higher group (Grayer et al. 1999) or by variation associated with geography (van Heerden et al. 2005) or ontogeny (Çirak et al. 2006).

Secondary compounds are economically important because they can be used as, or be the base for, active molecules in pharmaceutical chemistry (Young 1999, Butler 2004); they can also be important for the production of pesticides, perfumes, and other compounds. It is quite possible that the tropical floras being screened for new molecules today could provide the basis for tomorrow's new medicines (Butler 2004). Because it is now accepted that biodiversity is facing an important crisis, the economic value of that diversity may be a strong argument in favor of increasing conservation efforts on wild species (Balmford et al. 2002). If there is a significant correlation between chemical composition of species and their relatedness, then conservation strategies should be implemented to preserve the widest spectrum of the tree of life

[1] Manuscript accepted 27 March 2008.

[2] Corresponding author: Laboratoire de Botanique, Institut de Recherche pour le Développement, BP A5, 98848 Nouméa Cedex, New Caledonia (phone: (+687) 26 09 28; fax: (+687) 26 43 26; e-mail: pillon@noumea.ird.nc).

[3] Laboratoire Insulaire du Vivant et de l'Environnement, Université de la Nouvelle-Calédonie, BP R4, 98851 Nouméa Cedex, New Caledonia.

Pacific Science (2009), vol. 63, no. 1:97–103

(Mace et al. 2003) to ensure that the greatest diversity of molecules remains available.

The correlation between biochemistry and systematics has generally been accepted, but there has been little supporting evidence that it is statistically significant, probably because of the homoplasy of chemical characters. Few studies are available that have investigated a large number of taxa with multiple tests using a standard protocol. In this study we conducted statistical analysis on the results obtained by Fogliani et al. (2002*a*), who tested the activity of 50 species of Cunoniaceae against 10 pathogenic strains, and show that there is a significant and positive correlation between species bioactivity and their relatedness. Specifically, we used nonparametric tests such as Moran's *I* index that have proven their efficiency at revealing taxonomic patterns (Lockwood et al. 2002). New Caledonian Cunoniaceae have been the subject of extensive taxonomic work over the past decades (Hoogland et al. 1997, Hopkins 2005, 2006, 2007, Hopkins et al. 2007, Pillon et al. in press), and a tribal classification (Bradford et al. 2004) based on morphological (Hufford and Dickison 1992) and molecular (Bradford and Barnes 2001) phylogenies is available. This taxonomic scheme was used to provide rough estimates of phylogenetic distances between species.

MATERIALS AND METHODS

Data

Fogliani et al. (2002*a*) tested the bioactivity of 50 of the ca. 90 species of Cunoniaceae endemic to New Caledonia, representing six genera and four tribes following the most recent classification of Cunoniaceae (Bradford and Barnes 2001, Bradford et al. 2004). Activities against eight bacteria strains and two fungi were assessed. In most cases the activity of ethyl acetate and methanol extract of leaf and bark were tested. A disk containing the extract was placed in a petri dish containing a culture of each strain. In the presence of antimicrobial activities, an inhibition zone can be observed around the disk. The diameter of this zone is used as a measure of the strength of the activity.

Independence of Variables

The independence of variables needed to be assessed before further testing. Correlation of activities between bark and leaf extracts and ethyl acetate and methanol were tested by the mean of a chi-squared test. We tested if species showing activity in their bark were more likely to show activity in their leaves for a given solvent, and similarly we tested if species showing antimicrobial properties with the methanol extract of their leaves or bark were more likely to show properties with the ethyl acetate of the same part of the plant. Correlation between activities against the different pathogenic strains was also assessed (for an extract with a given solvent from a given part of the plant). Chi-squared tests were only conducted on a contingency table where all expected occurrences exceeded five.

Test of the Correlation between Systematics and Bioactivity

Moran's *I* index was used to test if there was a significant similarity in the bioactivity of species belonging to the same taxonomic group (genus or tribe). Moran's *I* index was calculated for each test against pathogenic strains for leaf or bark extracts with ethyl acetate or methanol. Only tests where at least 10 species showed activities were included. We established three distance classes (1, 2, and 3) between species depending on their phylogenetic distance: 1, if they belong to the same genus; 2, if they belong to the same tribe; 3, if they belong to different tribes. The Moran's *I* index tests whether species that are closely related phylogenetically have similar activity.

One single test against one pathogen provides little statistical power to detect significant correlation for several reasons. In most cases, few species showed activity against a given pathogen, and this activity does not follow a normal distribution. To increase the sensitivity of the tests we used another approach. We took into consideration the 10 pathogenic strains in a single statistical test known as the Mantel test. Activities were centered and reduced (i.e., for each species we

calculated the activity of the species minus the average over all species divided by the standard deviation over all). The aim of this procedure was to give a similar weight to each pathogenic strain. "Bioactive" distance was calculated for each pair of species as the geometric distance between their activities against all 10 pathogenic strains:

$$BD_{ij} = \sqrt{\sum_k (a_{ik} - a_{jk})^2}$$

BD is the "bioactive" distance between species *i* and *j*. a_{ik} and a_{jk} are the activity of species *i* and *j*, respectively, against the pathogenic strain *k*.

We used the same phylogenetic distance as for the Moran's *I* index. The Mantel test assesses the correlation between "bioactive" distance and phylogenetic distance. One thousand replicates were used to evaluate the significance of the test.

Both Moran's *I* index and the Mantel test were computed with the software PASSAGE (Rosenberg 2002).

RESULTS

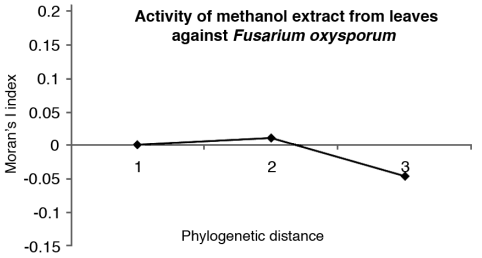
Independence of Variables

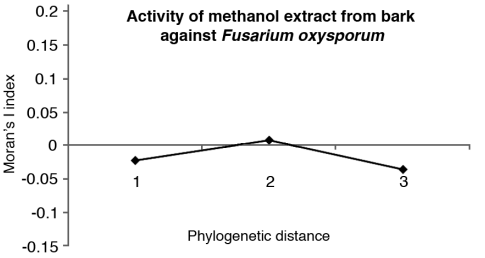
The conditions necessary to apply the chi-squared test (expected occurrence of at least five) to test the correlation between activities of leaf and bark extracts or of the two solvent extracts were satisfied in only a few cases. The chi-squared test always indicated a significant positive correlation between the activities of bark and leaf extract (test of methanol extract against *Fusarium oxysporum*, $P < .0001$; and *Corynebacterium accolans*, $P = .003$) but no significant correlation between ethyl acetate and methanol extract in the single case tested (test of bark extract against *Corynebacterium accolans*). Therefore activities of bark and leaves were considered as nonindependent variables.

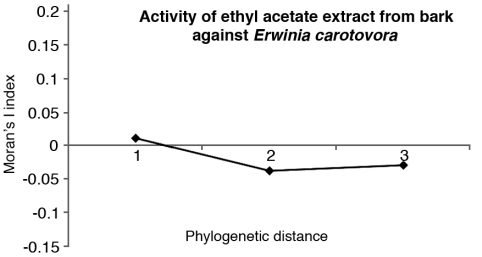
Correlation between activities against the different pathogenic strains was also assessed (for an extract with a given solvent from a given part of the plant). In four cases this could be tested; in three cases there was no significant correlation: activity of ethyl acetate extract from bark against *Staphylococcus aureus* and *Erwinia carotovora*, activity of methanol extract from bark against *Corynebacterium accolans* and *Fusarium oxysporum*, and activity of methanol extract from leaves against *Corynebacterium accolans* and *Fusarium oxysporum*. In only one case we found a significant correlation of activity against two microbial strains: activity of ethyl acetate extract from bark against *Corynebacterium accolans* and *Staphylococcus aureus* ($P = .006$). Therefore we assumed that the different microbial strains could be reasonably considered as independent measures of bioactivity.

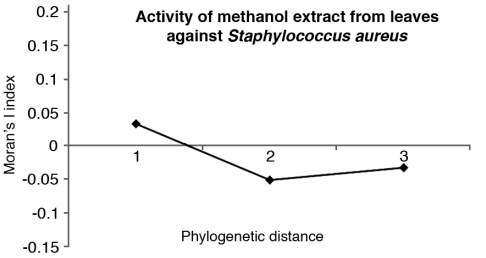
Test of the Correlation between Systematics and Bioactivity

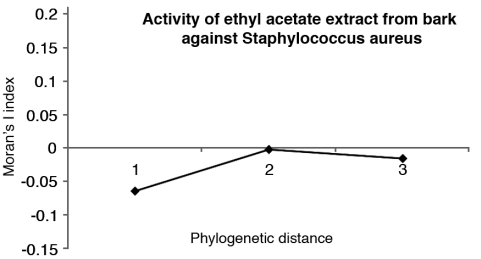
Nine tests against pathogenic strains could be tested with the Moran's *I* index; results for each test and the average over the nine tests are displayed in Figure 1. In the case of the activity of the methanol extract from leaves against *Corynebacterium accolans*, the value of the Moran's *I* index ($I = 0.059$) was significantly higher ($P = .008$) than expected by chance for species belonging to the same genus. Thus activities between congeneric species were closer than expected by chance. Concerning the activity of the ethyl acetate extract from bark against *Corynebacterium accolans*, we observed Moran's *I* index values that were significantly higher than expected for species belonging to the same genus ($I = 0.201$; $P < .001$) or to the same tribe ($I = 0.081$; $P = .028$). If a Bonferroni correction (Rice 1989) is applied to account for multiple tests (nine in total), the second *I* value would still be significant. In turn, species belonging to different tribes showed significantly different activities in two cases (Figure 1), with or without a Bonferroni correction. The seven other tests did not show values that were significantly different from expectation. Average Moran's *I* index for the nine tests showed that overall, *I* values are positive for species belonging to the same genus (similar activities), negative but close to zero for species belonging to the same tribe, and lower for species belonging to different tribes (dissimilar activities).

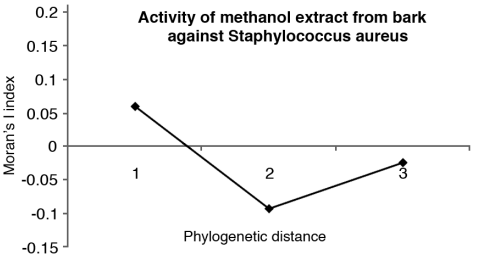

Activity of methanol extract from leaves against *Fusarium oxysporum*
Moran's I index
0.2
0.15
0.1
0.05
0
-0.05
-0.1
-0.15
1
2
3
Phylogenetic distance

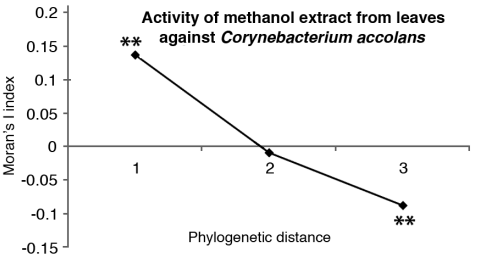

Activity of methanol extract from bark against *Fusarium oxysporum*
Moran's I index
0.2
0.15
0.1
0.05
0
-0.05
-0.1
-0.15
1
2
3
Phylogenetic distance

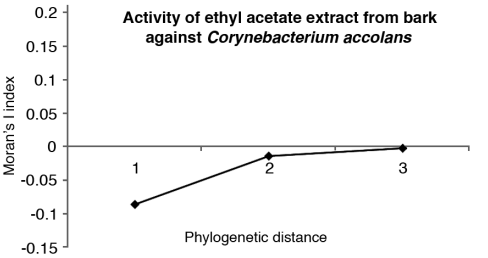

Activity of ethyl acetate extract from bark against *Erwinia carotovora*
Moran's I index
0.2
0.15
0.1
0.05
0
-0.05
-0.1
-0.15
1
2
3
Phylogenetic distance

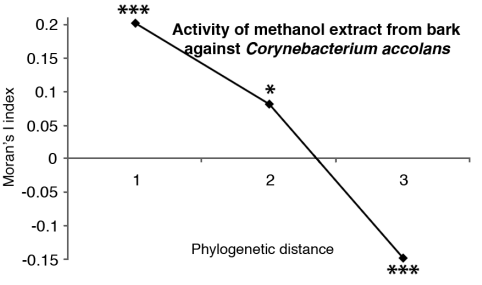

Activity of methanol extract from leaves against *Staphylococcus aureus*
Moran's I index
0.2
0.15
0.1
0.05
0
-0.05
-0.1
-0.15
1
2
3
Phylogenetic distance

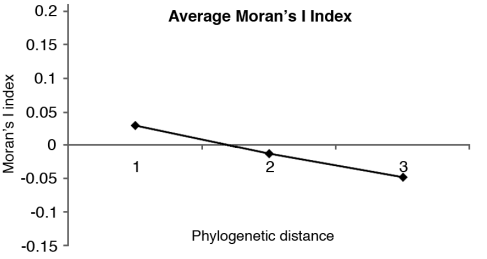

Activity of ethyl acetate extract from bark against Staphylococcus aureus
Moran's I index
0.2
0.15
0.1
0.05
0
-0.05
-0.1
-0.15
1
2
3
Phylogenetic distance


Activity of methanol extract from bark against Staphylococcus aureus
Moran's I index
0.2
0.15
0.1
0.05
0
-0.05
-0.1
-0.15
1
2
3
Phylogenetic distance


Activity of methanol extract from leaves against *Corynebacterium accolans*
Moran's I index
0.2
0.15
0.1
0.05
0
-0.05
-0.1
-0.15
**
**
1
2
3
Phylogenetic distance


Activity of ethyl acetate extract from bark against *Corynebacterium accolans*
Moran's I index
0.2
0.15
0.1
0.05
0
-0.05
-0.1
-0.15
1
2
3
Phylogenetic distance


Activity of methanol extract from bark against *Corynebacterium accolans*
Moran's I index
0.2
0.15
0.1
0.05
0
-0.05
-0.1
-0.15
***
*
***
1
2
3
Phylogenetic distance


Average Moran's I Index
Moran's I index
0.2
0.15
0.1
0.05
0
-0.05
-0.1
-0.15
1
2
3
Phylogenetic distance

Because the correlation between activities against different strains is limited, we considered that pooling them was reasonable and that potential addition effects would be limited. This justifies the use of the "bioactive" distance, which takes into account the activities against the different pathogenic strains altogether. On the other hand we tested separately the activities of bark and leaf extracts using ethyl acetate and methanol, considering that activities of bark and leaf extracts were at least partially correlated.

Mantel tests revealed a significant correlation between phylogenetic distance and "bioactive" distance for methanol extract of bark and leaves ($P = .001$ in both cases) and no significant correlation in the two other cases (ethyl acetate extract of bark and leaves). The first two tests were still statistically significant after applying a Bonferroni correction.

DISCUSSION

It is generally accepted that biochemistry is well correlated with systematics. However this has been difficult to test in a statistical manner because of the inadequacy of the data available. Chemical characters are often binary (presence/absence) and bear little information. It is relatively rare that a large number of characters are available for a large number of taxa. Data generated by different studies may not be comparable, because different protocols were used. Therefore few cladistics analyses have been carried out on chemical characters. Nevertheless when Bininda-Emonds et al. (2001) built cladistics using the lipids of the anal sac secretion of felids, they found good congruence with the existing taxonomic scheme. This was made possible by the large number of compounds included in their matrix (400).

Although several of our tests were nonsignificant, the few significant ones indicated a positive correlation between bioactivity and systematics. We believe that the lack of significant values is due to insufficient data, particularly the small number of species displaying activity against a given pathogenic strain, which limits the statistical power of the tests. The use of the Bonferroni correction shows that the few significant values observed are not due to chance as we multiplied statistical tests but to a genuine pattern. Also, because activities against pathogenic strains are not fully independent, there may be some addition effects that could enhance significance of the Mantel test for which we pooled the data. However, each activity was considered separately when we used the Moran's *I* index; therefore this cannot solely explain the significant correlation we observed.

Species that are closely related are more likely to have similar bioactivity. This can probably be explained by the fact that closely related species have the same active molecules. Nevertheless we cannot exclude other possibilities (e.g., species with similar activities may have similar amounts of the active molecules or have different molecules with the same bioactivity or other explanations). In turn, species that are not closely related are more likely to have different activities; this bears strong implications in screening and conservation.

Most of the world's biodiversity is found in the tropics, and the tropical floras are indeed being intensively screened for new active molecules (Butler 2004). Because many plant species are still to be discovered (Prance

FIGURE 1. Correlograms obtained for activity of different extracts (methanol or ethyl acetate extract from bark or leaves) against different pathogenic strains. Moran's *I* index is a measure of autocorrelation (i.e., it indicates whether species tend to have activities that are more similar [positive autocorrelation] or more different [negative autocorrelation] than expected by chance). It is calculated for species belonging to the same genus (phylogenetic distance = 1), for species belonging to the same tribe but different genera (phylogenetic distance = 2), and for species belonging to different tribes (phylogenetic distance = 3). Moran's *I* index expected value is −0.023; positive values for *I* indicate positive autocorrelation and negative values for *I* indicate negative autocorrelation. Asterisks indicate *I* values that are significantly different from those expected by chance (*, $P < .05$; **, $P < .01$; ***, $P < .001$).

et al. 2000), most species have never been studied for their bioactivity, and some are likely to disappear before they can be sufficiently surveyed. This is true for New Caledonian Cunoniaceae in particular, where many species remain undescribed and new genera are still being discovered (e.g., McPherson and Lowry 2004). The flora of the island is also being screened for new bioactivities (Bosisio et al. 2000, Fogliani et al. 2002*a*,*b*) and is threatened by fires, introduced species, and the continued development of nickel mining.

Thus, species included in screening should be spread out across the tree of life to increase the likelihood of finding new and useful bioactive molecules. Also, because the sampling for testing bioactivity is sometimes relatively destructive or because the discovery of interesting molecules may lead to unsustainable exploitation, red-listed and rare species could be avoided for screening purposes and instead be replaced by closely related and more abundant species.

Vane-Wright et al. (1991) argued that each species should not be given equal weight in conservation (e.g., the single species of *Welwitschia* versus one of the many species of *Taraxacum*). The fact that phylogenetically isolated taxa are more likely to have original bioactivity may be a good argument to prioritize them in conservation planning. The potentially high economic value of some natural products should be kept in mind, because it is a strong argument in favor of the preservation of biodiversity (Young 1999).

ACKNOWLEDGMENTS

We thank Jean-François Arnaud for assistance with the software. We are grateful to an anonymous reviewer and especially Zachary Rogers for carefully reviewing an early version of the manuscript.

Literature Cited

Albach, D. C., R. J. Grayer, G. C. Kite, and S. R. Jensen. 2005. *Veronica*: Acylated flavone glycosides as chemosystematic markers. Biochem. Syst. Ecol. 33:1167–1177.

Balmford, A., A. Bruner, P. Cooper, R. Costanza, S. Farber, R. E. Green, M. Jenkins, P. Jefferis, V. Jessamy, J. Madden, K. Munro, N. Myers, S. Naeem, J. Paavola, M. Rayment, S. Rosendo, J. Roughgarden, K. Trumer, and R. K. Turner. 2002. Economic reasons for conserving wild nature. Science (Washington, D.C.) 297:950–953.

Bate-Smith, E. C. 1977. Chemistry and taxonomy of the Cunoniaceae. Biochem. Syst. Ecol. 5:95–105.

Bininda-Emonds, O. R. P., D. M. Decker-Flum, and J. L. Gittleman. 2001. The utility of chemical signals as phylogenetic characters: An example from the Felidae. Biol. J. Linn. Soc. 72:1–15.

Bosisio, E., D. Mascetti, and P. Cabalion. 2000. Screening of plants from New Caledonia and Vanuatu for inhibitory activity of xanthine oxydase and elastase. Pharm. Biol. 38:18–24.

Bradford, J. C., and R. W. Barnes. 2001. Phylogenetics and classification of Cunoniaceae (Oxalidales) using chloroplast DNA sequences and morphology. Syst. Bot. 26:354–385.

Bradford, J. C., H. C. F. Hopkins, and R. W. Barnes. 2004. Cunoniaceae. Pages 91–111 *in* K. Kubitzki, ed. The families and genera of vascular plants. Springer, Berlin.

Butler, M. S. 2004. The role of natural product chemistry in drug discovery. J. Nat. Prod. (Lloydia) 67:2141–2153.

Çirak, C., B. Saglam, A. Kemal Ayan, and K. Kevseroglu. 2006. Morphogenetic and diurnal variation of hypericin in some *Hypericum* species from Turkey during the course of ontogenesis. Biochem. Syst. Ecol. 34:1–13.

Fogliani, B., S. Bouraïma-Madjebi, V. Medevielle, and R. Pineau. 2002*a*. Screening of 50 Cunoniaceae species from New Caledonia for antimicrobial properties. N. Z. J. Bot. 40:511–520.

Fogliani, B., S. Bouraïma-Madjebi, R. Pineau, and P. Cabalion. 2002*b*. Screening of fifty Cunoniaceae species from New Caledonia for inhibitors of xanthine oxidase and scav-

engers of superoxide anions. Pharm. Biol. 40:526–533.

Grayer, R. J., M. W. Chase, and M. S. J. Simmonds. 1999. A comparison between chemical and molecular characters for the determination of phylogenetic relationships among plant families: An appreciation of Hegnauer's "Chemotaxonomie der Pflanzen." Biochem. Syst. Ecol. 27:369–393.

Hoogland, R. D., J. Jérémie, and H. C. F. Hopkins. 1997. Le genre *Cunonia* (Cunoniaceae) en Nouvelle-Calédonie: Description de cinq espèces nouvelles. Adansonia 19:7–19.

Hopkins, H. C. F. 2005. Nomenclature and typification in the endemic genus *Codia* (Cunoniaceae) from New Caledonia. Adansonia 27:243–254.

———. 2006. Nomenclature and typification in *Geissois* (Cunoniaceae) in the South-West Pacific. Adansonia 28:311–327.

———. 2007. *Geissois bradfordii*, a new species of Cunoniaceae from New Caledonia. Kew Bull. 62:275–280.

Hopkins, H. C. F., B. Fogliani, and Y. Pillon. 2007. Four new species in the endemic genus *Codia* (Cunoniaceae) from New Caledonia. Kew Bull. 62:259–274.

Hufford, L., and W. C. Dickison. 1992. A phylogenetic analysis of Cunoniaceae. Syst. Bot. 17:181–200.

Lockwood, J. L., G. J. Russell, J. L. Gittleman, C. C. Daehler, M. L. McKinney, and A. Purvis. 2002. A metric for analyzing taxonomic patterns of extinction risk. Conserv. Biol. 16:1137–1142.

Mace, G. M., J. L. Gittleman, and A. Purvis. 2003. Preserving the tree of life. Science (Washington, D.C.) 300:1707–1709.

McPherson, G., and P. P. Lowry II. 2004. *Hooglandia*, a newly discovered genus of Cunoniaceae from New Caledonia. Ann. Mo. Bot. Gard. 91:260–265.

Petrakis, P. V., M. Couladis, and V. Roussis. 2005. A method for detecting the biosystematic significance of the essential oil composition: The case of five Hellenic *Hypericum* L. species. Biochem. Syst. Ecol. 33:873–898.

Pillon, Y., H. C. F. Hopkins, and J. C. Bradford. in press. Two new species of *Cunonia* (Cunoniaceae) from New Caledonia. Kew Bull.

Prance, G. T., H. Beentje, J. Dransfield, and R. Johns. 2000. The tropical flora remains undercollected. Ann. Mo. Bot. Gard. 87:67–71.

Rice, W. R. 1989. Analyzing tables of statistical tests. Evolution 43:223–225.

Rosenberg, M. S. 2002. PASSAGE: Pattern analysis, spatial statistics, and geographical exegesis, version 1.0. Department of Biology, Arizona State University, Tempe.

Vane-Wright, R. I., C. J. Humphries, and P. H. Williams. 1991. What to protect?—Systematics and the agony of choice. Biol. Conserv. 55:235–254.

van Heerden, F. R., B.-E. van Wyk, A. M. Viljoen, and P. A. Stoekamp. 2005. Phenolic variation in wild populations of *Aspalathus linearis* (rooibos tea). Biochem. Syst. Ecol. 31:885–895.

Wink, M., and G. I. A. Mohamed. 2003. Evolution of chemical defense traits in the Leguminosae: Mapping of distribution patterns of secondary metabolites on a molecular phylogeny inferred from nucleotide sequences of the *rbcL* gene. Biochem. Syst. Ecol. 31:897–917.

Young, R. N. 1999. Importance of biodiversity to the modern pharmaceutical industry. Pure Appl. Chem. 71:1655–1661.