THE POTTIACEAE S. STR. AS AN EVOLUTIONARY LAZARUS TAXON

RICHARD H. ZANDER

ABSTRACT. Genomic studies have shown that evolution can be based on clusters of genes that may be silenced and reactivated much later. Thus, apparent phylogenetic homoplasy may sometimes represent a one-time developmental response to selection, involving many genes or gene clusters and one or more regulators, and as such may be called homoiology. Evidence for such exaptations (preadaptations) in many organisms is cited. Distinctive features of the Pottiaceae (Bryopsida), the complex gametophytic morphology characterized by the pleisiomorphic Timmiella and the twisted peristome of 32 filamentous, spiculose teeth distal to a basal membrane are suggested to be likewise homoiological. The characteristic peristome may have reappeared in the family four to seven times. Inasmuch as exaptations may skip across cladistic events tracked by non-coding or neutral gene traits, this paper re-evaluates past evolutionary schema in light of an apparent, theoretically acceptable exception to Dollo's Law that complex traits are never re-evolved. An evolutionary diagram based in part on inferred process is offered in place of the usual cladogram based on inferred events of genetic isolation.

KEY WORDS: Exaption, Dollo's Law, evolution, homoiology, Pottiaceae, phylogenetics, Lazarus taxon.

INTRODUCTION

With a change during the 1700's and 1800's towards probabilistic thinking in science (Pap, 1962), philosophical or logical support for certainty or relative certainty about "truth" in science began to be replaced by a pragmatic attitude that science does not establish truth, but identifies theories that are so well supported by facts (well-corroborated observations) that they may be acted upon (used as a basis for additional research), while alternative theories may be ignored even though some facts support them. The new probabilistic science is centered around a phenomenon of distribution of random events in physics described by the Central Limit Theorem, and is the basis for probabilistic theory and statistics. Statistics as a field is well established today and, although there are conflicts between the different schools (Gigerenzer et al., 1989) of frequentist, Bayesian, and hypothesis testers, the basic requirements for assessing reliability are clear, including the important across-science convention of a 0.95 confidence limit or interval. There are, however, commonly today many phylogenetic studies that fail to provide adequate statistical evaluation, often focusing on only one or two of several requirements for establishing reliability.

Bayesian Philosophy and Science

Although the result of Bayes' Formula is the posterior probability and that is commonly regarded as the chance of a hypothesis being correct, the Bayesian philosophy also

1 Missouri Botanical Garden, PO Box 299, St. Louis, MO 63166-0299, U.S.A. E-mail: richard.zander@mobot.org
requires no bet (i.e., no confidence in the pragmatic value of the results) except with evaluation of risk. Bayesian betting in the phylogenetic context means taking the study out of the realm of speculation and deciding whether to view the solution as sufficiently reliable to base additional work (biogeography, etc.) on it. The risk of one’s science being wrong (requiring a high posterior probability of being right) also depends on the chance of any assumptions being wrong, which affects the final probability involved in the decision to make the bet or not. Thus, the posterior probability of a branch arrangement of interest must be multiplied by the chance that the arrangement is wrong because an assumption may be wrong.

There are two major methods of phylogenetic analysis that now compete with funding for taxonomy and biodiversity studies: analysis of morphological phenotypic expression commonly using maximum parsimony, optimality, Bremer support, and, recently, Markov chain Monte Carlo Bayesian analysis (“datatype = standard”); and DNA sequence analysis focusing on many apparently evolutionary neutral or nearly neutral independent and independently (i.e., randomly) distributed (I.I.D.) traits. Each has particular reliability problems. Morphological characters are relatively few and commonly convergent, optimality alone does not take into account nearly optimal alternative solutions, Bremer support needs branch length data to be even slightly meaningful, and Bayesian analysis shows too high posterior probabilities for short branch lengths (e.g., from typical morphological data) due apparently to a hard polytomy not being an alternative (Lewis et al., 2005). DNA sequence analysis suffers from major problems, e.g., (1) assumption that the data is independent and randomly distributed when it may not be, at least in part; (2) that the data is sufficiently copious to reflect a chi-squared distribution; (3) that the evolutionary model being used is correct or adequate; (4) that conclusions are not based on pre-selecting for further probabilistic study branch arrangements or groups with high probability scores from a set of many, i.e., multiple test problems; (5) ignoring the effect of implicit, often unaccounted assumptions in introducing uncertainty by an emphasis on the likelihood principle that the data provide all that is needed for evaluation of the truth of an hypothesis (because the likelihood of the hypothesis is the likelihood of the data or that the probability that a given tree is correct is fully given by Bayes’ Formula); (6) ignoring contrary data, such as branch arrangements from other studies that are less than 0.50 Bayesian posterior probability (BPP) or nonparametric bootstrap (BP); (7) conflicting results from different methods of analysis; and (8) falling back on philosophy or logic when statistical support is low. There are dozens of unaccounted assumptions that affect the reliability of phylogenetic analysis (Zander, 2005, and in preparation), and the present paper provides an across-the-board penalty of one percent to the BPP. This is fully 20 percent of the reliability window if 0.95 BPP is the accepted minimum level of reliability; also, a penalty of 0.01 on 6 chained internodes, each of 1.00 BPP, means that the chance of all branch arrangements being correct at once is reduced to 0.94. Thus, if corrected for unaccounted assumptions, a cladogram of seven or more terminal taxa cannot be entirely reliable regardless of published BPP.

Problems with statistics and phylogenetic estimation

Regarding using either total evidence (Allard & Carpenter, 1996; Eernisse & Kluge,
R. H. Zander: The Pottiaceae s. str. as an evolutionary lazarus taxon

1993; Nixon & Carpenter, 1996) or the comparison of individual gene studies (Chen et al., 2003), total evidence may be better because, in evaluating gene studies with Bayes’ Formula, a one-locus branch arrangement at high reliability in BPP by chance alone will never be refuted by a few studies of additional loci that show ca. 0.50 BPP because no matter how many 0.50 priors there are, the high score does not change with application of Bayes’ Formula. Only when a low BPP by chance alone occurs can the high randomly derived score be refuted. Therefore, many studies must be made to establish a chi-square distribution, and low scores retained (not discarded if below 0.50). On the other hand, consider a 4-taxon branch arrangement with the branches labeled A, B, C, and D, and possible branch lengths AB, AC, and BC when D is outgroup. With total evidence, in the 4-taxon case with random data or in the case of real data with a hard polytomy, it is possible to find or contrive an arrangement supported by chance alone with, for instance, a 0.98 confidence interval (CI) reliability of a branch length of 7 steps where \( AB + AC + BC = 10 \) (chi-squared with one degree of freedom for 7 correct out of 10 trials); doubling the data will reduce the chance down to 0.96 in a branch length of 11 with \( AB + AC + BC = 20 \). This assumes the additional data is random from additional studies of the hard polytomy (or from another contrived data set) and shared about equally by AB, AC and BC. Thus, for branch lengths of about 10 steps, highly reliable scores that may have occurred by chance alone are not corrected by total evidence involving a hard or soft polytomy until the data set relevant to that branch arrangement is more than doubled in size.

Multiple test problems (Felsenstein, 2004) occur when selection is made on the chance of being correct. In phylogenetic analysis, preselecting taxa for study based on morphological analysis can associate taxa with randomly established convergent gene traits that are linked and reflected in morphology, and reliability measures associated with molecular analysis of just that group are impacted. The same is true with all molecular data in that both maximum parsimony and likelihood analysis will group taxa, sometimes on the basis of imbalanced but random data, and each group is a preselection for succeeding studies. If the first molecular analysis resulted in a BPP of 0.95, then the second analysis including data from just this group is affected. Any preselection of taxa is a candidate for examination against introduction of multiple test problems. A totally artificial data set of 50 taxa and 50 random 2-state characters was contrived with RANDSET (Zander, 1999). Analysis with PAUP* under maximum parsimony (hs with 20 random sequence additions) produced 159 equally parsimonious trees from this totally random data set, and a largely unresolved strict consensus tree with, however, two distinct lineages (A,B) and ((C,D)E). An analysis of a subset of these 5 taxa, under maximum parsimony (bandb), produced one lineage of (C,D)E with 0.58 BP support for (C,D) and 0.81 BP for (C,D,E). Bayesian MCMC analysis (MrBayes 3.1, datatype = standard,ngen = 500000) of the 5 taxon data set provided 0.74 BPP and 0.94 BPP for the same groups. Clearly, preselection of a subset on the basis of a reliability measure, even intuitive taxonomic grouping, for further analysis introduces multiple test problems for linked data.

**Morphological and molecular analysis**

Morphological data is commonly discounted because they are relatively few and evi-
dentily commonly subject to convergence. On the other hand, molecular and morphological analytic results may be as similar as they are because molecular traits, even the apparently non-coding sites, may be affected to an important extent by the same evolutionary pressures as is morphology. Present theory indicates that some molecular data are subject to evolutionary pressures, but other data (most non-coding DNA and the neutral genes of legend) apparently do not contribute to phenotypic traits of the organism but may be considered baggage of randomly generated mutations, the nesting of which may show shared ancestry. As far as evolutionary process is concerned, however, adaptation and fitness has to do with the selected and fixed coding mutations not the non-coding mutations, and process may sometimes not follow inferred nesting.

Casual inspection of any cladogram, however, allows the following three kinds of apparently reliable information (given the data): (1) branch arrangements distant on a cladogram are probably reliably separated; (2) branch arrangements that are distal to an internode of 0.95 CI (or BPP) are probably reliably monophyletic; and (3) if the molecular result agrees with prior morphological analysis, the branch arrangement is probably reliable.

Contributions of genomic analysis

Homoiology is a kind of parallelism in which organs that are developmentally the same converge to the same kind of phenotype, e.g. wings of birds and bats are homoiologous as tetrapod forelegs but convergent as wings. It may be applied (e.g., Stuessy et al., 2003) to recent theorization based on data from evolutionary development ("evo-devo") work that indicates that clusters of genes, commonly controlled by one regulator gene or promoter sequence with expression fine-tuned by different copy numbers of tandem repeats, may be silenced (e.g., by DNA demethylation or histone hyperacetylation) then reactivated later in a phylogenetic lineage (Caporale, 2003a, b; Stuessy et al., 2003). Such gene clusters may determine the existence of major organs and be more highly or immediately adaptive than incremental selection and accumulation of traits (Caporale, 2002, 1999). Incremental evolution, a basic assumption of all present parsimony and likelihood analyses, is thus not necessarily always the case, e.g., very short branches signaling bursts of evolution or punctuated equilibria. There may be many "hopeful monsters" (Bateman & Dimichele, 1994) among our taxa, and much apparent homoplasy may be reactivation of silenced traits, not independent new mutation. As reviewed by Stevens (1980), there was early discussion of "evolutionary recall" (Robson, 1972), this possibly by developmentally based violation of Dollo's Law that complex traits are not re-evolved. Because the definitions of much terminology in phylogenetics include a theoretical first cause, the use of the word "homoiology" is various, as in: "Homoiologies are phylogenetically misleading resemblances among a group of taxa that can be ascribed to phenotypic plasticity. That is, homoiologies are homoplasies that result from the expression by a genotype of different phenotypes in response to different environmental conditions" (Lycett & Collard, 2005). In the present paper, on the other hand, homoiology is "evolutionary recall" or surprising parallelisms apparently from a deep homology.

Zander's (1993) maximum parsimony analysis of the morphology of all 76 then known genera of the Pottiaceae showed re-evolution of the twisted peristome complex of
32 filamentous, spiculose teeth arising from a basal membrane at four distant nodes on the cladogram. The molecular cladogram of Werner et al. (2004) based on rps4 chloroplast DNA showed scattered and apparent re-evolution of the twisted peristome complex. The latter cladogram also showed one pottiaceous genus, Timmiella, treated as basal by Zander (1993), to be phylogenetically distant from the Pottiaceae s.str. This has been confirmed by La Farge et al. (2000, 2002) and Heddderson et al. (2004). Sato et al. (2004) apparently confirm it but the relevant data is from a DNA database. Because Timmiella evinced a full range of characteristic, pleiomorphic Pottiaceae traits, its morphology may be determined by the same gene cluster that informs the Pottiaceae s.str. Intermediate nodes, however, shared lineages of quite different morphology, in the families Calymperaceae, Dicranaceae, Ditrichaceae, Fissidentaceae, and Grimmiaaceae. Although the Werner et al. (2004) cladogram was fully resolved, the reliability of phylogenetic distance was difficult to ascertain as only some nodes showed high values for Bayesian credibility intervals and the cladogram required reduction to show only reliable relationships.

Examples of phylogenetically distant but developmentally homologous traits include apparent re-evolution of shell coiling in snails (Collin & Cipriani, 2003); reactivation of wings in wingless walking sticks (Whiting et al., 2003); cyclopaean brine shrimp (Fryer, 1999); modes of vulva formation in nematodes (Srinivasan et al., 2001); ancestral features of the lateral lines, muscles and gill rakers of cichlid fishes (Stiassny, 1992); eyes in eyeless copepods (Dingle, 2003); teeth in chickens (Harris et al., 2006); and occurrence of a second molar in lynx (Kurtén, 1963). Not all similar atavisms are developmentally associated, as vestigial hips in snakes and whales have been shown to ultimately depend on quite different developmental pathways (Bejder & Hall, 2002). The Pottiaceae (Bryopsida) is a large and diverse family of mosses characteristic of harsh habitats. The Pottiaceae (less Timmiella) may well be a “Lazarus taxon” (Wignall & Benton, 1999), not in the geologic sense as a group that has skipped a long fossil epoch, but as a resurfacing in evolutionary time of a major developmental adaptive complex contrary to Dollo’s Law (Hall, 2003). Such silenced-gene reactivation has recently been estimated as restricted to 0.5–6.0 million years before full or partial DNA degradation to non-functionality (Marshall et al., 1994). According to a review by Stevens (1980), however, several studies have supported the idea that the “genetic capacity to produce structures now no longer visible” can be retained for 200 million years or more in vertebrates. There is also evidence that some process like DNA repair may act to preserve silenced genes (Fryer, 1999) perhaps 20–100 million years (Collin & Cipriani, 2003). If some traits evolve as silenced-reactivated clusters, then the independent and random distribution assumption critical to reliability measures (bootstrap, Bayesian posterior probability), and the idea of incremental evolution itself is affected at both morphological and coding DNA levels. It is possible that silenced genes may be harbored in the mosses by small effective population sizes together with asexual reproduction retaining silenced genes somatically. According to Lynch and Force (2000), evidence suggests that gain-of-function mutations may be as common as loss-of-function mutations, and may prolong the life of gene duplicates by resurrecting previously impaired copies.
Possible homoiologous traits in the Pottiaceae

There were many reversals in traits in Zander's (1993) maximum parsimony cladogram of Pottiaceae morphology. These involved length of stem; stem sclerodermis and hyalodermis presence; leaf stance when dry or wet; leaf shape and length; conformance of leaf ventral surface and costal groove; leaf base shape; number of rows of cells across ventral surface of costa; transverse section shape of dorsal stereid band; dorsal costal epidermis presence; costal hydroid strand presence; width of medial upper laminal cells; superficial wall width of upper laminal cells; sexual condition; perichaetial leaf shape; seta twisted or not; theca length; annulus type; peristome type; length of calyptra; and spore diameter.

Hypothetically, these traits could be changed from one state to another and back again via direct mutation of the gene or through a silencing-reactivation method, but if simple the traits can be treated as incrementally accumulated. If of definitely complex morphology, reversals are more likely to be homoiologous. Homoiologous clusters of traits can also be lumped by maximum parsimony into the same lineage. For instance, a multiple trait reversal through reactivation in a species that was pottiaceous but not a Tortula involving the major gametophytic traits of Tortula would place that taxon in Tortula or at least in the close relationship. Multiple trait reversal would thus be hidden to a large extent in morphological phylogenetic analysis. We know that genes are silenced during cell differentiation, so silencing is not debatable, and any reversal may involve reactivation. Evidence for multigene complex reactivation may be sought in identifying major phenotypic trait clusters that are at a distance on a molecularly based cladogram and apparently contravene Dollo's Law that complex traits cannot be expected to re-evolve.

METHOD

A Bayesian MCMC analysis of 48 exemplar species of the Pottiaceae and 24 species in other haplolepideous families was done by Werner et al. (2004), based on rps4 chloroplast DNA sequences. The results were here subjected to the five-step "operative transform" procedure of Zander (submitted), which creates a cladogram with only 0.95 or better BPP (here the same as CI) branch arrangements, all other branch arrangements being collapsed. The original cladogram was fully resolved, and thus "speculative," and the transform provides a reliable core on which a taxonomist can base classification and further research. First, the BPP is taken as equivalent to the chance that the arrangement is correct, even though the BPP of short branches are commonly expected to be too high (Lewis et al., 2005; Suzuki et al., 2002; Zander, 2004). If bootstrap or decay indexes are given instead, the CI is approximated from tables based on 4-taxon simulations (Zander, in preparation; 2004). Second, a small (0.01) reduction in probability is made to reflect unaccounted assumptions that contribute uncertainty, as a general penalty applicable to BPPs of each branch arrangement, and also additional individual penalties are made when necessary for suspicious branch arrangements. Third, morphological and molecular results are combined with Bayes' Formula when "uncontested" morphologically based branch arrangements, assigned a prior probability of 0.95 match molecularly based arrangements, and the posterior probability (BPP) must be therefore greater than 0.95 if the published molecular BPP is
greater than 0.50. Doubtful morphologically based branch arrangements are assigned a prior probability of 0.50, and, as is necessary with the Bayes’ Formula, the probability of the molecular arrangement is then accepted. Fourth, distance on the cladogram comprised of internodes of less than 0.95 BPP is evaluated for reliability in the following way: Chained internodes each with BPPs lower than 0.95 are combined into one implied reliable internode (IRR) using a formula that determines the chance that at least one internode among two or more is correct, by calculating an implied reliable confidence interval (IRCI). The formula for the IRCI is simply one minus the product of the chances of each of all concatenated arrangements being wrong (where the chance of being wrong is one minus their BPP). A spreadsheet is available for easy calculation of the IRCI (Zander, 2003). Fifth, at least two intervening internodes must be true at once for taxa to be distinguished as separate from a particular clade as opposed to being merely a basal branch, and the calculation used is similar to standard Bonferroni correction for multiple tests (i.e., the BPPs or IRCIs when multiplied together must reach 0.95). These five simple evaluations determine that one or more reliably resolved branch arrangements exist in a published cladogram, and that two or more are correct at once when distance is evaluated.

The morphologically based phylogenetic analysis of Zander (1993) was re-examined for homoiologous traits as complex traits that appear to re-evolve separately in distant parts of the cladogram. A cluster analysis (UPGMA) was performed (using PAST, Hammer & Harper, 2003) on the same data set (less Hypodontium, definitely excluded from the Pottiaceae by Werner et al., 2004), using the clustering method that kept uncontested or accepted groups together better than other methods (Dice, a.k.a. Sorensen). A subset of that data set containing Timmiella as outgroup and only the genera with at least some species with the complex pottiaceous twisted peristome was used to generate (with PAUP*, Swofford, 1998) a constraint tree with maximum parsimony (default settings, including ACCTRAN). Then, a neighbor-joining tree (NJ—Saitou & Nei, 1987) of all taxa (less Hypodontium) was made with PAUP* with genera with twisted peristomes constrained to the most parsimonious arrangement of that subset.

There are thus four sources, all somewhat different, of possible evolutionary relationships that can inform a pragmatic classification, (1) intuitive or Gestalt as discussed for most genera by Zander (1993); (2) the operative transform of the Werner et al. (2004) molecular cladogram (Fig. 1); (3) the NJ reanalysis of the original cladogram of Zander (1993) with genera with twisted peristomes constrained (Fig. 2); and (4) the UPGMA-Dice cluster analysis (Fig. 3). Hypothetical branch arrangements based on morphological traits are not assigned formal probabilities as these are simply optimal trees and there may be many barely less likely alternative trees, also significant nonparametric bootstrap proportions are usually impossible to get from morphological data sets with large numbers of taxa; on the other hand, intuitively uncontested groups for which alternative branch arrangements are apparently not supported by any morphological evidence may be informally assigned a high confidence level as a kind of Bayesian prior. The object was to create the most highly resolved tree that is both reliable and reflects all relevant data. In this case, given problems with (1) implicit assumptions of model and data integrity, (2) the work of Baptist et al. (2005), Jeffroy et al. (2006) and Salamin et al. (2005) critical of the ability of sequence
data to generate reliable phylogenetic trees, (3) the suggestion by Wortley et al. (2005) that as many as 10,000 base pairs are needed to fully resolve and support a difficult phylogeny, and by Rokas et al. (2003) that as many as 20 genes may be necessary to derive a supported, resolved phylogeny for such groups, then a much simplified evolutionary tree informed by a reduced number of apparently reliable branch arrangements in molecular studies is in this case the best to be hoped for. For taxa not treated in the molecular study or otherwise equivocal, the intuitive or Gestalt evaluations of Zander (1993) are followed, based on a strong a priori weighting of morphological traits through biologically well-founded character analysis as justified by Szalay et al. (1987).

Given the lack of adequate confirmation of the results of morphological study through fossils or statistics, and the number of unaccounted assumptions that plague molecular analysis, the taxonomist developing a classification must be prepared to accept at this time only a rather poorly resolved evolutionary tree (not that based on weakly founded speculation) as in any way reliable and suitable for basing further research.

RESULTS

The operatively transformed molecular tree is given in Fig. 1. The Pottiaceae s.str. is grouped terminally in the cladogram. The genus *Timmiella*, however, which was a basal branch in the Zander (1993) morphological cladogram, is reliably placed (given the data set) towards the base of the cladogram, separated from the Pottiaceae s.str. by three reliable (BPP 0.99) internodes. Details of the transform will be given in another publication. The genus *Timmiella* has a large array of the pleisomorphic traits of the Pottiaceae, including the filametous, twisted peristome with a basal membrane. Although it is distant by at least three speciation events (and probably many more that lack reliable data tracking them), it is a candidate for interpretation as having a homoiologous relationship with the Pottiaceae s.str., sharing the same many, developmental little-changed traits. Although both *Timmiella* and the Pottiaceae s.str. are theoretically Lazarus taxa, the absence of branching lineages connecting *Timmiella* with the main line of the cladogram indicates this genus may be frozen in time, a living fossil.

The transform shows to some extent those major infrafamilial groups of the Pottiaceae recognized by Zander (1993), with clustered taxa in lineages approximating his Timmieloideae, Trichostomoideae, Leptodontieae, and Barbuleae. A major group identifiable as the Pottieae including *Tortula* is clear, but *Syntrichia* and relatives are less well differentiated from the Barbuloideae (sensu Werner et al., 2004, which is the same as Merceyoideae sensu Zander, 1993). The Barbuloideae is somewhat exploded, with some members scattered among other lineages, and *Scopelophila*, exemplar of the type genus (*Merceya* is a synonym of *Scopelophila*) isolated as a basal branch. As *Scopelophila* is so reduced...
R. H. Zander: The Pottiaceae s. str. as an evolutionary lazarus taxon
morphologically that it may be a member of a different family, the Merceyoideae can be restricted nomenclaturally to that isolated genus and the more familiar Barbuloideae used for the remainder of the subfamily (this supported by Goffinet et al., 2001). The complex twisted peristome is scattered among the several apparently reliable lineages of Pottiaceae. Fig. 1 shows in bold face those genera known to have a twisted peristome in at least some member species, and those species that actually have such are marked with a square bullet. Clearly all ancestors of the reliable nodes would have to have twisted peristomes to avoid an interpretation as homoiologous gene complex reactivation; this requirement extends down to Pseudocrossidium hornschuchianum and even further to Timmiella. Given the various expressions of the peristome known for the Pottiaceae, it is improbable that no similar variation would occur in the ancestors, and that other families evolving from the line between Timmiella and the Pottiaceae s.str. would not have at least some twisted peristomes. Thus, an interpretation of multiplex silencing and reactivation is a good hypothesis. The phylogenetic distance between Timmiella and the Pottiaceae can be interpreted as events of genetic isolation associated with the Biological Species Concept (doubtless accumulating minor incremental trait changes) that do not affect the major evolutionary contribution of the transmission of the set of now activated pleisiomorphic traits, including the twisted peristome, of the ancestor of Timmiella to those of the Pottiaceae s.str.

The four instances of apparent re-evolution of the twisted peristome found in the original cladogram of Zander (1993) are reflected in the neighbor-joining tree with genera (in at least some species) with twisted peristomes constrained of Fig. 2. Genera with at least one species with a twisted peristome are given in bold face. The NJ tree is much like that of the original morphological tree of Zander (1993), but points out well that even though taxa are forced into a form that will best avoid contradiction of Dollo’s Law, NJ implies seven more or less separate groups of taxa with twisted peristomes, which indicates that the twisted peristome contributed developmentally by a direct ancestor of Timmiella may be been reactivated as many as seven times in the Pottiaceae. The neighbor joining method is fast and allows different branch lengths, but is not guaranteed to produce the shortest tree. The results, however, are commonly close to the maximum parsimony result, and a molecular clock is not assumed. Neighbor joining is used because this study is subject to the limitations of dubious independent and random distribution of data, a myriad unaccounted assumptions, a paucity of data, and inadequate models, and thus the expected degree of reliable resolution cannot demand the laborious inutility of exhaustive optimization.

The cluster analysis of Fig. 3 that best kept uncontested or accepted groups together is UPGMA with the Dice (or Sorensen) method, which emphasizes similarity over dissimilarity. By extension this method may best keep groups of unknown relationships in correct proximity, especially when evolutionary relationship may not be incrementally evolved. Seven groups of genera, identified with bold face, exhibit twisted peristomes in at least some member species.

*Reclassification of the Pottiaceae*

The molecular analysis, the neighbor-joining tree, and the cluster analysis can be used to develop a new classification of the genera, subfamilies and tribes of Pottiaceae that
Fig. 2. Neighbor-joining reanalysis of the original morphologically based cladogram (less Hypodontium) of Zander (1993) with genera with twisted peristomes constrained. Bold-faced are names of genera including at least some species with twisted peristomes.
Fig. 3. UPGMA cluster analysis of the morphological data set with Dice (Sorensen) method selected as best retaining uncontested (accepted) branch arrangements. Bold-faced are names of genera including at least some species with twisted peristomes.
R. H. Zander: The Pottiaceae s. str. as an evolutionary Lazarus taxon

reflects reliable features of the molecular work in light of the possibility of developmentally identical reactivated homoiologous trait clusters tracked on molecular cladograms only by general morphological similarity. A low resolution is warranted by these theoretical possibilities, and genera with twisted peristomes and no particular reduction in morphological complexity of the gametophyte are here considered central to a "halo" of related taxa with reduced peristomes and gametophytes, some being otherwise reduced genera with special, apparently secondary features adapted to arid climates. Included in this classification are taxa found by the molecular study of Werner et al. (2004) to be almost certainly members of the Pottiaceae, such as Clinclidotus (Clinclidotaceae) and Sphachnobryum (Sphachnobryaceae). On the other hand, Ephemerum (Ephemeraceae) and Goniomitrium (Funariaceae) though suggested as Pottiaceae by Werner et al. (2004) may be excluded by their large, non-papillose leaf cells (elongate in Ephemeraceae) and campanulate calypterae. Although these two taxa are highly reduced in morphological complexity, long-branch attraction and paucity of non-pottiaceous taxa in the data set, together with possible adaptive pressure on the coding sequences used in the analysis may have placed them in the Pottiaceae s.str. of the Werner et al. (2004) cladogram, while their correct placement may be lower in the cladogram (Funariaceae may not be well represented in its entirety by Funaria hygrometrica as outgroup). If further study indicates the contrary, however, the taxon closest to these is Bryoceuthospora. Scopelophila, representing a basal branch, may well belong to the Dicranaceae or Ditrichaceae (it lacks a peristome) and has equivocal morphological features.

Examination of the Werner et al. (2004) data set, graciously provided by Olaf Werner, shows that Anoectangium aestivum and Gymnostomum viridulum differ by seven sites (2 first position, 2 second, 3 third), but the two specimens of Sphachnobryum obtusum (the only species of which two specimens were analyzed) differ by 21 sites (4 first position, 7 second, 10 third). Although the rates of mutational change of sites, especially in the codon, surely differ, it is doubtful that the former two samples are not in the same genus, or the latter is a misidentification for A. handelii. Given that the word "exemplar" implies example representation of the molecular sequence characteristic of other specimens in the species and genus, one might wonder in view of the (unusually?) large internal variation in Sphachnobryum obtusum that the specimens studied might not be samples from a more heterogeneous molecular assemblage at both species and genus level than expected. Erthrophyllopsis fuscula and Erythrophyllastrum andinum differ by only six sites (3 first position, 1 second, 3 third), and, given the uncomfortable similarities of morphology, this second look requires the synonymy of Erythrophyllastrum R. H. Zander as a genus with Erthrophyllopsis Broth.

The best representation of a reliable hypothesis (or theory) of evolution in the Pottiaceae given both morphological and molecular study would be a low-resolution diagram (Fig. 4), similar to the speculative chart of Chen (1941), and based largely on the morphologically oriented NJ tree, but modified by reliable features of the molecular study of Werner et al. (2004), intuitive Gestalt evaluations of Zander (1993), and lastly the UPGMA-Dice cluster analysis. The classification, therefore, advanced here can be summarized as a combination of judicious melding of reliable ("operative") features of the analyses of the mole-
Fig. 4. Evolutionary diagram best combining reliable aspects of morphological and molecular analyses. Bold-faced are names of genera including at least some species with twisted peristomes. Straight lines indicate direct ancestral derivations of the characteristic pottiaceous morphology.

cular and morphological data. Those surprised that more definite lineages are not shown should remember that the molecular cladogram commonly had only one or two reliable internodes to distinguish genera, and the genera are often heterophyletic, i.e., species of one morphologically uncontested genus are scattered among the lineages of a molecular cladogram. Given the fact that there are commonly few informative characters in the Werner et al. (2004) molecular study of one locus, it is likely that reliable branching patterns of species within a genus are superposed across two or more generic branching patterns.
Thus, although it may well be that the genus *Streblotrichum* should be revived and transferred to the Trichostomoideae because *Barbula* species with plane leaf margins are placed there in the molecular cladogram, the separation of these from the *Barbula* species with recurved margins may be an artifact that is correctable with more data, i.e., more genes analyzed (but see Jeffroy et al., 2006). The diagram of Fig. 4 is not a speculation but summarizes reliable elements at the subfamily and tribal levels of past work. In the classification given below, genera with unclear relationships are marked with an asterisk. For full nomenclatural citations, see Zander (1993).

**POTTIACEAE** Schimp., *nom. fam. cons.*

**Timmielloideae** R. H. Zander

*Timmiella* (De Not.) Limpr.

**Trichostomoideae** (BSG) Limpr.


*Aschisma* Lindb.

* *Bryoceuthospora* H. A. Crum & L. E. Anderson

* Calypnerasstrum* Stone

* Calyptopogon* (Mitt.) Broth.

*Chionoloma* Dixon

*Eucladum* Bruch & Schimp.

* *Leptobarbula* Schimp.

* *Neophoenix* R. H. Zander

*Pachyneuropsis* H. Miller

* Pleurochaeae* Lindb.

*Potitopsis* Blockeel & A. J. E. Smith

*Pseudosymblepharis* Broth.

* *Quaesticula* R. H. Zander

**Barbuloideae** (Herz.) Hilp.

**Barbuleae** Herzog


*Anoectangium* Schweegr.

*Barbula* Hedw.

*Bellibarbula* Chen

*Dialytrichia* (Schimp.) Limpr.

*Cinclidotus* P. Beauv.

*Didymodon* Hedw.

* (includes *Kingiobryum* H. Robinson)

*Erythrophyllopsis* Broth.

*(Erythrophyllostomum* R. H. Zander, syn. nov.)*

*Gertrudiella* Broth.

*Gymnostomum* Nees & Hornsch.

*Gyroweisia* Schimp.

*Hymenostyliella* E. B. Bartram

*Hyophiladelphus* (Müll. Hal.) R. H. Zander
Molendoa Lindb. (includes Ozobryum Merrill)

**Bryoerythrophyllaceae** R. H. Zander

*Bryoerythrophyllum* Chen  
*Mironia* R. H. Zander  
*Mirocarpus* R. H. Zander  
*Sarcomeurum* Bryhn

**Bryoerythrophyllum** R. H. Zander

*Pseudocrossidium* Williams  
*Splachnobryum* Müll. Hal.  
*Teniolophora* Reese

**Hymenostylidiaceae** Herz.

*Leptodontioideae* (Herzog) Hilp.

*Hymenostylidium* Brid.  
*Leptodontium* (C. Müll.) Lindb.  
*Rhexophyllum* Herzog

**Pottiaceae** (Limpr.) Broth.

**Pottiaceae** (Limpr.) Dixon

*Tortuloideae* Visotska, *nom. inval.*; **Pottiaceae** Limpr., rank not given.

*Aloinella* Jur.  
*Corydorrhiza* Jur. (includes *Microcrossidium* Stegonia Vent.  
*Guerra* & *Cano*)  
*Globothyrium* Leptonia, syn. nov.


*Acaulon* Müll. Hal.  
*Microbryum* Schimp.

*Microbryum* Schimp.  
*Splachnobryum* Müll. Hal.  
*Teniolophora* Reese

**Acaulon** Müll. Hal.  
*Phascopsis* Stone  
*Pithecospermum* Müll. Hal.

**Chenia** R. H. Zander (includes *Leptophascum* Müll. Hal.) *Guerra* & *Cano*  
*Sagenotorta* R. H. Zander  
*Protobryum* R. H. Zander  
*Saitobryum* R. H. Zander  
*TrachyDock* Herzog  
*Streptopogon* Wils.

**Crumia** Schof.  
*Stonewort* Müll. Hal.

**Dolotortula** R. H. Zander  
*Streptopogon* Wils.

**Hennediella** Par.  
*Syntrichia* Brid.  
*Syntrichia ruralis* (Hedw.)Web. & Mohr.

**Hilpertia** R. H. Zander  
*Streptopogon* Wils.

**Mercyoeidae** Broth.

**Merceyeae** Casares-Gil; **Merceyeae** (Broth.) Chen

*Scopelophila* (Mitt.) Lindb.

**DISCUSSION**

There are two scenarios describing the relationship of *Timmiella* to the Pottiaceae s.str. First, following optimization thinking using maximum parsimony and Bayesian MCMC modeling incremental evolution of morphological and molecular traits, *Timmiella* was elaborated from ancestors that were much reduced, and the Pottiaceae was incrementally
re-elaborated. Second, in an evo-devo scenario, both *Timmiella* and the Pottiaceae are based on one multiplex genetic package the expression of which is controlled by a regulatory gene or promoter sequence, elaborated in the past, then switched off and on again, resulting in the two extant groups that have responded to at least similar selective pressures with the gene cluster that codes for pottiaceous phenotypy. The intermediate groups are phylogenetically discernable by tracking with non-coding sequences but are evolutionarily trivial, where the major evolutionary event is identified by developmentally identical adaptive phenotypes that are not explained by plasticity. Naturally, this hypothesis needs confirmation by developmental analysis of a possible genetic linkage and control.

Of 53 Pottiaceae genera in the Werner et al. (2004) molecular analysis, 12, or 1/4, have twisted peristomes, including *Timmiella* placed distantly near the cladogram base. To avoid a hypothesis of re-evolution of the complex peristome, all 9 reliable internodes from *Timmiella* through the remainder of the distal Pottiaceae clade of the operatively transformed cladogram would have to represent ancestors with twisted peristomes, 6 of these in the Pottiaceae-less-*Timmiella* group. Fig. 1 is the most highly telescoped cladogram, and the true tree or at least a more reliable tree from additional data would have many more internodes, all internodes connecting *Timmiella* and Pottiaceae s.str. necessarily representing ancestors with twisted peristomes. This seems improbable, unless we also postulate that the proportion of ancestral species with twisted peristomes is totally unlike that of modern taxa, and also that the ancestors of genera in the Calymperaceae, Dicranaceae, Ditrichaceae, Grimmiaaceae, and Ptychomitriaceae all had twisted peristomes.

Morphology and associated patterns of morphogenesis may be the key to real understanding of evolutionary relationships. Dollo's Law (Dingle, 2003; Hall, 2003) states that it is improbable (or probabilistically impossible) that particular apparently very highly adaptive complex combinations of traits would re-evolve through gradual accumulation of mutations a second time. This can be countered by demonstrations in the evo-devo literature that complexes of genes can be controlled by one or a few genes or a strongly conserved non-coding promoter sequence, which can control suppression or silencing, and later, perhaps after millions of years, initiate reactivation. Many genes also have been shown to be ancient, either active or potentially resurrected later in a lineage. There is thus an ability of gene complexes and their controlling elements to slide past speciation events, perhaps in part as paralogs or pseudogenes. Taxa with such exaptions (pre-adaptations) to selection pressure would have an evolutionary advantage (Caporale, 2002, 2003b). The many morphological characters of the Pottiaceae that are shared with the more basal *Timmiella* may well be developmentally homologous in that it is improbable that all these would re-evolve gradually and prove to be in combination the pleiomorphic traits of the Pottiaceae.

A parallel situation occurs with humans and orangutans, which share many critical morphological characters (Schwartz, 1988), yet humans are demonstrably more distant from orangutans than from chimpanzees and gorillas on the basis of reliable molecular data, and the morphological traits are apparently atavistic in humans. An explanation that would deal with conflicting morphological and molecular evidence is that there are two phylogenetic patterns that must be recognized: (1) a phylogenetic tree revealed by gradual accumulation of largely neutral mutations in introns, junk DNA, variable sites of exons, or
even evolutionarily neutral genes buffered by phenotypic plasticity, which should track by inference events of genetic isolation following the Biological Species Concept; and (2) an evolutionary tree or network incorporating evidence that massive convergence is probabilistically due to shared fellow-traveler (suppressed in the lineage but potent) homoiologous complexes, and is revealed by morphology determined by activated controlling genes or promoter sequences. At times, both trees may be the same, but contraindications of incremental evolution must be dealt with. The best way to evaluate morphology as revelatory of not-to-be-expected-by-chance-alone evolutionary relationships is to map morphological traits on a molecularly based cladogram and look for complex-trait homoplasy. Then, search for re-evolution of complex traits violating Dollo's Law. Thus, orangutan-homo could be viewed as a theoretical process-based evolutionary sister group while pan-homo is a phylogenetic sister group cast together by inferred nested events of genetic isolation. The question of which to assign greater phylogenetic valence needs further investigation, but should be guided by the need for a classification to be practical and helpful to the majority of users.

Additional context is provided by techniques of phenetic cluster analysis popular in the 1970's, which emphasize similarity alone. The present theoretical explanation of the phenomenon of molecularly based phylogenetic patterns (with high Bayesian posterior probabilities) versus contrary morphological patterns (massive convergence) may be supported by continuing research in development, and if so then phylogenetic analysis will no longer end with simply mapping morphological trait changes on a molecular tree.

Other studies appear to support the importance of silenced genes or gene complexes in evolution. The molecular study of pleurocarpous mosses of Buck et al. (2000) found the trait "peristome lacking" scattered among 78 exemplar taxa in many families. If ancestors were similarly diverse in this character, it is difficult to imagine no reversals. The molecular study of Didymodon (Pottiaceae) by Werner et al. (2005a) found the twisted peristome in several taxa scattered throughout the cladogram. Likewise, in a similar molecular study (Werner et al., 2005b) on the Trichostomoideae (Pottiaceae), among the many taxa involved, the characteristic pottiaceous twisted peristome was present in many of the outgroup taxa, and again at a distance (highly reliable given the many 1.00 BPPs) in the center of the tree in several species of Tortella. In these three studies, although it is possible, it is improbable that all of the deeper branches and up to half of the more recent ones were of peristomate taxa. Although viewing eperistomate taxa as evolutionary dead ends is an attractive explanation, it is not required by theory given the many instances of evolutionary Lazarus taxa in animals and plants listed above.

The functional effect of natural selection on cis-acting regulators on single genes have been recently demonstrated in human evolution (Rockman et al., 2005). A silenced gene cluster is thought to degrade over the passage of time, and if so, it may be that the trait complex may degrade in stages. If so, then one might expect to see a central group of taxa with the intact trait cluster, surrounded by a group of taxa with a gradually reduced trait complex. In fact, we do see this in published cladograms of Didymodon and the Trichostomoideae mentioned above, where taxa or groups of taxa with long, filamentous, twisted peristomes are deeply embedded in the cladogram among related taxa of short, long-trian-
The Pottiaceae s. str. as an evolutionary lazarus taxon

R. H. Zand

regular, straight peristomes, rudimentary peristomes, or none at all. Thus, it may be that the arrangement of a tree core with a halo of reduced forms is the evolutionary group in the Pottiaceae (not necessarily a lineage but a spray or a series of multifurcations).

Genes are more strongly conserved and their sequences vary less with time than do non-coding sequences. Given that genes can be significantly modified by selection or internal pressures such as codon bias (Kellogg & Juliana, 1997), molecularly based phylogenetic trees are here treated as unreliable in decisively selecting among problematic branching patterns with confidence intervals requiring independent and random segregation, at least in deeper branches best informed by genes, and in fact only represent events of genetic isolation. Even the additional data from multiple genes may not prove decisive (Jeffroy et al., 2006).

Given apparent genetic continuity in the lineage from Timmiella to the Pottiaceae s.s., but probabilistic support for the considerable patristic distance between that genus and the Pottiaceae s.s., then the pottiaceous complex of characters necessarily reappeared. Given Dollo’s Law regarding the improbability of major trait complexes re-evolving, it is possible that major portions of the pottiaceous gene complex survived genetic degradation or was repaired while epigenetically silenced on the tree between Timmiella and the Pottiaceae s.s., and was reactivated by selection, possibly involving regulator genes or promoter sequences. There is evidence that the twisted peristome, and perhaps the entire complex of pottiaceous traits, was re-evolved four to seven times. Such an explanation is needed and indeed necessary given the joint improbability of both phyletic distance and morphological similarity.

The results of molecular phylogenetics are in part rejected in this paper because relevant research papers to date (1) model incremental evolution, (2) require correction for unaccounted assumptions, (3) track traits only partly neutral and I.I.D., and (4) when speculative results are converted to operative following Bayesian philosophy of making a bet based on all evidence in light of risk, few reliable solutions to genuine problems at the genus level are demonstrable.

ACKNOWLEDGEMENTS

I appreciate conversations with Stuart McDaniel and Patricia Eckel regarding the topic of this paper. Special appreciation is extended to Olaf Werner for access to the Nexus file of the Werner et al. (2004) study. The Missouri Botanical Garden, Robert E. Magill, Vice-President for Research, is acknowledged for its considerable support over the past four years.

LITERATURE CITED


