



<https://doi.org/10.15407/ukrbotj81.02.087>

Minimally monophyletic genera are the cast-iron building blocks of evolution

Richard H. ZANDER 

Missouri Botanical Garden, 4344 Shaw Blvd, St. Louis, Missouri 63110

Address for correspondence: richard.zander@mobot.org

Abstract. Detailed evaluation is provided for the statistical methods intrinsic to interlocking Sequential Bayes analysis, which allows estimation of evidential support for stem-taxon dendrograms charting the macroevolution of taxa. It involves complexity functions, such as fractal evolution, to generate well-supported evolutionary trees. Required are data on trait changes from ancestral species to descendant species, which is facilitated by reduction of large genera to the smallest included monophyletic groups (one inferred ancestral species each). The genus is here defined as the smallest monophyletic unit, which turns out to be monothetic at least for the direct descendant species. The key fact is that the most-recently acquired traits of the single ancestral species are apparently selectively inviolate and passed on without change to each immediate descendant species. The details of sequential Bayesian analysis were clarified by comparing support of the optimal model with summed support of the alternative models. Because analysis is confined to optimal arrangements of only immediate branches from ancestral species to descendant species, conjugate priors were found to operate such that all alternative models are simply one minus the probability of the optimal model. Such analysis demonstrated that the optimum arrangement of ancestor and descendant species leads to high support values for fitting evolutionary theory, comparable to statistical support levels reported for molecular evolutionary trees, and conjugate priors may be assumed for similar model-building. The method is simple, free of special computer analysis, and well-suited to standard taxonomic practice.

Keywords: adaptation, bryophytes, conjugate priors, evolution, monothetic, minimally monophyletic, reserve ancestor, sequential Bayes

Introduction

Complexity analysis (Prigogine, 1978; Packard, 1988; Binning, 1989; Ito, Gunji, 1994; Ferriere, Fox, 1995; Lewin, 1999; Hilborn, 2000; Kaneko, Tsuda, 2000; Gershenson, 2004; Mesarovic et al., 2004; Liu, Bassler, 2006; Abel, 2009; Bennett, 2010; Doebeli, Ispolatov, 2014; Kondepudi et al., 2020) includes chaos theory, dissipative structure, fractal self-similarity, and self-organization aspects of evolutionary

analysis. Complexity analysis addressed the overwhelming diversity of data on the natural world, and comprises the emergent processes that sustain and constrain that data, presenting taxonomists with species and higher taxa. Given the past range and substance of research on complexity, it is surprising that there have been few (Notale et al., 2000; Lv et al., 2014) direct applications to biodiversity study, one of the most informationally complex fields of scientific endeavor.

ARTICLE HISTORY. Submitted 10 December 2023. Revised 29 March 2024. Published 27 April 2024

CITATION. Zander R.H. 2024. Minimally monophyletic genera are the cast-iron building blocks of evolution. *Ukrainian Botanical Journal*, 81(2): 87–99. <https://doi.org/10.15407/ukrbotj81.02.087>

© M.G. Kholodny Institute of Botany, NAS of Ukraine, 2024

© Publisher PH "Akademperiodyka" of the NAS of Ukraine, 2024

This is an open access article under the CC BY license (<https://creativecommons.org/licenses/by/4.0/>)

Complexity functions, e.g. the edge of chaos, fractal evolution, and logistic map (Pimm, 1984; Packard, 1988; Nicolis, Prigogine, 1989; Schroeder, 1991; Kauffman, 1993, 2000; Ito, Gunji, 1994; Lewin, 1999), are not deduced from axioms or rounded up by reductionist techniques, but are emergent phenomena in the mesocosm. An important recent paper developing the use of complexity analyses in the study of the natural environment is that of Wong et al. (2023), who propose the ubiquity in macro- and microcosmic systems of selection on functional information in a complexity context. They asserted that functional information must increase with degree of function, from zero for no function (or minimal function) to a maximum value corresponding to the number of Shannon informational bits that are both necessary and sufficient to specify any system configuration. Functional information must have three critical characteristics: (1) there is multiple interacting components, (2) elements occur in combinatorically large numbers of different configurations, and (3) selection processes differentially support configurations that display useful functions. These strictures apply well to the present paper, which similarly uses Shannon information bits to specify model configurations.

Interlocking sequential Bayes is a way to judge Bayesian support for the order of a series of elements, and has been used for taxonomic study by Zander (e.g. 2013, 2016, 2018, 2019a, 2019b, 2021a, 2021b, 2023a, 2023b). In the present paper, we want to know how well the data fit evolutionary theory (e.g. Pianka, 2000; Gould, 2002; Barraclough, 2010; and standard works such as that of Grant, 1985), in particular, how well traits of taxa support different orderings of ancestor and descendant species. Evolutionary theory is, of course, complex, but most simply, as here applied, ancestor-descendant order is expected to reflect reasonable interpretation of adaptations (Mayr, 1983) involving reduction and elaboration, given Darwinian gradualism (no or few major jumps in numbers of trait combinations, as in *natura non facit saltus*), reflective of Dollo parsimony (Gould, 1970). Analytic methods include outgroup comparison, in which species of a group may be ordered assuming that a related group, the outgroup, shares traits very similar to those of the ancestor of the ingroup, and these shared traits are primitive or plesiomorphic in cladistic terms. A second method is ingroup comparison such that species with rare or unusual traits are more probably

derived than primitive, and are those of advanced descendant species.

The simplest model for analysis of trait changes is a genus consisting of one ancestral species and its direct descendant species. In the past, a genus was termed monothetic if all species share the same diagnostic traits (Humphreys, Linder, 2009; Sokal, Sneeth, 1963: 113). I have found (Zander, 2023a) that reducing polythetic genera (all species each share by overlap a portion of the generic diagnosis) to a one or more sets each with only one ancestral species makes such sets *both monophyletic and monothetic* (that is, the same diagnosis applies to all species). It does the latter because new traits of the ancestral species are preserved entire in each immediate descendant species, which may have a selective advantage. In the present paper, the genus is narrowly defined as the smallest monophyletic unit.

Van Valen (1973) early pointed out that ancestral species were both extant and common. Ancestral species are mostly ignored in modern taxonomic work because: (1) classical taxonomy commonly assigns species to polythetic genera and relegates trait changes to the intellectual domain of evolutionists, (2) systematics in the cladistic context focuses on clustering taxa by relative degree of shared ancestry, without identifying particular species as ancestors, i.e., all species are terminal on a cladogram, and (3) molecular systematics depends on relative degree of shared apparently non-expressed and apparently randomly fixed molecular sequences, and simply maps expressed traits to the molecular cladogram assumed to track expressed-trait evolution.

The monothetic genus is the minimally monophyletic group, and is most easily dealt with if named as a separate genus, but for convenience a subgenus or informal name may be used. The work of Zander (2023a) summarizes several papers dealing with monothetic genera and trait changes between ancestral species and descendants, and also between ancestral monothetic genera and descendant monothetic genera.

Zander (2023a) found that monothetic genera in the groups studied (in the moss families *Pottiaceae* and *Streptotrichaceae*) were usually of four descendant species, and each species in the genus usually had four newly fixed traits. The genera were fractal, being self-similar across scales. The constraint around the number four (actually averaging about 3.6) was explained using NK-analysis with a random Boolean network model (Kauffman, 1993: 218;

McKelvey, 1999; Gershenson, 2004). The complexity-based optimal one ancestor and four descendants curiously also reflects the Pareto Principle that 20 percent of causes generate 80 percent of effects. This is more precisely reflected in the power law Pareto distribution (Newman, 2005; Hardy, 2010) of $\log_4 5 = \log 5 / \log 4 \approx 1.16$, which is the same as its fractal dimension. A power law is a distribution based on a negative exponent, which results in a "hollow curve" that is the mirror of an exponential curve. The number four was interpreted as the optimal edge of chaos (Packard, 1988) for interactions of competition and mutualism for each genus for survival across geologic time. This includes major perturbations, such as boloid impacts at the KT boundary, the late Cretaceous and early Eocene temperature maxima, continent-level volcanism, sea level fluctuations, and Pleistocene glaciations, and less catastrophic climate change such as Milankovitch events (Behrensmeyer, 1992; Bender, 2013).

Trait changes are grouped as the *novon*, the set of new traits of a descendant species, and the *ancestron*, the set of traits of the ancestral species. The *immediate ancestron* is critical, defined as the set of new traits provided to the ancestral species by its own ancestral species. Of importance is that *the immediate ancestron is passed on as an identical set to all immediate descendant species*. The fact of this latency of the advanced traits of the ancestor is the stabilizing concept for fractal evolution (four species descending from one species gives $\log 5 / \log 4$, or a fractal dimension of 1.16, following Zander, 2023a). This becomes the solid evolutionary substance of a monothetic, monophyletic genus. A definition of a genus as those species sharing an immediate ancestron in the smallest monophyletic unit is effective and productive, and supports the thesis of Wong et al. (2023) that multi-scale complexity is closely associated with selection on functionality.

This latency of the immediate ancestron is the key to the fractal nature of a genus, and provides a clue to the natural-selection-based process supporting survival through adaptation. The immediate ancestron ensures that a descendant species is equipped with tested traits for local or sympatric (Artzy-Randrup, Kondrashov, 2006) and peripatric survival, and the novon is a probe into a constantly changing environment including speciation that is allopatric in geography and across geological time. Theoretically, the immediate ancestron generates a

burst of punctuated evolution, a time-wise stable cluster of strongly adapted species similar to fossil punctuated equilibrium (Eldredge, Gould, 1972). The species as an entity is not, then, the central actor in evolution, but it is the monothetic genus that is a tiny, working Spaceship Earth, one of an integrated fleet of lineages comprising the ecosphere in space-time (Zander, 2023a).

A series of connected monothetic genera or stem-taxon dendrograms (e.g., Fig. 1) is termed a *caulogram*, and is obtained with Shannon-Turing analysis (Zander, 2023a, 2023b). This method assigns each newly evolved trait one informational bit, which is given a probability using an odds table (Table 1). The bits may be added because they are exponents. Log likelihoods are additive; to get likelihood ratios, one simply subtracts the log likelihoods, and if a Bayesian prior is available, these ratios are called Bayes factors. Treating the bits of the posterior of an ancestor as the prior of the next is entirely equivalent to concatenating instances of Bayes' formula, thus we have sequential Bayes. The method (Good, 2011) was pioneered by A. Turing in breaking German codes during World War 2, but is now used (Zander, 2013, 2018, 2021a, 2021b) with Shannon informational bits rather than decibans. The essential statistical elements in the present paper are one ancestor, one or more descendant species, and an outgroup species; and the three elements are rendered as a minimum sequential Bayesian posterior probability (min SBPP), the outgroup providing the prior. The method is interlocking because all elements in a caulogram support all others. This is why a new species can be easily inserted in a proper position in an established caulogram. A more detailed explanation is given by Zander (2023a).

Analyses assign one bit per theoretically advanced trait, a negative bit for a reversal, and no bits for no information pertinent to evolutionary theory involving outgroup selection, and rarity or specialization of traits. Each bit is a power of two; value is the decimal equivalent; odds ratio compares the success of a model over an alternative model; fraction is a value converted from odds ratio; probability is the fraction in decimal form. A number with a negative exponent is the reciprocal of the corresponding number with a positive exponent. Zero bits has a decimal value of zero, an odds ratio of 1:1, a fraction of 1/2, and a probability of 0.500. See spreadsheet (Zander, 2023b) for other values through plus or minus 32 bits.

Table 1. Chart of conjugate priors using Shannon informational bits, odds, and probabilities

Bits-Positive	1	2	3	4	5	6	7	8
Value	2	4	8	16	32	64	128	256
Odds ratio	2:1	4:1	8:1	16:1	32:1	64:1	128:1	256:1
Fraction	2/3	4/5	8/9	16/17	32/33	64/65	128/129	256/257
Probability	0.667	0.800	0.889	0.941	0.970	0.985	0.992	0.996
Bits-Positive	9	10	11	12	13	14	15	16
Value	512	1024	2048	4048	8096	16192	32768	65536
Odds ratio	512:1	1024:1	2048:1	4048:1	8096:1	16192:1	32768:1	65536:1
Fraction	512/513	1024/1025	2048/2049	4048/4049	8096/8097	16192/16193	32768/32769	65536/65537
Probability	0.99805	0.99902	0.9995	0.99975	0.999876	0.999938	0.9999695	0.9999847
Bits-Negative	-1	-2	-3	-4	-5	-6	-7	-8
Value	0.500	0.250	0.125	0.063	0.031	0.016	0.008	0.004
Odds ratio	1:2	1:4	1:8	1:16	1:32	1:64	1:128	1:256
Fraction	1/3	1/5	1/9	1/17	1/33	1/65	1/129	1/257
Probability	0.333	0.200	0.111	0.059	0.030	0.015	0.007	0.004
Bits-Negative	-9	-10	-11	-12	-13	-14	-15	-16
Value	0.002	0.001	0.0005	0.000025	0.0000124	0.00006	0.000031	0.000015
Odds ratio	1:512	1:1024	1:2048	1:4048	1:8096	1:16192	1:32768	1:65536
Fraction	1/513	1/1025	1/2049	1/4049	1/8097	1/16193	1/32769	1/65537
Probability	0.00195	0.00098	0.00049	0.000025	0.000124	0.000062	0.000031	0.000015

The Bayes Formula combines a prior probability with a likelihood (Winkler, 1972; Bernardo, Smith, 1994). The prior gives known information on the model and may be well conceived, or be a flat prior of 0.50. The likelihood is the actual data (flips of coins, numbers of traits). The prior is the initial estimate of the chance that the model explains the data. The likelihood is the chance of the new data explaining the model, i.e. that the model is supported by the data. A posterior distribution (a distribution between 0.00 and 1.00) is generated by combining the prior and likelihood with a normalizing factor

The answer obtained by the Formula is the Bayesian posterior probability (BPP), essentially an updated prior. The normalizing factor, the denominator of Bayes' formula is not needed when the priors are conjugate, as in Table 1, in which Bayesian posterior probabilities of optimal model and all alternative models add to 1.00 (Etz, 2015). Sequential Bayes analysis uses the posterior probability of one instance of the Bayes' formula as the prior of the next instance, sequentially updating the priors. Any Bayesian statistics manual or treatments on the

Web will provide well-illustrated explanations of the use of Bayes' formula in statistics.

A caulogram presents series of monothetic genera with species arranged as best representing evolutionary theory. Because it was found (Zander, 2023a) that the advanced traits of the ancestral species are donated entire to each and every immediate descendant species (i.e., as the latency of the immediate ancestor), there are very few instances where the optimal arrangement of ancestral and descendant species must include exceptions to theory, such as surprising reversals or non-parsimonious trait changes. One assumes there need be a minimum of two correlated new traits to identify a population as a distinct species, otherwise one new trait may be a simple mutation that does not imply a degree of genetic isolation.

Materials and Methods

The methods of recent papers, particularly those of Zander (2013, 2018, 2021a, 2021b) associated with fractal evolution were evaluated by comparing optimal and alternative models of monothetic genera in

terms of Bayesian analysis. The steps of interlocking sequential Bayes analysis involve concepts rather different than those of cladistical analysis. Speciose genera are reduced to monophyletic genera of one ancestral species each. The ancestral species is that which is most similar to an outgroup species in an evolutionary nearby group, and which is also generalist in relation to other members of the genus, which have relatively advanced traits. This is a simple version of cladistic parsimony. Advanced means relatively specialized or uncommon. Traits means those expressed character states that a classical taxonomist may use in a key to or diagnosis of species. Here > means "generates a descendant species", X is an outgroup, A is inferred ancestral species, and B, C, D, etc. are descendant species. A caulogram is a stem-taxon dendrogram given considerable structure by interpolation of ancestral species and their descendant species, the whole is treated as a multichotomous second-order Markov chain (decisions based on the last two nodes).

The number of traits different between the outgroup and the inferred ancestral species is converted to Shannon informational bits, one bit per trait. The same is calculated for the number of traits different between the ancestral species and each of the immediate descendants. The support for a monothetic genus is calculated by adding the bit count of the outgroup to ancestor to the sum of bits distinguishing each immediate descendant from the ancestor (e.g. $X > A$, $+ A > B$, $+ A > C$, etc.), the sum of the bits is converted to Bayesian posterior probabilities using the odds table (Table 1), or a simple formula (1). The bit count of outgroup-to-ancestor is treated as Bayesian prior, and is used just once per monothetic genus (that is, any branch in a dendrogram) but may be summed for any series of species in the same sequence.

$$\text{BPP for positive bits (new traits) in optimal model} = (2^p / 1+2^p), \quad (1)$$

where p = bits supporting $A > B$ or other non-reversal speciation events.

Support for the optimal model — the most parsimonious arrangement of ancestor and descendant species — is calculated by comparing the probability of the optimal model with the summed probabilities of all alternative models.

The alternatives to the most parsimonious model of a monothetic genus are devised by switching the inferred true ancestral species with one of the

descendants, evaluating support for each alternative switch. Such replacement means that the outgroup to false ancestor bits are negative bits, and that one false ancestor to true ancestor bits are also negative because they represent non-parsimonious reversals. Negative means their probabilities are less than 0.50, and can be read from Table 1, or calculated by formula (2). All the bits of all the models are added to provide the probability of all alternative models. A number to a negative exponent is simply the reciprocal (divide into 1) of the number to the same but positive exponent, formula (2).

$$\text{BPP for negative bits (reversals) of alternative models} = (1/2^q) / (1+1/2^q), \quad (2)$$

where q = number of bits at one per trait reversal.

If the priors deal with only alternative models involving switching of ancestor and descendant, then they are conjugate priors and are equivalent to one minus the optimal probability. This can be established by the actual computation of the support for each alternative model. The Bayesian support for the optimal model can be quite high. For examples, see Figs. 1 and 2. Support for secondary generation of descendants from descendants themselves are calculated separately using the ancestor as outgroup. Secondary ancestry may be considered an early stage in formation of a new genus, and is treated as such.

The BPP assigned to a single descendant is the sum of the outgroup to ancestor plus ancestor to descendant bits ($X > A > B$), and is called the minimum sequential Bayesian posterior probability (min SBPP), as given for analyses by Zander (2023a). The monothetic genus as a whole, however, has the outgroup-to-ancestor bits added only once to the genus total. When there are long branching lines of monothetic genera, the bits of all descendant species are added (Etz, 2015) fully across the caulogram because any one determination of ancestor-descendant status supports any other in the lineage, both backwards and forwards. Thus, all bits of each outgroup to ancestor and ancestor to descendant are summed, often resulting in the Bayesian equivalent of a six-sigma standard deviation (Zander, 2023: 19) for a large caulogram.

If an ancestor cannot be selected that is both generalist to the ingroup and similar to the outgroup, then an artificial, working temporary ancestor can be constructed that (1) is intermediate between the outgroup and the extant descendants and (2) lacks the advanced traits of the descendants.

Results

The results of the full Bayesian analysis of interlocking Bayesian sequential analysis using Shannon-Turing statistics are in large part summarized in Figs. 1 and 2. Statistical support for optimal and alternative contrived ancestor-descendant models are compared. Each double lump represents a species with colors representing its novon and immediate ancestor. Set up in Fig. 1 are models of the following scenarios:

1-1. *Two species in a genus*. Each species is composed of two critical sets of traits, a novon of most recent new traits (yellow), and an immediate ancestor composed of the ancestor's new traits (blue). The descendant's immediate ancestor is the same as the ancestor's novon (blue). The alternative model is simply switching ancestor with descendant, and considering trait changes contrary to theory as reversals. The ancestor is generalist and the descendant is rare or more specialized.

1-2. *Two species in a genus plus an outgroup species*. The outgroup provides additional support for the position of the ancestral species of the genus by being more similar to it than to any other species, in this case species A as opposed to species B.

1-3. *Optimal five-species genus with outgroup*. Blue denotes ancestor's novon shared by all four descendants (fractal dimension 1.16, see Zander, 2023: 93). Trait changes occur at every change in color, all in line with evolutionary theory. Alternative model switches ancestor with one of descendants (they all have same number of traits) and there are trait reversals non-parsimoniously contrary to evolutionary theory between outgroup and false ancestor and between false ancestor and true ancestor.

The species in the models are assigned data in Fig. 2. Positive trait changes as positive bits are shown in the optimal model that is evolutionarily parsimonious (no reversals). Negative bits are assigned to reversals of those traits in alternative models. For simplicity, species of the models have four trait changes each in the novon, including the ancestor with four trait changes relative to the outgroup.

2-1. *Two species, data*. Optimal model has +4 bits supporting the model. Alternative model has 4 reversals from false ancestor B (yellow novon, blue immediate ancestor) to true ancestor (blue and green), or -4 bits. Likelihood ratio is likelihood of +4 (0.941) divided by likelihood of -4 bits (0.059),

or 16. Likelihoods add to 100 so these may be taken as conjugate with a flat prior, and scored as BPPs.

2-2. *Two species in a genus plus an outgroup species, data*. There are two sets of trait reversals in the alternative model, between outgroup and descendant switched with ancestor, and between descendant and switched ancestor. Both are critical to evaluation of Bayesian support. Green shows traits shared by outgroup and ancestor of monothetic genus. Optimal model has +4 bits from outgroup-to-ancestor, and +4 bits from true ancestor and true descendant. By sequential Bayes, ancestor-descendant relationship totals +8 bits (0.996). Alternative has -4 bits from outgroup to false ancestor, which adds to -4 bits from false ancestor to true ancestor, totaling -8 bits (0.004). The BPP of the optimal model is 0.0996, which added to the BPP of 0.004 of the alternative model, yields probability 1.00. That the priors are conjugate is demonstrated.

2-3. *Optimal five-species genus with outgroup, data*. This more complex model adds 4 bits for each ancestor-to-descendant plus 4 bits for outgroup-to-ancestor, or 20 bits. There are four alternative arrangements of the main ancestor-descendant order. The one shown (Fig. 2-3) results in -8 bits for the reversals between outgroup and false ancestor, and false ancestor and misplaced correct ancestor. To this is added -12 bits of reversals from the other three possible switches of descendant and ancestor to yield -20 bits. The very high BPP of the optimal model (0.999999046) adds to the very low BPP of the summed alternative models (0.000000954) to get probability 1.00, thus the priors are conjugate for analysis of this and similar models.

A test case from nature

The interlocking Shannon-Turing analysis may have its Bayesian statistics further clarified with an actual monothetic genus in the moss family *Pottiaceae*. *Tainoa* R.H. Zander is a small genus of six moss species endemic to the West Indies and adjacent Central America and Mexico. Trait details are given by Zander (2023a). Five species were investigated using the present bit-summing method. The species involved are here assigned a letter and number of bits reflecting its number of newly evolved traits:

X (outgroup) is *Neotrichostomum crispulum* (Buch) R.H. Zander;

A (putative progenitor species) is *Tainoa pygmaea* (E.B. Bartram) R.H. Zander 3 bits;

B is *T. sinaloensis* (E.B. Bartram) R.H. Zander 5 bits;

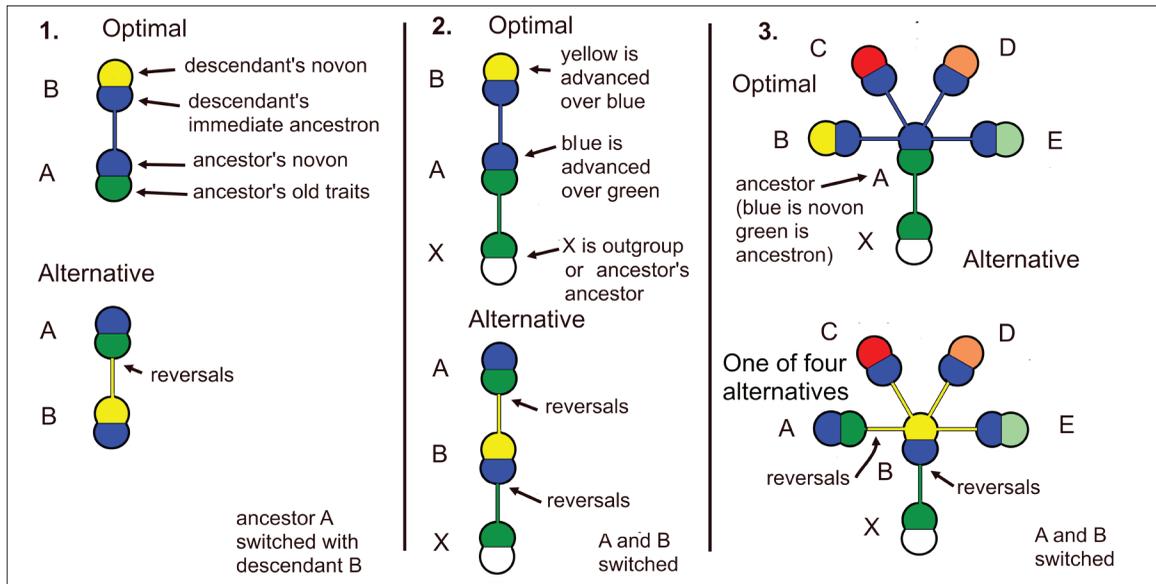


Fig. 1. Optimal and alternative models of one-ancestor genera: 1. Genus of two species. Each species consists of two evolutionarily effective parts. Shared sets of traits shown in color. 2. Genus of two species with added outgroup, reversal of traits occurs twice. 3. Optimal five-species genus plus outgroup compared with one of four alternative models involving switching ancestor with descendant

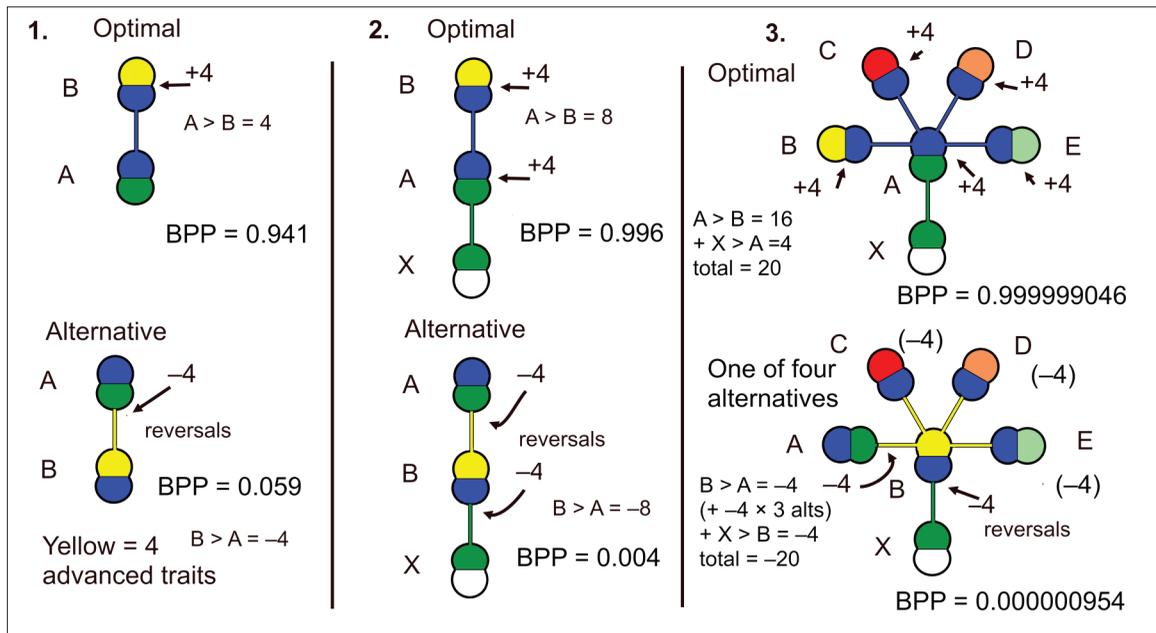


Fig. 2. Calculations of BPP demonstrating conjugate priors (add to 1.00) with four trait changes assigned to all species: 1. Four traits inferred as new in descendant for +4 bits, four reversals in alternative model are -4 bits. Likelihoods add to 1.00. 2. Two species in genus plus outgroup. New traits add to +8, reversals in alternative model add to -8, BPPs add to 1.00, and the priors are conjugate. 3. Optimal five-species model plus outgroup adds 4 bits for each ancestor to descendant plus 4 bits for outgroup to monothetic genus (represented by ancestor) results in +20 bits, or very high BPP supporting this model. Alternative model is one of four switches of ancestor with descendant, with two -4 bit reversals. Adding reversals from three other ancestor-descendant switches, plus the outgroup-ancestor switch, gives -20 bits. The BPP add to 1.00

C is *T. subangustifolia* (Thér.) R.H. Zander 3 bits;
D is *T. subcucullata* (R.S. Williams) R.H. Zander
4 bits;

E is *T. bartramiana* (Steere) R.H. Zander 3 bits.

The evolutionary formula for the above monothetic genus *Tainoa* is $X > A > (B, C, (D > E))$, as given by Zander (2023a).

The analysis of Bayesian support for the model presented by the evolutionary formula is the same as given in Fig. 2-3, except for the bit count and one descendant (E) is a secondary ancestor to its own descendant (D). The evolutionary transition from X to A is +3 bits. This number is added to the sum of the bits distinguishing the three immediate descendant species, that is, bits for $A > B$, $A > C$ and $A > D$, or $5 + 3 + 4$ bits, 12 bits. Total for the monothetic genus *Tainoa* is 15 bits, or a BPP of 0.9999695 (from Table 1), as restricted to the genus represented by the ancestor and immediate descendants. Given conjugate priors, the probability of all alternative models based on re-arrangements of the immediate ancestor-descendant species is one minus the BPP of the optimal model.

The secondary ancestry of $D > E$ is calculated separately, as if it were a separate genus (which it may become in time). With the ancestor A as outgroup prior, $A > D$ is 3 bits (A is zero in this case of min SBPP), $D > E$ is 3 bits, total bits for the secondary ancestry is then 6 bits, or a BPP of 0.985. We assume conjugate priors. The secondary ancestry contributes some uncertainty to the accuracy of the complete model of *Tainoa*. This is calculated by multiplication of the probabilities, or $0.9999695 \times 0.985 = 0.985$. The final BPP of *Tainoa* including the secondary ancestry of $D > E$ is apparently then much dependent on the uncertainty of the order of evolution of the secondary ancestry. The immediate branching is statistically certain, however, and the secondary is acceptably dependable.

In words, species A is very well supported as ancestral to the remaining species by its A's strong distinction and yet close similarity to the outgroup X, while the other species are removed from this relationship by advanced traits of 3 to 5 bits. Interestingly, in the case of secondary ancestor D, first ancestor A becomes the prior for $D > E$. Secondary ancestors apparently serve to distance their own descendants from the phyletic constraint of the outgroup, and probably signal a genus changing through selection. However, its immediate ancestor does not change and remains of survival

advantage. An example of such a transitioning genus is *Anoetangium* Hedw., with two secondary ancestral species (Zander, 2019b), to be discussed in a future paper.

Discussion

The Bayesian formula has two parts, one is the likelihood (the actual data) and the other is the prior (reflecting previous knowledge of probability of the model, or if none, then 0.50 probability). If there are two concatenated monothetic genera, then the first is the prior of the second. Bit values for the first may be added to those of the second giving logically the same result as does the Bayes formula (formula 3), and the likelihood is then called the conditional probability. The bit count in the above analyses are equivalent to conditional probabilities or $P(\text{data}|\text{model})$, see formula (3).

$$P(\text{model}|\text{data}) = \\ = P(\text{model}) \times P(\text{data}|\text{model}) / P(\text{data}) \quad (3)$$

In Bayes' formula the probability of the model given the data (here the particular evolutionary diagram) equals the prior (initial probability of the model) times the probability of the data given the model (the likelihood), that divided by the probability of the data (a normalizing function that scales the value between zero and one). In short, the posterior probability equals the likelihood times the prior divided by the normalization constant. The posterior probability, $P(\text{model}/\text{data})$, is how likely is the model given the data. The likelihood, $P(\text{data}|\text{model})$, is the probability of seeing the data given the evidence. The normalizing constant is unnecessary if the prior and likelihood are conjugate priors, that is, having the same statistical distributions. One can then use Bayes' Rule, that the posterior is the likelihood multiplied by the prior (Etz, 2015), to simplify calculations. This is done in sequential Bayes by adding bits (exponents of 2). Bayesian statistical analysis in general can be difficult and mind-bending in logical and mathematical complexity, and has always been a battleground between Fisherian, Neyman-Pearson and Bayesian schools of statistics (Gigerenzer et al., 1989), but the present method is extremely simplified.

Fate of survival-neutral traits

This paper has emphasized the evolutionary importance of the novon of new traits and immediate

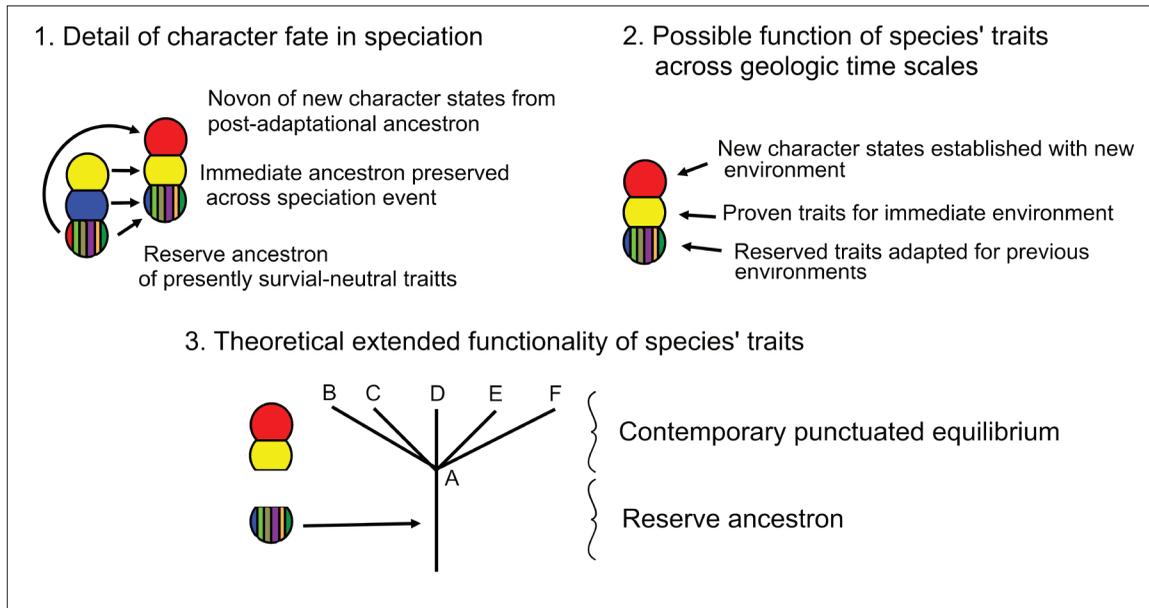


Fig. 3. 1. Detail of character states in speciation. A third element, the reserve ancestor, added to each species representing trail of traits valuable for past survival. 2. Possible function of species' traits across geologic time scales. 3. Theoretical extended functionality of species' traits. Survival optimality of novon and immediate ancestor sustains bursts of long-stable speciation. Remainder of ancestor (parti-colored) acts as reserve for ecosystem survival, and is (under) represented by the thin lines connecting species in a caulogram

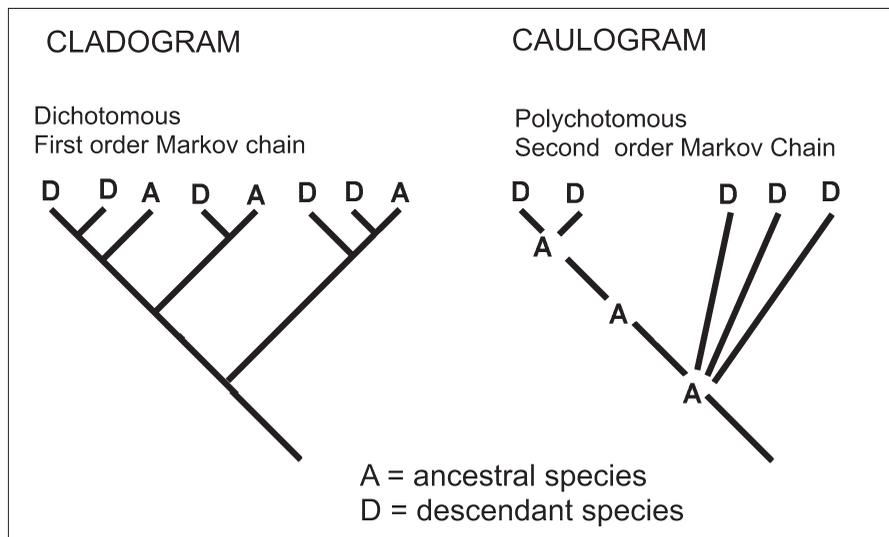


Fig. 4. Comparison of cladogram and caulogram. The cladogram (left) has both descendant and ancestral species placed terminal on the branches, and is dichotomous and a first-order Markov chain. A cladogram allows only likelihood analysis or the essentially identical Bayesian analysis with a flat prior. The caulogram (right) inserts inferred ancestral species at the nodes, which allows informative priors in sequential Bayesian analysis. This evolutionary dendrogram is multichotomous and a second-order Markov chain

ancestron of shared new traits of the ancestral species. Here it is conceived that the narrow lines in a dendrogram connecting species and genera (Fig. 3) represent all the traits of a species that are apparently not critical for survival in the present regional or local environment. One may speculate, however, that relic traits are important to the ecosystem as having been active in survival during environmental perturbations of the geologic past. As subordinate character states, they are the source of all new character states in newly evolved descendant species. Such highly adaptive but presently inactive traits in what we might call the *reserve ancestron* are not used up lightly.

If the branching order of evolution were indeed dichotomous, as in a cladogram, each speciation event would eliminate, by state changes, two to four traits from the post-adaptational reserve ancestron of the lineage. The turn-over of reserve traits is different with monothetic genera, however, in that the lineage links series of ancestral species, optimally (from complexity analysis) each with four descendants. In this case, the new traits of the descendants are sampled from the reserve traits but all new traits are different in each descendant.

Caulograms of West Indian genera (Zander, 2023a) show clearly bursts of speciation apparently intact throughout the 25 million years of existence of this island archipelago (Ricklefs, Berminham, 2007). The caulograms are only one or two linked ancestral species in depth, implying that recent geographically restricted genera have changed little since derivation from globally widespread ultimate ancestors at the caulogram base. One might theorize that monothetic genera are frozen in stasis after an initial burst of speciation, which would be a contemporary equivalent to fossil punctuated equilibrium. Survival of the descendants as a lineage of genera in stasis ensures that at least a portion of the reserve ancestron remains long in reserve states such that it is the lineage of the genus that is the important unit of evolution over geologic time.

Problems with modern phylogenetic systematics

The present paper offers interlocking Bayesian sequential analysis as an alternative to modern phylogenetic methods of evolution-based taxonomy, even if the statistical analysis is similar. Likelihood assumes the model — the cladogram as informative about evolution — is correct. To a major extent, molecular phylogenetic analysis does uncover plausible

shared ancestry. On the other hand, no species are singled out as ancestral to others, and relative shared ancestry is rather empty of clues to evolutionary processes. Evolutionary systematics holds, on the contrary, that the model must reflect ancestor-descendant relationships as determined by studies that reflect the established premises of evolutionary theory. Appropriate theory (Artzy-Randrup, Kondrashov, 2006; Barraclough, 2010; Lewontin, 1978; Mayr, 1983; Schneider, 2000) involves adaptation, chance of reversals, rarity and specialization of traits, non-saltational changes, and other elements used in generating optimal evolutionary trees.

I have critiqued phylogenetic systematics at length in past papers (e.g. Zander, 2013, 2019a, 2019b, 2023a). A short summary of the most important unaddressed problems is provided here. (1) A cladogram is not an evolutionary tree. It was intended as a clean slate on which to present the results of cluster analysis, both overall similarity and by synapomorphies. (2) All ancestral species are considered extinct and all taxa are placed at the ends of cladogram branches; this leaves little room for differently diagnosed implied shared ancestors at cladogram nodes (Fig. 4). (3) In past work I have found about half of species studied are ancestral to one or more other species, that is, only half are actually terminal on an evolutionary dendrogram, the remainder are identical with a more basal node. This means that cladograms are evolutionarily inaccurate by at least one node half the time, with one species not the sister of the nearest. (4) Because molecularly segregated families, genera and species are not clearly delimited by evolutionarily correlated expressed traits, there is no evidence of process-based evolution. (5) The principle of holophyly is used to lump some large and small taxa otherwise well-defined morphologically, while some other taxa are split when mapped in pieces to a molecular cladogram. Thus stability of nomenclature is now much compromised.

Conclusions

Complexity entails envisioning new, over-arching processes not easily derived from known physics, and analogic conceptions can clothe poorly understood trends and biases in nature with form and function. Sets of monothetic genera may be analogically linked together as strongly coherent and well-fitting jigsaw puzzle pieces. The illustration

locked in place on the analogous jigsaw puzzle is that of evolutionary theory regarding adaptive speciation. Complexity analysis of evolution, given the power of the latency of the immediate ancestor (advanced traits of ancestral species transferred entire to descendant species), is also like solving a maze, a NP-complete (non-deterministic polynomial-time-complete) problem (Garey, Johnson, 1979; Poundstone, 1988: 164), by exploring all paths at once to find the exit. Analogically this may be accomplished either by modeling the maze as branching tube, flooding with a hose and sending a cork through, or as stream channels and following the fastest flow in a boat. That flood is metaphorically the immediate ancestor.

The present environmental crisis needs advice from the systematic community. The public funds our large multi-million-specimen herbaria and faunal collections. Molecular cladograms cannot predict the edge-of-chaos actions of natural processes. Ancestor-descendant cladograms based on actual trait changes that probably reflect adaptations to environmental perturbations can do so. Lineages are multimillion-year data sets, and expressed trait changes might be mapped as adaptations to major perturbations such as changes in global temperature and extinction events in the past, as per discussion by Zander (2023a).

The critical fact enabling ease of interlocking sequential Bayesian analysis is the observation (Zander, 2023a), in the smallest monophyletic group, that the new traits of the descendant are not derived from the important, most recent traits of the ancestor (the immediate ancestor) but from older, long-established traits in the train of characters of the ancestor (Fig. 3), that is, the results of ancestral selection across multiple perturbation events. In other words, the new traits of the ancestral species are preserved in the descendant species, and the new traits of the descendant species are

modified character states that the ancestral species apparently now no longer needs for differential survival. An evolutionary dendrogram reflecting outgroup-to-ancestor prior information of allows statistical analysis that more accurately evaluates the coherence of the resulting evolutionary relationships with full respect for modern, hard-won evolutionary theory.

Wong et al. (2023) have raised the role of complexity analysis into importance at micro-, meso- and macrocosmic levels, emphasizing selection on functional attributes as primary in sustaining and constraining evolving systems. The present study focuses on one complexity function, the latency of the immediate ancestor, which is hypothesized as extensible across scales by fractal evolution throughout evolving life.

Acknowledgements

I acknowledge with gratitude the valuable comments of anonymous reviewers. The Missouri Botanical Garden (St. Louis, MO, USA) is saluted as it continues to support major research in biodiversity studies.

ETHICS DECLARATION

There is no actual or potential conflict of interest with other persons or institutions.

No funding supported this work.

SUPPLEMENTARY MATERIAL

This article includes supplementary material, Table S1: [ukrbotj81-02-087-S1.xlsx](https://doi.org/10.3390/ijms10010247) (13 KB).

ORCID

R.H. Zander:  <https://orcid.org/0000-0003-3676-2667>

REFERENCES

- Abel D. 2009. The capabilities of chaos and complexity. *International Journal of Molecular Science*, 10: 247–291. <https://doi.org/10.3390/ijms10010247>
- Artzy-Randrup Y., Kondrashov A.S. 2006. Sympatric speciation under incompatibility selection. *Proceedings of the National Academy of Sciences*, 103: 11619–11624.
- Barracough T.G. 2010. Evolving entities: Towards a unified framework for understanding diversity at the species and higher levels. *Philosophical Transactions of the Royal Society B*, 365: 1801–1813. <https://doi.org/10.1098/rstb.2009.0276>
- Behrensmeyer A.K. 1992. *Terrestrial Ecosystems through Time*. Chicago: Chicago University Press, 588 pp.
- Beneder M.L. 2013. *Paleoclimate*. Princeton: Princeton University Press, 306 pp.
- Bennett K. 2010. The chaos theory of evolution. *New Scientist*, 13 October 2010. Available at: <https://www.newscientist.com/article/mg20827821-000-the-chaos-theory-of-evolution/> (Accessed 30 June 2022).

- Bernardo J.M., Smith A.F.M. 1994. *Bayesian Theory*. New York: John Wiley & Sons, 586 pp.
- Binning G. 1989. The fractal structure of evolution. *Physica D: Nonlinear Phenomena* 38: 32–36.
- Doebeli M., Ispolatov I. 2014. Chaos and unpredictability in evolution. *Evolution*, 68: 1365–1373.
- Eldredge N., Gould S.J. 1972. Punctuated equilibria: An alternative to phyletic gradualism. In: *Models in Paleobiology*. Ed.: T.J.M. Schopf. San Francisco: W. Freeman, pp. 82–115.
- Etz A. 2015. Understanding Bayes: Updating priors via the likelihood. *The Etz-Files*. Available at: <https://alexanderetz.com/2015/07/25/understanding-bayes-updating-priors-via-the-likelihood/> (Accessed 12 October 2023).
- Ferriere R., Fox G.A. 1995. Chaos and evolution. *Trends in Ecology and Evolution*, 10: 480–485.
- Garey M.R., Johnson D.S. 1979. *Computers and Intractability: Guide to the Theory of NP-Completeness*. New York: W.H. Freeman, 338 pp.
- Gershenson C. 2004. Introduction to random Boolean networks Workshop and Tutorial Proceedings. In: *Ninth International Conference on the Simulation and Synthesis of Living Systems (ALife IX)*. Eds: M. Bedau, P. Husbands, T. Hutton, S. Kumar, H. Suzuki pp. 160–1733. Also arXiv:nlin/0408006, <https://arxiv.org/abs/nlin/0408006> (Accessed 14 November 2022).
- Gigerenzer G., Swijtink Z., Porter T., Daston L., Beatty J., Krüger L. 1989. *The Empire of Chance*. Cambridge: Cambridge University Press, 360 pp.
- Good I. 2011. *A List of Properties of Bayes-Turing Factors*. National Security Agency Report DOCID: 3838681. Declassified 9 March 2011. Available at: <https://www.nsa.gov/portals/75/documents/news-features/declassified-documents/tech-journals/list-of-properties.pdf> (Accessed 17 November 2021).
- Gould S.J. 1970. Dollo on Dollo's Law: irreversibility and the status of evolutionary laws. *Journal of the History of Biology*, 3: 189–212.
- Gould S.J. 2002. *The Structure of Evolutionary Theory*. Cambridge: Belknap Press of Harvard University Press, 1464 pp.
- Grant V. 1985. *The Evolutionary Process*. New York: Columbia University Press, 487 pp.
- Hardy M. 2010. Pareto's Law. *Mathematical Intelligencer*, 32(3): 38–43. <https://doi.org/10.1007/s00283-010-9159-2>
- Hilborn R.C. 2000. *Chaos and Nonlinear Dynamics*. 2nd ed. Oxford, U.K.: Oxford University Press, 672 pp.
- Humphreys A.M., Linder H.P. 2009. Concept versus data in delimitation of plant genera. *Taxon*, 58: 1054–1074.
- Ito K., Gunji Y.P. 1994. Self-organization of living systems towards criticality at the edge of chaos. *Biosystems*, 33: 17–24.
- Kaneko K., Tsuda I. 2000. *Complex Systems. Chaos and Beyond, A Constructive Approach with Applications in the Life Sciences*. New York: Springer, 273 pp.
- Kauffman S.A. 1993. *The Origins of Order, Self-Organization and Selection in Evolution*. New York: Oxford University Press, 728 pp.
- Kauffman S.A. 2000. *Investigations*. Oxford: Oxford University Press, 300 pp.
- Kondepudi D.K., De Bari B., Dixon J.A. 2020. Dissipative structures, organisms and evolution. *Entropy*, 22(11), 1305. <https://doi.org/10.3390/e22111305> (Accessed 30 June 2022).
- Lewin R. 1999. *Complexity: Life at the Edge of Chaos*. 2nd ed. Chicago: University of Chicago Press, 242 pp.
- Lewontin R. 1978. Adaptation. *Scientific American*, 239(3): 213–230.
- Liu M., Bassler K.E. 2006. Emergent criticality from coevolution in random Boolean networks. *Physical Review E*, 74(4 Pt 1): 041910. <https://doi.org/10.1103/PhysRevE.74.041910>. Epub 17 Oct 2006.
- Lv X., Wu Y., Ma B. 2014. The fractal dimension of the tree of life. *PeerJ Preprints*, 23 January 2014. Available at: <https://peerj.com/preprints/198/> (Accessed 30 June 2022).
- McKelvey B. 1999. Self-organization, complexity catastrophe, and microstate models at the edge of chaos. In: *Variations in Organization Science: In Honor of Donald T. Campbell*. Eds: J.A.C. Baum, B. McKelvey. SAGE Publications, Thousand Oaks, pp. 279–307.
- Mayr E. 1983. How to carry out the adaptationist program? *American Naturalist*, 121: 324–334.
- Mesarovic M.D., Sreenath S.N., Keene J.D. 2004. Search for organizing principles: understanding in systems biology. *Systems Biology (Stevenage)*, 1: 19–27. <https://doi.org/10.1049/sb:20045010>
- Newman M.E.J. 2005. Power laws, Pareto distributions and Zipf's law. *Contemporary Physics* 46: 5, 323–351. <https://doi.org/10.1080/00107510500052444>
- Nicolis G., Prigogine I. 1989. *Exploring Complexity: An Introduction*. New York: W.J.H. Freeman and Company, 313 pp.
- Nottale L., Chalime J., Grou P. 2000. On the fractal structure of evolutionary trees. In: *Fractals in Biology and Medicine, Vol. III*. Eds: E.G. Lusa, T. Merlini, T. Nonnenmacher, E. Weibel. Switzerland, Basel: Weibel Birkhäuser Verlag, pp. 247–258.
- Packard N.H. 1988. *Adaptation towards the Edge of Chaos*. Illinois, Urbana: University of Illinois at Urbana-Champaign, Center for Complex Systems Research.
- Pianka E.R. 2000. *Evolutionary Ecology*. Ed. 6. San Francisco: Addison Wesley Longman, Inc., 528 pp.
- Pimm S. 1984. The complexity and stability of ecosystems. *Nature*, 307: 321–326.
- Poundstone W. 1988. *Labyrinths of Reason*. New York: Doubleday, Anchor Books, 274 pp.
- Prigogine I. 1978. Time, structure and fluctuations. *Science*, 201: 777–785.
- Ricklefs R., Bermingham E. 2007. The West Indies as a laboratory of biogeography and evolution. *Philosophical Transactions of the Royal Society B*, 363: 2393–2413.

- Schneider C. J. 2000. Natural selection and speciation. *Proceedings of the National Academy of Sciences*, 97: 12398–12399.
- Schroeder M. 1991. *Fractals, Chaos, Power Laws: Minutes from an Infinite Paradise*. New York: W.H. Freeman and Company, 429 pp.
- Sokal R.R., Sneath P.H.A. 1963. *Principles of Numerical Taxonomy*. San Francisco: W.H. Freeman, 359 pp.
- Van Valen L. 1973. A new evolutionary law. *Evolutionary Theory*, 1: 1–30.
- Winkler R.L. 1972. *An Introduction to Bayesian Inference and Decision*. N.Y.: Holt, Rinehart and Winston, Inc. 563 pp.
- Wong M.L., Cleland C.E., Arend Jr. D., Bartlett S., Cleaves II, H.H., Demarest J., Prabhu A., Lunine J.L., Hazen R.M. 2023. On the roles of function and selection in evolving systems. *Proceedings of the National Academy of Sciences*, 120(43): e2310223120. <https://doi.org/10.1073/pnas.2310223120> (Accessed 17 October 2023).
- Zander R.H. 2013. *Framework for Post-Phylogenetic Systematics*. U.S.A., St. Louis: Zetetic Publications, 209 pp.
- Zander R.H. 2016. Macrosystematics of *Didymodon* sensu lato (*Pottiaceae*, *Bryophyta*) using an analytic key and information theory. *Ukrainian Botanical Journal*, 73: 319–333. <https://doi.org/10.15407/ukrbotj73.04.319>
- Zander R.H. 2018. *Macroevolutionary Systematics of Streptotrichaceae of the Bryophyta and Application to Ecosystem Thermodynamic Stability. Edition 2*. U.S.A., St. Louis: Zetetic Publications.
- Zander R.H. 2019a. Macroevolutionary versus molecular analysis: Systematics of the *Didymodon* segregates *Aithobryum*, *Exobryum* and *Fuscobryum* (*Pottiaceae*, *Bryophyta*). *Hattoria*, 10: 1–38.
- Zander R.H. 2019b. Macroevolutionary evaluation methods extended, consolidated, and exemplified with *Anoetangium* (*Pottiaceae*) in North America and the Himalayas. *Annals of the Missouri Botanical Garden*, 104: 324–338.
- Zander R.H. 2021a. Evolutionary leverage of dissilient genera of *Pleuroweisiaceae* (*Pottiaceae*) evaluated with Shannon-Turing analysis. *Hattoria*, 12: 9–25.
- Zander R.H. 2021b. Synopsis of *Ozobryum* (*Pottiaceae*, *Bryophyta*), and sequential Bayes evaluation of genus integrity. *Annals of the Missouri Botanical Garden*, 106: 458–468.
- Zander R.H. 2023a. *Fractal Evolution, Complexity and Systematics*. U.S.A., St. Louis: Zetetic Publications. Available at: https://www.researchgate.net/publication/372885638_Fractal_Evolution_Complexity_and_Systematics
- Zander R.H. 2023b. Spreadsheet for Likelihood Ratios in Interlocking Sequential Bayesian Analysis. *Res Botanica Technical Report*, 02 October 2023. Available at: <https://www.preprints.org/manuscript/202310.1653/v1>. <https://doi.org/10.20944/preprints202310.1653.v1>

Мінімально монофілетичні роди є основоположними структурними блоками еволюції

Р.Г. ЗАНДЕР

Міссурійський ботанічний сад, Сент-Луїс, Міссурі, США

Реферат. У статті наведено детальну оцінку статистичних методів взаємопов'язаного послідовного байєсівського аналізу, що дозволяє оцінити доказову підтримку для дендрограм зв'язків таксонів, які відображають макроеволюцію цих таксонів. Цей аналіз включає функції складності, такі як фрактальна еволюція, для створення еволюційних дерев з високою підтримкою. Для нього потрібні дані про зміни ознак від предкових видів до видів-нащадків, що сприяють зменшенню великих родів до найменших складових монофілетичних груп (по одному гіпотетичному предковому виду для кожної). Рід визначається тут як найменша монофілетична одиниця, яка виявляється монотетичною принаймні для видів, які є безпосередніми нащадками предкового виду. Ключовим є те, що нещодавно набуті ознаки одного предкового виду є, очевидно, вибірково непорушеними і передаються незмінними кожному виду його безпосередніх нащадків. Деталі послідовного байєсівського аналізу уточнювали шляхом порівняння підтримки оптимальної моделі з сумарною підтримкою альтернативних моделей. Оскільки аналіз був обмежений оптимальним розміщенням лише безпосередніх відгалужень від предкових видів до видів-нащадків, було виявлено, що для спряжених апіорних розподілів усі альтернативні моделі дорівнюють одиниці мінус ймовірність оптимальної моделі. Такий аналіз продемонстрував, що при оптимальному розміщенні видів-предків і їхніх нащадків показники підтримки еволюційної теорії є високими, співставними зі статистичними рівнями такої підтримки, що наводились для молекулярних еволюційних дерев, і спряжені апіорні розподіли можуть бути обрані для побудови подібних моделей. Цей метод є простим, не потребує спеціального комп'ютерного аналізу і добре підходить для стандартних таксономічних досліджень.

Ключові слова: адаптація, еволюція, мохоподібні, мінімально монофілетичний, монотетичний, об'єднані попередники, послідовний Байєс, резервний предок