

MISCELLANEOUS PUBLICATIONS
MUSEUM OF ZOOLOGY, UNIVERSITY OF MICHIGAN, NO. 209(3): 1-26

DIMENSIONS OF SEDIMENTARY LITHOTOPES
AND TAXONOMIES OF FISHES

By

BRUCE H. WILKINSON, CARL N. DRUMMOND
AND LINDA C. IVANY

JOHN LUNDBERG, EDITOR
MACKENZIE SCHONDELMAYER, COMPOSITOR

Department of Ecology and Evolutionary Biology and Museum of Zoology
University of Michigan
Ann Arbor, Michigan 48109-1079, USA

Ann Arbor, December 23, 2021
ISSN 0076-8406

DIMENSIONS OF SEDIMENTARY LITHOTOPES AND TAXONOMIES OF FISHES

By

BRUCE H. WILKINSON¹, CARL N. DRUMMOND²
AND LINDA C. IVANY³

ABSTRACT

The size of subgroups among larger taxonomic units, as measured by the number of taxa within them, is a metric of fundamental importance to the appreciation of causes of change in biodiversity in both time and space. Central to such evaluations is an understanding of the expected and observed variation in the numbers and sizes of groups comprising various taxonomic levels. Here we show that numbers of fish taxa within subdivisions (memberships) of any supertaxon in a Linnaean taxonomy are virtually identical to areas of patches of like sediment (lithotopes) that are distributed across various depositional surfaces. Both sedimentary surfaces and Linnaean taxonomies are closely approximated by functions that generally describe random divisions of geographic and/or shape-space. We describe a 'broken plate' model for taxonomic membership that is akin to Robert MacArthur's (1957) classical 'broken stick' model for abundance distributions, where species abundances in an ecosystem are described by an exponential function of abundance (segment length) frequencies reflecting the random subdivision of resources. In a taxonomic context, the broken plate presumes that the amount of morphospace realized at any taxonomic level is proportional to the numbers of subtaxa of which it is comprised. A hypothetical transect across the morphospace associated with any higher taxon would comprise a 'broken stick', or exponential, distribution of square roots of the number of contained subtaxa. Taxonomic membership (occupied morphospace) within the higher taxon is therefore randomly partitioned among subtaxa, analogous to the sizes of fragments of the broken plate. Thus, just as the broken stick distribution is well-described using only the length of the stick and the number of segments into which it is broken, the partitioning of taxa into subtaxa within any supertaxon is random and adequately described using only the number of taxa and the number of subtaxa into which they are partitioned.

Such 'broken plate' functions yield excellent agreement for membership partitioning among classes, orders, families, and genera of fishes. Quantification across all taxonomic levels provides several insights related to the biodiversity of this important group: (1) Membership of taxonomic groups of fishes is self-similar among all levels of Linnaean division (e.g., families per order, genera per family, species per genus) and is almost entirely independent of levels of taxonomic separation between groups being considered, with an average of seven to eight members within any taxonomic group. (2) The 'broken plate' representation implies that divisions within one taxonomic level are independent of all other divisions; a similar partitioning of species among genera belonging to both diverse and depauperate families supports the supposition that little 'memory' exists between levels of taxonomic membership. (3) Special explanations for the generation of apparently extreme polytypy may be largely unnecessary; taxonomic diversities expected from the 'broken plate' model suggest that observed disparity in numbers of fish species comprising many clades is no greater or less than one would expect from a random fragmentation of morphospace.

¹Department of Earth and Environmental Sciences, University of Michigan, Ann Arbor, MI 48109-1005, eustasy@umich.edu

²Department of Physics, Purdue University Fort Wayne, Fort Wayne, IN 46805

³Department of Earth and Environmental Sciences, Syracuse University, Syracuse, NY 13244

INTRODUCTION

This paper is more about fish classification than about sizes of sedimentary units, but strives to illuminate a number of philosophic and process-based linkages between areas occupied by bodies of sediment and sedimentary rock, and memberships of Linnaean taxonomic units. The history, current state, and immediate future of biodiversity is obviously a subject of fundamental importance to paleontologists and neobiologists alike, as this aspect of the history of life directly reflects the nature of biologic responses to states and changes in the Earth's environment. Taxonomic memberships have also been widely used to estimate numbers of yet-to-be-described groups (e.g. Joppa et al. 2011; Mora et al. 2011) as well as the biodiversity of species and genera across a variety of earth-surface biotopes (e.g. Bertrand, et al., 2006; Ricotta et al., 2002; Mazaris et al., 2010; Callaghan et al., 2021, Moura and Jetz 2012).

The principal metric for the study of such subjects has largely been based on the numbers and distributions of groups comprising various levels of the Linnaean system of taxonomic classification. Although the history and state of biodiversity has been measured at the family (e.g., Benton, 1995), genus (e.g., Alroy et al, 2008), and species (e.g.,

Kallimanis et al., 2012) level, a considerable diversity of opinion exists as to the validity and utility of various levels of classification, as well as to the rationality and usefulness of such taxonomic divisions as metrics of clade success. In light of the span of these perspectives, this communication begins by briefly reviewing the current state of fish taxonomy as summarized in one of the current databases on modern organisms. It then proceeds to apply a size metric derived from surface mapping of modern sedimentary lithotopes to data on membership abundances within various taxonomic units of fishes, and proposes that similarities in size of lithotope as measured by area, and amount of taxonomic membership as measured by numbers of subtaxa within any taxonomic group, largely reflect the more-or-less random division of geographic and morphologic spaces, respectively. It concludes by considering several issues related to the biologic diversity of fishes that might be addressed from this analysis and understanding of taxonomic practice.

SOURCES OF DATA

Data on taxonomic memberships among clades of fishes is derived from FishBase (<https://www.fishbase.se/home.htm>), a global biodiversity

Table 1 — Data on fishes from FishBase. Shaded diagonals are numbers of taxonomic groups at that rank and percentages of those groups that are monotypic (e.g., 38% of the eight classes contain only one order; 31% of 176 families contain only one genus). Italicized cells (upper right) contain r^2 values for observed versus modeled taxonomic memberships (Fig. 3) of subtaxonomic (top row) and suprataxonomic (left column) groups (e.g. the r^2 of observed versus modeled numbers of families per order is 0.96). Bold cells (lower left) are values of p , the membership inclusion probability determined from the total numbers of subtaxonomic (left column) members of each suprataxonomic (top row) group (eq. 2). Note similar p values between similar levels of taxonomic separation (upper left top to lower right bold diagonals) and that values of p decrease with increasing degree of taxonomic separation (columns).

	Classes	Orders	Families	Genera	Species
Classes	8 (38%)	<i>0.58</i>	<i>0.68</i>	<i>0.63</i>	<i>0.66</i>
Orders	0.401	22 (28%)	<i>0.96</i>	<i>0.99</i>	<i>0.98</i>
Families	0.150	0.468	176 (31%)	<i>0.99</i>	<i>0.99</i>
Genera	0.050	0.155	0.415	5,140 (34%)	<i>0.92</i>
Species	0.019	0.060	0.161	0.486	33,912 (-)

information system containing data on all known extant species: their taxonomy, biology, trophic ecology, life history, and uses, as well as historical data reaching back to 250 years. As accessed on February 2, 2020, FishBase lists 33,912 fish species, distributed among 5,104 genera, 560 families, 78 orders, and 8 classes compiled from over 52,000 references (Table 1).

Two of eight classes of fishes comprise the bulk of the group: (1) the Actinopterygii, or ray-finned fishes, is the largest, containing sturgeons and paddlefishes, bowfins, gars, and teleosts, collectively comprising 59 (75.6%) orders, 489 (87.3%) families, 4,876 (95.5%) genera, and 32,587 (96.1%) species; and (2) the Elasmobranchii or cartilaginous fishes, which includes sharks, rays, skates, and sawfish comprising 13 (16.7%) orders, 60 (10.7%) families, 205 (4.0%) genera, and 1,162 (3.4%) species. The remaining classes are: (3) the Myxini (hagfish), (4) the Petromyzonti (lampreys), (5) the Cladistii (bichirs), (6) the Holocephali (chimaeras), (7) the Dipneusti (lungfish) and (8) the Coelacanthi (coelacanth). Collectively, these six groups comprise a total of six (7.7%) orders, 11 (2.2%) families, 26 (0.5%) genera, and 162 (0.5%) species (Figure 1).

Perhaps the most striking aspect of these data is the fact that taxonomic memberships at all hierarchical levels are typically ‘heavy-tailed’, in that very few taxonomic groups contain the great majority of subtaxonomic units, and that a very large number of groups contain very few members. While the largest class of fishes contains between about 96% and 76% of all orders, families, genera, and species; 28% to 38% of these taxonomic units are monotypic with respect to each immediately subordinate taxonomic level (e.g. 31% of 176 families contain only one genus; Table 1).

Not surprisingly, this aspect of taxonomic diversity has been the subject of much study following that of Willis (1922), who pointed out the “hollow-curve” nature of taxonomic membership frequencies. Different authors have variably interpreted such curve shapes as representing hyperbolic, logarithmic, log-

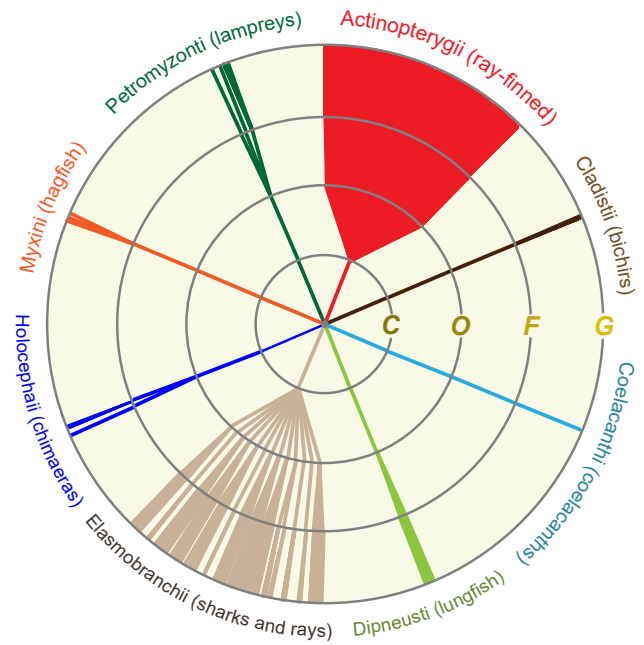


Figure 1: Classes, orders, families, and genera of fishes. Memberships of each hierarchical group represented by the number of radial lines intersecting that level’s bounding circle.

normal, exponential, geometric, and/or power law functions (e.g. Chamberlin 1924, Fisher et al. 1943, MacArthur 1957, Williams 1964, Burlando 1990, Minelli et al. 1991, and Nee et al. 1992); Anderson (1974) provides an excellent review. In addition, the origins of such “hollow curves” have been variably ascribed to deterministic versus random or stochastic processes during biological diversification, and/or historical artifacts of taxonomic classification (e.g. Yule 1924; Wright 1941; Kendall 1948, Reddingius 1971, Walters 1961, Chu and Adaimi 1999, and Scotland and Sanderson 2004).

SEGMENTATION

Membership at all taxonomic levels of fish is closely simulated with a ‘broken plate’ function that describes the abundance of subtaxa within any higher taxon. Succinctly stated, it defines the sizes of entities whose diameters describe an exponential density function. Because taxonomic ‘size’ is the same as group membership, here ‘size diameter’

approximately corresponds to the square root of numbers of members within the taxon. However, except for mathematical erudites, this declaration conveys little in the way of understanding of taxonomic associations or the reasons that such a function might better serve as a metric for Linnaean memberships. It is therefore useful to develop this approach in some greater detail.

Segmentation of Lines

As a geologic example, consider a stratigraphic section along which bedding planes separating different types of rock occur at random. That is, the spatial location of any horizon of lithologic change is independent of the position of any other horizon of change. As is the case for differences between ordered random numbers, the stratigraphic separations between such horizons describe an exponential distribution of lithofacies thickness. In probability theory and statistics, the exponential distribution is the distribution that describes the frequencies of distances (or waiting times) between horizons (or events) where points of change occur independently (randomly), but at a constant average rate. In these cases, the distribution of thicknesses of rock between horizons of lithologic change (as well as the average and standard deviation of those distances) is only dependent on the total length (thickness) of the outcrop and the number of lithologic change horizons distributed along that length.

More generally, when discrete events (such as lithologic change) occur randomly along a continuum (such as a sedimentary sequence) with some mean rate, such that the number of events occurring within some span of stratigraphic interval has a Poisson distribution, then the average length of intervals between these horizons is only dependent on the total length (L) of that sequence divided by the total number of changes (N). Moreover, the probability (p) that a lithologic change occurs per unit thickness is the reciprocal of the average thickness. When the spatial locations of horizons of lithologic change are random in

space there are many thinner intervals than thick, and the number of intervals equal to or greater than (the exceedance of) some thicknesses (E_T) is expressed as:

$$E_T = N e^{-p E_T}$$

Here, slope (N/L) is the inclusion probability (p), the y intercept (N) is the number of transitions, and the reciprocal of the slope (L/N) is mean thickness (e.g., Davis, 1986; Swan and Sandilands, 1995). An excellent illustration of such stratigraphic division seen in exposures of Lower Ordovician peritidal carbonates in the Franklin Mountains of west Texas (Goldhammer et al. 1993). There, 304 lithologic transitions occur over a stratigraphic distance of some 282 meters; frequencies of separation distances define an almost perfect exponential distribution (Figure 2). The 304 transitions along this thickness of section yield an inclusion probability (p) of 0.011 /cm (Figure 2). That is, if horizons of change in rock type occur randomly, numbers of transitions (N) and total section length (L) alone would predict that 98.9% of 1-cm steps up this outcrop will remain in the same lithofacies unit, and that the remainder (1.1%) would cross some lithofacies boundary. Comparison of this expectation with data on unit thicknesses yields a Pearson's r^2 of 0.961; along this exposure, locations of lithofacies transitions are generally independent. Similar exponential distributions of distances between lithofacies boundaries have now been reported from a wide range of stratigraphies (e.g. Drummond and Wilkinson, 1993,1996; Burgess, 2008).

The "broken stick" model for niche apportionment by MacArthur (1957) noted above is an excellent example of exponential distributions from the biologic sciences. In this classic paper, MacArthur studied ecological reasons for abundance distributions of bird species, and argued that ecological niches within some pool of resources are randomly divided; that the amount of the environment occupied by various species was analogous to a stick of some length (L) which

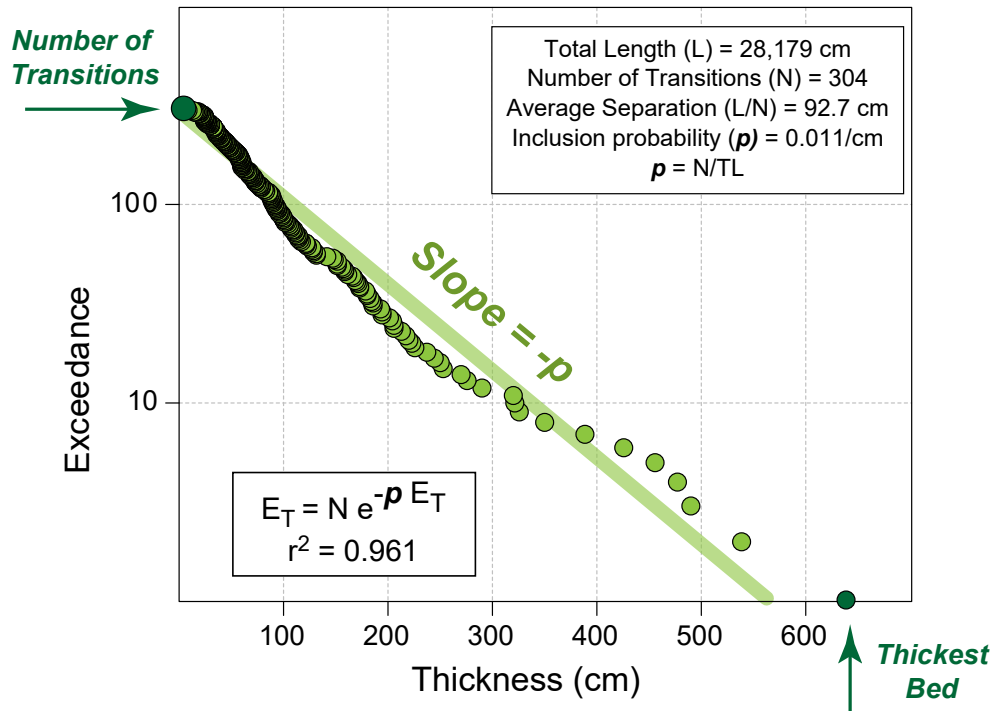


Figure 2: Succession of platform carbonate lithofacies comprising the Lower Ordovician El Paso Group exposed in the Franklin Mountains of west Texas (Goldhammer et al., 1993). Exceedance (Y axis) is the number of transition separations that are equal or greater to some X axis value. The model (green line) y-intercept is at the number of transitions (all are equal to or greater than the thinnest; here at 4.4 cm); the X intercept (~6 m) is the largest separation; it exceeds all other separation of horizons of lithologic change.

is randomly broken into a number of segments (N), the lengths of which are proportional to the abundances of species.

Segmentation of Areas

The observation that the thicknesses of lithofacies comprising stratigraphic columns are similar to exponentially-distributed lengths of segments along randomly-broken sticks serves to raise an analogous question about areas of sediment accumulation over a depositional surface: do similar distributions characterize areas of lithotopes (patches of like sediment) in space? One source of appropriate data to address this question is available from mapping by Weiss and Muller (1988), who delineated areas of Holocene carbonate sediment bodies around the Island of Antigua in the West Indies (Wilkinson and Drummond, 2004). As is the case with thicknesses

of ancient platform carbonate lithofacies (e.g., Figure 2), transects across these modern carbonate lithotopes describe an exponential distribution of lithotope diameter frequencies (Figure 3A). In this example, the 220 diameters of carbonate sand bodies comprising a net transect length of about 99.4 km yield an inclusion probability (p) of 0.0021 per meter of transect length. That is, if edges of these lithotopes occur independently across this island-bounding platform, then numbers of edges (N) and the sum of lithotope diameters along the transect (N) alone would predict that 99.89% of meter-long steps would end in the same lithofacies unit in which they began, and that only 0.21% of steps will transition into a new lithofacies. This lateral ‘inclusion’ parameter is exactly the same as that derived from vertical lithofacies transitions in the El Paso Group (Figure 2). Comparison of this theoretical expectation of lateral facies changes

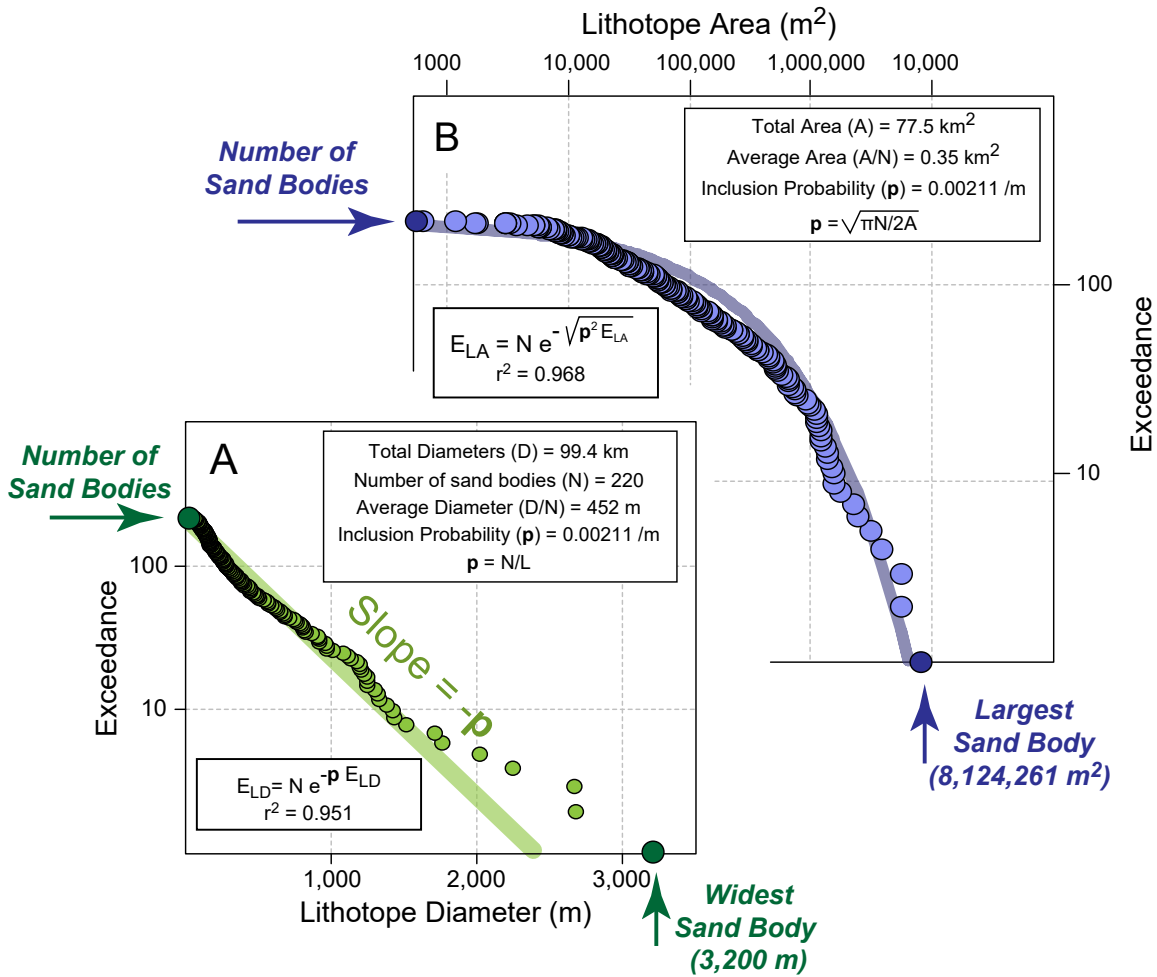


Figure 3: Sizes of Holocene carbonate sand bodies surrounding the island of Antigua; West Indies. A- frequencies of distances across 220 sandy lithotopes; this distribution has exactly the same function at that defining distances between Lower Ordovician lithofacies transitions in Figure 1. B- Frequencies of areas of these same 220 sandy lithotopes. Each presumes that lateral boundaries of sand bodies occur independently (randomly) in space. In both cases, the transition parameter, p (~2/km), is the likelihood of crossing the edge of some sand body per kilometer of transect length; a probability that gives rise to the broken stick (A) and broken plate (B) approximations of lithotope diameter (A) and area (B) with average lengths and areas of 0.45 km and 0.35 km², respectively.

from data on areas of Holocene sand bodies around Antigua yields a Pearson r^2 of 0.961 (Figure 3B). Across the Antigua platform, edges of sandy carbonate lithotopes occur randomly.

Moreover, the frequency distribution of these lithotope areas define a curvilinear trend in log area versus log exceedance space (Figure 3B); this trend reflects the nature of the area frequency of map units whose diameters exhibit an exponential distribution of element sizes. Similar to the dependency of linear inclusion probabilities (p)

on lengths and numbers of facies changes along vertical (Figure 2) and lateral (Figure 3A) transects, the areal inclusion probability p is only dependent on the number of lithotopes in question (N) and the total area (A) that they occupy. When considering areas rather than lengths, the areal inclusion probability p takes the form:

$$p = \sqrt{\frac{\pi N}{2A}}$$

and the exceedance of some lithotope area (E_{LA} ; the number of lithotopes that equal exceed some area value) is expressed as:

$$E_{LA} = N e^{-\sqrt{p^2 E_{LA}}}$$

Conceptually, this inclusion probability, p , for (the diameter of) an area is the same as that for a linear; with any additional step taken in some transect across an area, p represents the probability that a lithotope transition will occur. If a boundary is not crossed, then element diameter increases by one length unit, and element area increases in proportion to that increase in diameter. Similar distributions of sizes of sediment patches are now reported for a range of Holocene sediment areas, including tidal flats on northwest Andros Island, Bahamas (Rankey, 2002), coral and seagrass communities in the Arabian Gulf (Purkis et al., 2005), carbonate sand bodies on Great Bahama Bank (Harris et al., 2011), and patch reefs and ponds on Alacranes Reef (Purkis et al., 2015).

In short, because the locations of boundaries of carbonate sand bodies are generally independent in space, the density functions for patch sizes measured as average diameters of areas are readily expressed in the form of the broken stick (Figure 3A) and broken plate (Figure 3B) functions, respectively. Size frequency distributions for either are only dependent on numbers of lithotopes under consideration and their sizes (diameters, areas). Agreement between these exponential distributions and those observed in the real world arise largely because of the random division of surfaces of accumulation into the mosaic of lithotopes that make up this and other depositional surfaces.

Membership of Taxonomic Units

Taxonomic memberships at all levels of Linnaean division are closely approximated with the broken plate function. By analogy with broken sticks, the square root of the number of members of some subtaxon (e.g. species) within some intermediate taxonomic group (e.g. orders)

that comprise some suprataxonomic group (e.g. fish) is taken as being conceptually equivalent to the number of length units in each of the lithologic units that makes up some stratigraphic section. As an example, consider the 33,912 species that make up the 78 orders of fishes, analogous to the 28,179 cm of stratigraphic section divided into 304 lithologic units comprising the Lower Ordovician El Paso Group of west Texas (Figure 2). We might presume that any species within one of these groups is defined on the basis of some aggregate collection of shape (or possibly ecologic) characters; that is, each species is a manifestation of some realized amount of n-dimensional morphospace. If each species indeed represents some amount of shape variation about some mean, and if the number of species in each order is then taken as a measure of order ‘size’, then the frequency distribution of species per order of fishes is exactly like that exhibited by lengths of transects across, and areas of, lithotopes on depositional surfaces. That is, the frequency distribution of species memberships among the 78 orders of fishes is readily described by a broken plate function for areas (Figure 4A) in which ‘diameters’ of order-level morphospace (diameter units = $2 \times ((\text{number of species}/\pi)^{0.5})$) are exponentially distributed; such diameters are equally well described by the broken stick function for linear units (Figure 4B).

Although this representation for the taxonomic division of fishes is the same as that for linear divisions of stratigraphic units and for the areal division of depositional surfaces, an important distinction is that the p , inclusion metric for linear transects (Figure 2) is the reciprocal of the average length, and represents the probability that 1 ‘step’ will or will not result in crossing some lithofacies boundary. Similarly, for lines across some area (Figure 3), it represents the probability that a lateral ‘step’ might result in leaving one sedimentary lithotope and passing into another. However, with respect to taxonomy (Figure 4), the metric of a ‘step’ represents the membership of some subtaxonomic unit (e.g. the diameters of morphospace are taken as sizes of the smaller

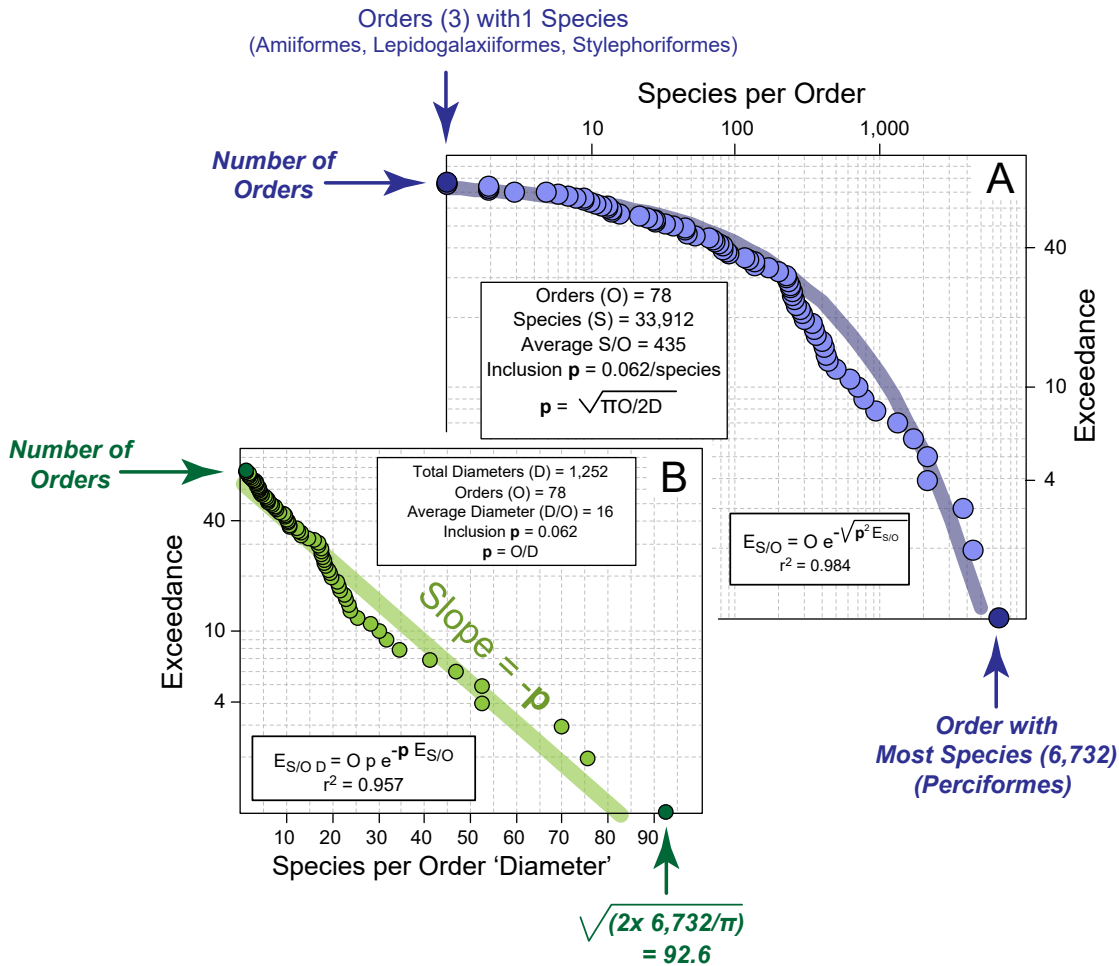


Figure 4: Sizes of orders of fishes as measured by numbers of species per order. A- Number of species per order among the 78 orders of fishes. B- frequencies of distances across the ‘shape-space’ represented by each of the orders’ ‘morphospace’; this distribution has exactly the same function as that defining areas of carbonate sand lithotopes surrounding the island of Antigua in Figure 3-B. Functions describing area (A) and diameter (B) both presume that lateral boundaries of subtaxonomic units (species) occur independently (randomly). In both cases, the transition parameter, p (~6%), is the likelihood of increasing the number of orders with the addition of one new species; a probability that gives rise to the broken stick (B) and broken plate (A) approximations of order size.

taxonomic units). Therefore, the taxonomic inclusion metric p represents a probability that the addition of a smaller taxonomic unit (e.g. a species) might or might not result in the crossing of some theoretical ‘shape-space’ boundary into that represented by some other larger taxonomic unit (e.g. order). In a taxonomic sense, ‘addition’ refers to the identification of a new taxon and the attendant probability that that description would also give rise to membership in a higher taxonomic group.

The value of inclusion p with respect to species membership in orders, for example, is ~6.2% (Figure 4). By analogy with a broken stick, that value means that if one were to begin in the ‘shape-space’ represented by a given order, and if one were to cross an area represented by that order along a transect (diameter, D) with a length represented by the square root of the number of species in that order (that represented by $\pi 0.5D^2$), there is a 93.8% chance of remaining in that morphospace (i.e., of not engendering a new order) and a

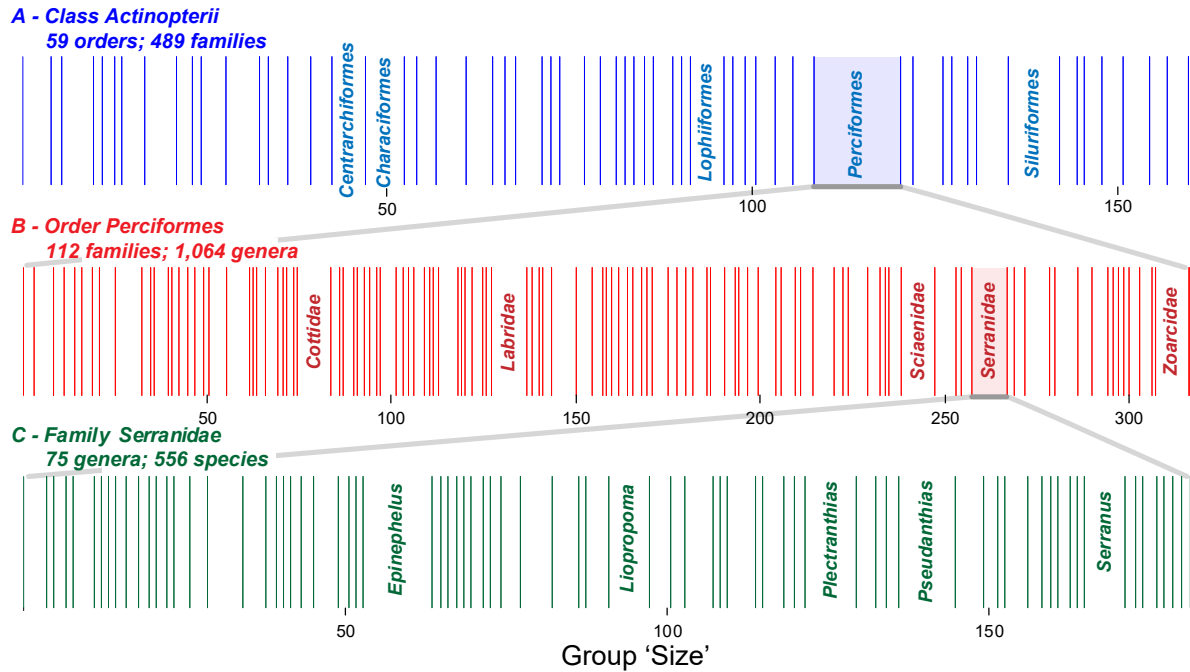


Figure 5: Broken stick illustration of the hierarchical nature of fish taxonomy where membership ‘size’ is taken as the diameter of represented morphospace, here calculated as the $2 \left(\frac{Mn}{\pi} \right)^{0.5}$ where Mn is the number of component subtaxa (families, genera, and species for orders, families, and genera, respectively). A - Sizes of the 59 orders (arranged alphabetically) comprising the class Actinopterygii (ray-finned fishes; the largest class of fishes). Perciformes (perch-like ray-finned fish) is the largest order. B - Sizes of the 112 families comprising the order Perciformes; the Serranidae (includes sea basses and groupers) is one of the larger families. C - Sizes of the 75 genera making up the family Serranidae. *Epinephelus* (groupers) is the largest genus with 6,556 species). Membership dimensions, calculated as inferred diameters of morphospace represented by each level of division, comprise an exponential distribution of clade sizes (Figure 6).

6.2% probability of ‘entering’ the morphospace represented by some other order.

As further illustration, consider numbers of members belonging to successively smaller Linnaean divisions of fishes. Among the eight classes that make up this group, Actinopterygii (ray-finned fish) is the largest. It consists of 59 orders and 489 families (Figure 5A). If the ‘size’ of a stick segment is taken as (the square root of) the number of families it contains, and if the area of morphospace realized by that class (the total ‘stick’) is taken as being the sum of morphospaces realized by numbers of each family, then lengths of order-level shape variation can be represented as $D_o = 2 \left(\pi Nf \right)^{0.5}$ where D_o is a linear measure (diameter) of morphospace occupied by each family and Nf is the number of families in that order. When represented as ‘distances’ across

contained family memberships, this linearization of size gives rise to a broken stick representation of Actinopterygii memberships (Figure 6A).

A similar quantification of memberships is derived when taking a similar approach to the order Perciformes (perch-like fish) which, with 1,064 genera, is the largest order in the class Actinopterygii. The length of the ‘stick’ represented by this class can be similarly broken, where linear segmentation is expressed by roots of numbers of genera in each family (Figure 5B). This linearization of sizes gives rise to an exponential distribution of Perciformes memberships that is also exponential in form (Figure 6B). Perhaps not surprisingly, when this approach to division (that is, when genus ‘size’ is taken as the abundance of species within each contained genus) is applied to one of the largest groups of Perciformes, the

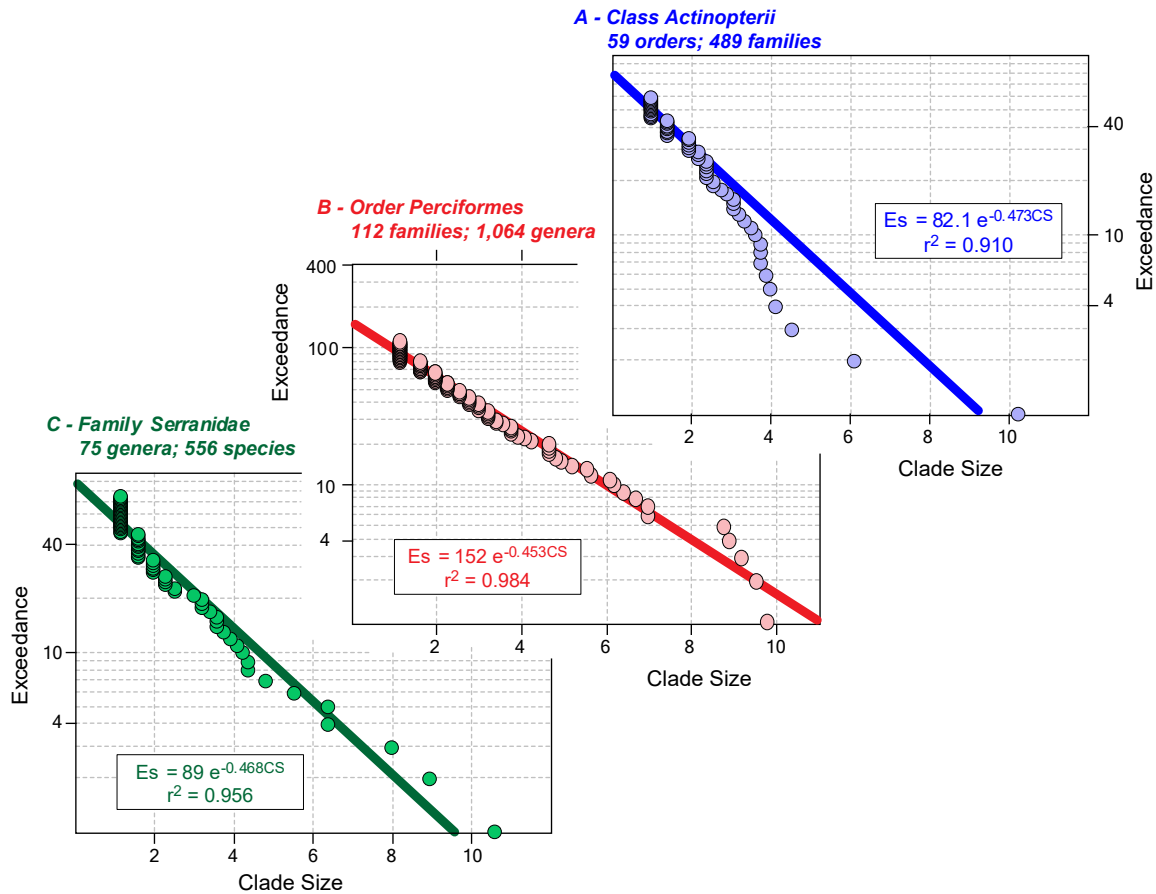


Figure 6: Sizes of clades comprising: A- the class Actinopterygii, B - order Perciformes, and C- the family Serranidae. Membership magnitudes, calculated as inferred diameters of morphospace represented by each level of division, comprise exponential distributions of clade sizes. Note that each yields a slope value (an inclusion p) of about 46%, the probability of leaving some intermediate taxonomic level (e.g. a family in B) with the addition of morphospace represented by a subtaxon (e.g. a genus in B).

family Serranidae with 75 genera (includes sea basses and groupers), it also gives rise to a broken stick segmentation (Figure 5C) with exponentially-distributed stick (square roots of numbers of genera) lengths (Figure 6C).

Memberships predicted by the broken plate model among the six possible taxonomic groupings of fish (species per genus, family, order; genera per family, order; family per order) are in good agreement with available data (Figure 7). Model and observed memberships plotted as numbers of groups (Y axes) exceeding some membership value (X axes) yield Pearson's r^2 values of 0.923, 0.989, 0.984, 0.936, 0.989, and 0.984, respectively (Table 1). Numbers of subtaxa making up taxonomic units of fishes closely fit those expected

if the division of organism morphologic characters is analogous to the random division of areas of sediment accumulation.

The 'broken plate' representation of such division is also in agreement with data on sizes of entities among other geographic and taxonomic areas, suggesting that the inclusion metric (p) serves as a good measure of sizes of many 'space' entities when positions of lateral boundaries occur independently. Examples include areas of Earth craters (Figure 8A, Earth Impact Database), calderas (Figure 8B, Worldwide Collapse Caldera Database), countries (Figure 8C), continents (Figure 8E, Mortimer, 2007), and brachiopods and modern mammals (Figures F and G; Wilkinson, 2011).

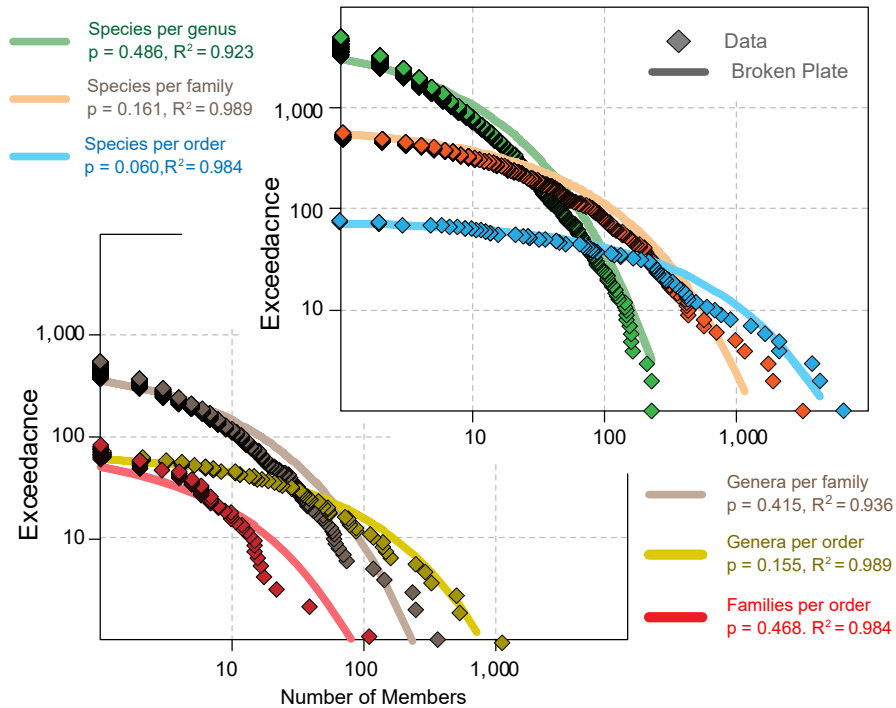


Figure 7: Broken plate functions fit to taxonomic data on fishes plotted as number of subtaxa among each hierarchically higher level of supertaxa (X-axes) relative to numbers of subtaxa (Y-axes) that equal to or are greater than some X-axis membership number. Metrics for each membership curve are listed in Table 1.

Division of Shape-space

Membership in taxonomic groups of fishes, at all levels of consideration, are closely approximated by the broken plate function in which the square root of numbers of constituents comprise an exponential distribution, the numerical equivalent to the broken stick model of MacArthur (1957). Moreover, Linnaean taxonomy is generally recognized to reflect the phylogeny of organisms (their descent by evolution), but is largely based on morphologic and DNA sequence similarity among different organisms. Higher levels of taxonomic classification therefore represent greater morphologic disparity. By analogy with sizes of sedimentary lithotopes measured as their lateral extents, this depiction of fish taxonomy suggests that sizes of taxonomic units represent amounts of shape variation. The numbers of members in any taxonomic unit therefore should be proportional to the amount of morphologic variation or disparity among the organisms comprising that group.

Although historically, the practice of taxonomy has been largely a qualitative exercise, taxonomists with expertise in the affinities among any larger group of organisms are cognizant of the general form and amount of variation in form that is intrinsic to any group. With the discovery of some hitherto unreported form, intuition derived from knowledge of similar, and presumably related groups, often leads to correct suppositions about the uniqueness and Linnaean level of classification appropriate to that newly-discovered organism. Given that newly-discovered forms are readily placed into current classifications and acknowledging that much shape-space has been vacated by extinction, it seems apparent that the amount of potential morphospace represented by any modern group is (perhaps largely) unfilled.

This vacuity is also alluded to when evolutionary relationships combined with taxonomic memberships are represented in the ‘tree of life’ iconography (e.g. Figure 9A). Although largely

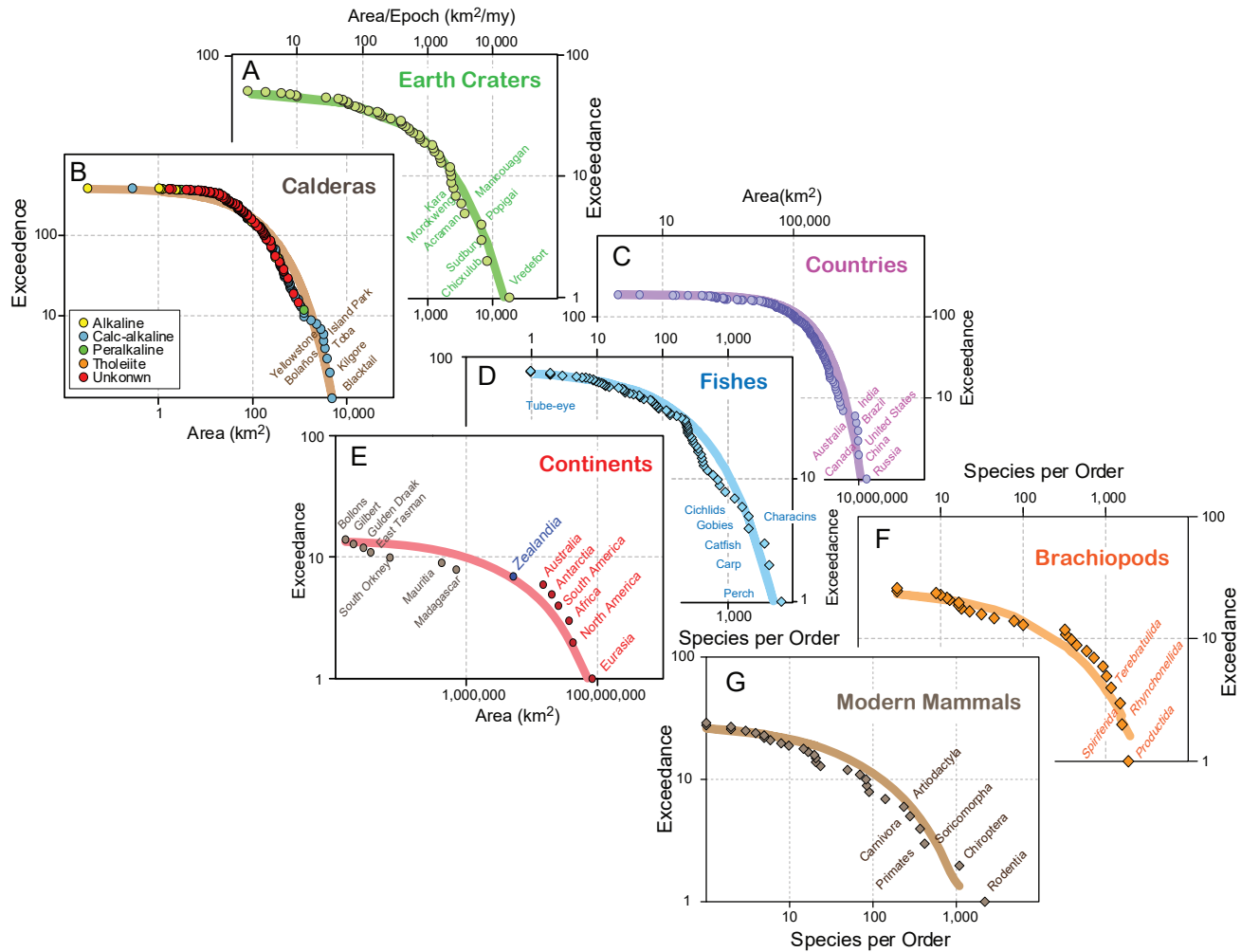


Figure 8: Broken plate function fit to data on areas/numbers of: A) Earth craters, B) calderas, C) countries, D) species per order of fishes (this study), E) continents, F) species per order of all brachiopods, and G) species per order of modern mammals.

a phylogenetic metaphor (e.g., Morrison, 2014), irrespective of its representation as a tree, network, or some other shape, and regardless of whether constructed from Linnaean categories, genomes, or some other criteria, most depict occupied ‘morphologies’ (now or in the past) as branches and twigs, with some proportion of morphospace that has been vacated during extinction, and some proportion to be ‘claimed’ during evolution. Moreover, as noted above, numerical aspects of group membership are similar among different levels of classification (e.g. Figure 9B), and all are well-described if sizes of subtaxa making up any suprataxonomic group are perceived as realized

areas with exponentially-distributed diameters (Figures 9D-G) suggesting a self-similarity among memberships at different levels of consideration. Although amounts of unrealized fish morphospace are unknown, the broken plate function allows for a visualization of sizes of taxonomic units if the present time plane can be envisaged as a two-dimensional section across the time-shape continuum represented by that group (Figure 9).

Why Random Division?

At this point it is perhaps appropriate to ask two interrelated questions: (1) Why is a model

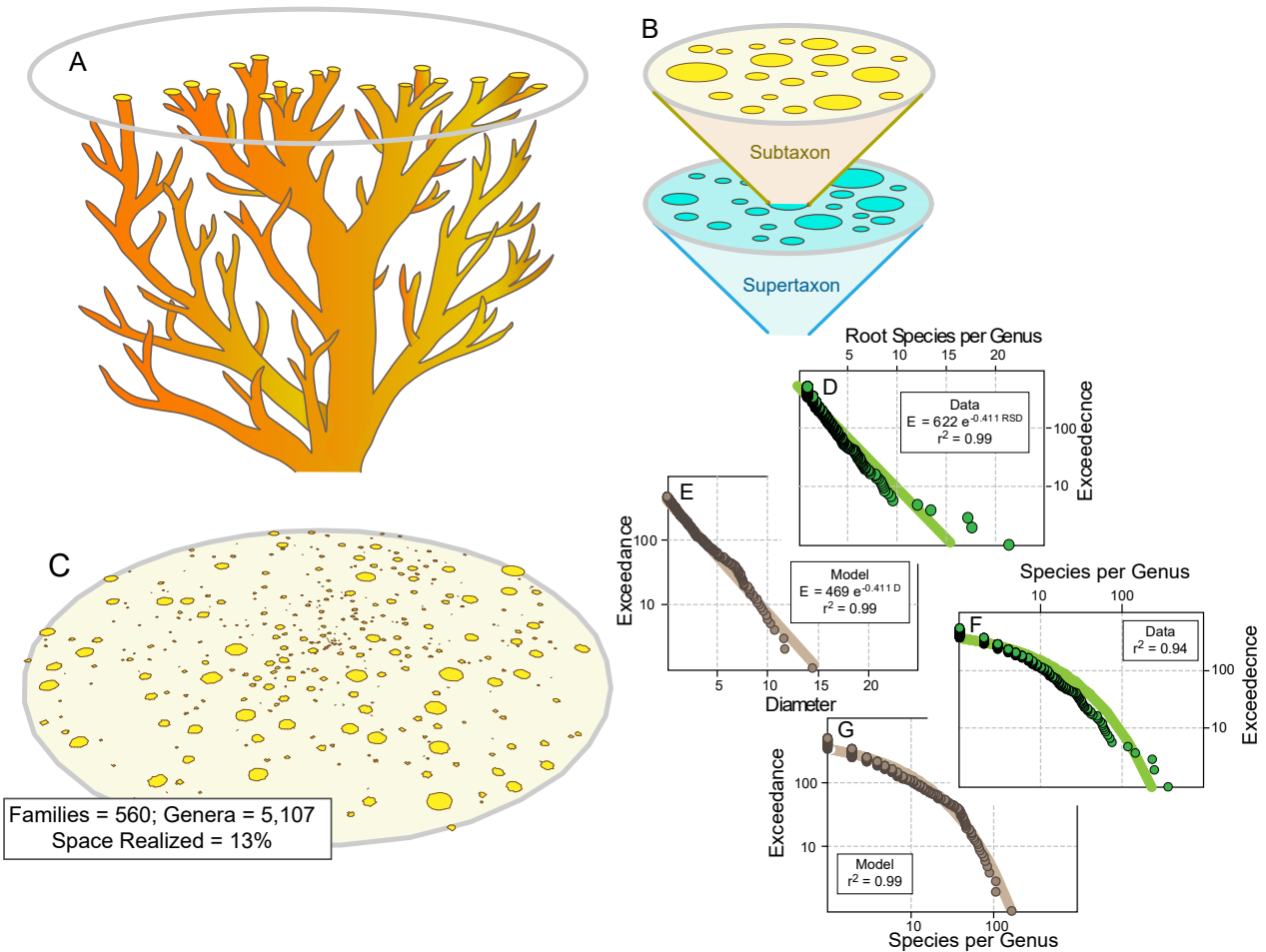


Figure 9: A- Cartoon of paleobiologic change depicting origination and extinction as limbs in the “Tree of Life” (Weller, 1969). Grey oval represents a cross section through this tree as a time plane like the present, intersection of extant groups as yellow circles within that oval. B- Two nested taxonomic hierarchies; grey oval represents a time plane through potential shape space represented by one taxonomic level; C- Oblique view of size frequencies (generic membership) of 560 families (yellow circles) of finned fish (grey circle) where each family size (yellow circle area) is represented as numbers of contained genera. Here, the total cross-sectional area through the model time-plane space (grey oval) is arbitrarily scaled at ~40,000 ‘genus units’ such that the presently known number of genera (5,107) occupy ~13% of potential ‘morphospace’. D and E- exponential length frequencies of ‘diameters’ of 560 family memberships for data (D) and modeled (E) family sizes. F and G- broken plate area frequencies of family memberships for data (F) and modeled (G) family sizes.

representing the random division of morphospace appropriate for the description of taxonomic memberships? and (2) Why does a 2-dimensional model suffice for division of something that is surely more complex such as the n-dimensional variance in shapes of organisms (e.g. Gerber et al., 2008)? That is, why is n-dimensional morphospace well-described by a second order relationship such as that represented by a broken plate?

The observation that hierarchical classifications

of fishes (and many other groups) typically give rise to a ‘many small - few large’ distribution of memberships have engendered two general types of explanations (Anderson, 1974). These are: (1) group sizes reflect natural process; such patterns reflect the importance of either deterministic (e.g. Willis, 1922) or stochastic (e.g. Reddingius, 1971) process of isolation, selection, and evolution within the group; and (2) current classifications are largely an artifact of past taxonomic practices and

are generally unrelated to any natural processes (Walters, 1961).

It is therefore perhaps not surprising that a model of random division closely replicates observed patterns of classification. In the case of historical happenstance, it must be acknowledged that a (perhaps) significant part of taxonomic practice is an empirical exercise, in that criteria for division have largely relied on observation and an inheritance of an existing system of organization. As noted by Bertrand et al. (2006) the placement of any organism into any Linnaean pigeon-hole is decided by some taxonomist on a case-by-case basis. The reasons for making such choices are complex, but include perceptions about the 'appropriate' sizes of groups (small enough to be learned but large enough to allow for generalization) as well as a desire to preserve traditional names and associations but still signify monophyletic groups. The rank assigned to any given taxon therefore reflects perceived phylogenetic relations, but is also influenced by existing ranks of related taxa. Walters (1961) for example makes a strong case that the major aspects of Angiosperm classification are primarily artifacts of the botanical literature available to Linnaeus in 17th and 18th century Europe; that is, Angiosperm classification would be substantially different if botany had developed in New Zealand in the 19th century. In this view, current taxonomic choices are largely predicated by decisions of earlier generations.

Why two dimensions?

Occupied morphological space serves to describe characteristics of a group of organisms. Although these classifications may largely reflect a random division of shape continua by taxonomists, the form of organisms is surely a reflection of the deterministic and/or random evolutionary processes by which they evolved. As noted by Foote (1997), morphological divergence generally increases with taxonomic rank. Therefore, while taxonomic classifications may indeed be subjective, biased by historical artifacts, and/or phenetic rather than phyletic considerations, they still serve as a basis

for relating taxonomic richness and disparity.

If Linnean classifications are based on morphology, and if morphology is effectively n-dimensional, why does a simple, two-dimensional model like a broken plate serve to describe taxonomic richnesses across different Linnean levels? Numerically, the cumulative density function for 1-dimensional division (i.e., a 'broken-stick exponential'; Figure 2) is significantly different than that for two-dimensional fragmentation (i.e., a 'broken-plate'; Figure 3), but two-dimensional fragmentation ($\sim \text{length}^2$) is little different than that for random division in three dimensions (like a 'broken-sphere'; $\sim \text{length}^3$).

In this context, shape disparity is usually quantified with reference to the axes of some form of morphospace; an n-dimensional space in which the distances between species or other operational taxonomic units are proportional to some measure of the morphological distances between them. A common approach to the quantification of organismal shape utilizes landmark-based geometric morphometric methods in order to create multivariate space, and principal component analysis in order to define major axes of shape variation (Zelditch et al. 2012). In a context of shape dimensions, this methodology typically yields only two or three meaningful principal components (e.g. Claverie and Wainwright, 2014). By analogy with dimensions of morphospace, the two-dimensional broken plate model may therefore adequately capture and represent shape variation manifest as numbers of morphologically-defined taxonomic units.

TAXONOMIC INCOMPLETENESS

One aspect of fish classifications potentially impacts the appropriateness and applicability of the broken plate representation of Linnean memberships - this is changes in memberships associated with the historical evolution of fish taxonomy. The numbers of members included in the various taxonomic levels now understood to comprise fish taxonomy has been completely

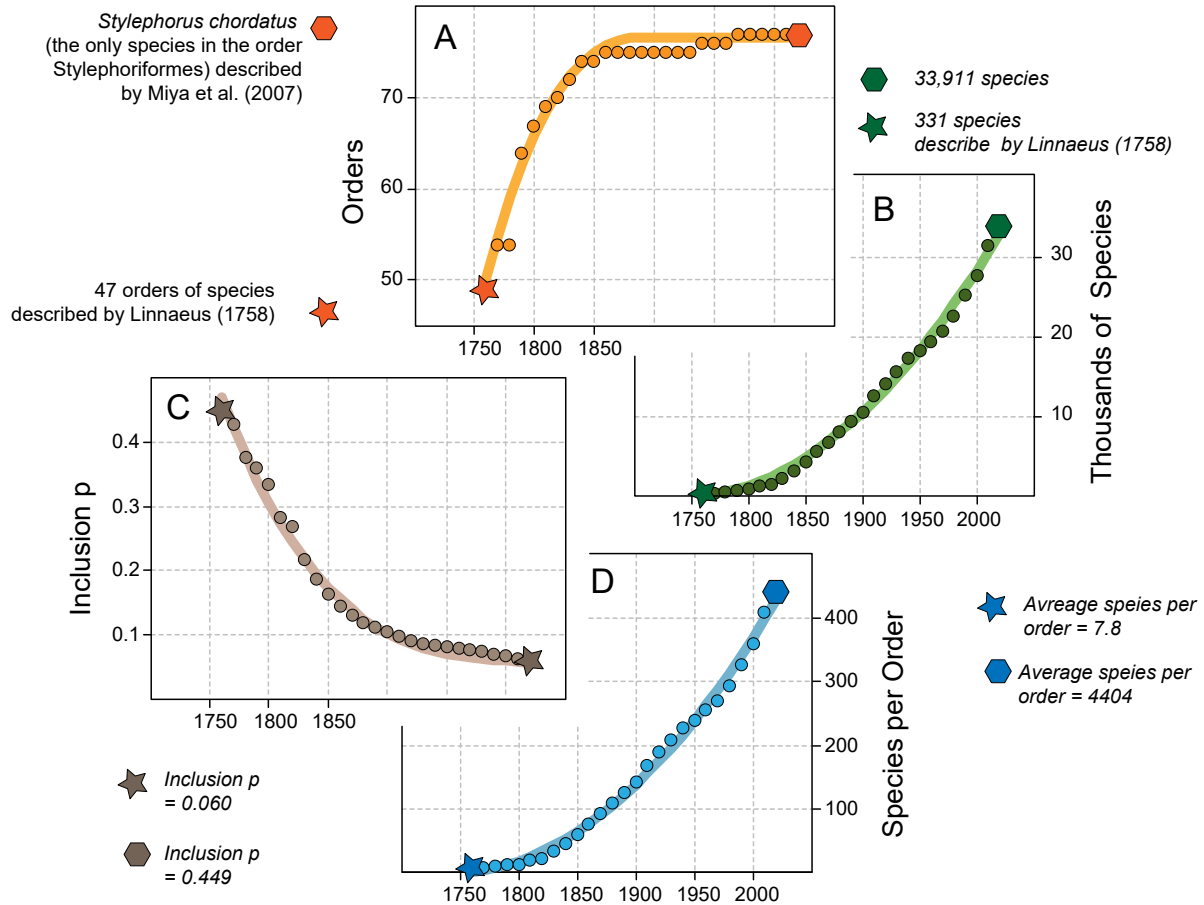


Figure 10: Historical evolution of fish taxonomy. A- cumulative number of now-recognized orders as a function of the year during which member species were described. Star locates 47 orders represented by 331 species described by Linnaeus (1758). Numbers of orders reached (74 out of 77 that are now recognized) within the first century. B- cumulative number of species as a function of date of description. Since 1758, the rate of description of new species has increased by about 1.6% per year to a current rate (2019) of ~29 per year. C- Historical order-of-magnitude decreasing values of the inclusion parameter p , from 0.449 in 1760 to a present value 0.060. D- Increase in species per order, an increase largely attendant with increasing numbers of species.

reformulated and developed over the past ~270 years, and metrics of group size change as our understanding of evolutionary relationships among these organisms continues to evolve. The sizes of different groups of fish must be dependent on the history of group descriptions. That history begins with the description of 331 species by Linnaeus (1758) that now belong to 47 of the 77 currently recognized orders. Over the past 2½ centuries, the number of recognized orders increased abruptly and then changed little, with the most recent being established in 1961 (Figure 10A). Numbers of recognized species, on the other hand,

have increased at a rate of about 1.6% per year, and several hundred new species of fish are now being named each year (Figure 10B), an increase foretelling a significant number of fish species yet to be described.

In this context, Mora et al. (2008) found that rate of description of new species of marine fish began to decrease toward the end of the 20th century, and suggested that classification of these organisms is about 80% complete. If so, much of the continuing description of fish will be of those clades adapted to freshwater habitats. In contrast, Freitas et al. (2020) evaluate the classification

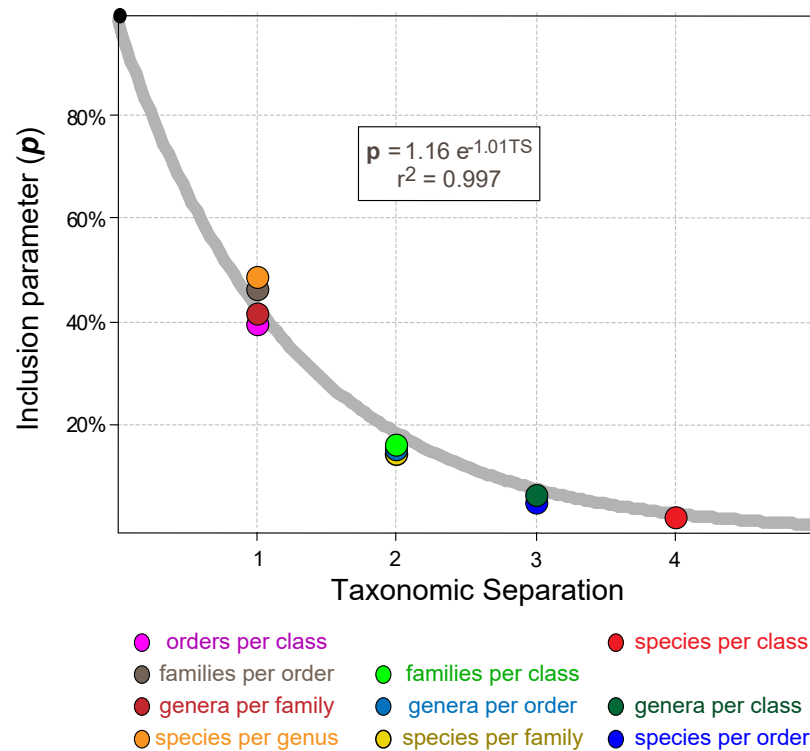


Figure 11: Values of the inclusion parameter p as a function of degree of taxonomic separation among fish groups. Here, the exponent of the slope (-1.01) is about 0.364, and represents the rate of decrease in p values for each increase in level of separation (0 = 100%; 1 = 36.4%, 2 = 13.6%, 3 = 4.8%, and 4 = 1.8%). This rate of change corresponds to about a 7-fold increase in membership with each increase in Linnaean taxonomic level.

of members of the family Auchenipteridae (driftwood catfishes) by comparing data from the Global Biodiversity Information and the Brazilian SpeciesLink databases. They found that invalid combinations of specific and generic names together with synonymy resulted in about 25% of species with inaccurate names. Similar overcounting results when now-recognized taxa have names that eventually will be synonymized with other groups. Alroy (2002) for example, estimates that about a quarter of named North American fossil mammal species will become invalid, thus reducing current diversity estimates.

Given these very different rates and caveats of group recognition, it is not surprising that metrics of order membership have changed over this same time period. The average number of species per order in 1758 was about seven; it has now risen to ~440 (Figure 10D) and is presently increasing by two to three species per order per year. Over

the same time, values of the inclusion parameter have decreased from about 45% to a present value of about 6% (Figure 10E). More importantly, although values of p have decreased by almost an order of magnitude, the asymptotic nature of this decrease is because of change in numbers of orders (now largely stable), and because p is calculated as the square root of membership number. Further change in the membership parameter therefore is largely insensitive to further increases in numbers of species. Given these considerations, we conclude that the broken plate approximation of group membership will be largely insensitive to future changes in the development of fish taxonomies.

IMPLICATIONS OF BROKEN PLATES

We conclude with a brief discussion of several applications to, and ramifications of, the broken

plate function and fish taxonomy. These include: (1) understanding that taxonomic memberships among fish (and other organism) groups are highly self-similar, (2) providing a robust metric for gauging degrees of lumping and splitting when effecting Linnaean classification, (3) assessing degrees of memory among nested levels of taxonomic membership, and (4) evaluating the necessity for explanations for monotypy, polytypy, and adaptive radiants during the evolution of fishes.

Self-similarity of Taxonomic Hierarchies

One of the more interesting attributes of the broken plate representation of taxonomic memberships is that different Linnaean hierarchical levels of classification exhibit strikingly similar membership metrics regardless of the level of taxonomic consideration. As noted above, the principal metric of membership is the value of inclusion p , the probability that the addition of some amount of morphospace through the discovery of a taxonomic subunit might result in the definition of a new larger taxonomic unit. These inclusion values are nearly the same among groups with like amounts of taxonomic separation, and they decrease predictably with increasing degrees of separation.

Similar inclusion metrics are apparent for data on order membership of the largest class of fishes (the actinopterygians), on familial membership of its largest order (Perciformes), and on generic membership of one of its largest families (Serranidae) with values of 47.3%, 45.3% and 46.8%, respectively (Figure 6). These are about the same as p statistics derived from data for orders among classes, families among orders, genera among families, and species among genera (40.1%, 46.8%, 41.3%, and 48.6%, respectively; Figure 7, Table 1) of all fishes. In addition, values of inclusion p decrease predictably with increasing separation of taxonomic hierarchies (Figure 11; Table 1). Specifically, the value of p decreases with increasing taxonomic separation (TS) as:

$$p = e^{-0.977 TS}$$

The exponent of this decrease, -0.977 (Figure 11) is the natural log of 37.6%, which represents the average decrease in p values with increasing separation among Linnaean levels as $p = 0.0376^{TS}$. The observed average and predicted p values for orders per class, families per order, genera per family, and species per order ($TS = 1$) are 44.3% and 37.6 (0.376^1); for families per class, genera per order, and species per family ($TS = 2$) are 15.5% and 14.2% (0.376^2), for genera per class, and species per order ($TS = 3$) are 5.5% and 5.3% (0.376^3); and for species per class ($TS = 4$) are 1.9% and 2.0% (0.376^4). Taxonomic memberships are largely the same, regardless of taxonomic levels of consideration (Figure 11).

Memberships of Fish Taxonomies

Data on numbers of subtaxa per taxon allow for calculation of inclusion parameters p (e.g. Figure 11). Because p is derived from data on numbers of taxa and subtaxa, rates of change in p with changes in degrees of taxonomic separation (Figure 11) also allow for calculation of changes in mean taxonomic memberships with amount of separation. Specifically, the membership (M) of subtaxa (Sb) in any suprataxonomic group (Sp) is:

$$M_{(Sp/Sb)} = \pi/2p^2$$

Data for fishes describe an exponential increase in membership with (linear) increase in separation of taxonomic levels (Figure 12A). The exponent of this slope (1.94) is 7.0; for each increase in taxonomic separation, average membership increases sevenfold. Average numbers of species per genus, genera per family, families per order and/or orders per class of fish are all 7 to 8 times larger than that of the next smaller group.

Similar compilations of taxonomic memberships for modern and ancient brachiopods and bivalves (as downloaded from the Paleobiology Database on 09-08-2009) and for modern mammals (downloaded from Mammal Species of the World on 11-20-2009; Wilson and Reeder, 2005) describe nearly identical trends (e.g.

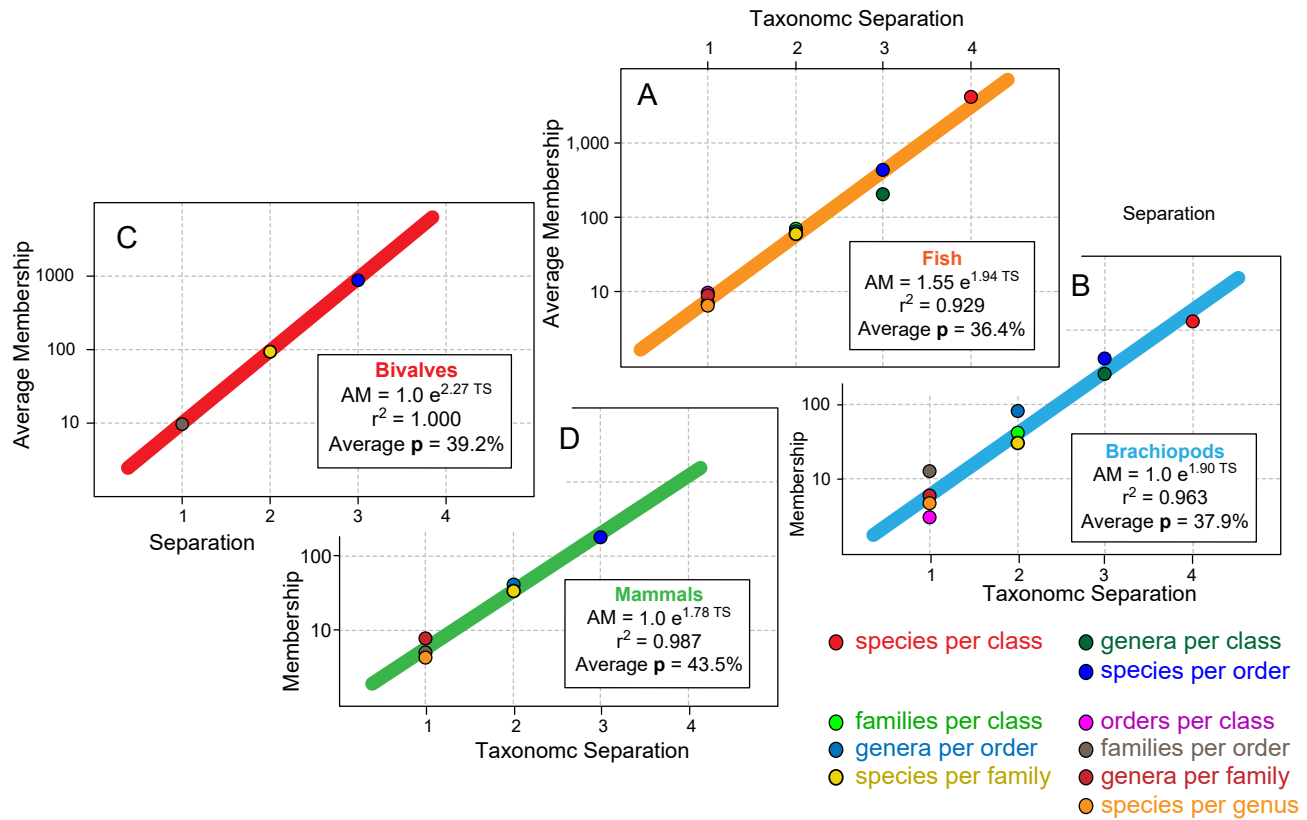


Figure 12: Average memberships of Linnaean hierarchies among: A) living fishes, B) extant and fossil brachiopods, C) extant and fossil bivalves, and modern mammals (D). Note similarity of fish, brachiopod, bivalve, and mammal membership exponents of -1.94, -1.90, 2.7, and -1.78, respectively. Exponents of these slopes (7.0, 6.7, 9.7, and 5.9, respectively) reflect proportional increases in memberships with increasing taxonomic separation. Memberships among all Linnaean taxonomic levels increases 6- to 10-fold for each increase in taxonomic separation.

Wilkinson 2011), with taxonomic separation versus log membership slopes reflecting 6- to 10-fold increases in membership (Figure 12). On the basis of these data it seems apparent that the taxonomies of invertebrate and vertebrate organisms are quite similar. Regardless of group, successively higher branches of the Linnaean classification comprise a similar increase in numbers of included subgroups. Similar relations among taxonomic rank and numbers of members are also reported for data on modern floras from several localities in Italy (Ricotta et al., 2007).

Broken Plates, Lumpers, and Splitters

Until the genomes of all extant organisms have been fully sequenced, the assignment of membership in any clade must rely primarily on

qualitative perceptions about the importance of morphological differences and/or similarities among different organisms. As a result, utilization of the Linnaean system of classification has resulted in a range of attitudes as to the degree of difference and/or similarity that suffices to include or exclude some particular group into some particular category. Taxonomists who tend to focus on like characters might therefore ‘lump’ organisms that share a few major characteristics in the same group, while those who value dissimilarity might ‘split’ groups on the basis of even a small disparity. The nature and importance of this spectrum of taxonomic “lumpers” and “splitters” has been the subject of much study and discussion since at least 1847 when Edward Newman used the terms while making a plea for “discarding imaginary species”.

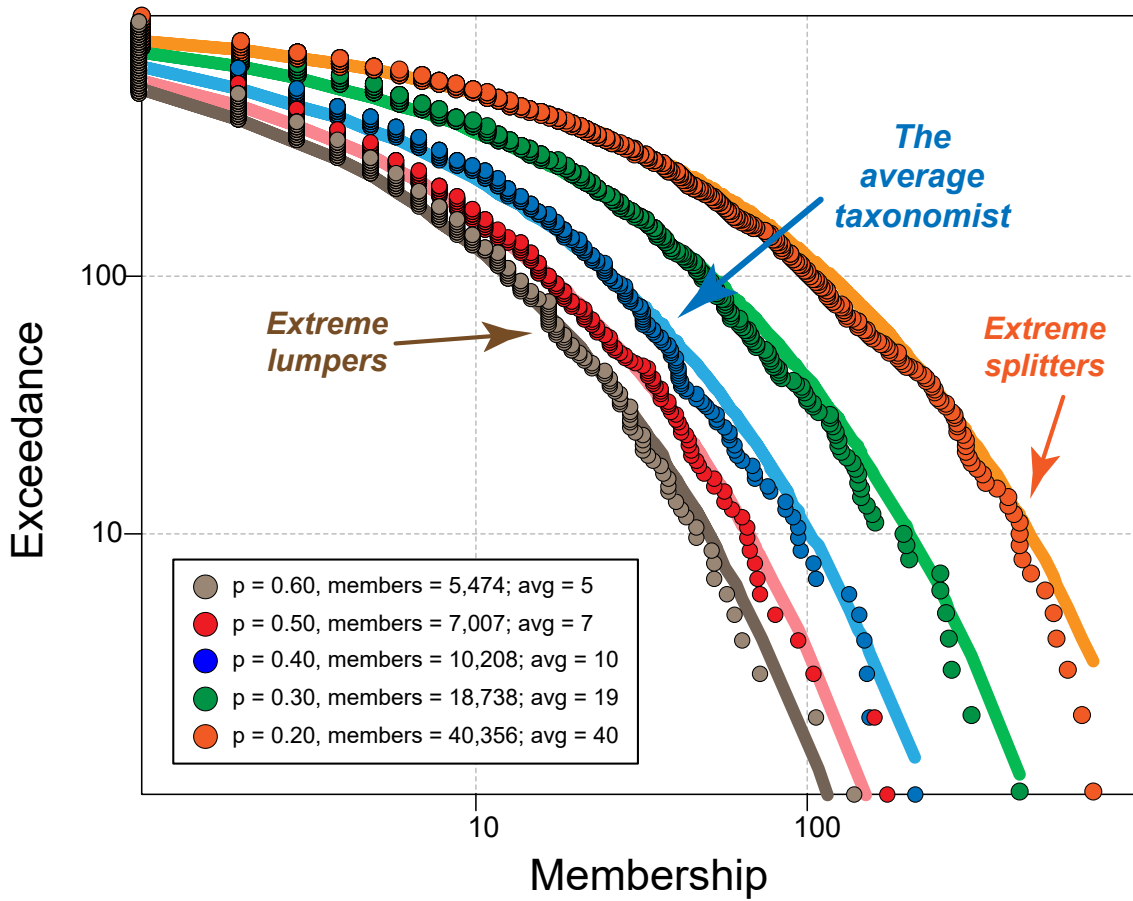


Figure 13: Range of memberships in a lumpers-to-splitter continuum for a hypothetical supertaxon (family) containing 1,000 intermediate taxa (genera) composed of variable numbers of some subtaxa (species). Each formulated from the broken plate function with inclusion parameters varying from -0.2 (low inclusion probability) to 0.6 (high probability). All curves represent the same amount of morphological variation (that represented by the family), but division ranges from 5,464 ‘species’ (extremely lumped) to 40,356 ‘species’ among the 1,000 genera in the family. Division represented by p value of 0.40 (with an average of 10 species per genus) is close to that realized in the classification of fishes (~ 0.376 ; Figure 6).

Perhaps obviously, values of the inclusion parameter are intimately related to discussions of the ‘appropriate’ degree of division at different levels of taxonomic classification, and serve to allow for some degree of quantification of the inclusion-exclusion spectrum. The broken plate model of membership allows for a quantification of the ‘most-accepted’ degree of division. As an example, consider some supertaxon (the ‘stick’) that is divided into some number of taxa (here, 1000 ‘segments’). Variation in the number of subtaxa per taxon (the number of ‘steps’ per segment) will depend on variation in the inclusion parameter. In a linear context, greater values of

p will result in higher probabilities of crossing a segment boundary along a broken stick (e.g. Figure 2). In an areal context, greater values of p result in higher probabilities of crossing a boundary (or remaining in that lithotope) along some lateral transect; average area is $\pi/2p^2$. In a taxonomic context, greater values of p should result in higher probabilities of crossing a morphospace boundary into a new taxon (Figure 7).

In a context of available data on fishes, brachiopods, bivalves, and mammals (Figure 12), at 1 degree of taxonomic separation, average values of p are about 36%, 40%, 39%, and 43%, respectively. That is, the ‘average taxonomist’

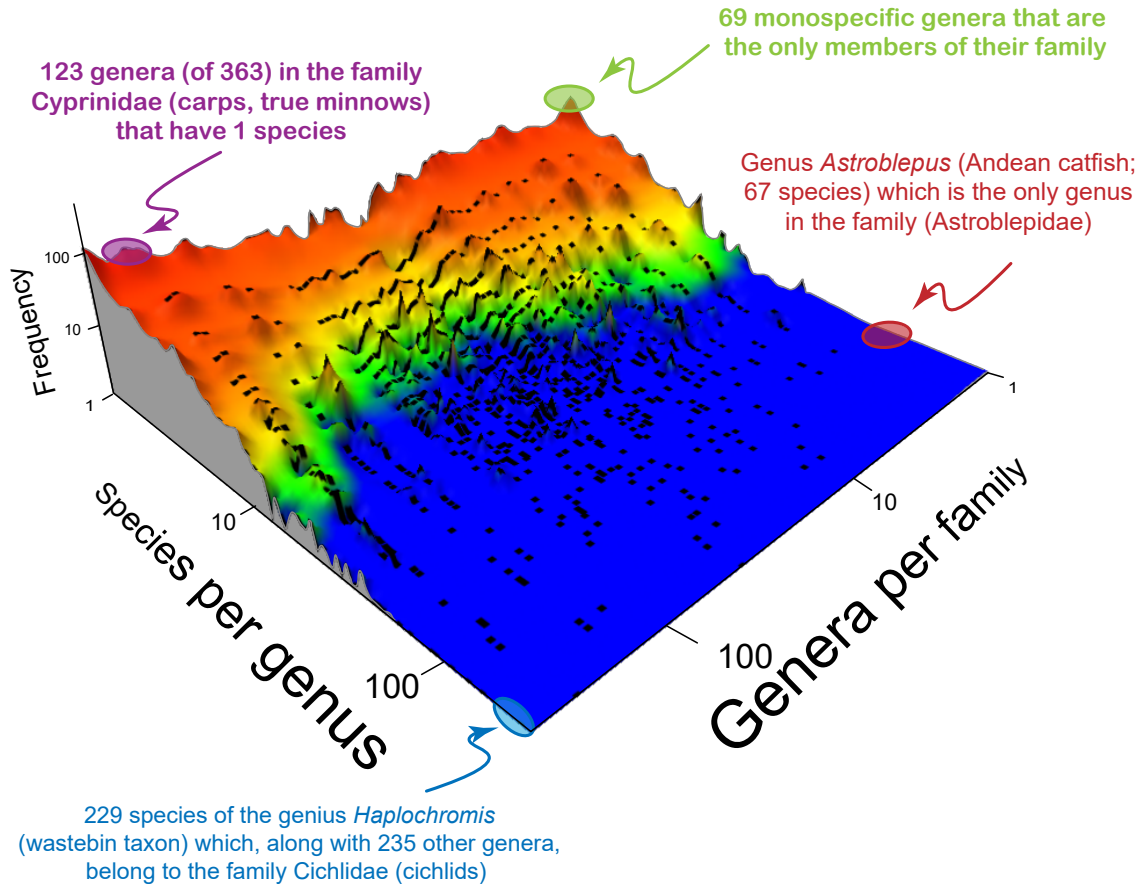


Figure 14: Relations between numbers of species per genus (lower left axis) and numbers of genera per family (lower right axis) among fishes. Vertical shading (Z axis) reflects membership numbers. Note that no relation exists between numbers of genera in some family compared to the numbers of species in that family. Those genera that contain many species (running along the lower right axis) embody the same range (few to many) of genera per family as do those genera that contain many species (running along the upper left axis). That is, variation in genera per family (the ‘strike’ of the species per genus versus genera per family slope) is largely unrelated to numbers of genera per family among fishes.

effects classifications at a p of about 40%; a value that corresponds to an average clade membership of between six and seven subgroups (Figure 13).

Memory and Taxonomic Membership

Agreement between fish memberships and broken plate functions suggests that fish classification may largely reflect a random division of morphospace among different groups. Such independence of the location of taxonomic boundaries also suggests that different taxonomic ranks might be similarly unstructured; that membership at any particular taxonomic level is unrelated to the degree of division at higher and/or

lower taxonomic ranks. Do, for example, genera with larger numbers of species tend to belong to families with larger numbers of genera? Are degrees of polytypy carried through different levels of taxonomic division?

In order to address these questions, we can compare numbers of species in genera with numbers of genera in families (Figure 14). From this it is clear that generic and familial memberships are unrelated. Numbers of genera in families are unconnected to numbers of species in genera; species per genus are largely independent of numbers of genera in families. The same range of species per genus (few to many) is uncorrelated

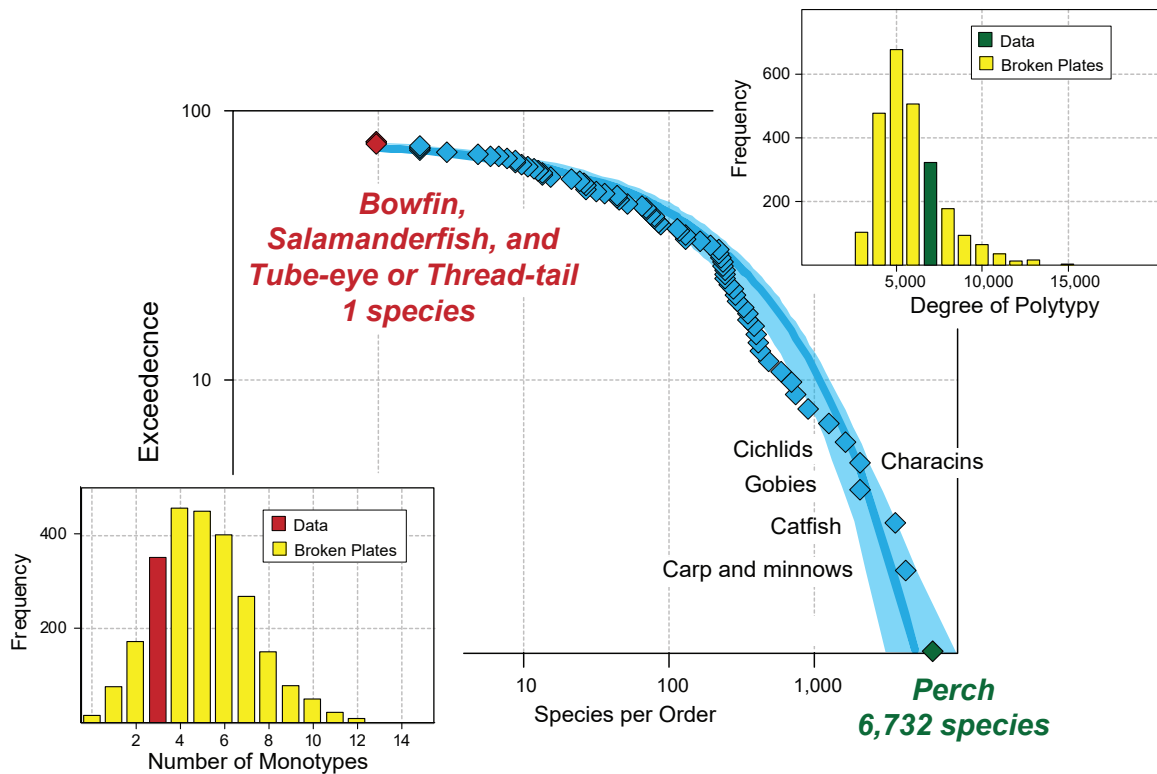


Figure 15: Data and expected numbers of species per order of fish. Diamonds are data on order memberships, ranging from 6,732 species belonging to the order Perciformes (perch) to one species within the three orders that contain one species (Amiiformes- the bowfin; Lepidogalaxiiformes- the salamanderfish, and Stylephoriformes- the tube-eye or thread-tail). The blue line is one iteration of the model fragmentation of the morphospace ‘plate’ (comprising all fishes) into 78 orders that collectively include 6,732 species. The blue field represents 5 to 95% confidence limits from 1,000 iterations of this model fragmentation. Degrees of mono- and polytypy among such iterations (lower left and upper right graphs, respectively) completely span the ranges (6,732 to 1) of memberships observed for fishes. That is, biodiversity among these orders is no greater (or less) than that expected for the random fragmentation of the total morphospace now occupied by fish.

with a similar range of genera per family (few to many). That is, genera containing greater or fewer numbers of species are equally distributed among the various families, irrespective of species memberships (Figure 14). This independence of memberships among successively higher levels of Linnaean taxonomy is in accord with inferences derived from the broken plate model.

Monotypy, Polytypy, and Adaptive Radiations

Similarity between observed fish memberships and those expected for a random division of fish morphospace also bears on interpretations of monotypy and polytypy at different levels of taxonomic consideration. Virtually all membership frequencies are characterized by abundance at

some taxonomic levels and scarcity at others (e.g. Figure 7). This abundance of some groups (e.g. beetles, Evans, 1975; teleost fish, Kochmer and Wagner, 1988) and scarcity of others has led to speculation that different factors may serve to amplify or dampen disparity in taxonomic richness (e.g. Dial and Marzluff, 1989). Moreover, many have suggested that high abundances are driven by various processes of “adaptive radiation” directly or indirectly related to some key innovation (Allmon, 1992) such as changes in body size (May, 1986), reproductive schemes (Stuart-Fox and Owens, 2003), ecological adaptation (MacArthur et al., 1966), and geographical range (Cardillo et al., 2003).

In spite of a rich literature on, and descriptions

of, adaptive radiation, the broken plate analogy implies that such explanations may be overstated. Differences in memberships are to be expected during the random partitioning of any higher taxonomic group into some number of lower taxonomic units. If observed degrees of monotypy and polytypy among fishes (and probably other groups) are no different than that expected from random division, then seemingly large memberships (i.e. adaptive radiants) are not unexpected and therefore necessitate no validation. Extreme radiation only requires rationalization if one assumes that membership numbers should be subequal.

As noted above, Perciformes (perch-like fish) is the most diverse of the 78 orders of fishes, comprising 20% of families, 21% of genera, and 20% of species. Why this number of polytypes? Of the 78 orders, three (~4%) contain a single species (Amiiformes- the bowfin; Lepidogalaxiiformes- the salamanderfish, and Stylephoriformes- the tube-eye or thread-tail). Why this number of monotypes? Although a literature exists on the nature of radiation of among the most polytypic of these groups (e.g., Schaeffer and Rosen, 1961, Matschiner et al., 2015; Aguilar-Medrano et al., 2015), we might first ask if these seemingly small (monotypy) or large (polytypy) numbers are any more or less than would be expected from a random taxonomic division.

The broken plate formulation allows us to address that question. If 33,912 species of fishes were repeatedly (but randomly) parsed among 78 orders, how many orders would contain only three species, and where would an abundance of 6,732 species that comprise the Perciformes fall within this continuum? Simulation of such a pattern of allotment of species among orders of fishes is readily accomplished by bootstrapping relations in equations (2) and (3), where lower taxonomic levels (the ‘steps’ = 33,192 species) are parsed among the 78 intermediate taxonomic groups (the ‘segments’ = orders) that make up this (the ‘stick’) group of organisms. Repeated bootstrapping shows that observed numbers of monotypes (three) and

the greatest polytypy (6,732 species) each fall well within ranges expected from the random division of the fish morphospace (Figure 15).

The main lesson to be learned from this exercise is that the general nature of taxonomic membership is such that large proportions of groups contain few members, and small proportions contain many. Moreover, the presence of monotypy in three orders and 6.732 in one order is no more or less than would be expected from effecting such a taxonomic practice. In this regard, it is probably no more fruitful to ponder the reasons for small or large memberships, or to speculate about which “key innovation” gave rise to presumed adaptive radiation, than it is to deliberate on the reasons for many short and few long segments that comprise a randomly broken stick.

ACKNOWLEDGEMENTS

An early version of this paper benefited from excellent reviews by Michael Foote, Shanan Peters, and Gerald Smith. Long ago and far away, Smith took an enthusiastic but naive fledgling professor to Idaho to study rocks and fossils. Over a dozen summer months and as many years, he narrowed the lad’s focus and sharpened his interest until the young man eventually became a half-assed Earth Scientist. We thank Smith for all those kindnesses. Really close friends are as rare as a good salmon skull.

REFERENCES

- Aguilar-Medrano, R., Reyes-Bonilla, H., and Polly, D.P., 2015, Adaptive radiation of damselfishes (Perciformes, Pomacentridae) in the eastern Pacific: *Marine Biology*, 162:2291–2303, DOI 10.1007/s00227-015-2759-9
- Allmon, W. D. 1992, A causal analysis of stages in allopatric speciation: *Oxford surveys in evolutionary biology* 8:219-219.
- Alroy, J., 2002, How many named species are valid?: *PNAS*, 99:3706–3711, /doi/10.1073/pnas.062691099.
- Alroy, J., M., Aberhan, D. J. Bottjer, M. Foote, F. t. Fürsich, P. J. Harries, A. J. W. Hendy, S. M. Holland, L. C. Ivany, W.

- Kiessling, M. S. Kosnik, S. R. Marshall, A. J. McGowran, A. I. Miller, T. D. Olszewski, M. E. Patzkowsky, S. E. Peters, L. Villier, P. J. Wagner, N. Bonuso, P. S. Borkow, B. Brenneis, M. E. Clapham, L. M. Fall, C. A. Furguson, V. L. Hanson, A. Z. Krug, K. M. Layou, E. H. Leckey, S. Nürnberg, C. M. Powers, J. A. Sessa, C. Simpson, A. Tomašových, and C. C. Visaggi. 2008. Phanerozoic trends in the global diversity of marine invertebrates. *Science* 321:97-100.
- Anderson, S. 1974. Patterns of faunal evolution. *Quarterly Review of Biology* 49:311-332.
- Benton, M., 1995, Diversification and Extinction in the History of Life: *Science*, 268:52-58.
- Bertrand, Y., Pleijel, F., and Rouse, G.W., 2006, Taxonomic surrogacy in biodiversity assessments, and the meaning of Linnaean ranks: *Systematics and Biodiversity* 4:149-159.
- Burgess, P.M., 2008, The nature of shallow-water carbonate lithofacies thickness distributions: *Geology*, 36:235–238, DOI: <https://doi-org.proxy.lib.umich.edu/10.1130/G243326A.1>
- Burlando, B. 1990. The fractal dimension of taxonomic systems. *Journal of Theoretical Biology* 146:99-114.
- Callaghan, C.T., Nakagawa, N., and Cornwell, W.K., 2021, Global abundance estimates for 9,700 bird species. *PNAS*, 118:e2023170118, <https://doi.org/10.1073/pnas.2023170118>.
- Cardillo, M., J. S. Huxtable, and L. Bromham. 2003. Geographic range size, life history and rates in Australian mammals. *Journal of Evolutionary Biology* 16:282–288.
- Chamberlin, J. C. 1924. The hollow curve of distribution. *American Naturalist* 58:350-374.
- Chu, J., and C. Adami. 1999. A simple explanation for taxon abundance patterns. *Proceedings of the National Academy of Sciences USA* 96:15017-15019.
- Claverie, T., and Wainwright, P.C., 2014, A morphospace for reef fishes: elongation Is the dominant axis of body shape evolution: *PLOS ONE* 9:e112732.
- Davis, J. C., 1986, *Statistics and data analysis in geology*: New York, John Wiley and Sons, 646 p.
- Dial, K. P., and J. M. Marzluff. 1989. Nonrandom diversification within taxonomic assemblages, *Systematic Zoology* 38:26-37.
- Drummond, C.N., and Wilkinson, B.H., 1993, Aperiodic accumulation of cyclic peritidal carbonate: *Geology*, 21:1023-1026.
- Drummond, C.N., and Wilkinson, B.H., 1996, Stratal Thickness Frequencies and the Prevalence of Orderedness in Stratigraphic Sequences: *Journal of Geology*, 104:-18.
- Evans, G. E. 1975. *The life of beetles*. Hafner, New York.
- Foote, M., 1997, The Evolution of Morphological Diversity: *Annual Review of Ecology and Systematics*, 28:129-152, <https://doi.org/10.1146/annurev.ecolsys.28.1.129>.
- Fisher, R. A., A. S. Corbet, and C. B. Williams. 1943. The relationship between the number of species and the number of individuals in a random sample of an animal population. *Journal of Animal Ecology* 12:42-58.
- Freitas, T.M.S, Montag, L.F.A., DeMarco, P., AND Hortal, J., 2020, How reliable are species identifications in biodiversity big data? Evaluating the records of a neotropical fish family in online: *Systematics and Biodiversity*, 18:181–191.
- Goldhammer, R.K., Lehman, P.J., and Dunn, P.A., 1993, The origin of high-frequency platform carbonate cycles and third-order sequences (Lower Ordovician El Paso Gp, West Texas)- constraints from outcrop data and stratigraphic modeling: *Journal of Sedimentary Petrology*, v. 63, p. 318–359.
- Harris, P.M., Purkis, S.J., and Ellis, J., 2015, Analyzing spatial patterns in modern carbonate sand bodies from Great Bahama Bank: *Journal of Sedimentary Research*, 81:185–206.
- Joppa, L.N., Roberts, D.L., and Pimm, S.L., 2011, How many species of flowering plants are there? *Proceedings Royal Society – Biology*, 278:554–559, doi:10.1098/rspb.2010.1004
- Kallimanis, A.S., Mazaris, A.D., Tsakanikas, D., Dimopoulos, P., Pantis, J.D., and Sgardelis, S.P., 2012, Efficient biodiversity monitoring: Which taxonomic level to study?: *Ecological Indicators* 15:100–104.
- Kendall, D. G. 1948. On some modes of population growth leading to R.A. Fisher's logarithmic series distribution. *Biometrika* 35:6-15.
- Kochmer, J. P., and R. H. Wagner. 1988. Why are there so many kinds of passerine birds? Because they are small; A reply to Raikow. *Systematic Zoology* 37:68-69.
- Linnaeus, C., 1758, *Systema naturæ per regna tria naturæ, secundum classes, ordines, genera, species, cum characteribus, differentiis, synonymis, locis*: 10th ed., Stockholm: Laurentius Salvius, 1–824.
- MacArthur, R. H. 1957. On the relative abundances of birds. *Proceedings of the National Academy of Sciences USA* 43:293-295.
- MacArthur, R. H., H. F. Recher, and M. L. Cody. 1966. On the relation between habitat selection and species diversity. *American Naturalist* 100:319–332.
- Matschiner, M., Colombo, M., Damerau, M., Ceballos, S., Hanel, R. < and Salzburger, W., 2015, The Adaptive Radiation of Notothenioid Fishes in the Waters of Antarctica: in R. Riesch et al. (eds.), *Extremophile Fishes*, p. 35-57, DOI 10.1007/978-3-319-13362-1_3
- May, R. M. 1986. The search for patterns in the balance of nature—advances and retreats. *Ecology* 67:1115–1126.
- Mazaris, A.D., Kallimanis, A.S., Tzanopoulos, J., Sgardelis, S.P., and Pantis, J.D., 2010, Can we predict the number of plant species from the richness of a few common genera, families or orders?: *Journal of Applied Ecology* 47:662–670. doi: 10.1111/j.1365-2664.2010.01814.x.
- Mees, G.F. 1961, Description of a new fish of the family

- Galaxiidae from Western Australia: *Journal Royal Society Western Australia*, 44: 33–38.
- Minelli, A., G. Fusco, and S. Sartori. 1991. Self-similarity in biological classifications. *BioSystems* 26:89–97.
- Miya, M., Holcroft, N.I., Satoh, T.P., Yamaguchi, M., Nishida, M. and Wiley, E.O., 2007, Mitochondrial genome and a nuclear gene indicate a novel phylogenetic position of deep-sea tube-eye fish (Stylephoridae): *Ichthyological Research*, 54:323–332, doi.org/10.1007/s10228-007-0408-0.
- Mora, C., Tittensor, D.P., and Myers, R.A., 2008, The completeness of taxonomic inventories for describing the global diversity and distribution of marine fishes: *Proceedings Royal Society – Biology*, 275:149–155, doi:10.1098/rspb.2007.1315.
- Mora, C., Tittensor, D.P., Adl, S., Simpson, A.G.N., and Worm, B., 2011, How many species are there on Earth and in the ocean?: *PLoS Biology*, 9- e1001127.
- Moura, M.R., Jetz, W. 2021, Shortfalls and opportunities in terrestrial vertebrate species discovery. *Nat Ecol Evol*, <https://doi.org/10.1038/s41559-021-01411-5>
- Morrison, D.A., 2014, Is the Tree of life the best metaphor, model, or heuristic for phylogenetics?: *Systematic Biology*, 63:628–638.
- Nee, D., A. O. Mooers, and P. H. Harvey. 1992. Tempo and mode of evolution revealed from molecular phylogenies. *Proceedings of the National Academy of Sciences USA* 89:8322–8326.
- Newman, E., 1947, Query concerning the first section of Mr. Babbington's genus *Centaurea*: *The Phycologist*, 2:925.
- Preston, F.W., 1948, The commonness and rarity of species, *Ecology*, v. 29, n.3., p. 254–283.
- Purkis, S., Riegl, B.M., and Andréfouët, S., 2005, Remote Sensing of Geomorphology and Facies Patterns on a Modern Carbonate Ramp (Arabian Gulf, Dubai, U.A.E.): *Journal of Sedimentary Research*, 75:861–876.
- Purkis, S., Casini, G., Hunt, D., and Colpaert, A., 2015, Morphometric patterns in Modern carbonate platforms can be applied to the ancient rock record: Similarities between Modern Alacranes Reef and Upper Palaeozoic platforms of the Barents Sea: *Sedimentary Geology*, 321:49–69
- Rankey, E.C., 2002, Spatial Patterns of Sediment Accumulation on a Holocene Carbonate Tidal Flat, Northwest Andros Island, Bahamas: *Journal of Sedimentary Research*, 72: 591–601.
- Reddingius, J. 1971. Gambling for existence: a discussion of some theoretical problems in animal population ecology. *Acta Biotheoretica* 20(Suppl.):3–208.
- Ricotta, C., Ferrari, M., and Avena, G., 2002, Using the scaling behavior of higher taxa for the assessment of species richness: *Biological Conservation* 107: 131–133
- Schaeffer, B., and R. E. Rosen. 1961. Major adaptive levels in the evolution of the actinopterygian feeding mechanism. *American Zoologist* 1:187–204.
- Scotland, R. W., and M. J. Sanderson. 2004. The significance of few versus many in the tree of life. *Science* 303:643.
- Stuart-Fox, D., and I. P. F. Owens. 2003. Species richness in agamid lizards: chance, body size, sexual selection or ecology? *Journal of Evolutionary Biology* 16:659–669.
- Swan, A. R. H., and Sandilands, M., 1995, *Introduction to geological data analysis*: London, Blackwell Science, 446 p.
- Tutin, S.L., and Butler, R.J., 2018, The completeness of the fossil record of plesiosaurs, marine reptiles from the Mesozoic: *Acta Palaeontologica Polonica*: 62:563–573, doi:10.4202/app.00355.
- Walters, S. M. 1961. The shaping of angiosperm taxonomy. *New Phytologist* 60:74–84.
- Weiss, M.P., and Multer, H.G., 1988, *Modern Reefs and Sediments of Antigua, West Indies*: Department of Geology, Northern Illinois University, DeKalb, Illinois.
- Weller, J.M., 1969, *The Course of Evolution*, McGraw-Hill, 696 p.
- Wilkinson, B. H., and C. N. Drummond. 2004. Facies mosaics across the Persian Gulf and around Antigua-Stochastic and deterministic products of shallow-water sediment accumulation. *Journal of Sedimentary Research* 74:513–526.
- Wilkinson, B.H., 2011, On taxonomic membership: *Paleobiology*, 37(3), 2011, pp. 519–536.
- Williams, C. B. 1944. Some applications of the logarithmic series and the index of diversity to ecological problems. *Journal of Ecology* 32:1–44.
- Williams, C. B. 1964. *Patterns in the balance of nature*. Academic Press, London.
- Willis, J. C. 1922. *Age and area*. Cambridge University Press, Cambridge.
- Wilson, D.E., and Reeder, D.M., 2005, *Mammal Species of the World. A Taxonomic and Geographic Reference* (3rd ed), Johns Hopkins University Press, 2,142 pp.
- Wright, S. 1941. The “age and area” concept extended (Review of “The course of evolution by differentiation or divergent mutation rather than by selections,” by J. C. Willis). *Ecology* 22:345–347.
- Yule, G. U. 1924. A mathematical theory of evolution, based on the conclusions of Dr. J. C. Willis. *Philosophical Transactions of the Royal Society of London A* 213:21–87.
- Zelditch, M.L., Swiderski, D.L., and Sheets, D.H., 2012, *Geometric morphometrics for biologists: a primer*: Academic Press, 488 p. <https://doi.org/10.1016/C2010-0-66209-2>.