The fractal dimension of the tree of life

Bin Ma[a], Xiaofei Lv[a,b], Jun Gong[a,*]

[a] *Laboratory of Microbial Ecology and Matter Cycles , Yantai Institute of Coastal Zone Research, Chinese Academy of Sciences, Yantai, 264003, China*
[b] *University of Chinese Academy of Sciences, Beijing 100049, China*

Abstract

The structure pattern of the tree of life clues on the key ecological issues; hence knowing the fractal dimension is the fundamental question in understanding the tree of life. Yet the fractal dimension of the tree of life remains unclear since the scale of the tree of life has hypergrown in recent years. Here we show that the tree of life display a consistent power-law rules for inter- and intra-taxonomic levels, but the fractal dimensions were different among different kingdoms. The fractal dimension of hierarchical structure (D_r) is 0.873 for the entire tree of life, which smaller than the values of D_r for Animalia and Plantae but greater than the values of D_r for Fungi, Chromista, and Protozoa. The hierarchical fractal dimensions values for prokaryotic kingdoms are lower than for other kingdoms. The D_t value for Viruses was lower than most eukaryotic kingdoms, but greater than prokaryotes. The distribution of taxa size is governed by fractal diversity but skewed by overdominating taxa with large subtaxa size. The proportion of subtaxa in taxa with small and large sizes was greater than in taxa with intermediate size. Our results suggest that the distribution of subtaxa in taxa can be predicted with fractal dimension for the accumulating taxa abundance rather than the taxa abundance. Our study determined the fractal dimensions for inter- and intra-taxonomic levels of the present tree of life. These results emphases the need for further theoretical studies, as well as predictive modelling, to interpret the different fractal dimension for different taxonomic groups and skewness of taxa with large subtaxa size.

Introduction

Understanding the pattern of the tree of life has long been a driving force for system biologist. Since the dramatic development in molecular technology, there has been an exponential growth in the number of clades in tree of life each year (Ciccarelli et al., 2006). The end point of the tree of life is the construction of the single phylogenetic tree linking all species living and extinct (Benton and Ayala, 2003). The hierarchical structure of tree of life contains valuable clues on the key issue of realizing the modern diversification of life (Mora et al., 2011; Tittensor et al., 2010), accessing the shape of evolution (Doolittle, 2009), and determining where the diversity is threatened (Mace, Gittleman, and Purvis, 2003) and the underlying mechanisms that constrain ecological complexity (Solow, 2005). Fractal phenomena, which is a mathematical object that has a fractal dimension that usually exceeds its topological dimension and may fall between the integers, are widespread in nature (Brown et al., 2002). The tree of life has long been recognized as a fractal structure, including the diversity of life and taxonomic systems in the tree of life, which have self-similar features that look the same when there is a change in scale (Burlando, 1990, 1993; Chaline, Nottale, and Grou, 1999). Recently released OneZoom visualise the tree of life based on an adaptation of fractal mind (Rosindell and Harmon, 2012).

The fractal property of the tree of life shows a emergent feature that scaling relationship are self-similar over a wide range of taxon scales (Rabosky, Slater, and Alfaro, 2012; Lane, 2011; Foote, 2012; Chaline, Nottale, and Grou, 1999; Solow, 2005; Marquet et al., 2005; Burlando, 1990, 1993; Herrada et al., 2008). Power-law curve of size-frequency distributions of taxa, derived from a number of checklists and catalogues of species concerning protists, fungi, plants, and animals, pointed out a very large number of taxa with one or few subtaxa and a very small number of taxa with many subtaxa (Mora et al., 2011; Burlando, 1990). The fractal dimension of the taxonomic assemblages represents their diversity characterization, which is viewed as an

*Laboratory of Microbial Ecology and Matter Cycles , Yantai Institute of Coastal Zone Research, Chinese Academy of Sciences, Yantai, 264003, China. Tel: 86-535-2109123
Email address: jgong@yic.ac.cn (Jun Gong)

PeerJ PrePrints | http://dx.doi.org/10.7287/peerj.preprints.198v2 | CC-BY 3.0 Open Access | received: 13 Jan 2014, published: 13 Jan 2014

evolutionary pattern related to scaling evolutionary processes (Villarreal, 2006; Marquet et al., 2005; Bapteste et al., 2009; Foote, 2012). The investigation in the fractal geometry of the taxonomic system from both fossil record and phylogenetic systems indicated that arrangement of life taxonomy generally show fractal properties reflects evolutionary feature (Rabosky, Slater, and Alfaro, 2012). The branching patterns of a large set of phylogenetic tree follow allometric rules conserved across the different levels in the Tree of Life (Herrada et al., 2008). The universal patterns of phylogenetic differentiation suggests that similar evolutionary forces drive diversification across the broad range of scale, shaping the diversity of life in the planet (Brown et al., 2002).

The fractal dimension of taxonomic systems have been previously estimated based on the size-frequency distributions of taxa with different number of subtaxa (Burlando, 1990). Non-random occurrence of fractal dimension values among groups suggests a relationship with true biologic diversity patterns. The largest checklist used in this study contained 70000 species, and catalogues of species concerning protists, fungi, plants and animals. At present the number of species in tree of life, however, reach to more than 200000 species, and with different kingdom system (Delsuc, Brinkmann, and Philippe, 2005). Understanding the newly pattern of tree of life require knowing the fractal properties of the tree of life. Here we analyze the hierarchic structure of the global tree of life are obtained, allowing a characterization of the tree of life through the estimation of its fractal dimension. This emphasizes the self-similar relationship for size-frequency distributions of both the hierarchic size among taxon levels and the subtaxa diversity in each taxon.

Materials and Methods

Data sets

The data sets used in this paper were based on the classification of currently valid species from the Taxonomy Database of NCBI (`http://www.ncbi.nlm.nih.gov/taxonomy`) and the Catalogue of Life (`www.sp2000.org`). The eukaryotic species in the former is largely contained within the latter, whereas the prokaryotic species in the latter is largely contained within the former. These databases were screened for homonyms and the classification of taxa into multiple clades. The combined data sets included five eukaryotic dominations and two prokaryotic dominations, with 126 phylum, 259 class, 4163 order, 14939 family, 2031430 genus, and 2031430 species.

The self-similar hierarchy of tree of life

To describe the fractal property of tree of life, we probe the self-similarity pattern among hierarchical level of tree of life and within each taxonomic level. First, we related the logarithmic number of taxa against their numerical rank and estimated the parameters of linear models with least squares regression models:

$$lg(N_r) = -D_r + \mu \tag{1}$$

where N_r denotes the number of taxon in taxonomic level r, $\mu = lg(N_1) + r_1 D_r$ is the proportionality coefficient, and D_r is the fractal dimension among taxonomic levels in the tree of life. Since data are not strictly independent across hierarchically organized taxa, we used models based on generalized least squares assuming autocorrelated regression errors.

Secondly, for each taxonomic rank from phylum to genus, we represent the frequency distribution of taxa with different subordinate taxa abundance in each taxa as a rank-abundance curve. The taxa were arranged in increasing order of the abundance of its subordinate, and taxa frequency were plot as a function of the abundance of subordinate taxa. The probability distribution of the frequency of taxa with different subordinate taxa abundance, $P(k)$, can be represented by a power-law (scale free) with the subordinate taxa abundance:

$$P_k = P_1 k^{-D_t} \tag{2}$$

in which P_1 refer to the number of the taxon with only one subordinate taxa, k to the size rank of the subordinate taxa abundance, and D_t to the scaling exponent, which also called ”fractal dimension” for . We used logarithmic scales for both axes of the rank-abundance curve, so that the power-law abundance distribution is represented as a straight line, and the slope is equal to the power-law exponent. We estimated the scaling exponent of power law using maximum likelihood method with BFGS optimization to eliminate the influence of the long tail at low frequency. In order to include the influence of long tail, we converted the taxa abundance curve into the accumulating taxa abundance based on the rank of subordinate taxa:

$$N_k = N_1 k^{-D_{ac}} \tag{3}$$

in which N_1 to the total number of taxa, and D_{ac} to the fractal dimension for accumulated taxa frequency distribution. Combining equation 2 and 3 yield such a relation:

$$\left(\frac{P1}{P_k}\right)^{\frac{1}{D_t}} = \left(\frac{N1}{N_k}\right)^{\frac{1}{D_{ac}}} \tag{4}$$

PeerJ PrePrints | http://dx.doi.org/10.7287/peerj.preprints.198v2 | CC-BY 3.0 Open Access | received: 13 Jan 2014, published: 13 Jan 2014

In term of equation 4, we have:

$$D_a c = D_t \times \log_{\frac{P1}{P_k}}(\frac{N1}{N_k}) \quad (5)$$

The distribution of taxonomic abundance is:

$$M_k = k \times P_1 k^{-D_t} = P_1 K^{1-D_t} = P_1 K^{-D_a} \quad (6)$$

in which M_k refer to the taxa abundance at different rank, $D_a = 1 - D_t$ to the slope of taxa abundance distribution.

Results

Fractal property of hierarchical structure in the tree of life

We complied 2 million currently valid species of the tree of life from publicly accessible database. The power-law relation between abundance in each taxa level and the rank of each taxa level indicated the fractal property of hierarchical structure in the tree of life. For the entire tree of life we find fractal dimension of hierarchical structure D_r=0.873 (Figure 1). Figure 2 shows the power-law property of hierarchical structure for eight kingdoms. It reveal that D_r=1.004 and 0.889 for Animalia and Plantae, respectively. Among all eukaryotic kingdoms, only the D_r for Animalia and Plantae were greater than the entire tree of life. The values of D_r for Fungi, Chromista, and Protozoa were all lower than the entire tree of life, with hierarchical fractal dimension D_r(fungi)=0.8, D_r(chromista)=0.586, and D_r(protozoa)=0.573, respectively. For archaea and bacteria, we find D_r values were 0.444 and 0.521, respectively. This results indicates that hierarchical fractal dimensions for prokaryotic kingdoms are significantly lower than for other kingdoms. The Viruses have hierarchical fractal dimension D_r=0.596, which is lower than most eukaryotic kingdoms, but greater than prokaryotes.

The value of fractal dimension for hierarchical taxa level indicates the universal scaling between the number of taxa (N_t) and subordinate taxa (N_{st}): $N_{st} = N_t \times 10^{D_r}$. For the entire tree of life, the number of N_{st} is approximately 7.5 times of N_t since D_r value is 1.133. The D_r values for bacteria and archaea denote that the number of N_{st} is approximately 2.8 times and 3.3 times of N_t, respectively. The number of N_{st} is approximately 6.3 times to 10.1 times of N_t for Animalia, Plantae and Fungi. For Chromista, Protozoa and Viruses, the number of N_{st} is approximately 3.8 times of N_t.

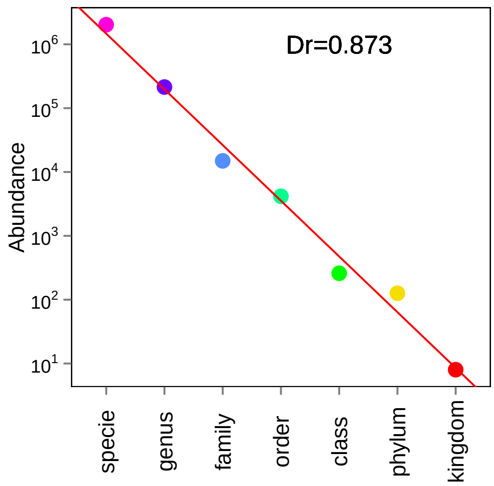

Figure 1: Relationship between the number of taxa and hierarchy of each taxonomic rank for the entire tree of life

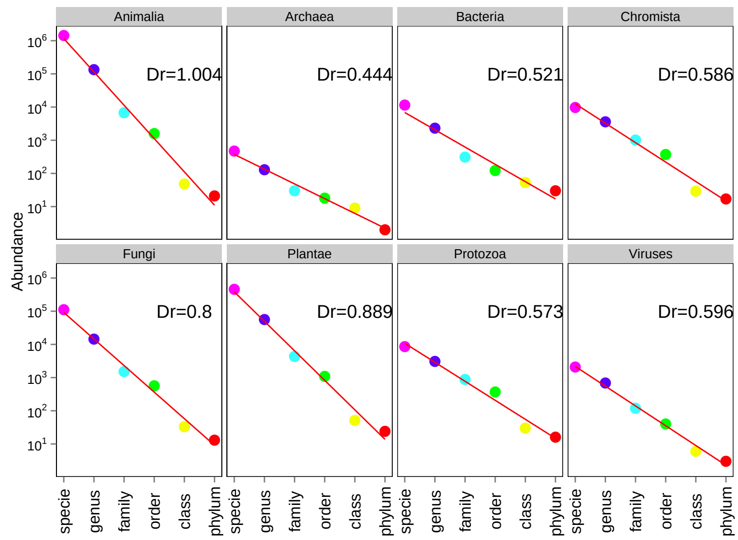

Figure 2: Relationship between the number of taxa and hierarchy of each taxonomic rank for each kingdom

Fractal property of each taxa level in the tree of life

An assessment of the size-frequency distribution of subtaxa in any taxa shows a consistent power-law relation between the number of subordinate taxa in each taxa and the frequency of taxa with same number of subordinate taxa at any taxonomic rank. Figure 3 shows the fractal dimension (power-law exponent D_t) for each taxa level of the entire tree of life. They reveals the existence of self-similarity in each taxa level with fractal dimension $D_t(genus)$=1.62, $D_t(family)$=1.42, $D_t(order)$=1.56, $D_t(class)$=1.57, and $D_t(phylum)$=1.72, respectively. The D_t values were estimated with power law model and can not represent the influence of the long tail of low frequency taxa with very large subtaxa size. We then converted the rank abundance plot into rank accumulating abundance plot in order to reflect the long tail. The fractal dimension (power-law exponent D_{ac}) for accumulating frequency plot is related with rank abundance fractal dimension D_t from Equation 5. The values of D_{ac} for different taxa level are $D_{ac}(genus)$=1.41, $D_{ac}(family)$=1.05, $D_{ac}(order)$=1.35, $D_{ac}(class)$=0.67, and $D_{ac}(phylum)$=1.11.

Tendency of D_t and D_{ac} values were different among eukaryote, prokaryote, and Viruses. Among eukaryote kingdoms, the variation tendency of fractal dimension D_t and D_{ac} is similar at same taxa levels. D_t and D_{ac} values at class and family levels were obviously smaller than at other taxa levels. For Animalia, Plantae and Fungi, the fractal dimension values were similar to values for entire tree of life. Except at class level, the D_t values for Protozoa and Chromista were greater than values for entire tree. The tendency of fractal dimension among prokaryotic taxa levels is different with eukaryotic kingdom. The fractal dimensions D_t for bacteria and archaea in all taxa levels were all greater than values for entire tree of life. D_t for genus and family level were smaller than other taxa level, but D_{ac} for different bacterial taxa levels were all closely to 1.3. D_t values for archaea were similar to bacteria, but D_{ac} values for archaea varied in wide range. The fractal dimension D_t for viruses were obviously greater than other kingdoms.

We also accounted for the distribution of subtaxa abundance in each taxa level (Fig 3). The mathematical description of the proportion distribution (Equation 4) indicated that the relation between abundance of subtaxa and size of subtaxa in taxa could be also represented by a power-law relationship. For the entire tree of life, the abundance of subtaxa in taxa with small size of subtaxa was greater than in taxa with large size of subtaxa. However, the abundance of subtaxa in taxa with same

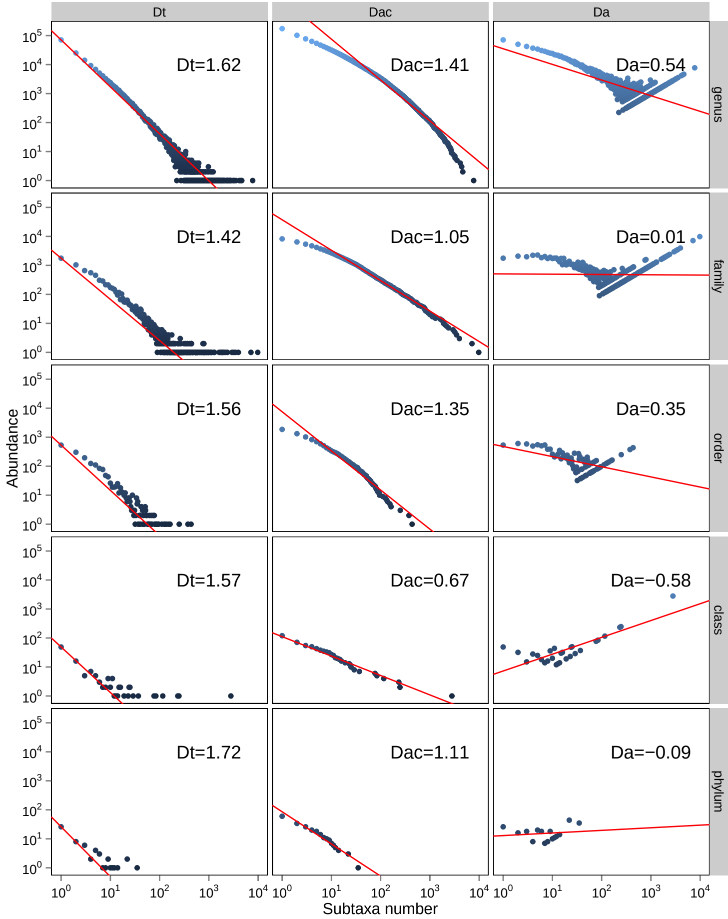

Figure 3: Relationship between the number of taxa and subtaxa number for the entire tree of life

PeerJ PrePrints | http://dx.doi.org/10.7287/peerj.preprints.198v2 | CC-BY 3.0 Open Access | received: 13 Jan 2014, published: 13 Jan 2014

frequency was increase linear with the size of subtaxa. Accordingly, the abundance of subtaxa in taxa with small and large sizes was greater than in taxa with intermediate size. Since $D_a = 1 - D_t$, the proportion for taxa with small size of subtaxa number should similar with the proportion for taxa with large size of subtaxa number when $D_t \approx 1$. However, the long tail of taxa with large size of subtaxa would skewed the slope of lineage regression for subtaxa abundance distribution. Accordingly, D_{ac} values could roughly refer D_a values because D_{ac} values can reflect the influence of long tail. For genus and order with $D_{ac} > 1$, the abundance of species and family in genus and order with small species and family number was greater than in genus and order with great species and family number, respectively. For class with $D_{ac} < 1$, the abundance of order in class with small order number was greater than in class with great order number. For family and phylum with $D_{ac} \approx 1$, the abundance of genus and class in family and phylum with small genus and class number was greater than in family and phylum with great genus and class number, respectively.

The distribution of subtaxa abundance was also different among eukaryote, prokaryote, and Viruses (Fig S1-S8). The D_a values of eukaryotic kingdoms were most similar with the entire of tree of life. The D_a values for genus and order of eukaryotic kingdoms were all greater than zero, whereas the D_a values for class of eukaryotic kingdoms were all smaller than zero. The D_a values for family of eukaryotic kingdoms were close to zero, except D_a values of Chromista and Protozoa which were approximately -0.5. The D_a values for genus, order, and class of archaea were greater than zero, and for family and phylum were less than zero. All D_a values of bacteria were greater than zero, indicating that the abundance of subtaxa in all bacteria taxa with small subtaxa number was greater than in taxa with great subtaxa number. On the contrary, all D_a values of Viruses were greater than zero.

Discussion

Knowing the fractal property of tree of life has been a question of great interest motivated in part of our collective curiosity about the evolution of different types of life, and in part by the need to provide a reference point for current and future pattern of biological diversity. In this paper, we describe two type of fractal property in the tree of life, the fractal dimension among different taxa levels (D_r) and between taxa and subtaxa (D_t).

The values of fractal dimension D_r were different among different kingdoms. The D_r value represents the scales among different taxa levels, which may promote the discrete evolutionary forces drive diversification across different domination. The evolutionary of organisms is driven with both natural variation and natural selection Foote (2012). We now know that the mechanisms of natural variation entailing recombination in its various forms differ starkly between prokaryotes and eukaryotes Bapteste et al. (2009); Drake (1999). Among eukaryotes, meiosis ensures reciprocal recombination among homologous chromosomes and reassortment of alleles within lineages that recombine within or very near (in the case of hybridization) species boundaries Ramesh, Malik, and Logsdon Jr (2005). But for prokaryotes, the mechanisms of natural variation are quantitatively, and many would say fundamentally, different from what goes on in sexual eukaryotes Bapteste et al. (2009); Lake (2009). These mechanisms include , gene transfer agents and integrons Frost et al. (2005). In eukaryotes, the D_r values for Fungi, Chromista and Protozoa were greater than Animalia and Plantae. The Animalia and Plantae kingdoms were multicellular organisms, but some organisms in the Fungi and Chromista kingdoms and all organisms in Protozoa were single celled organisms Woese, Kandler, and Wheelis (1990). It is already known that in single-celled eukaryotes, endosymbiosis and gene transfer are important processes for evolution Hotopp et al. (2007); Moustafa et al. (2009). Increasing knowledge of the genomes of protists may thus in the future expand our conclusion: not only are the tree of life and prokaryotic evolution are two different things, but all microbial evolution (that of prokaryotes and protists) may also be poorly described if addressed in an exclusively tree-like framework Adl et al. (2007). The fractal dimension for Virus was smaller than prokaryote, was greater than Animalia, Plantae and fungi in eukaryote, and was similar with Chromista and Protozoa in eukaryote. Nevertheless, fractal dimension value did not definitely indicate evolutionary force. The fractal dimension values for Virus was almost equal to the fractal dimension values for Chromista, but the evolutionary force for Virus and Chromista is obviously different Villarreal (2006); Drake (1999).

In each taxa levels, the frequency of taxa with same subtaxa sizes also follow the power-law, which indicated the fractal property of subtaxa numbers in each taxa level. Fractal dimension (D_t) indicates the distribution of taxa with rare and frequent subordinate taxa. The fractal dimension for genera calculated 30 year ago based on 70 000 species was 1.59 Burlando (1990), which is very similar to 1.62 reported in our study based on more than 2 million species. This results might suggests that the fractal pattern of diversity in genera is reli-

PeerJ PrePrints | http://dx.doi.org/10.7287/peerj.preprints.198v2 | CC-BY 3.0 Open Access | received: 13 Jan 2014, published: 13 Jan 2014

ably, although species number increase about 30 times. The D_t values for each taxa level were similar, whereas the D_{ac} values for class was extremely lower than for other taxa levels because of the scatting of points at the lower end. Similar fractal dimension at different taxa levels might implies that tree of life is scale free at both the pattern of both macro and micro evolution Raff (2000). The long tail in size frequency distribution appears that power law model under estimate the size of large taxa Burlando (1990). However, the distinctly difference between D_t and D_{ac} might suggests Burlando (1993). The skewness of distribution caused by the presence of large size taxa has been regard as an evidence of Blum and Francoise (2006). However, the skewness can also be viewed as a scaling cutoff, which shows a transition from fractal to non-fractal diversity Burlando (1990). Consequently, D_t values and scatting points denote frequency distribution of fractal to non-fractal distributing taxa, and D_{ac} values show the feature of both the fractal to non-fractal diversity. Subtaxa abundance distribution is also influence by both the fractal to non-fractal distributing taxa. Accordingly, D_a values represents the subtaxa abundance distribution is related to D_{ac} rather than D_t.

Fractal dimension of size frequency distribution of subtaxa in each taxa varied with kingdoms. The D_t values for multiple cellular organisms (e.g. Animalia, Plantae, and Fungi) were smaller than the D_t vales for singular cellular organisms (Bacteria, Archaeal, Pro, Protozoa and Chromista). The D_t values for Virus were greater than all other kingdoms. The D_t vales yield a diversity measure, since high D_t values indicates that taxa with one or a few-taxa are more numerous. In other words, the kingdom with high D_t values has proved with high diversity pattern. The results in present study might suggest that evolutionary scaling is closely related to the morphologic scaling of organisms Marquet et al. (2005). The size of an organism affects its all aspects of life, including metabolic, growth, mortality, and other vital rates Gouws, Gaston, and Chown (2011); Coetzee, le Roux, and Chown (2013). The body size spectra in natural community are scale-free, which is the product of intra- and interspecies regulation of the relative abundance of organisms of different sizes Giometto et al. (2013).

We recognize a number of factors that can influence the fractal property.

Taxa definitions. Different taxonomic communities use different levels of differentiation to define taxonomic levels Mora et al. (2011). This difference implies that the numbers of taxa for different taxonomic communities are not directly comparable. For example, the species concept for prokaryotes tolerates a much higher degree of genetic dissimilarity than in most eukaryotes opez Garcia and Moreira (2008). Species take longer to isolate in prokaryotes than in eukaryotes due to horizontal gene transfers among phylogenetic Ochman, Lawrence, and Groisman (2000). Thus, implication of estimated fractal dimensions are different for different taxonomic communities. Nevertheless, the aim of the present study is describing the hierarchical structure of the Tree of Life but not the topological property. We found that in any taxonomic communities, there is a constantly follows power-law rule for rank-abundance relationship between taxa number and diversity subordinate taxa.

Completeness of the tree of life. It is obviously that the tree of life is still incomplete at present Benton and Ayala (2003). The number of eukaryotic species have been estimated to be 8.7 million on earth, but the catalogued species is just 1.2 million at present Mora et al. (2011). Although the catalogued species number is approximately 10 thousand, it is believe that isolated prokaryotic species is only 1% of entire prokaryotic species at present Gich et al. (2012). Although the rate of catalogued species varied from 1% (prokaryotes) to 70% (Plantae), there is a constantly follows power-law rule for rank-abundance relationship between taxa number and diversity subordinate taxa. These results indicated that new discovered species might influence the fractal dimension but not the power-law relationship tendency. Furthermore, increase in the number of higher taxa will distort the shape of the current tree of life. Increasing number of new discovered higher taxa in ongoing for prokaryotes, but the number of catalogued higher taxa is almost reach the entire number of higher taxa for eukaryotes, except Chromista, Protozoa and fungi Mora et al. (2011). These results suggest that our fractal property analysis for prokaryotes and Chromista, Protozoa and Fungi in eukaryotes should be interpreted with that caution in mind.

Self-similarity of the tree of life. The tree of life is widely accept to be self-similarity. Sub-fractal structure for different taxonomic communities varied in a wide range due to various self-similarity property. However, in this study we do not concern the topological of tree of life, but the diversity number in each taxonomic level. The consistent patterns for entire tree of life imply that the different self-similarity in sub-fractal structure do not obscure the robust underlying relationship for inter- and intra-taxonomic levels.

In summary, the diversity for each level of the Tree of Life display a consistent power-law rules for inter- and intra-taxonomic levels. The discrepancy of frac-

PeerJ PrePrints | http://dx.doi.org/10.7287/peerj.preprints.198v2 | CC-BY 3.0 Open Access | received: 13 Jan 2014, published: 13 Jan 2014

tal nature indicates different evolutionary force for various kingdoms. The distribution of taxa size is governed by fractal diversity but skewed by overdominating taxa with low frequency. The distribution of subtaxa abundance is influence by both fractal and non-fractal overdominating taxa. The use of fractal geometry provides a unified view of diversity in tree of life and might therefore give clue to the evolutionary of tree of life.

References

Adl SM; Leander BS; Simpson AG; Archibald JM; Anderson OR; Bass D; Bowser SS; Brugerolle G; Farmer MA; Karpov S; et al. 2007. Diversity, nomenclature, and taxonomy of protists. *Systematic Biology* 56:684–689.

Bapteste E; O'Malley M; Beiko R; Ereshefsky M; Gogarten J; Franklin-Hall L; Lapointe FJ; Duprus J; Dagan T; Boucher Y; Martin W. 2009. Prokaryotic evolution and the tree of life are two different things. *Biology Direct* 4:34.

Benton MJ; Ayala FJ. 2003. Dating the tree of life. *Science* 300:1698–1700.

Blum MG; Francoise O. 2006. Which random processes describe the tree of life? a large-scale study of phylogenetic tree imbalance. *Systematic Biology* 55:685–691.

Brown J; Gupta V; Li BL; Milne B; Restrepo C; West G. 2002. The fractal nature of nature: power laws, ecological complexity and biodiversity. *Philosophical transactions of the Royal Society of London Series B, Biological sciences* 357:619–626.

Burlando B. 1990. The fractal dimension of taxonomic systems. *Journal of Theoretical Biology* 146:99–114.

Burlando B. 1993. The fractal geometry of evolution. *Journal of Theoretical Biology* 163:161–172.

Chaline J; Nottale L; Grou P. 1999. Is the evolutionary tree a fractal structure? *Comptes Rendus de l'Academie des Sciences Series IIA Earth and Planetary Science* 11:717–726.

Ciccarelli FD; Doerks T; Von Mering C; Creevey CJ; Snel B; Bork P. 2006. Toward automatic reconstruction of a highly resolved tree of life. *Science* 311:1283–1287.

Coetzee BWT; le Roux PC; Chown SL. 2013. Scale effects on the body size frequency distributions of african birds: patterns and potential mechanisms. *Global Ecology and Biogeography* 22:380–390.

Delsuc F; Brinkmann H; Philippe H. 2005. Phylogenomics and the reconstruction of the tree of life. *Nature Reviews Genetics* 6:361–375.

Doolittle W. 2009. The practice of classification and the theory of evolution, and what the demise of charles darwin's tree of life hypothesis means for both of them. *Philosophical transactions of the Royal Society of London Series B, Biological sciences* 364:2221–2228.

Drake J. 1999. The distribution of rates of spontaneous mutation over viruses, prokaryotes, and eukaryotes. *Annals of the New York Academy of Sciences* 870:100–107.

Foote M. 2012. Evolutionary dynamics of taxonomic structure. *Biology Letters* 8:135–138.

Frost LS; Leplae R; Summers AO; Toussaint A. 2005. Mobile genetic elements: the agents of open source evolution. *Nature Reviews Microbiology* 3:722–732.

Gich F; Janys MA; König M; Overmann J. 2012. Enrichment of previously uncultured bacteria from natural complex communities by adhesion to solid surfaces. *Environmental Microbiology* 14:2984–2997.

Giometto A; Altermatt F; Carrara F; Maritan A; Rinaldo A. 2013. Scaling body size fluctuations. *Proceedings of the National Academy of Sciences* 110:4646–4650.

Gouws EJ; Gaston KJ; Chown SL. 2011. Intraspecific body size frequency distributions of insects. *PLoS ONE* 6:e16606.

Herrada E; Tessone C; Klemm K; Eguluz V; Hernndez-Garca E; Duarte C. 2008. Universal scaling in the branching of the tree of life. *PloS ONE* 3:e2757.

Hotopp JCD; Clark ME; Oliveira DC; Foster JM; Fischer P; Torres MCM; Giebel JD; Kumar N; Ishmael N; Wang S; et al. 2007. Widespread lateral gene transfer from intracellular bacteria to multicellular eukaryotes. *Science* 317:1753–1756.

Lake JA. 2009. Evidence for an early prokaryotic endosymbiosis. *Nature* 460:967–971.

Lane N. 2011. Energetics and genetics across the prokaryote-eukaryote divide. *Biology Direct* 6:35.

Mace GM; Gittleman JL; Purvis A. 2003. Preserving the tree of life. *Science* 300:1707–1709.

Marquet P; Quiåones R; Abades S; Labra F; Tognelli M; Arim M; Rivadeneira M. 2005. Scaling and power-laws in ecological systems. *The Journal of Experimental Biology* 208:1749–1769.

Mora C; Tittensor D; Adl S; Simpson A; Worm B. 2011. How many species are there on earth and in the ocean? *PLoS biology* 9:e1001127.

Moustafa A; Beszteri B; Maier UG; Bowler C; Valentin K; Bhattacharya D. 2009. Genomic footprints of a cryptic plastid endosymbiosis in diatoms. *Science* 324:1724–1726.

Ochman H; Lawrence JG; Groisman EA. 2000. Lateral gene transfer and the nature of bacterial innovation. *Nature* 405:299–304.

opez Garcia P; Moreira D. 2008. Tracking microbial biodiversity through molecular and genomic ecology. *Research in Microbiology* 159:67–73.

Rabosky D; Slater G; Alfaro M. 2012. Clade age and species richness are decoupled across the eukaryotic tree of life. *PLoS biology* 10:e1001381.

Raff RA. 2000. Evo-devo: the evolution of a new discipline. *Nature Reviews Genetics* 1:74–79.

Ramesh MA; Malik SB; Logsdon Jr JM. 2005. A phylogenomic inventory of meiotic genes: evidence for sex in giardia and an early eukaryotic origin of meiosis. *Current Biology* 15:185–191.

Rosindell J; Harmon L. 2012. Onezoom: a fractal explorer for the tree of life. *PLoS biology* 10:e1001406.

Solow AR. 2005. Power laws without complexity. *Ecology Letters* 8:361–363.

Tittensor DP; Mora C; Jetz W; Lotze HK; Ricard D; Berghe EV; Worm B. 2010. Global patterns and predictors of marine biodiversity across taxa. *Nature* 466:1098–1101.

Villarreal LP. 2006. How viruses shape the tree of life. *Future Medicine* 1:587–595.

Woese CR; Kandler O; Wheelis ML. 1990. Towards a natural system of organisms: proposal for the domains archaea, bacteria, and eucarya. *Proceedings of the National Academy of Sciences* 87:4576–4579.

PeerJ PrePrints | http://dx.doi.org/10.7287/peerj.preprints.198v2 | CC-BY 3.0 Open Access | received: 13 Jan 2014, published: 13 Jan 2014