Article

An introduction to phylogenetic analyses and modelling in ecology

H.O. Minoarivelo[1], **G. Diedericks**[2], **C. Hui**[1,3]

[1]Centre for Invasion Biology, Department of Mathematical Sciences, Stellenbosch University, Matieland 7602, South Africa

[2]Centre for Invasion Biology, Department of Botany and Zoology, Stellenbosch University, Matieland 7602, South Africa

[3]Mathematical and Physical Biosciences, African Institute for Mathematical Sciences, Muizenberg 7945, South Africa

E-mail: chui@sun.ac.za

Received 25 July 2015; Accepted 3 August 2015; *Published online 1 December 2015*

Abstract

Phylogenetic systematics seeks to describe and reconstruct the evolutionary relationships among and between organisms making use of molecular data. This field has become immensely popular in recent years, with the associated computational demands growing in leaps and bounds. Here, we review the progress made in statistical phylogenetics, compare the various methods and highlight the recent trends and pitfalls. Furthermore, we delve into the mathematical models associated with these methods to understand the underlying assumptions, while tracking the improvements made. Lastly, we look at the impact and use of phylogenetics in ecology.

Keywords Markov process; Bayesian inference; substitution matrix; likelihood-based approach; parsimony.

Computational Ecology and Software
ISSN 2220-721X
URL: http://www.iaees.org/publications/journals/ces/online-version.asp
RSS: http://www.iaees.org/publications/journals/ces/rss.xml
E-mail: ces@iaees.org
Editor-in-Chief: WenJun Zhang
Publisher: International Academy of Ecology and Environmental Sciences

1 Introduction

Phylogenetics seeks to understand the evolutionary relationships among organisms making use of an array of data types, such as morphological traits or molecular sequence data (Hastings and Gross, 2012; Holland, 2013). Historically, morphological traits were commonly used to explain these relationships, but with the dawn of molecular sequencing techniques in the 1970's, DNA phylogenies became commonplace (Morrison, 2012). Homologous traits, i.e. traits that have a common origin but have been modified by decent, are the fundamental units of phylogenetic data, though these are easier to acquire for molecular data as each base pair is considered a homologous character (Holland, 2013). Initially, mitochondrial DNA (mtDNA) was used to infer relationships among taxa, using parsimony based methods and analyses. Gene histories, however, are strongly influenced by chance events, enhancing the probable inference error when looking at a single locus (Knowles and Maddison, 2002; Ronquist and Sanmartín, 2011). Furthermore, as the mitochondrial and nuclear genomes are independent from one another and thus, evolve independently, the notion to use only mtDNA quickly became obsolete. Hence, incorporating multiple loci spanning both the mitochondrial and nuclear

genomes has become the method of choice, supplanting the parsimony-based methods with likelihood-based statistical approaches and Bayesian inference techniques (Cutter, 2013; Morrison, 2012).

1.1 Parsimony

Parsimony-based approaches use the 'path of least resistance' when searching for the most probable tree (Hastings and Gross, 2012; Ronquist and Sanmartín, 2011), while maximum parsimony (MP) is the optimality criterion associated with selecting the tree that requires the least number of mutations to explain the observed sequence data (Holland, 2013). Even though it has been shown that likelihood-based methods are more reliable than those obtained through the parsimony-based approach (Sorhannus, 2003), it is still common to incorporate parsimony-based methods to validate likelihood results (Robinson et al., 2014). Recently, Ronquist and Sanmartín (2011) even argued that "the power of the parsimony approach still has not been fully explored". Furthermore, haplotype networks are often used to infer relationships at population level, indicating the number of base pair changes between individuals (Teacher and Griffiths, 2011), however, the most widely used haplotype network program, TCS (Clement et al., 2000), still uses a parsimony-based approach (Table 1). Hence, parsimony-based approaches may remain viable alternatives to maximum likelihood and Bayesian approaches, as predicted by Ronquist and Sanmartín (2011).

Table 1 The most common Parsimony, Likelihood and Bayesian methods and programs currently used in phylogenetic studies. All citation values are based on Google Scholar information available January 2015.

Inference method	Program	Author	Used for	Output	Release date	Citations
Parsimony	PAUP*	Swofford	Phylogenetic Analysis	Trees	2002	9168
	TCS	Clement et al.	Estimating gene genealogies	Haplotype network	2000	5472
Likelihood	MEGA5	Tamura et al.	Inferring evolutionary trees etc.	Trees	2011	18259
	RAxML-VI-HPC	Stamatakis	Inference of large phylogenies	Trees	2006	6086
	PhyML 3.0	Guindon et al.	Estimating phylogenies	Trees	2010	1012
Bayesian Inference	MrBayes	Huelsenbeck & Ronquist	Phylogeny inference	Trees	2001	14063
	BEAST	Drummond & Rambaut	Dated phylogeny	Trees	2007	6003

1.2 Likelihood

The maximum likelihood (ML) concept was first developed by Fisher in the early 1900's (Morrison, 2012), and refers to the maximum likelihood of producing the observed data (Nielsen and Beaumont, 2009; Ronquist and Sanmartín, 2011). Likelihood-based methods employ an explicit model of evolution, selecting the model parameters and the phylogenetic tree (topology with branch length) that maximize the possibility of obtaining the observed data (Hastings and Gross, 2012). Indeed, the likelihood of a tree is the probability of observing

the given n data sequences (S_i) placed at the tips of the tree T, given the tree structure and its branch lengths and parameters of the evolutionary model (θ): $P(S_1,\ldots,S_n|T,\theta)$. These models can have the same parameters across sites or be allowed to differ across sites, i.e. codon positions, taxa or loci (Hastings and Gross, 2012), with recent publications following the latter trend (Diedericks and Daniels, 2014). This is most often due to the fact that different codon positions evolve at different rates, incorporating more biological 'realism' into the model, as different parts of the genome will evolve at different rates due to selection (Hastings and Gross, 2012; Kumar et al., 2011). Although ML is considered the most challenging inference method, it has been shown to have the most robust confidence intervals (Ronquist and Sanmartín, 2011), employing nonparametric bootstrapping to evaluate the congruence between the data and the resulting tree (Holland, 2013).

1.3 Bayesian approach

Currently Bayesian inference (BI) is the most popular approach for analysing phylogenetic data. Although both likelihood and Bayesian approaches use the same set of models, the Bayesian approach, contrastingly, seeks to render the tree with the highest posterior probability (Pp) rather than find a tree that maximizes the data's probability (Hastings and Gross, 2012). To calculate this, the Markov chain Monte Carlo (MCMC) algorithm is used, which maximizes P(T|S) given by the Bayes' rule:

$$P(T \mid S_1,\ldots,S_n) = \frac{P(S_1,\ldots,S_n \mid T)P(T)}{P(S_1,\ldots,S_n)}.$$

Though ML and BI are quite similar in practice, ML is currently computationally more demanding, making it slower and technically more challenging than BI (Ronquist and Sanmartín, 2011). Furthermore, most studies incorporating both ML and BI analyses have found that the BI P(T|S) supersedes the ML bootstrapping values (Kumar et al., 2011), rendering more support to the same branch (e.g. Diedericks and Daniels, 2014). This is due to the fact that they are measuring different things. The BI approach seeks to estimate the P(T|S) of all the generated trees, while the ML bootstrapping is measuring the point estimate sensitivity (Hastings and Gross, 2012). Furthermore, P(T|S) is more sensitive to model violation, leading to inflated values, while bootstrap values are thought to be over sensitive, rendering conservative results (Kumar et al., 2011).

1.4 Models

Deciding on which model to use depends on the questions being asked, as each model has its own assumptions. Generally, ML and BI methods employ an Akaike Information Criterion (AIC) (Akaike, 1973) to select an appropriate model, however, an AIC corrected for small sample sizes (AICc), or Bayesian Information Criterion (BIC) (Schwarz, 1978) may also be used. The simplest DNA model, namely the Jukes-Cantor (JC) model, assumes that DNA substitutions are equally likely to occur across pyrimidines (C and T) and purines (A and G) and that all base pairs occur in equal frequencies. More complex models, such as the general time reversible (GTR) model, are more realistic and include six rate parameters describing the relative substitution rates between nucleotide pairs, and three nucleotide frequency parameters describing the base pair (A, T, G, C) proportions (Holland, 2013). However, an array of models varying in transition and transversion rates, shapes and distributions, span the continuum (Posada, 2006), with model selection usually being calculated via the goodness-of-fit test (Hastings and Gross, 2012), as implemented in programs such as MODELTEST (Posada and Crandall, 1998) and jModelTest (Posada, 2006).

2 Modelling Techniques in Phylogenetics

Models of evolution in phylogenetics differ depending on the methods and algorithms that are used for tree inference. Depending on the type of available data, models of evolution can be classified into two categories: distance matrix based models such as models used in UPGMA (Unweighted Pair Group Method with Arithmetic Mean) (Sokal and Michener, 1958) or NJ (Neighbour Joining) methods (Saitou and Nei, 1987), and

substitution models such as those used in ML and BI.

2.1 Distance matrix based models

In molecular genetics, a distance between a pair of sequences can be viewed as a measurement of their dissimilarity. It expresses the expected number of substitutions per site that have occurred between the pair of sequences and their common ancestor, and are often used as the length of the branches in phylogenetic trees. There are several ways to derive a distance matrix. It can directly be derived from the alignment of multiple sequences (Durbin et al., 1999). Other sources of genetic distance include the measurement of similarity between immunological data or a DNA-DNA hybridation method (Sibley and Ahlquist, 1984). The basic pairwise distance between two sequences (also called the p-distance) is the proportion of site that differs among them (fig 1). However, the p-distance often underestimates the real genetic distance because some of the characters in the sequence may have undergone multiple substitutions. Consequently, some models of molecular evolution (see section 2.2) offer a way to correct the p-distance. For example, the oldest attempt for an adjustment of the p-distance was made by Jukes and Cantor in 1969 (Jukes and Cantor, 1969). They defined the phylogenetic distance to be $d = -3\ln(1-4p/3)/4$ in order to model an increase of the effect of the correction as the p-distance is increasing and a saturation effect above a certain p-distance (Fig 1). Later on, more complicated molecular models of evolution have also proposed their own adjustment to the p-distance.

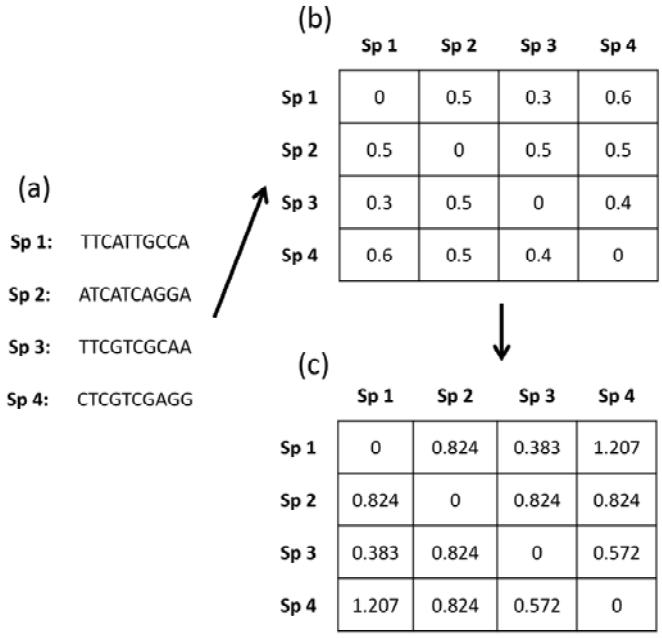

Fig. 1 Simple example of distance matrices computed from the sequence alignment of four species. (b) Standard p-distance or the proportion of dissimilarity between the sequences; (c) Jukes-Cantor distance matrix.

2.2 Substitution matrix based models

Phylogenetic tree inference methods that are frequently used nowadays, such as ML and BI make use of a substitution matrix offered by molecular evolution models. A molecular evolution model describes probabilistically the process at which a sequence of characters (DNA or protein) is substituted into another sequence of characters within a certain amount of time. The substitution process is usually modelled by a continuous time Markov process (Liò and Goldman, 1998), in which the probability of a character to be substituted into another character depends on the rate at which the substitution takes place and the time needed for substitution. Generally, the absolute time is not used and is in practice scaled by the expected number of substitutions per site. The rate at which a character is substituted into another one is stored in a substitution

matrix which characterizes the evolutionary model.

A substitution matrix can be empirical or parametric. On the one hand, empirical substitution matrices have been mainly used for amino acid sequences. They are built by direct statistical analysis on the observed frequency of each character substitution in an empirical sequence dataset. Common examples of empirical substitution matrices for amino acid sequences are the BLOSUM (Block Substitution Matrix) or the PAM (Point Accepted Mutation) matrices.

On the other hand, parametric substitution matrices are commonly used for DNA sequences. Parametric substitution matrices usually present two different types of key parameters. First, the equilibrium frequencies (π_A, π_C, π_G and π_T) describe the probability of having each of the nucleotide character when the Markov process is at equilibrium (Liò and Goldman, 1998). For them to be elements of a probability, the four equilibrium frequencies should sum up to one. Secondly, a set of rate parameters control the rate of substitution. The simplest substitution model is called JC69 and was developed by Jukes and Cantor in 1969 (Jukes and Cantor, 1969). Their model is simple in the sense that it only has one rate parameter because it assumes equal substitution rate for all nucleotides and has equal equilibrium frequency ($\pi_A = \pi_C = \pi_G = \pi_T = 1/4$). Hence, the Jukes-Cantor substitution matrix is given by:

$$Q = \begin{array}{c} \\ A \\ C \\ G \\ T \end{array} \begin{array}{c} \begin{array}{cccc} A & C & G & T \end{array} \\ \begin{bmatrix} -3\alpha & \alpha & \alpha & \alpha \\ \alpha & -3\alpha & \alpha & \alpha \\ \alpha & \alpha & -3\alpha & \alpha \\ \alpha & \alpha & \alpha & -3\alpha \end{bmatrix} \end{array} \tag{1}$$

in which α is the rate parameter.

Later on, since it was observed that transitions (substitutions between nucleotides A – G and C – T) occur more frequently than transversions (substitutions between nucleotides A – T, C – G, A – C and G – T), the assumption of equal substitution rate was found to be too simplistic. Thus, the Kimura model, also called K80, (Kimura, 1980) with one transition and one transversion rate parameters was proposed. Like the JC69 model, this model assumes equal frequencies of the nucleotide bases. The substitution rate matrix of the Kimura model has the form:

$$Q = \begin{array}{c} \\ A \\ C \\ G \\ T \end{array} \begin{array}{c} \begin{array}{cccc} A & C & G & T \end{array} \\ \begin{bmatrix} -2\beta-\alpha & \beta & \alpha & \beta \\ \beta & 2\beta-\alpha & \beta & \alpha \\ \alpha & \beta & 2\beta-\alpha & \beta \\ \beta & \alpha & \beta & 2\beta-\alpha \end{bmatrix} \end{array} \tag{2}$$

in which α is the transition rate parameter and β the transversion rate parameter.

Even more complicated parametric substitution models were proposed later on. For instance, the F81 model (Felsenstein, 1981) was inspired by the Jukes-Cantor model with only one rate parameter, but the equilibrium frequencies are not assumed to be equal. In the HKY85 model (Hasegawa et al., 1985), the principles of the F81 and the K80 models were combined: the rate of transversion and the rate of transition are

distinct and equilibrium frequencies are not assumed to be equal.

Hence, depending on the level of parameterization of the substitution matrix, one can have a substantial number and type of evolutionary models, the majority of which are not named. The GTR (general time reversible) model (Tavaré et al., 1986) generalizes all the models by assuming different rates and different frequencies for all nucleotides. Its substitution rate matrix is then given by:

$$Q = \begin{array}{c} \\ A \\ C \\ G \\ T \end{array} \begin{array}{c} \begin{array}{cccc} A & C & G & T \end{array} \\ \begin{bmatrix} . & \pi_C\alpha & \pi_G\beta & \pi_T\gamma \\ \pi_A\alpha & . & \pi_G\delta & \pi_T\varepsilon \\ \pi_A\beta & \pi_c\delta & . & \pi_T\theta \\ \pi_A\gamma & \pi_C\varepsilon & \pi_G\theta & . \end{bmatrix} \end{array} \tag{3}$$

in which α, β, γ, δ, ε and θ are all rate parameters. The diagonal elements of the matrix are set such that each row sums to zero.

Standard Markov process (Liò and Goldman, 1998) assumes independence of the evolution of each site in the sequence data. Consequently, the probability of a DNA sequence to evolve into another one would just be the product over all sites of the probability of substituting a character into another one. However, a significant improvement on phylogenetic reconstruction methods was the consideration of heterogeneity of evolutionary rates among different sites in the sequence. Due to different amount of selection imposed on the different sites, this approach was considered more realistic (Hasting and Gross, 2012). Evolutionary rate heterogeneity is implemented in most current phylogenetic software. The simplest of this approach considers that a proportion of sites are invariable while others evolve at a constant rate (e.g., Rannala and Yang, 1995). In this case, the Markov process is applied only to the sites that are evolving. It can also happen that the sequence data is partitioned into sets of sites (according to each gene for instance) so that different evolutionary model is applied to each of the set of sites. The most common approach nowadays is the use of a discrete gamma distribution to model rates of substitution across sites. Hence, the discrete gamma distribution adds one more parameter to the considered evolutionary model. Sometimes, the model can present both a discrete gamma distribution of evolutionary rates and a proportion of invariant sites. The most complex evolutionary model is then the GTR+I+G (GTR model with a proportion of invariant sites and gamma distributed rates).

In the context of studying the evolutionary history of a set of organisms, having the right estimates of parameters governing the evolutionary history can be as interesting as getting the right phylogenetic tree. These parameters, such as the rate of substitution or the time of divergence (given by the tree branch length) are usually incorporated into the considered evolutionary model. Hence, they are estimated together with the phylogenetic tree topology in the ML inference or the BI. However, in some cases when the phylogenetic tree is given, one can be interested in knowing the rate and number of evolutionary changes that have occurred from the most common ancestor to the studied organisms. In these cases, the maximum likelihood approach can also be used.

2.3 Model selection methods and support for the phylogenetic tree

Since a large number of evolutionary models exist, selecting the most appropriate one has become a central task. Most available model selection methods are based on the likelihood of using the model. A popular statistical method for model selection in phylogenetics is the likelihood ratio test (LRT) (Felsenstein, 1981; Swofford et al., 1996). We notice that some evolutionary models are nested, in the sense that a model is a constrained version of another one. For example, if the transition rate parameter in the K80 model is fixed

equal to the transversion rate parameter, then the K80 model will not differ from the JC69 model. LRT can only be applied to such nested models. Since the model with more parameter will always have a higher likelihood score compared to the simpler one, LRT aims at choosing whether the more complex model is significantly better than the simpler model. The LRT test statistic, giving the level of significance of using the more complex model, is $\delta = 2(\ln L_1 - \ln L_0)$ in which L_1 is the maximized value of the likelihood for the complex model and L_0 the maximized value of the likelihood for the simpler model. LRT statistics approximately follow a chi-square distribution. Unfortunately, using LRT requires a subsequent amount of computation as it compares the model two by two.

Recently, the Akaike Information Criteria (AIC) test has become the most common method for model selection in phylogenetics. Pioneered by Akaike (1973), the AIC test has been used first in phylogenetics by Hasegawa (1990). Like LRT, AIC is also based on comparing likelihoods for each model. However, it has the advantage of being suitable for both nested and non-nested models and of considering all the models simultaneously so as to make it computationally less intensive than LRT. The basic idea of AIC test is a likelihood comparison between models with a penalization for over-parameterization. The AIC score is given by: $2k - 2\ln L$ in which k is the number of parameters in the model and L denotes the maximized value of the likelihood for the model. The model having the smallest AIC score is selected to be the best model. Some phylogenetic studies make use of a corrected version of AIC (the AICc, Brunham and Anderson, 2004) which takes into account the bias caused by sample size.

For phylogenetic inference using Bayesian method, the Bayesian information criteria (BIC) can also be used for model selection. BIC has been developed first by Schwarz (1978) in order to overcome computational difficulties offered by previous model selection methods used in Bayesian inference, such as the use of Bayes factors (Kass and Raftery, 1995) or the use of the P(T|S) (see section1.3). As the AIC test, BIC rewards models with high likelihood and penalizes models with high number of parameters. The BIC score is given by $-2\ln L + k\ln n$ in which k is the number of parameters in the model, L denotes the maximized value of the model likelihood and n is the number of sites in the sequence data.

Once the evolutionary model has been selected and the phylogenetic tree obtained, one can still assess the confidence of the resulting tree. Two approaches are commonly used for this end. On the one hand, the nonparametric bootstrapping method was first proposed by Felsenstein (1985), and is still very common nowadays. It consists of rearranging the order of characters in the sequence many times and inferring a tree from each rearrangement. Then, phylogenetic trees obtained from the bootstrapping process are compared with the one previously inferred from the data sequence. This process will give a support value assigned to each node of the phylogenetic tree. The support value represents how frequently the clade under the node is represented in the replicate trees.

On the other hand, when using a Bayesian inference for tree construction, one can also use a Bayesian approach to assess the reliability of the selected tree in representing the sequence data (Rannala and Yang, 1995; Huelsenbeck et al., 2001). As in the case of the nonparametric bootstrapping approach, the Bayesian approach also assigns a support value to each node in the selected tree. This value represents the percentage of observation of that clade within the set of trees sampled from the posterior probability distribution. Generally, a support value of more than 70% is considered acceptable (considered as well resolved) for a clade, for both of the two approaches.

2.4 Phylogenetic modelling in ecology

As phylogenetics aims at providing information on the past evolutionary history of organisms, phylogenetic models and methods are sometimes used by ecologists. With the rapid increase of the availability of phylogenetic data and computation power, the study of the importance of evolution in ecological assemblages

has been boosted. It is important to notice that most ecological studies using phylogenetics make use of species level phylogeny and chronological time branch length often estimated from fossil records of know ancestors (Morlon et al., 2011).

In community ecology or the study of interactions between species and their environment, phylogenetic techniques are generally used to explore phylogenetic relatedness of species that are ecologically close to each other (i.e. ecologically interacting). Indeed, species phenotypes which reflect the evolutionary history of species drive the way species interact. For example, Rezende et al (2007a, 2007b) were among the first to explore the role of phylogenetic history in the structure of mutualistic interaction networks. They showed the substantial importance of phylogeny in explaining some patterns observed in mutualistic networks with statistical support from empirical dataset of pollination and frugivory communities. A study by Minoarivelo et al. (2014) suggested a model to quantify the importance of evolutionary history in shaping mutualistic interactions. Not only did they analyse the phylogenetic trees which assembled species involved in the interactions, but they also made use of a modelling approach similar to evolutionary models (section 2.2) (i.e. using a Markov process) in their approach. Their model simulates the possible evolution of species interactions along the branches of the phylogenetic trees, and estimates the rate at which ecological interactions are changing (Fig. 2).

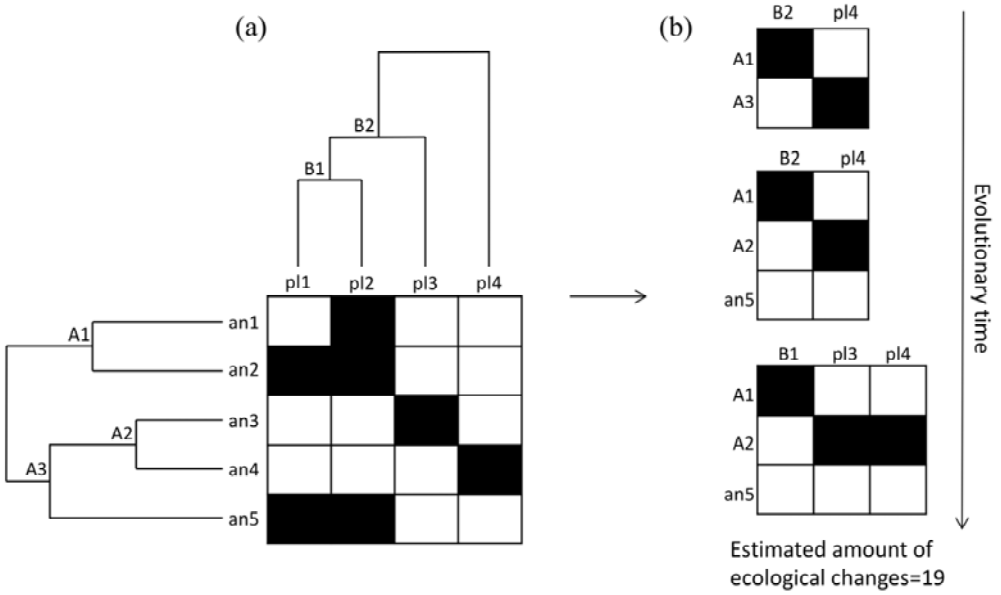

Fig. 2 Case example of the inference of the evolution of ecological interactions based on the phylogenetic history of species, as obtained from the model of Minoarivelo et al. 2014. (a) Model inputs: binary matrix of interaction between the two sets of species and phylogenetic trees of the two sets; (b) Model outputs: probable interactions between ancestors and inferred amount of interaction state changes.

Conservation ecologists have also taken advantage of phylogenetics in exploring conservation prioritization based on phylogenetic relatedness of species (Rolland et al., 2012; Winter et al., 2013). Indeed, communities with high phylogenetic diversity (Cardotte et al., 2008) should be given conservation priority as they enhance ecosystem stability (Cardotte et al., 2008; Gravel et al., 2011). Phylogenetics is also widely used by evolutionary ecologists. A common focus of their study is to explain the emergence and maintenance of

diversity (Richardson et al., 2011, 2014; Hui et al., 2013; Morlon, 2014), estimating the rate and time of phenotypic diversification (Ackerly, 2009; Tringali et al., 1999), or the rate of speciation or extinction within a set of species (Hey, 1992; Pyron and Burbrink, 2013; Rolland et al., 2014).

3 Conclusion

Reconstructing the phylogenetic history shared by a group of organisms mainly consists of building a phylogenetic tree and inferring an appropriate model of evolution on which the phylogenetic tree reconstruction is based. From simple insightful methods such as the parsimony algorithm or other methods based on phylogenetic distances, phylogenetic reconstruction techniques have evolved to complex probabilistic methods based on substitution models such as maximum likelihood and Bayesian inference. With an increase in the size of datasets to be analysed, probabilistic methods have incorporated complex statistical techniques to validate the resulting inferences. While the computational and statistical advances in the phylogenetic field continue to escalate, we can safely assume that the field will continue its upward trajectory, boosting its appeal to an even broader audience.

Acknowledgements

We acknowledge funding received from the South African National Research Foundation (No. 76912 and 89967) to CH, the PhD Scholarship from the Deutscher Akademischer Austausch Dienst (DAAD; German Academic Exchange Service) to HOM and the PhD Scholarship from the DST-NRF Centre of Excellence for Invasion Biology to GD.

References

Ackerly D. 2009. Conservatism and diversification of plant functional traits: Evolutionary rates versus phylogenetic signal. Proceedings of the National Academy of Sciences of the United States of America, 106: 19699-19706

Akaike H. 1973. Information theory and an extension of the maximum likelihood principle. In: Second International Symposium on Information Theory (Petrov PN, Csaki F, eds). 67-281, Adad. Kiado, Budapest, Hungary

Berthouly-Salazar C, Cassey P, van Vuuren BJ, van Rensburg BJ, Hui C, Le Roux JJ. 2012. Development and characterization of 13 new, and cross amplification of 3, polymorphic nuclear microsatellite loci in the Common myna (Acridotheres tristis). Conservation Genetics Resources, 4: 621-624

Berthouly-Salazar C, Hui C, Blackburn TM, Gaboriaud C, van Rensburg BJ, van Vuuren BJ, Le Roux JJ. 2013. Long-distance dispersal maximizes evolutionary potential during rapid geographic range expansion. Molecular Ecology, 22: 5793-5804

Burnham KP, Anderson DR. 2004. Multimodel inference. Sociological Methods and Research, 33: 261-304

Cardotte MW, Cardinale BJ. 2008. Evolutionary history and the effect of biodiversity on plant productivity. Proceedings of the National Academy of Sciences of the United States of America, 105: 17012-17017

Clement M, Posada D, Crandall KA. 2000. TCS: a computer program to estimate gene genealogies. Molecular Ecology, 9: 1657-1659

Cutter AD. 2013. Intergrating phylogenetics, phylogeography and population genetics through genomes and evolutionary theory. Molecular Phylogenetics and Evolution, 69: 1172-1185

Diedericks G, Daniels SR. 2014. Ain't no mountain high enough, ain't no valley low enough? Phylogeography of the rupicolous Cape girdled lizard (Cordylus cordylus) reveals a generalist pattern. Molecular Phylogenetics and Evolution, 71: 234-248

Drummond A J, Rambaut A. 2007. BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evolutionary Biology, 7: 214

Durbin R, Eddy S, Krogh A, Mitchison G., 1999.Multiple sequence alignment methods. In: Biological Sequence Analysis: Models of Proteins and Nucleic Acids. 134-159, Cambridge University Press, USA

Felsenstein J. 1981. Evolutionary trees from DNA sequences: a maximum likelihood approach. Journal of Molecular Evolution, 17: 368-376

Felsenstein J. 1985. Confidence limits on phylogenies: An approach using the bootstrap. Evolution, 39: 783-791

Gravel D, Bell T, Barbera C, Bouvier T, et al. 2011. Experimental niche evolution alters the strength of the diversity-productivity relationship. Nature, 469: 89-92

Guindon S, Dufayard JF, Lefort V, Anisimova M, et al. 2010. New Algorithms and Methods to Estimate Maximum-Likelihood Phylogenies: Assessing the Performance of PhyML 3.0. Systematic Biology, 59: 307-321

Hasegawa M, Kishino H, Yano T. 1985. Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. Journal of Molecular Evolution, 22: 160-174

Hasegawa M. 1990. Phylogeny and molecular evolution in primates. The Japanese Journal of Genetics, 65: 243-266.

Hastings A, Gross LJ. 2012. Encyclopedia of Theoretical Ecology. 550-556, University of California Press, USA

Hey J. 1992. Using phylogenetic trees to study speciation and extinction. Evolution, 46: 627-640

Holland BR. 2013. The rise of statistical phylogenetics. Australian and New Zealand Journal of Statistics, 55: 205-220

Huelsenbeck JP, Ronquist F. 2001. MRBAYES: Bayesian inference of phylogenetic trees. Bioinformatics, 17: 754-755

Huelsenbeck JP, Ronquist F, Nielsen R, Bollback J. 2001. Bayesian inference of phylogeny and its impact on evolutionary biology. Science, 294: 2310-2314

Hui C, Richardson DM, Pyšek P, Le Roux JJ, Kučera T, Jarošík V. 2013. Increasing functional modularity with residence time in the co-distribution of native and introduced vascular plants. Nature Communications, 4: 2454

Jukes TH, Cantor CR. 1969. Evolution of Protein Molecules. In: Mammalian Protein Metabolism (Munro HN, ed). 22-126, Academy Press, New York, USA

Kass RE, Raftery AE. 1995. Bayes factors. Journal of the American Statistical Association, 90: 773-795

Kimura M. 1980. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. Journal of Molecular Evolution, 2: 111-120

Knowles LL, Maddison WP. 2002. Statistical phylogeography. Molecular Ecology, 11: 2623-2635

Kumar S, Filipski AJ, Battistuzzi FU, Pond SLK, Tamura K. 2011. Statistics and truth in phylogenomics. Molecular Biolology and Evolution, msr202

Liò P, Goldman N. 1998. Models of molecular evolution and phylogeny. Genome Research, 8: 1233-1244

Minoarivelo HO, Hui C, Terblanche JS, Pond SLK, Sheffler K. 2014. Detecting phylogenetic signal in mutualistic interaction networks using a Markov process model. Oikos, 123: 1250-1260

Morlon H, Parsons TL, Plotkin JB. 2011. Reconciling molecular phylogenies with the fossil record. Proceedings of the National Academy of Sciences of the United States of America, 108: 16327-16332

Morlon H. 2014. Phylogenetic approaches for studying diversification. Ecology Letters, 17: 508-525

Morrison DA. 2012. Book review: Phylogenetics: The theory and practice of phylogenetic systematics, 2nd edn. Systematic Biology 61: 1087-1088

Nielsen R, Beaumont MA. 2009. Statistical inferences in phylogeography. Molecular Ecology, 18: 1034-1047

Posada D, Crandall KA. 1998. Modeltest: testing the model of DNA substitution. Bioinformatics, 14: 817-818

Posada D. 2008. JModelTest: phylogenetic model averaging. Molecular Biology and Evolution, 25: 1253-1256

Pyron RA, Burbrink FT. 2013. Phylogenetic estimates of speciation and extinction rates for testing ecological and evolutionary hypotheses. Trends in Ecology and Evolution, 28: 729-732

Rannala B, Yang Z. 1995. Probability distribution of molecular evolutionary trees: a new method of phylognetic inference. Journal of Molecular Evolution, 43: 304-311

Rezende EL, Jordano P, Bascompte J. 2007. Effects of phenotypic complementarity and phylogeny on the nested structure of mutualistic networks. Oikos, 116: 1919-1929

Rezende EL, Lavabre JE, Guimarães PR, Jordano P, Bascompte J. 2007. Non-random coextinctions in phylogenetically structured mutualistic networks. Nature, 448: 925-928

Richardson DM, Carruthers J, Hui C, Impson FAC, Maslin B, Robertson MP, Rouget M, Le Roux JJ, Wilson JRU. 2011. Human-mediated introductions of Australian acacias – a global experiment in biogeography. Diversity and Distributions, 17: 771-787

Richardson DM, Hui C, Nuñez MA, Pauchard A. 2014. Tree invasions: patterns, processes, challenges and opportunities. Biological Invasions, 16: 473-481

Robinson TJ, Cernohorska H, Diedericks G, Cabelova K, Duran A, Matthee CA. 2014. Phylogeny and vicariant speciation of the Grey Rhebok, Pelea capreolus. Heredity, 112: 325-332

Rolland J, Cadotte MW, Davies J, Devictor V. et al. 2012. Using phylogenies in conservation: new perspectives. Biology Letters, 8: 692-694

Rolland J, Condamine FL, Jiguet F. 2014. Faster speciation and reduced extinction in the tropics contribute to the mammalian latitudinal diversity gradient. PLoS Biology, 12: e1001775

Saitou N, Nei M. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Molecular Biology and Evolution, 4: 406-425

Schwarz G. 1978. Estimating the dimension of a model. The Annals of Statistics, 6: 461-464

Sibley CG, Ahlquist JE. 1984. The Phylogeny of the Hominoid Primates, as Indicated by DNA–DNA Hybridization. Journal of Molecular Evolution, 20: 2-15

Sokal R, Michener C. 1958. A statistical method for evaluating systematic relationships. University of Kansas Science Bulletin, 38: 1409-1438

Sorhannus U. 2003. The effect of positive selection on a sexual reproduction gene in Thalassiosira weissflogii (Bacillariophyta): results obtained from maximum-likelihood and parsimony-based methods. Molecular Biology and Evolution, 20: 1326-1328

Stamatakis A. 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics, 22: 2688-2690

Swofford DL, Olsen GJ, Waddell PJ, Hillis DM. 1996. In: Molecular Systematics, 2nd edition, Sinauer Associates, Sunderland, Massachusetts, 407-514

Swofford DL. 2002. PAUP*: Phylogenetic Analysis Using Parsimony (and Other Methods). Version 4 Sinauer Associates, Sunderland, MA, USA

Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. 2011. MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Molecular Biology and Evolution, 28: 2731-2739

Tavaré S. 1986. Some Probabilistic and statistical problems in the analysis of DNA sequences. In: Some Mathematical Questions In Biology: DNA sequence analysis (Miura RM, ed). American Mathematical Society, New York, USA, 17: 57-86

Tringali MD, Bert TM, Seyoum S, Bermingham E, et al. 1999. Molecular phylogenetics and ecological diversification of the transisthmian fish genus Centropomus (Perciformes: Centropomidae). Molecular Phylogenetics and Evolution, 13: 193-207

Winter M, Devictor V, Schweiger O. 2013. Phylogenetic diversity and nature conservation: where are we? Trends in Ecology and Evolution, 28: 199-204