Case study: molecular phylogenetics

Phylogenetics is the study of evolutionary associations involving diverse groups of organisms. Such evolutionary relationships are discovered through molecular and morphological data matrices. The term phylogenetics is a self descriptive term of Greek origin which means " tribe, or race related by birth". Identification, classification, and naming of organisms have been opulently informed by phylogenetics. Phylogenetics and taxonomy overlap in the science of Phylogenetic systematic. In Biology, most importantly, Phylogenetic analysis has become an essential feature in researching the evolutionary tree of life, Larget, B., & Simon. D. (1999).

The evolution of mammals was a gradual process stretching approximately fifty million years, between 280 and 230 million years ago (Permian Period). Moreover, between 230 million and 190 million years ago (Triassic Period), many organisms which resembled mammals came into existence. However, it was only until 190 and 135 million years ago (Jurassic Period) that the first true mammals appeared. This also marked the period of the end of ancient reptiles, the appearance of flowering plants and modern insects, Murphy W. J., Davis B., David V. A., Agarwala R, Schaffer A. A., Wilkerson A. J, Neelam B, O'Brien S. J, Menotti-Raymond M. A. (2207). The earliest mammals, the eutherian and pouched mammals, acquired new life forms and established biological environments through which all modern mammals have appeared. While examining mammalian evolution, a number of techniques have been employed. Some of these methods include: Maximum Parsimony; Minimum Evolution; Maximum Likelihood and Bayesian Analyses. In this study, I shall conduct a literary review of these methods, their applicability in Phylogenetic

Reconstruction while pointing out their strengths and weaknesses, Kohn M. H., Murphy W. J, Ostrander E. A., Wayne R. K. (2006).

LITERARY REVIEW

Maximum Parsimony

Parsimony refers to a non-parametric statistical method used universally in computational phylogenetics in the estimation of phylogenies. It is a component of a category of character-based tree estimation methods using a matrix of discrete Phylogenetic characters to infer one or more of the best possible Phylogenetic trees for a set of taxa, usually a set of class or reproductively-isolated populations of a single class. This method was the most common character-tree estimation until recently when other methods were discovered, Miller C. M., Murphy W. J., O'Brien S. J., Jacobs D. S., Springer M. S., Teeling E. C. (2007). Nonetheless, the most preferred Phylogenetic tree under parsimony is one that requires slightest evolutionary transformation to explain observed data. This openly follows the branching outline of evolution. Thus it can be inferred that if two organisms possess a mutual character, then they are closely linked to each other as compared to another organism which lacks this character. For instance, it is believed that monkeys are more closely related to bats as opposed to fish. This is simply because they both posses hair as their body covering, Mau, B., & Newton M. (1997).

Advantages and Disadvantages

The most notable advantage of Maximum Parsimony lies in the fact that its study runs in an incredibly simple approach. Trees are scored in line with the degree to which they imply a prudent allocation of quality data. The most

prudent tree for the dataset represents the preferred hypothesis of interaction between the taxa in the analysis, Wible J. R, Rougier G. W, Novacek M. J, Asher R. J. (2007). Additionally, there are other potential Phylogenetic trees that can be searched comprehensively for more than eight taxa. Furthermore, the trees resulting from parsimony search are un-rooted hence display all likely associations of the integrated taxa.

Nonetheless, it also has a number of disadvantages. To begin with, it lacks statistical consistency. This implies that data for the true tree with elevated likelihood is not guaranteed. Consistency here implies monotonic convergence on the accurate response with the accumulation of extra statistics, is a desirable property of any arithmetical technique. Its second disadvantage is that finding the most economical tree is an NP-Hard problem, Capelli C, MacPhee R. D, Roca A. L, Brisighelli F, Georgiadis N, O'Brien S. J, Greenwood A. D. (2006). The single presently obtainable, resourceful way of obtaining a way out, given randomly bulky set of taxa, is by use of heuristic methods with no guarantee of the most economical tree. These methods use hill-climbing algorithms to gradually draw near the best tree, Nielsen, R. (2002).

Minimum Evolution

Minimum Evolution was developed to recreate molecular evolutionary Phylogenetic trees. This is dealt with from the viewpoint of an inductive inference and hence applies to Rissanen's J. (1986) concept of extracting the minimum density in a phylogenetic tree through his minimum description length principle. It describes the density of molecular Phylogenetic tree by three expressions which are related to the tree topology, the account of

branch lengths and the disparity between the copy and the data calculated by logarithmic probability, Posada, D., & Crandall K. A. (1998). To investigate the validity of this method, five mitochondrial DNA sequences from common chimpanzee, human beings, gorilla, orangutan, and pygmy chimpanzee are used, Kim, S., K. M. Kjer, & Duckett C. N. (2003).

Advantages and Disadvantages

There is one main advantage of the minimum evolution; it is superior to the traditionalized methods since it is more accurate even near the root of Phylogenetic trees. This implies that Minimum Evolution is able to give credible data even from the root of Phylogenetic trees where most methods have failed to do. It produces reliable and consistent data that can help build meaningful evolutionary relationships. Nonetheless, as noted by Rannala, B., & Yang Z. (1996), this method exhibits an aspect of inconsistency. This is evident when we try to show the connection between minimum evolution and ordinary-LS since it does not establish its final consistency, Driscoll C. A, Menotti-Raymond M, Roca A. L, Hupe K, Johnson W. E, Geffen E, Harley E. H, Delibes M, Pontier D, Kitchener A. C, Yamaguchi N, O'brien SJ, Macdonald D. W. (2007).

Maximum Likelihood

Maximum Likelihood is a technique for the deduction of phylogeny. It evaluates a theory about evolutionary account in terms of the likelihood that the projected form and the hypothesized account would give rise to the experiential data set. The theory is that a history amid an advanced likelihood of getting the observed state is ideal to a history with a lower likelihood. The technique searches for the tree with the maximum prospect,

Asher R. J. (2007). Recent developments in applying Bayesian probabilistic and statistical thoughts toward professional systems have established the significance of Bayesian analysis. By means of a real, fairly multifaceted, remedial illustration a researcher can demonstrate how qualitative and quantitative understanding can be outlined in a directed graphical representation, commonly recognized as a principle arrangement in this framework, Roca A. L, O'Brien S. J. (2005). Precise probabilistic supposition on entity cases is then probable by means of a universal dissemination formula. Once statistics on a chain of cases are accessible, Bayesian statistical techniques can be used to update the inventive subjective quantitative inputs, after which a researcher can present a set of diagnostics for identifying inconsistencies connecting the facts and the preceding measurement.

Advantages and Disadvantages

Maximum Likelihood has a number of advantages. These advantages include frequent minor inconsistencies than other methods, that is, it is often the inference technique that is least affected by sampling related errors. Secondly, it is inclined to be immune to several violations of the assumptions in the evolutionary representation. Thirdly, even with very undersized sequences they tend to surpass substitute methods such as parsimony or distance techniques. Finally, the technique is statistically sound and uses all the chain sequence in addition to assessing dissimilar tree topologies. In spite of these advantages, this method has some disadvantages. This technique is heavily dependent on the Control Processing Unit (CPU) hence relatively slow as compared to other techniques. Secondly, the end result is

only reliant on the form of evolution technique applied by the researcher,
Lee, M. S. Y. (1999).

Bayesian Analyses

Bayesian analysis is a statistical system which seeks to approximate
parameters of a fundamental allocation based on the experiential circulation.
The system starts with a " prior distribution" which is based on anything.
Moreover, an appraisal of comparative likelihoods of parameters in a non-
Bayesian observation is also essential elements in this method. In practice, it
is common to suppose a consistent allocation over the proper range of
standards for the prior distribution. It is also used in the drawing of software
filters to mechanically identify and cancel debris in the electronic mail, spam.

Data collected to be evaluated using this method is normally treated as
random variables. Its data and parameters are described by the probability
of the posterior distribution in cases where actual values, such as the
phylogenetic tree, are not certain. This method differs from Maximum
Likelihood in the sense that it simplifies uncertain parameters for easy and
clear description. This is achieved through the provision of the probability
distribution thereby making available some of the variable specifics. In this
case, specifics of variance, mean and range containing most of the posterior
probability parameters are directly calculated, Jones, D. T., W. R. Taylor, & J.
M. (1992). Nonetheless, in the recent times, a hybrid mechanism making use
of the Bayesian and Maximum Likelihood have been used. For instance,
Nielson and Yang (1998) established that with the help of the Bayesian
method in analyzing the sequence of DNA protein coding, they could locate
the position of amino acids in the protein. Not only did they discover the

position of each amino acid but also established their posterior probability. Moreover, with the help of Maximum Likelihood, they approximated all the parameters of the prototype. This method was referred to as the empirical Bayes approach it based the ascertaining the probability under positive selection on the Bayesian reasoning and determining the nuisance parameters using Maximum Likelihood, Springer M. S., Burk-Herrick A, Meredith R, Eizirik E, Teeling E, O'Brien S. J., Murphy W. J. (2007).

Advantages and Disadvantages

Bayesian models by far contain hidden variables like one's true status on a particular disease in case a diagnostic error occurred. The use of prior likelihood dispersions represent a great mechanism for integrating data from earlier studies and for controlling confusing discoveries. Latest developments in Markov chain Monte Carlo have assisted in the accomplishment of Bayesian analysis of complex statistics sets containing missing comments and multidimensional outcomes.

However, a number of limitations are associated with this method. One of its limitations rests in the fact that it does not tell a researcher how to select a prior. Furthermore, Bayesian inferences necessitate skills to interpret one-sided prior approach into a scientifically formulated prior. If a researcher does not advance with prudence, he or she can produce disingenuous results. In addition to this, Bayesian analysis can create subsequent distributions greatly inclined by the priors, Murphy W. J., Pringle T. H., Crider T. A., Springer M. S., Miller W. (2007). Beginning with a realistic point of view, it is sometimes complicated to induce subject matter experts with contrary view to the validity of the chosen prior. Moreover, Bayesian analysis often

comes with an elevated computational rate, particularly in models with a large number of parameters. Simulations therefore present somewhat dissimilar answers unless the same unsystematic process is used. Slight variations in replication outcome also do not challenge the early claim that Bayesian inferences are exact, Springer M. S., Murphy W. J. (2007).

SUMMARY AND CONCLUSION

The scientific study of evolutionary relations must proceed from the understanding of the history of the evolution of mammals since this understanding presents a unique opportunity for any researcher to comprehend the intricate details of such a research. Additionally, the three techniques used to undertake this scientific analysis, though different and with somewhat distinct advantages and disadvantages, offer a precise understanding of the evolutionary relations. Phylogenetic Reconstruction through the maximum parsimony, minimum evolution, and Bayesian analysis provides a complex yet somewhat clear apprehension of the evolutionary process. In deed the various significance of each system of analysis provides understanding for the present and the future thereby creating avenues through which molecular phylogenetics and the evolution of mammals can be evaluated. The facts, statistics and evolutionary relationships presented by these methods can help heighten better comprehension of the mammalian evolutionary process. This will in turn help reduce isolated and ambiguous facts associated with the evolution of mammals.

It has been ascertained that stochastic evolution methods contribute greatly towards molecular evolution. They have been used in discovering the patterns of DNA substitution, phylogeny estimation, detection of amino acids

and inferring histories of populations. These models continue getting sophisticated overtime thereby drawing us closer towards realizing evolutionary relationships among different organisms. With the use of Maximum Parsimony, Minimum Evolution, Maximum Likelihood and Bayesian Analyses, the biological field of molecular phylogenetics has been opened up and provided valuable assistance towards describing various aspects of the mammalian evolution. In conclusion, such methods have and continue to influence molecular studies thereby offering new insights into different aspects of our existence.