ELECTRONIC APPENDIX A

1. Phylogenetic relationships of dinosaurs

A number of constraints and special requirements, imposed by cladistic biogeographic methods, place certain restrictions on the taxon cladograms that can be successfully analysed. Such limitations, therefore, provide a set of criteria which can assist the investigator in selecting between multiple, often conflicting, competing phylogenetic hypotheses. These criteria include:

(1) Tree resolution. Cladistic biogeography has not yet established a rigorous method for coping with taxon cladograms that contain polytomies. As a result, computer programs such as Component and TreeMap will only analyse fully-resolved trees.

(2) Taxon sampling. Ideally, biogeographic analyses should be based on cladograms that contain as many taxa as possible. This is because, at least in the case of tree reconciliation, 'signal' detection depends on the presence of the same area relationships repeated many times across different clades. A large number of taxa in the initial cladogram also facilitates effective analysis even when taxon numbers are reduced by time-slicing. Unfortunately, computer programs such as TreeMap are currently unable to analyse cladograms with more than 100 taxa.

At present, no single 'global' analysis of dinosaur relationships fulfils both of these criteria. In addition, the strict avoidance of polytomies rules out a number of approaches that might otherwise have been used to build a single dinosaur cladogram from several constituent trees. For example, one standard method for combining cladograms with differing topologies is to produce a consensus tree. However, both Strict (Schuh & Farris 1981) and Adams (Adams 1972) consensus methods produce polytomies. Reduced consensus (Wilkinson 1994) is capable of producing a single fully resolved tree, but only at the expense of a decrease in the number of taxa available for biogeographic analysis. Another approach would be to repeat the biogeographic analysis for each of the alternative cladogram topologies available and then examine to what extent these different trees support the same biogeographic signals. Unfortunately, the large number of alternative cladograms for various dinosaur groups would make analysis very time-consuming and unwieldy. For example, if there were two alternative topologies for each of five different dinosaur clades, one could generate 32 alternative cladograms for dinosaurs as a whole.

Consideration of the various constraints and limitations outlined above led to the following strategy for the creation of the 'global' dinosaur cladogram shown in Figure 1: (1) Sereno's (1997, 1999*a*) detailed and well-resolved cladograms for dinosaurs as a whole were adopted as a basic 'framework'; (2) the relationships within particular clades in Sereno's cladograms were then modified in order to enhance taxon sampling and eliminate polytomies. Details of these modifications, and their justification, are provided below. In particular, it should be noted that Sereno's cladograms include several 'aggregate' terminal taxa (e.g. 'Ornithomimidae', 'Titanosauria' and 'other ankylosaurines'). Such aggregates frequently contain taxa from several different

geographic areas and stratigraphic units. Retention of aggregate taxa has the potential to seriously distort the results of cladistic biogeographic analyses, largely through their impact on time-slicing and the introduction of many apparently 'widespread' taxa. If relevant biogeographic patterns are preserved at lower taxonomic levels, then the use of aggregates may obscure such signals altogether. Modifications to Sereno's cladograms include:

(1) Sauropodomorpha. Sereno's (1997, 1999*a*) cladograms represent Sauropodomorpha by a relatively small number of taxa (six prosauropods and no more than 12 sauropods). In addition, these taxa include some large aggregate terminals, such as 'Titanosauria'. Sereno's entire sauropodomorph clade has therefore been replaced by cladograms which provide increased taxon sampling for Prosauropoda (Galton & Upchurch submitted *a*) and Sauropoda (Upchurch 1998). Geographic/stratigraphic sampling was improved further for sauropods by adding *Austrosaurus*, *Cedarosaurus* and *Pleurocoelus* (Texan material only) to the analysis (Upchurch, unpublished data).

(2) Abelisauridae. Although this clade contains taxa with a potentially important biogeographic distribution (Sampson *et al.* 1998), there is currently no fully-resolved and well-supported cladogram for this group. Abelisauridae has therefore been provisionally deleted from the cladogram, though it will be considered in future, expanded, analyses.

(3) Coelophysidae. This aggregate taxon has been replaced by the sister-taxa *Syntarsus* and *Coelophysis*.

(4) Spinosauroidea. This aggregate taxon has been replaced by the genus-level cladogram of spinosaurs and torvosaurs proposed by Sereno *et al.* (1998). The relationships of this clade have been further modified by making *Baryonyx* and

Suchomimus a single widespread taxon (based on A. C. Milner, pers. comm. 2000). It should be noted that other workers have produced cladograms which do not contain a monophyletic Spinosauroidea.

(5) Allosauroidea. This aggregate taxon has been replaced by the genus-level analysis produced by Sereno *et al.* (1996). The latter, however, contains a polytomy involving *Allosaurus*, *Cryolophosaurus*, *Monolophosaurus*, Sinraptoridae, and a lineage leading to a fully-resolved Cretaceous clade. This polytomy was removed by deleting taxa, but preferentially retaining *Cryolophosaurus* because it potentially provides information on the relationships of the poorly sampled 'Antarctica' area. Future analyses might extract more detailed biogeographic data by utilising the well-resolved allosauroid cladogram obtained by Holrz (2000).

(6) Therizinosauridae. This aggregate taxon has been replaced by the genus *Therizinosaurus*. Future analyses could utilise the more highly resolved topology found by Xu *et al.* (1999).

(7) Alvarezsauridae. This aggregate taxon has been replaced by the genus-level cladogram produced by Novas (1997). The position of this clade within Theropoda or basal Aves is rather unstable.

(8) Ornithomimosauria. Sereno's 'global' cladograms show *Pelecanimimus* as the sister-taxon to a monophyletic Ornithomimidae. The latter aggregate taxon has been replaced by the genus-level cladogram proposed by Barsbold and Osmolska (1990). It should be noted that no data-matrix was provided by Barsbold and Osmolska, and it is therefore not possible to evaluate the strength of support for the cladogram topology they propose.

(9) Tyrannosauridae. This aggregate has been replaced by the genus-level cladogram presented by Sereno (1999*c*).

(10)   Oviraptorosauria. This aggregate taxon appears in the cladograms presented

by Sereno (1999a), though the accompanying data-matrix provides greater

resolution (i.e. Oviraptorosauria includes *Caudipteryx*, Caenagnathidae and

Oviraptoridae). Therefore, the Oviraptorosauria has been replaced by the genus-

level cladogram produced by Barsbold *et al.* (1990). It should be noted that

Barsbold *et al.* do not provide a data-matrix and it is therefore not possible to

evaluate the strength of support for the cladogram topology proposed.

(11)   Dromaeosauridae. This aggregate taxon has been replaced by the genus-level

cladogram provided by Sereno (1999*c*).

(12)   Aves. Avian taxa, above the level of *Archaeopteryx*, were excluded from the

cladogram because their enhanced dispersal abilities may have enabled them to

effectively ignore marine barriers between continents. Future analyses will

attempt to test this hypothesis by incorporating birds and assessing to what extent

they display the same area relationships as terrestrial forms.

(13)   Stegosauria. This aggregate taxon has been replaced by the cladogram

produced by Galton and Upchurch (submitted *b*). The latter stegosaur cladogram

is congruent with those produced by Sereno (1997, 1999*a*) concerning the basal

positions of *Huayangosaurus* and *Dacentrurus*.

(14)   Nodosauridae. Sereno's cladogram for the Nodosauridae shows *Sarcolestes*,

Hylaeosaurus and 'Nodosaurinae' forming a trichotomy. This trichotomy was

removed by deletion of *Hylaeosaurus*, on the basis that *Sarcolestes* provides

information on a more poorly sampled stratigraphic range (i.e. the Middle

Jurassic). The aggregate taxon 'Nodosaurinae' has been replaced by the genus-

level cladogram produced by Lee (1996).

(15)   Ankylosauridae. Sereno's cladograms indicate that *Gargoyleosaurus* is the most basal ankylosaurid. *Minmi* and *Shamosaurus* form a trichotomy with 'Ankylosaurinae'. A fully-resolved topology was obtained by deleting *Shamosaurus*, on the basis that *Minmi* provides information on the relationships of the poorly sampled 'Australia' area. The aggregate taxon 'Ankylosaurinae' has been replaced by the genus-level cladogram produced by Sullivan (1999). The latter cladogram includes a trichotomy between *Tarchia*, *Saichania* and *Nodocephalosaurus*. This trichotomy was removed by deleting one of the two Late Cretaceous Asian forms, in this case *Saichania*.

(16)   Marginocephalia. The marginocephalian relationships in Fig. 1 are essentially those in Sereno's (1997, 1999*a*) global dinosaur cladograms. A more recent analysis by Sereno (2000), however, incorporates more genera (especially among the pachycephalosaurs) and has therefore been used in the current study. Sereno (2000) also provides genus-level topologies which replace the aggregate taxa 'Protoceratopsidae', 'Centrosaurinae' and 'Chasmosaurinae'.

(17)   Ornithopoda. Although Sereno provides a detailed cladogram for ornithopods, we prefer the cladogram produced recently by D. B. Norman (unpublished data). The two competing cladograms are broadly similar in topology, although there are some important differences regarding the relationships of various 'iguanodontian' and basal hadrosaurian forms, such as *Iguanodon*, *Ouranosaurus* and *Probactrosaurus*. The Hypsilophodontidae is poorly represented in our dinosaur phylogeny because of the controversial nature of their relationships at present. The Hadrosauridae was originally represented by a larger number of taxa in Norman's cladogram, but several generawere removed prior to analysis. These deletions were carried out using 'Rosen's Rule' (Rosen 1979): i.e. a monophyletic clade

containing taxa found in only a single area can be represented by a single taxon form that clade without decreasing the information content on area relationships. Given the upper limit on taxon numbers accepted by TreeMap, we implemented Rosen's Rule to reduce the size of the Hadrosauridae.

2. Information requird for replication of tree reconciliation analyses

(1) Area designation. This study employs eight 'continental' areas, representing key regions of interest that were probably separate entities during at least part of the Cretaceous.

(2) Widespread taxa. A 'widespread' taxon is any organism that is present in two or more of the designated geographic areas. Such distributions may arise for several different reasons, including missing data and dispersal. Cladistic biogeographers have modelled these different interpretations as three assumptions ('0', '1' and '2') (Nelson & Platnick 1981; Zandee & Roos 1987) which allow widespread taxa to be included in analyses. *Brachiosaurus*, *Camptosaurus*, *Dryosaurus* and *Elaphrosaurus* are represented by morphologically distinct species in each of the areas they occupy. In these cases, each genus has been divided into two sister-species, each of which is endemic to one of the two areas occupied by that genus (i.e. assumption 0). *Baryonyx* is usually regarded as being found solely in Europe. However, the 'autapomorphic' features characterising the African genus *Suchomimus* appear to be insufficient to adequately distinguish it from *Baryonyx* (A. C. Milner, pers. comm. 2000). These two taxa are therefore provisionally treated as a single widespread taxon present in both Africa and Europe (i.e. assumption 2). *Iguanodon bernissartensis* has been reported from both Europe and Asia (Norman 1996). Since there are currently no grounds for separating the

remains from these two areas into separate taxa, *I. bernissartensis* was treated under assumption 0.

(3) Searching for optimal area cladograms. The optimal area cladogram topologies were found by analysing the time-sliced dinosaur cladograms using the heuristic search in TreeMap (Page 1995). The optimality criterion used to find the area cladograms and test them against randomised data was 'maximal codivergence'.

(4) Randomisation tests. Normally, such tests randomise the area ('host') cladogram and reconcile each random replicate with the original taxon ('associate') cladogram (Page 1991). One problem with this approach is that, in continent-level historical biogeography, the number of areas is usually small (a maximum of eight in this study). In fact, if only three areas were available, the randomisation test would break down: there are only three different bifurcating rooted trees for three terminals, and it is therefore impossible to generate a sufficiently large number of topologically distinct random area cladograms. Fortunately, the number of taxa in each analysis is usually somewhat larger than the number of areas, and the randomisation tests provide a more effective measure of statistical support if the taxon cladogram is randomised. Thus, for each time-slice analysis, the dinosaur cladogram was used to generate 10,000 randomised replicates, each of which was reconciled with the optimal area cladogram. The randomised trees were generated using the 'proportional –to-distinguishable' method available in TreeMap.

APPENDIX REFERENCES

The following references appear in the Appendix but not in the main text

Adams, E. N. 1972. Consensus techniques and the comparison of taxonomic trees. *Syst. Zool.* **21**, 390-397.

Holtz, T. R. Jr. 2000. A new phylogeny of the carnivorous dinosaurs. *Gaia* **15**, 5-61.

Norman, D. B. 1996. On Mongolian ornithopods (Dinosauria: Ornithischia). 1. *Iguanodon orientalis* Rozhdestvensky 1953. *Zool. J. Linn. Soc. London* **116,** 303-315.

Page, R. D. M. 1991. Random dendrograms and null hypotheses in cladistic biogeography. *Syst. Zool.* **40,** 54-62.

Rosen, D. E. 1979. Fishes from the uplands and intermontane basins of Guatamala: revisionary studies and comparative geography. *Bull. Am. Mus. Nat. Hist.* **162,** 267-376.

Schuh, R. T. & Farris, J. S. 1981. Methods for investigating taxonomic congruence and their application to the Leptopodomorpha. *Syst. Zool.* **30**, 331-351.

Wilkinson, M. 1994. Common cladistic information and its consensus representation: reduced Adams and reduced cladistic consensus trees and profiles. *Syst. Biol.* **43**, 343-368.

Yu, X., Tang, Z.-L. & Wang, X.-L. 1999. A therizinosauroid dinosaur with integumentary structures from China. *Nature* **399**, 350-354.

Zandee, M. & Roos, M. C. 1987. Component-compatibility in historical biogeography. *Cladistics* **3**, 305-332.