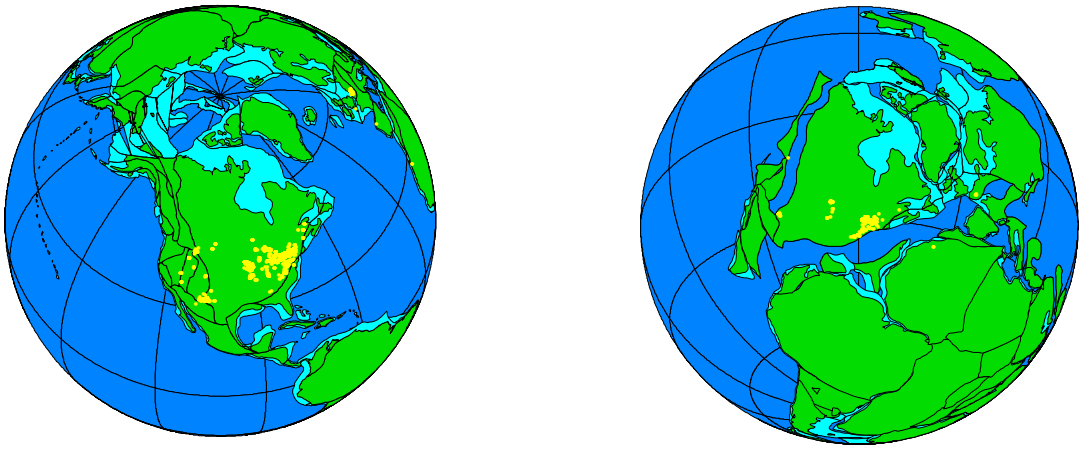
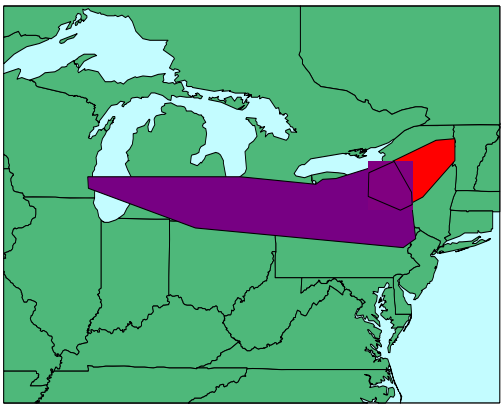
Getting Started with GIS for Paleobiogeographic Reconstruction





Using Excel, PaleoGIS, Point Tracker, ArcView 3.x, and ArcGIS 9.x

Alycia L. Stigall

Introduction

The use of Geographic Information Systems (GIS) in paleontology offers great potential for increased quantification within paleontological analyses. The primary advantage of GIS analysis is that has the capacity to utilitize large amounts of species occurrence data to produce quantitatively constrained geographic range reconstructions that are amenable to statistical analysis. GIS analysis offers several additional advantages over traditional range mapping methods. First, range maps are constructed dynamically, so they can be updated whenever new data become available. In addition, fine temporal and spatial patterns can be discerned, which can then be related to evolutionary or environmental events. GIS data are amenable to both traditional statistical analyses and spatial statistics. In addition, data for GIS analysis can be both extracted from and donated to large database projects such as the Paleobiology Database (www.pbdb.org). Most importantly, the combination of these attributes results in the ability to quantitative test biogeographic hypotheses! Paleobiogeographers can now quantify patterns and improve upon prior "story telling" scenarios.

The basic steps in GIS range reconstruction are data base assembly (including taxonomic, geographic, and stratigraphic information for each specimen), mapping of species localities on modern continental configuration, rotation of species occurrence data onto paleocontinental reconstructions, and reconstructions of geographic ranges. This guide is intended as a brief introduction to the steps involved with working through these stages. Further information, particularly related to theoretical aspects of using GIS in paleobiogeography, can be found in the references listed in the final section.

My hope is that this guide will facilitate new paleontologic studies using GIS, so that the list of papers in the final section can be expanded! Please contact me with questions and suggestions for improvement.

Alycia L. Stigall
Department of Geological Sciences
Ohio University
316 Clippinger Laboratories
Athens, OH 45701
stigall@ohio.edu

Please cite this guide as: Stigall, A.L. 2006. Getting Started with GIS for Paleobiogeographic Reconstruction Web page available at: http://oak.cats.ohiou.edu/~stigall/; accessed [day-month-year].

Table of Contents

Sections Page

Section 1. Database construction

Database construction is the considerably simpler to do prior to entry into the GIS. While the database portion of ArcGIS has more advanced features than ArcView 3.x, working with Excel or a similar program is still more efficient. Currently, ESRI programs require databases to be saved in database IV format (.dbf), but the newer versions of ArcGIS (e.g. 9.2) will be able to accept Excel files directly.

The critical items that MUST be included in your database are:

- Taxon identification
- Latitude and longitude of collection
- Temporal interval of occurrence

Section 2. Exporting Excel data to GIS

Preparing data format for export

- Create database in Excel (see section 1).
- Make sure the data immediately follows the column labels (no extra row between the column labels and the data).
- Save file as a DBF IV.
- Close Excel.

Import data into ArcGIS 9.x

- Open ArcMap 9.x and the map file of interest.
- Under the TOOLS menu, select 'Add XY Data'.
- Browse for the Excel file.
- The X Field will display 'LONGITUDE' and the Y Field 'LATITUDE'.
- Hit OK.
- The point data will be added to the Data Frame as a new layer.

Import data into ArcView 3.x

- Open ArcView 3.x and the project of interest.
- Step 1: Add the Table
    - Under the Project window, click on Tables
    - Click Add, browse for the DBF file, then click OK.
    - Check the table attributes and close the attribute table
- Step 2: Add data to view
    - Activate view of interest
    - Under the VIEW menu, select 'Add Event Theme'
    - Browse for the Excel file
    - For the X Field choose 'LONGITUDE' and the Y Field 'LATITUDE'
    - Hit OK
    - The point data will be added to the View as an event theme
    - Click on event theme to make active, then click THEME menu, 'convert to shapefile'.
    - Save shapefile with descriptive name in known location.

***Note: where you save your files is CRITICAL!!!! Make sure you are well organized and keep clear records. If you move the position of a file between creation and reopening of a GIS session, you will need to explain to the program were EACH coverage has moved in the next session. It is recommended that you create a folder directly on the C drive to hold your files. This can be transferred directly to the C drive of other computers if you are transferring data.**

Section 3. Paleocontinental reconstruction

Section 3.1 PaleoGIS

PaleoGIS functions as an extension within the ArcView 3.x. It cannot be used directly with ArcGIS. Chris Scotese discontinued the licensing of this program in 2004, but if you have a copy you can still use it.

The primary advantage of PaleoGIS is that it allows the construction of customizable reconstructions. The user can specify both the time interval to reconstruct and can rotate point, line, and polygon coverages onto the paleocontinental topology. Also, the entire data table remains intact during paleocontinental reconstruction. Furthermore, it is possible to manually adjust the positions of the reconstructed plates themselves.

The key disadvantage to PaleoGIS is that is cannot be used directly with ArcGIS 9.x. All layers must be created within the ArcView 3.x environment and then transferred to ArcGIS.

<u>Getting started</u>

- Open PaleoGIS
- Create a new PaleoGIS project
  *****The default location to save PaleoGIS project files is: C:\paleogis\APR*projectname*
- Give the new directory a name

***Note: where you save your files is CRITICAL!!!!—see note above.**

<u>Choose base map layers (e.g. states, countries, counties)</u>

- Select themes to be added (they ask twice)
  - C:\esridata\usa for US files
  - OK to hit cancel 2x
- The present day earth will appear with the default and selected layers in Orthographic projection. Lines across globe are an error with ArcView 3.2 drawing, can be corrected by using ArcView 3.0.
- Remove unnecessary layers
  - *Keep the Grid30.shp and Wcnt.shp layers*

*ArcView 3.2 includes drawing errors. Edges of continents will connect across the front of the globe rather than behind it creating bars. Don't worry, this will go away with reprojection in ArcGIS.

<u>Import lat/long data for rotation</u>

Option 1: Import data previously imported into ArcView 3.x (see section 2)

- Select ADD THEME
- Navigate to shapefile and add.

Option 2: Import data from directly from database

Step 1. Add table to project:

- Under the Project window, click on Tables
- Select 'Tables' and depress the 'Add' button
- Browse for the Excel database file (.dbf)
- Check the table attributes
- Close the attribute table

Step 2. Add data to Present Day view:

- Click on the Present Day window for the menu items to appear
- Select 'View' and 'Add Event Theme'
- Select the Excel database file in the Table heading
- X field remains 'Longitude' and Y field 'Latitude'
- Press OK for the points to appear in the Present Day view

- Make shapefiles of all layers to be rotated (for event theme files)
  - Select the 'Theme' menu and 'Convert to Shapefile'
  - Save the shapefiles with appropriate names. *Keep track of where you save this file!

<u>Assign data to tectonic plates</u>

*This step will add three columns to the data base: (1) code for tectonic plate, (2) first appearance, and (3) last appearance.

*First and last appearances can be altered using table editing to customize WHEN a point will be reconstructed in the rotation step.

- Open the 'Builder' menu, and select 'Cookie Cutter'
- Add the shapefiles to the 'Cookie Cutter' view using the 'Add Theme' button
- Turn on all layers
- Depress the 'intersect' button in the toolbar (scissors)
- Select the theme to be cut up (point data, states, county)
- The Cookie.shp theme will be the cookie cutter theme
- Save the merged shapefile output. *Keep track of where you save this file!
- Add shapefile as theme to a view
- Add Theme to 'Present Day'
- Repeat for all shapefiles you wish to rotate
- Close the 'Cookie Cutter'

<u>Rotate coverages onto Paleocontinental positions</u>

- Remove the old themes from the 'Present Day' view
- Under the 'Tables' menu, open the 'agedata' table and find an age value to reconstruct
  * Note: make sure you find the correct RELATIVE target age and then determine the numeric value from this map that will conform to WHEN you want to be.
- Close table
- Depress the 'reconstruct' button in the toolbar (R)
- Enter the age of reconstruction

- Hit OK
- A new view will open with a title that reflects the chosen age data
  *ArcView 3.2 includes drawing errors. Edges of continents will connect across the front of the globe rather than behind it creating bars. Don't worry, this will go away with reprojection in ArcGIS.
- Save project

*The files created during the rotation will be saved under:
C:\paleogis\APR*projectname*\data*year*master

*Note: You can only save ONE project for a given time. For example, you cannot reconstruct 570.1 twice….but you can reconstruct 570.1 and 570.01, which would be essentially identical paleocontinental configurations.

***Shapefiles cannot be edited within the PaleoGIS environment, so all created files must be imported into ArcView or ArcGIS for manipulation.**

Section 3.2 Point Tracker

With the discontinuation of PaleoGIS, point tracker functions the primary rotation program for point data. Only point data (no lines or polygons) may be rotated with this program. Point tracker requires that data be entered into the program from a notebook or wordpad document including ONLY location ID number, latitude, and longitude of your sites to be reconstructed. The remaining data must be excluded and later reattached to the rotated latitude and longitude values. Only a certain set of time slices are available for reconstruction, currently: 10,20,30,40,50,60,70,80,90,100,120,140,160,180,200,220,240,260,280,300,320, 340,360,380,400,420,440,460,480,500,520,540,560, and 600 million years.

Point Tracker must be installed in the Master Files folder on the C drive.

1. Create a text file with the siteid#, lat, long's of your sites
to be reconstructed (use Notepad or Wordpad). See example in Master Files folder called "sample.txt" for the format of the file.

2. Double click the Point Tracker icon.
Follow steps 1 - 5 in the program.
  Step 1: Enter time (in millions of years)
  Step 2: Hit the Load button (if message "No rotations available" appears, then choose a different time)
  Step 3: Open text file with lat & longs to be rotated
  Step 4: Choose an output format, either PlateTracker format or ArcView format
  Step 5: Save rotated coordinates in a new file. If you are using PlateTracker format, then name the output file "userdefined.dat"
Quit

Once you have determined paleocoordinates for your data, you can reappend this information to your original datafile either in Excel or through a database join in GIS. The rotated lat/long can be used to create an event theme as outlined in Section 2 and placed onto an ArcView shapefile of the appropriate paleogeography obtained from Chris Scotese. These can be manipulated in either ArcView 3.x or ArcGIS.

Section 4. Working with PaleoGIS coverages in ArcView 3.x

ArcView 3.x is directly compatible with all aspects of the PaleoGIS program, since they operate within the same environment. While ArcView 3.x is adequate for use in paleobiogeographic reconstruction, it is recommended that ArcGIS 9.x be used instead due to its increased capacity for manipulation of data and programming capabilities.

Initial considerations

- Construct a project design on paper before committing to it on the computer!
- Consider how to set up the view—new view for each time slice, new view for each phylum, everything in one view
    - Data in different views cannot be compared directly. To use spatial statistics within ArcView all coverages must be in the same view.
    - However, TOO many coverages in a view can be very cumbersome to work with.
- Remember, ArcView draws coverages in order from bottom to top…if you can't see your file, move it up higher in the list.

Getting started

- Start up ArcView 3.2. Choose 'Create a new project' 'with a new view'.
- Add data to the view from your saved project files.
    - Add all data files created during rotation saved within C:\paleogis\APR*projectname*\data*year*master
    - Add lat/long grid from c:\paleomap\data\grids
- Reset the working directory to save your project, shapefiles, etc within a contained folder on the C drive for organized storage and retrieval.
  FILE menu -> Set working directory ->specify an appropriate location

Creating new point shapefiles from a query

The query builder can be a very useful tool for selecting portions of the data set to form the basis for creation of new shapefiles.

- Open the Query builder (from icon or THEME -> Query…)
- Use the tools to build a statement (e.g. [Genus] = "*Acrothele*")
- Use the 'add to' or 'select from' feature to further limit or expand selection
- The selected data will be highlighted on the map.
- To create a new shapefile of the selected data, choose THEME -> 'convert to shapefile. Name and save the file to the correct drive; add to view if you wish.

Creating new polygon shapefiles

To create geographic ranges as polygon features to bound occurrence points, you can use either highlighted points from a whole database query or form a polygon to bound a specifically created point shapefile as created above.

- VIEW -> New theme -> choose polygon. Save appropriate name in appropriate location.
- Push the Draw button, hold down to select the polygon option.
- Use the mouse to click a digital boundary around the area. Double click to end and enclose polygon.
- Reorder the coverages so the points overlie the polygon layer.
- Adjust the position of the vertices with the Vertex Edit tool (open arrow button).
- To save edits, choose THEME -> 'save edits' or 'stop editing'.

*Note: Be sure to create polygon using consistent, repeatable rules (map scale, rounding of corners, etc.)

<u>Quantifying areas: Use the Spatial Analyst Extension</u>

- Spatial analyst can calculate areas, spatial overlap, shifts, etc.

Section 5. Working with PaleoGIS coverages in ArcGIS 9.x

The ArcGIS 9.x system offers a diverse array of features within an incredibly powerful platform that outperforms ArcView 3.2 every way. Therefore, it is recommended that new paleobiogeography projects begin their work within the ArcGIS rather than ArcView platform as early in the analysis as possible. If using Point Tracker for point rotation with prepared paleogeographic base maps, there is no need to use the ArcView 3.x in any stage of the project. If PaleoGIS is used for rotation, then shapefiles should be imported into ArcGIS for manipulation.

Initial considerations

- Construct a project design on paper before committing to it on the computer!
    - This is especially critical with ArcGIS! ArcMap and ArcCatalog are powerful tools and can create a very efficient and organized system of data entry when geodatabases are well designed.
    - It is often simpler to organize data in ArcCatalog prior to working with the data spatially in ArcMap.
- You should organize your data within a single Geodatabase for your project. This will keep all data constrained and allow easier transportation/archiving of the files in a way that will maintain database functionality.

Create a geodatabase

*For projects up to 2 gigabytes of data, you can utilize a personal geodatabase.

- Start ArcCatalog. In the ArcCatalog tree, navigate to the location where you intend to store your data. (It is recommended this is in a folder directly attached to the C drive).
- Right click folder, point to New, and click Personal Geodatabase. Name the database.

Import shapefiles from PaleoGIS to ArcCatalog

- In ArcCatalog, right click on *personalgeodatabase*. Point to import, and click feature class (single). (You can also bring in multiple features classes at once).
- Either click on the folder icon and navigate to the PaleoGIS shapefile OR find the file the catalog tree then click and drag the shapefile into Input Features box of the dialog.
    - ArcCatalog should correctly assume that you want to import the data into your geodatabase as the output location.
    - Rename the shapefile for your geodatabase—names must NOT include spaces.
    - Click ok. When the operation is completed, click close on the progress report.

    *Your rotated shapefiles will be saved at C:\paleogis\APR*projectname*\data*year*master

*Also import the lat/long grid from c:\paleomap\data\grids

- You can preview the data (both tabular and spatial) within ArcCatalog by clicking on the feature class, then its preview tab.

Adding data to ArcMap

- Launch ArcMap (from within ArcCatalog or from icon). Choose Start ArcMap with a 'New Empty Map'.
- Right click on 'Layers', choose Add Data. Navigate to *personalgeodatabase* and click on layers to add.
- Right click on 'Layers', choose Properties. Rename for clarity.
- Within Properties, click on the Coordinate system tab. Reset to a new system (will remove edge overlap drawing errors).
- Save file. *Save to same folder as geodatabase!

***Note:** You could add data into ArcMap without organizing in ArcCatalog, but this creates a difficult bookkeeping system. To import project DIRECTLY from PaleoGIS, FILE -> Import from ArcView Project -> navigate to the file. This will create workable files within ArcMap, but will NOT reorganize the files within the hard drive. Hence, saving/archiving/transporting for the map document will be difficult. This method is NOT recommended.

Changing display options

- ArcMap will display the added layers with a default projection and symbology.
- To *change the map projection*, right click on the data frame and choose 'properties'.
    - Choose the tab labeled 'coordinate system'.
    - Change the projection to your preferred coordinate system.
    - **Note:** Projected coordinate systems will allow you to reorient the globe to a preferred orientation, but will not allow digitizing of polygons. Non-projected coordinate systems will allow polygon creation, but cannot be rotated.
- To *change the layer symbology* (marker or fill pattern, etc.), right click on the layer and choose 'properties.'
    - Customize using the symbology tab.

Additional Data Frames

One of the great features of ArcGIS 9.x is that you can keep ALL time slices within a single document and file storage system. Use additional Data Frames for each distinct time slice. Eventually, these individual data frames can be combined into a single layout for presentation.

- Choose the INSERT menu -> Data Frame.
- This creates a new data frame. The same features classes can be added here, but modified separately in different layers.
- Add layers and customize data frame as before. You can also just drag and drop all data from previous data frames onto the new one.

Creating new point shapefiles from a query

The select by attributes feature can be a very useful tool for selecting portions of the data set to form the basis for creation of new layers.

- From the SELECTION menu -> Select by Attributes
- In the selection attributes dialog box, click the Layer of interest from the drop down list. Chose the appropriate selection method.
- Use the tools to build a statement (e.g. [Genus] = "*Acrothele*"). After entering the first part of the syntax, click the Get Unique Values button for choices.
- The selected data will be highlighted on the map.
- To create a new layer from the selected data, right click on the queried layer. Choose Selection -> Create layer from selected features.
- A new, temporary layer will appear on the data frame named '*Layer* selection'.
- To save the layer, right click on its name, choose 'Save as Layer File…' and save with appropriate name. ArcMap will save the layer by default in the same folder as the map document.

Creating new polygon feature classes

In ArcGIS, the creation of new polygons is done within a new feature class. For each time slice, one new feature class could be constructed to include all polygon ranges for the time slice (data frame) OR each individual range could be saved as its own feature class.

Step 1: Create a new feature class

- Within ArcCatalog, right click on the geodatabase where you want to create the feature class. Point to New -> Feature Class.
- Choose the type of feature you wish to create, then click next.
- Specify the storage configuration (default is fine), then click next.
- Click Shape under Field Name. Under field properties, choose Geometry Type and click the option you want (polygon is default).
- Click Finish.
- To add this layer to ArcMap, click and drag to the map or use Add data in ArcMap.

Step 2: Adding polygons to the feature class

- Make sure the Editor toolbar is visible (if not, go under VIEW menu to activate)
- Under Editor drop down, click Start Editing.
- The Task box should read: Create New Feature.
- The Target box should list the layer to edit.
- Make sure the sketch tool (pencil) is selected.
- Use the mouse to click a digital boundary around the area. Double click to end and enclose polygon.
- Reorder the coverages so the points overlie the polygon layer.
- Adjust the position of the vertices with the Modify Feature options.
- To save edits, choose Editor -> 'save edits' or 'stop editing'.

- The attribute table associated with this new feature class will automatically calculate area and length of the polygon.

*Note: Be sure to create polygon using consistent, repeatable rules (map scale, rounding of corners, etc.)

Quantifying areas: Use the Spatial Analyst or Geostatistical Extensions

- Spatial analyst can calculate spatial overlap, shifts, etc.

Use of ArcToolbox and Modeling

- ArcGIS is a very powerful program and includes an easy to use modeling program in ArcToolbox as well as use of the Python programming language. It is possible to automate polygon construction for large datasets (although we haven't worked it out yet.)

Section 6. References and Further Reading

Program References

Environmental Systems Research Institute, Inc. (ESRI). 1999. ArcView GIS 3.2. Redlands, California.

Environmental Systems Research Institute, Inc. (ESRI). 2006. ArcGIS 9.2. Redlands, California.

Ross, M. I., and C. R. Scotese. 2000. PaleoGIS/Arcview 3.5. PALEOMAP Project, University of Texas, Arlington.

Scotese, C.R. 2004. Point Tracker. PALEOMAP Project, University of Texas, Arlington.

GIS Reference/Starter Books

Ormsby, T, E. Napoleon, R. Burke, C. Groessl, L. Feaster. 2004. Getting to know ArcGIS desktop, 2nd Edition. ESRI Press, Redlands, CA, 572 p. *Great place to learn the basics of how ArcGIS operates.*

Environmental Systems Research Institute, Inc. (ESRI). 1999. Getting to know ArcView GIS: The geographic information system (GIS) for everyone, 3rd Edition. ESRI Press, Redlands, CA,572 p. *Great place to learn the basics of how ArcView operates.*

Environmental Systems Research Institute, Inc. (ESRI). 2004. ArcGIS 9: Using ArcCatalog. ESRI Press, Redlands, CA, 293 p. *Good reference manual for program functionality.*

Environmental Systems Research Institute, Inc. (ESRI). 2004. ArcGIS 9: Using ArcMap. ESRI Press, Redlands, CA, 585 p. *Good reference manual for program functionality.*

Environmental Systems Research Institute, Inc. (ESRI). 2004. ArcGIS 9: Using ArcGIS Geostatistical Analyst. ESRI Press, Redlands, CA, 300 p.

Environmental Systems Research Institute, Inc. (ESRI). 2004. ArcGIS 9: Using ArcGIS Spatial Analyst. ESRI Press, Redlands, CA, 232 p.

***ESRI.com is an excellent website with additional resources!**

GIS in Paleobiogeography

Graham, R. W., E. L. Lundelius, Jr., M. A. Graham, E. K. Schroeder, R. S. Toomey, Iii, E. Anderson, A. . Barnosky, J. A. Burns, C. S. Churcher, D. K. Grayson, R. D. Guthrie, C. R. Harington, G. T. Jefferson, L. D. Martin, H. G. Mcdonald, R. E. Morlan, H. A. Semken, Jr., S. D. Webb, L. Werdelin, And M. C. Wilson., 1996. Spatial response of mammals to late Quaternary environmental fluctuations. Science, 272:1601-1606.

Mensing, S.A., Elston, R.G., Jr., Raines, G.L., Tausch, R.J. and Nowak, C.L., 2000. A GIS model to predict the location of fossil packrat (Neotoma) middens in central Nevada. Western North American Naturalist, 60(2): 111-120.

Kovar-Eder, J. and Kvacek, Z. (2002) Towards vegetation mapping based on the fossil plant record. *Acta Universitatis Carolinae Geologica,* 46: 7-13.

Lacruz, R., Berger, L.R., Brink, J.S., Hancox, P.J., and Ungar, P. (2003) Gladysvale; fossils, strata and GIS analysis. *South African Journal of Science,* 99: 283-285.

Manship, L.L. (2004) Pattern matching: Classificaiton of ammonitic suters using GIS. *Palaeontologica Electronica,* 7: 1-15.

Markwick, P.J. (2002) Integrating the present and past records of climate, biodiversity and biogeography; implications for palaeoecology and palaeoclimatology, In *Palaeobiogeography and biodiversity change: the Ordovician and Mesozoic-Cenozoic radiations* (eds. J. A. Crame and A. W. Owen), Geological Society of London, London, pp. 179-199.

Markwick, P.J. and Lupia, R. (2002) Palaeontological databases for palaeobiogeography, palaeoecology and biodiversity; a question of scale, In *Palaeobiogeography and*

biodiversity change: the Ordovician and Mesozoic-Cenozoic radiations (eds. J. A. Crame and A. W. Owen), Geological Society of London, London, pp. 169-178.

Mensing, S.A., Elston, R.G., Jr., Raines, G.L., Tausch, R.J., and Nowak, C.L. (2000) A GIS model to predict the location of fossil packrat (Neotoma) middens in central Nevada. *Western North American Naturalist,* 60: 111-120.

Rayfield, E.J., Barrett, P.M., McDonnell, R.A., and Willis, K.J. (2005) A geographical information system (GIS) study of Triassic vertebrate biochronology. *Geological Magazine,* 142: 327-354.

Rode, A.L. and Lieberman, B.S. (2004) Using GIS to unlock the interactions between biogeography, environment, and evolution in Middle and Late Devonian brachiopods and bivalves. *Palaeogeography, Palaeoclimatology, Palaeoecology,* 211: 345-359.

Rode, A.L. and Lieberman, B.S. (2005) Integrating evolution and biogeography; a case study involving Devonian crustaceans. *Journal of Paleontology,* 79: 267-276.

Stigall Rode, A.L. (2005) The application of Geographic Information Systems to paleobiogeography: Implications for the study of invasions and mass extinctions. *Paleontological Society Memiors,* 11: 77-88.

Vermeersch, P.M. (2005) European population changes during marine isotope stages 2 and 3. *Quaternary International,* 137: 77-85.

Stigall, A. L., and B. S. Lieberman. 2006. Quantitative Paleobiogeography: GIS, Phylogenetic Biogeographic Analysis, and Conservation Insights. Journal of Biogeography, 33: in press.

Stigall, A.L. In review. Combining GIS and phylogenetic biogeography in paleontology: Implications for the study of invasions and mass extinctions, p. xxx. *In* P. Upchurch, A. McGowan, C. Slater (eds.), Palaeogeography and Palaeobiogeography: Biodiversity in Space and Time. CRC Press.