PhyloGeoViz v.2.4.4 User’s Manual
04/13/2011
Yi-Hsin Erica Tsai

PhyloGeoViz is a web-based program that plots frequency data on a map.

Table of Contents

What is PhyloGeoViz?

PhyloGeoViz is a web based application that plots any sort of frequency data with a geographic component. It was designed to create geographic visualizations of DNA haplotype data that are often used in the course of phylogeographic analysis. These visualizations demonstrate the spatial distributions of each haplotype, the frequency of each haplotype in each population, and the number of samples included per population. However, the program can be used for any kind of data.

Technical specifications

PhyloGeoViz has been tested and runs successfully in Firefox v.3.6 (some limited functionality in Internet Explorer). The latest version is available at http://phylogeoviz.org.

Example 1: A simple tutorial

This tutorial walks you through how to input a simple dataset and generate a map using PhyloGeoViz.

Starting PhyloGeoViz:

In your Mozilla Firefox browser, go to http://phylogeoviz.org. Click on “Start PhyloGeoViz”.

Data input

PhyloGeoViz input is a simple data matrix containing information for each locality in a different row. Data must be input in this order: population name, latitude, longitude, then the frequency of each category. Frequency information can be input as counts (can sum to greater than 1); in that case, pie charts will be scaled to the total sample size per population. All values should be separated by whitespace (space or a tab). For instance, for four populations with three types, data should be formatted like this:

```
popA 35.929673 -78.948237     3      2      5
popB 38.889510 -77.032000     5      1      4
popC 38.032120 -78.477510     0      1      9
popD 36.379450 -75.830290     7      2      1
```

Generating a Visualization

Click on “Draw Map!”. PhyloGeoViz will convert the haplotype frequency data into pie-charts overlaid on an interactive Google Map (Figure 1). You can move the pie-charts, zoom the map, and re-center the map to customize the display of your data. Colors can be changed by expanding the menu “Haplotype colors” and using the colorpicker. Change pie sizes under the “Sample sizes” tab. Outlines can be added under “Pie options”. The total map area can be expanded or shrunk to fit your browser window under “Map options”, this can be useful when trying to generate figures.

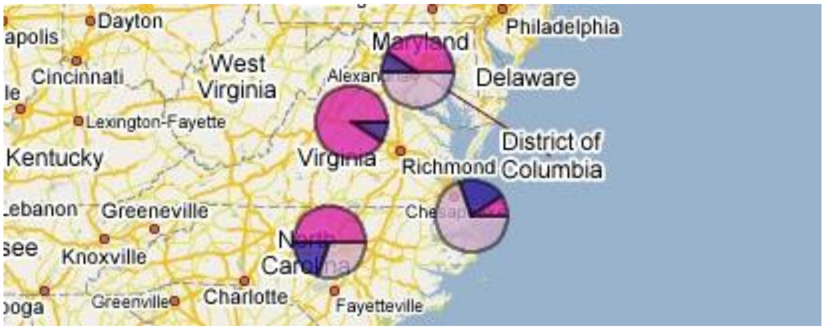

Figure 1: Example PhyloGeoViz output in Google Maps

Exporting to Google Earth

Under “Export options”, click “Export to Google Earth” to export the visualization data in the KML file format. Once in Google Earth, you can still change the colors.

Example 2: Investigation of phylogeographic data

This section discusses how you might use PhyloGeoViz to explore multi-locus genetic data.

Example data:

This example utilizes microsatellite data from:
Tsai Y-HE & Manos PS (2010) Host density drives the postglacial migration of the tree parasite, Epifagus virginiana. *Proc. Natl. Acad. Sci. USA* 107:17035-17040.
Download the example phylogeoviz input file here: MS BAPS.phylogeoviz.txt
Download example Google Earth output here: Tsai and Manos.MS pies.kmz

Plotting data marker by marker:

Many studies of population genetic variation involve numerous markers. PhyloGeoViz can help you visualize their spatial distribution in several ways. First, you can plot each marker's information separately (see Tsai and Manos.MS pies.kmz). For instance, here are the maps for three of the microsatellite loci with each allele colored differently (Figs 2-4):

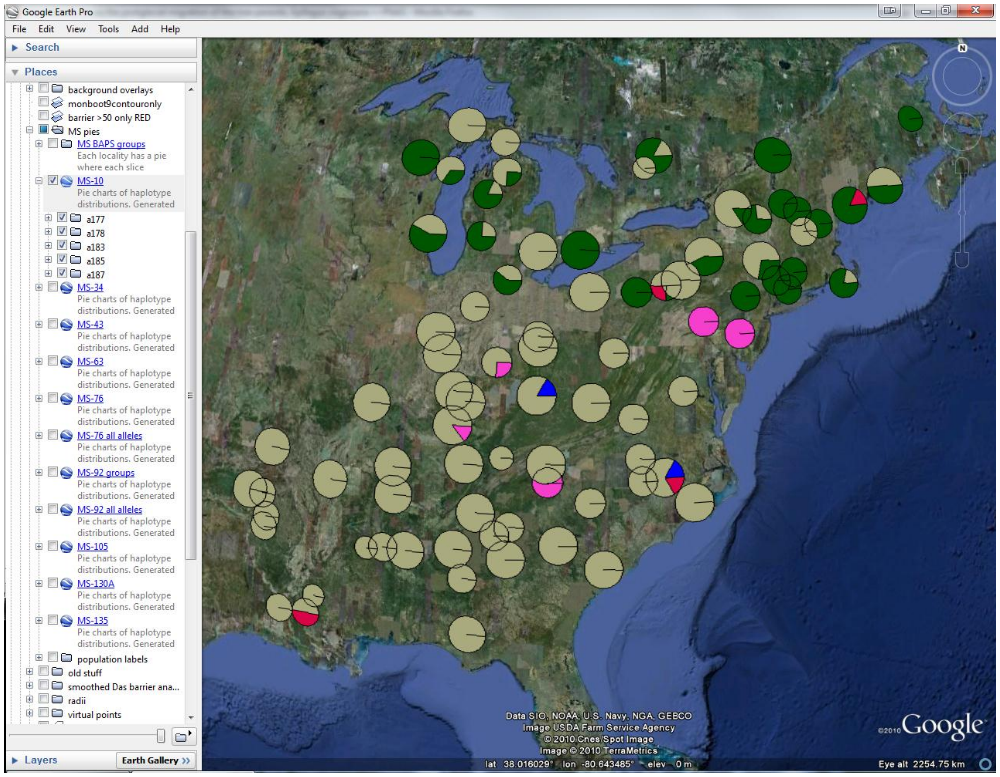

Figure 2: Alleles of microsatellite locus MS-10

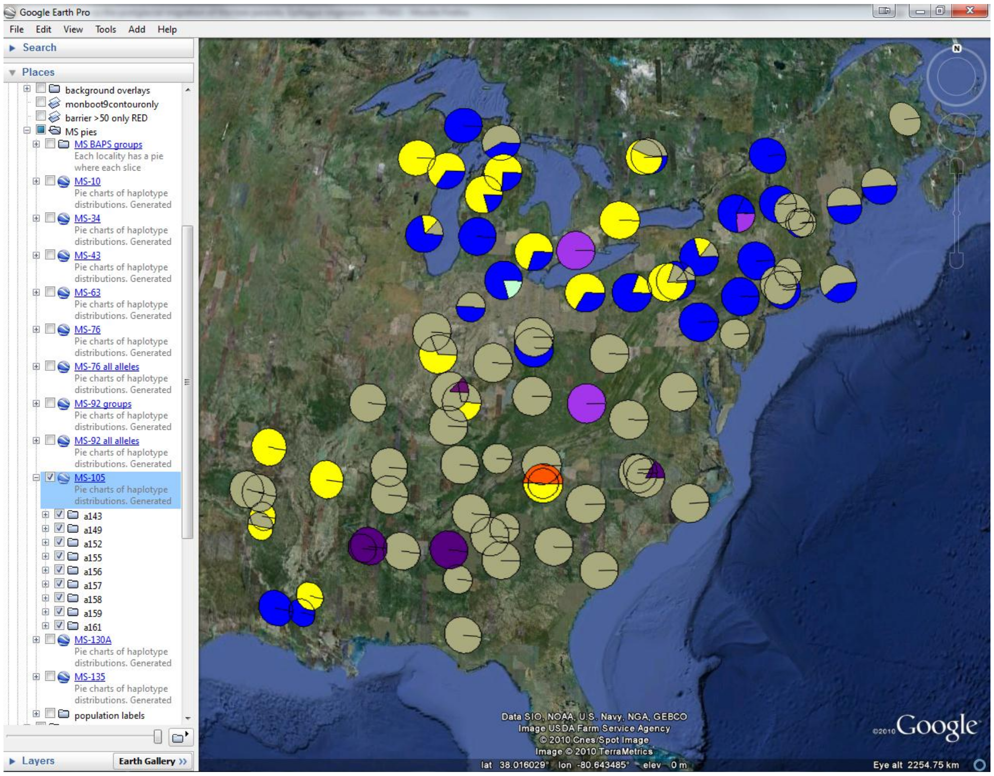

Figure 3: Alleles of microsatellite locus MS-105

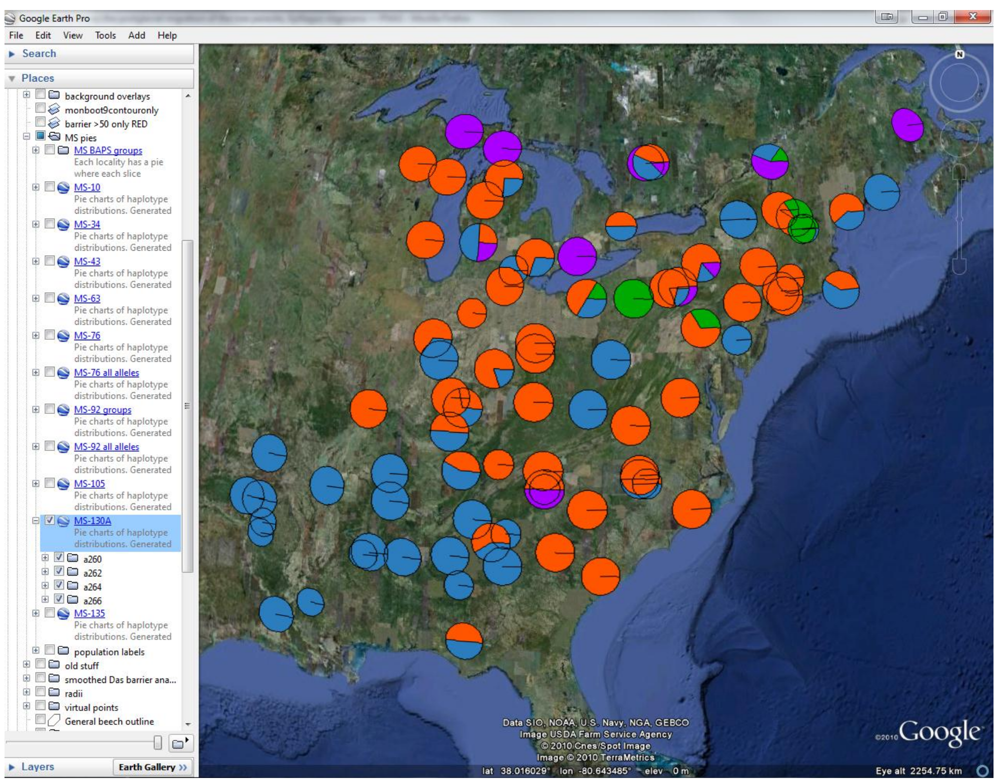

Figure 4: Alleles of microsatellite locus MS-130A

Plotting data summarized over all loci:

To better understand the overall signal across loci, Tsai and Manos used the Bayesian clustering program, BAPS (http://web.abo.fi/fak/mnf/mate/jc/software/baps.html). The cluster group assignments were then visualized in PhyloGeoViz. The input file is available here (MS BAPS.phylogeoviz.txt). Other ways of summarizing across loci could be through a STRUCTURE (http://pritch.bsd.uchicago.edu/software.html) type assignment test, population specific Fsts, a moving window analysis of allelic richness or endemism, etc.

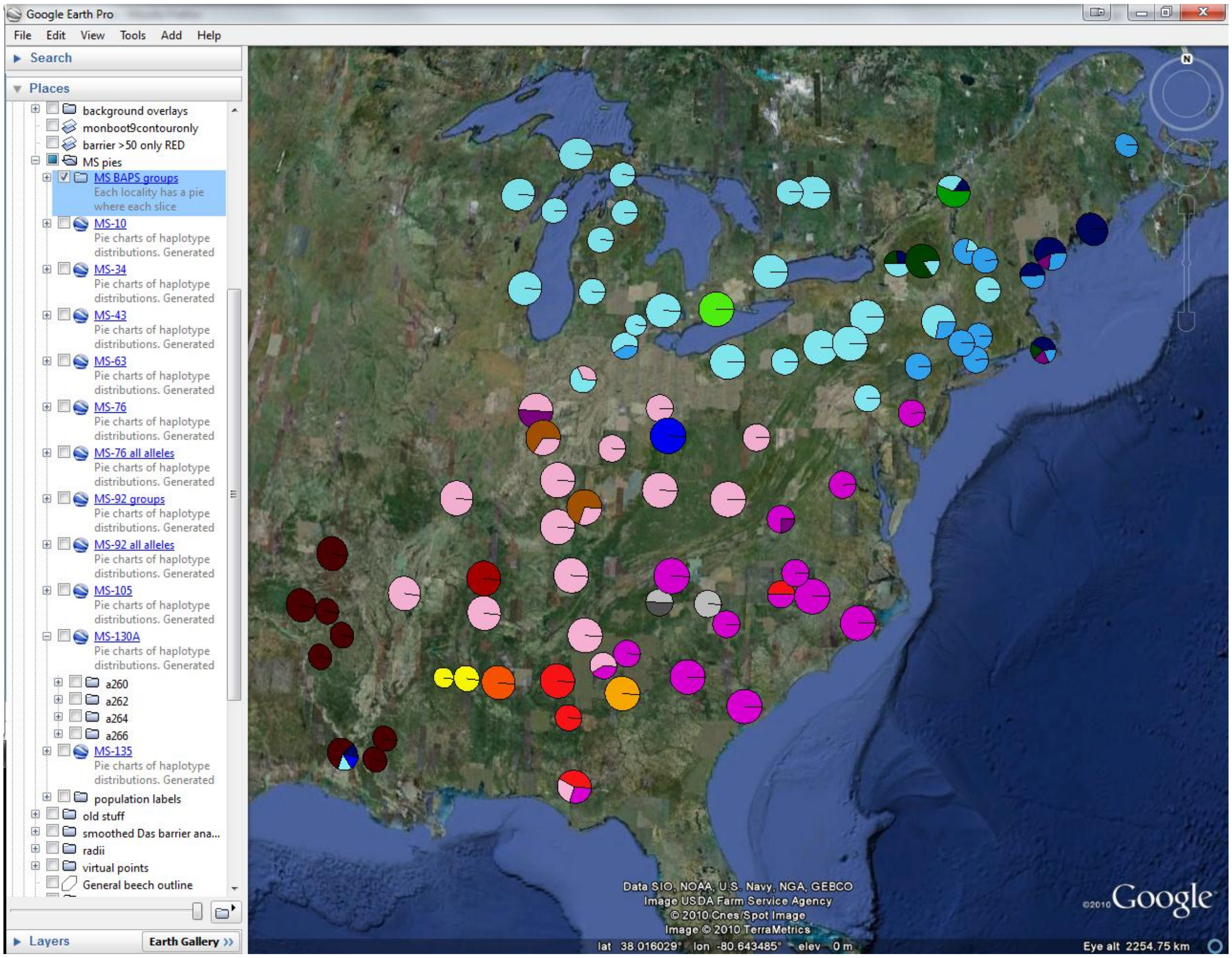

Figure 5: BAPS cluster assignments based on 9 microsatellite loci

Where to go from here?

Once the data are in Google Earth many spatial layers become available that may help generate hypotheses about the processes determining the distribution of genetic variation in your system. For instance, historical imagery is available at many locations and scales (in Google Earth enable View->Historical Imagery), the USGS provides many maps of land cover, watershed information, human population data, etc. (see http://edna.usgs.gov/watersheds/), GBIF has downloadable species occurrence information (http://data.gbif.org/welcome.htm), and many more user generated overlays are available.

PhyloGeoViz options

Haplotype colors:

By default, each haplotype/category is assigned a random color and is placed within its own haplotype group. If you have many haplotypes, it can become very difficult to assign colors that are easily distinguishable. One solution to this is to assign each haplotypes to a haplotype group. Then, the haplotype colors of an entire group can be manipulated together as a unit. On the input screen, you can add this information in the second text box. The format is 2 columns, first is the haplotype name, followed by the haplotype group. Each row is a different haplotype. For example, for a 3 haplotypes in 2 groups system:

```
hapA  high
hapB  low
hapC  high
```

To change the haplotype color, click on the box next to the haplotype name or the haplotype group's name. A color picker window should pop up, and you can use the slider to browse colors, or use the field to input a 6-digit web color.

Sample Sizes

Clicking on this option produces a legend of sample sizes. You can increase or decrease all the pie sizes by 10% here.

Pie options

From here you can also change the pie sizes, and you can choose to outline the pie wedges. Also, if you would like to make any changes to the underlying data, follow the "Edit Data" link. From here, you can also assign haplotypes to haplotype groups and choose what haplotypes to display.

Map options

You can enlarge or shrink the size of the map. This can be very useful if you are trying to take a screenshot.

Export options

Click on "Export to Google Earth" to produce a kml file readable by Google Earth.

Tips for producing publication quality figures

Resolution considerations:

When producing a figure, you must consider what your goal size and resolution are. All figures generated via Google Earth or Google Maps are raster based – there is no vector based output format, so the resolution of images becomes very important. The following table gives my recommendations on what procedure to follow to generate a quality figure. I assume that a typical screenshot results in a 1280 px by 1024 px image. If you have a very large, high resolution monitor, you can make larger images with screenshots and Google Earth.

For use in:	Resolution	Size (approx.)	Recommendation
Web	72-200 dpi	NA	Screenshot or Google Earth -> Save as Image
Powerpoint/presentations	96 dpi	Up to 13" x 10.5"	Screenshot or Google Earth -> Save as Image
Print	300+ dpi	Up to 4" x 3.25"	Screenshot or Google Earth -> Save as Image
Print	300+ dpi	Larger than 4" x 3.25"	Must use Google Earth Pro

Screenshots or Google Earth -> Save as Image

The first option is to simply take a screen shot (On a PC press the button "PrntScr", then paste into a graphics editor like Adobe Photoshop or MS Paint, and crop as needed) or use the "Save Image" option in Google Earth (File->Save->Save Image). This process works well if you are generating either a small image or one of low resolution. Typically, this method is fine for presentation images, because most projectors can only output at 96 dpi, and also this works well for figures destined for the web (96 dpi is fairly standard).

Google Earth Pro -> Save as Image

Most print journals and posters require higher resolution images of 300 dpi or larger. For instance, my computer's screen shot results in a 1280 x 1024 pixels image. At 300 dpi, this image is only 4.27" x 3.41". At 600 dpi, the image is 2.13" x 1.7". To get larger higher resolution images, you must have access to Google Earth Pro. For instance, on my screen using Google Earth Pro I can produce an image of 4800 x 3758 pixels, which at 300 dpi = 16" x 12.5".

Producing a high quality image

1. I assume that you have finalized color and other choices at this point and are working in Google Earth.
2. Sometimes the default background satellite images in Google Earth can be too "busy" for publication.

   a. You can download simplified background overlays showing for instance, just political boundaries at GPSVisualizer.com. If they don't have what you're looking for, try searching for the keywords: "KML", "overlay", and your study region.
   b. If the background overlay is almost right, but you want to change a few colors, follow these steps.
      i. Locate the actual image associated with the overlay. Right click on the overlay -> Properties ->Link shows the path of the image.
      ii. Open the image in a graphics editor like Adobe Photoshop. (Note you may want to copy the image in case of mistakes!).
      iii. Using the graphics editor make whatever color changes or manipulations you want. See specific documentation from your graphics program for help. Here's a tutorial for Photoshop to help you get started:
      iv. Right-click -> Refresh to reload the overlay in Google Earth to see the changes.
   c. If you have a background map you'd like to use, import that image as an overlay in Google Earth (Add->Image Overlay). The best case is that you have lat/lon information on the image, but if not, you can still stretch and manipulate the image to roughly fit on the Google Earth globe. Beware – this only works well for certain projection systems. For instance, Equal-Albers projection maps may be very difficult to fit appropriately.
   d. Are your pies obscured by the overlay? Change the altitude to raise them over the overlay (right click on polygon->Properties->Altitude), but beware, read the next section("For each color") carefully.
3. For each color change the altitude and set an appropriate outline width:
   a. Ideally, you would perform the following steps BEFORE you have finalized all of your color selections. If not, be sure to write down all the RGB values for your colors before continuing. Oftentimes these color selections are erased during this process if you accidentally share properties across multiple colors.
   b. Set all pie wedges to be "relative to ground" at some altitude (right click on a folder or polygon->Properties->Altitude). Usually 500m works well, but you may have to play around with this value. Why do you have to do this? Unfortunately, there is a bug in Google Earth Pro (follow the link for more information, and star it if you want to see it changed!).
   c. After changing the altitude you may notice that all your pie wedges are now darkened as if by a black filter. This is unfortunately another bug in Google Earth. Please follow the link and star the error to encourage Google Earth developers to fix this problem. The work around for this problem is unfortunately very tedious. You must click within each individual pie wedge's altitude field for the opacity to be corrected. I highly suggest downloading a macro recorder (for example, AutoHotkey) to help you with this.
   d. Note: When saving images through Google Earth, the outlines tend to be much thinner than they appear on the screen. I suggest outline widths of 5 as working well, even though they look terrible on the screen. Save a few test images as you work to judge if the altitude and outline settings are reasonable.
4. Dealing with pie collisions/overlapping pies:
   a. Before exporting your data to Google Earth (i.e. when you were in the Google Maps interface), click and drag to separate the worst offenders.
   b. Adjust the altitudes of individual pie wedges (right click on wedge->Properties->Altitude) so that one pie is entirely above another. Note, when doing this, have the map at the same viewpoint as you want the finished product to be. Sometimes different zoom levels can show different overlapping patterns.

5. Occasionally, pies that look fine on the screen do not print well. For example, a pie may appear to have a black spot in the middle. This generally means that the altitude property of the pie should be adjusted. Go back to the "altitude" tab and increase the altitude, that should fix the problem.
6. Save often!

FAQ

How to cite this program

Tsai, Y.-H. E. (2011). "PhyloGeoViz: a web-based program that visualizes genetic data on maps." *Molecular Ecology Resources.* 11(3): 557-561. [html] [pre-review pdf]

Do I need to cite Google?

Absolutely you need to cite Google Earth and the data layers underlying your maps. Check Google's permissions website for full details.

How do I get same sized pies?

Eventually, I'd like to add this functionality to PhyloGeoViz, but for now, here's a work around. In a spreadsheet, convert all of your count data to frequency data (ranging from 0-1), then input the newly formatted data.

How do I change the bin size?

Look for this to be fixed in the next version of PhyloGeoViz. Right now the bin sizes are hard-coded in, and there are a maximum of 3 bins. To work around this, decide what 3 bins you would like. Scale all your count data to fit these three values: 1, 5, and 10. For example, say your original data looks like this:

```
popA 35.929673 -78.948237     3     2     5
popB 38.889510 -77.032000     1     0     0
popC 38.032120 -78.477510     0     1     9
popD 36.379450 -75.830290     2     2     1
```

You would like the new bins to be 0-5, 6-9, 10+. For small populations (0-5), scale the count data so that the total sample size is 1. For medium populations (6-9), scale the count data to 5. For large populations, scale the count data to 10. Resulting in this new data matrix:

```
popA 35.929673 -78.948237     3     2     5
popB 38.889510 -77.032000     1     0     0
popC 38.032120 -78.477510     0     1     9
popD 36.379450 -75.830290     0.4   0.4   0.2
```

How do I deal with overlapping pies?

1. Before exporting your data to Google Earth (i.e. when you were in the Google Maps interface), click and drag to separate the worst offenders.
2. Adjust the altitudes of individual pie wedges (right click on wedge->Properties->Altitude) so that one pie is entirely above another. Note, when doing this, have the map at the same viewpoint as you want the finished product to be. Sometimes different zoom levels can show different overlapping patterns.

Where do I submit a bug report? Can I request a feature?

Yes, please do submit bug reports and request new features! Email me (Erica Tsai) at etsai@lsu.edu. I cannot promise to implement everything requested, but I certainly try my best to think of workarounds at the very least.