**Population Genetics Analysis using Haplotype Network**

Haplotype mapping is useful way of graphical representation of genetic variations (SNPs) among DNA sequences within the populations or species.

Tool Requirement and Installation:

1. **Geneious** (<https://www.geneious.com>) or any other Multiple Sequence alignment tool.
2. **DNASP**: DNA Sequence Polymorphism. System requirement is Windows. (<http://www.ub.edu/dnasp/>).
3. **Popart**: Population Analysis with Reticulate Trees is a free Population genetics software. It works on Windows/Mac OSX (10.6 or greater). (http://popart.otago.ac.nz/index.shtml)

Method

1. **Geneious:**

* **Input file**: fasta file of two more sequence
* For pairwise alignment, fasta files need to be in same length. User can also modify sequences in Geneious to make them same length.
* **Pairwise/Multiple Alignment**: In Geneious from Align/Assemble toolbar, choose Pairwise alignment. To get faster results, choose Muscle alignment with default parameters or Geneious multiple alignment option.
* Export the alignment into fasta file format from Geneious.

1. **DNASP:**
   * **Input file**: Alignment file in fasta format.
   * Import the alignment in fasta file: From file Manu Open data file.

Note: Fasta alignment will not work if input data file has any invalid character (ambiguity). To correct it, from file menu select **convert a fasta file with Ambiguity codes to Ns** Option. Then select input file and output file. And click Run. It will convert any ambiguity character to Ns and save the results to the folder User has selected. File import has been successfully done. Now user can proceed to Haplotype analysis.

* + To create a haplotype in nexus file format, Select **Generate** Menu and choose Haplotype Data file. It will Create the data file with haplotype information. It will open a dialog box with different options or use the default. Click ok button. It will start the analysis and ask the user to save the file. Save the file in nexus format to be able to use in Popart.

Note: DNASP won't take any ambiguity in alignment and it will convert any ambiguity into Ns. But to avoid false haplotypes calls, it is recommended not to consider any sequences with Ns/gaps.

1. **Popart:**

* **Input file:** Popart reads Nexus alignment file. The Nexus file generated from the DNASP needs some modifications before running into Popart to color the nodes of haplotype map.

**Example and description of Nexus file is below.**

* Nexus file begins with #NEXUS command. Any text within the bracket is comment and will be ignored by the program.
* All the blocks starts with a keyword **Begin block\_name;** and close with keyword **End;** The first two blocks are **taxa block**, followed by **Character block** which user can copy from DNASP alignment file. The taxa block is the simple sequence labels and character block is the sequence alignment.

***Example of nexus file:*** (Text in color is the example of nexus file. The text in bracket is the description for file**)**

#NEXUS

*[This is an example of a Nexus file, modified from a sample file distributed*

*with DNASP. It should run in PopART.]*

**BEGIN TAXA**;

*[You need either taxa and characters blocks, or else a data block.]*

DIMENSIONS NTAX=7;

TAXLABELS

seq\_1

seq\_2

seq\_3

seq\_4

seq\_5

seq\_6

seq\_7

;

END;

**BEGIN CHARACTERS**;

[You can either use the matchchar option or else include all characters for

all sequences.]

DIMENSIONS NCHAR=56;

FORMAT DATATYPE=DNA MISSING=? GAP=- MATCHCHAR=.;

MATRIX

seq\_1 ATATACGGGGTTA---TTAGA----AAAATGTGTGTGTGTTTTTTTTTTCATGTGG

seq\_2 ......--..A..---...C.----.G...C.A...C..C...C............

seq\_3 ..........A..---...T.----.G.............................

seq\_4 ..........A..---G...T----..............................A

seq\_5 ..........A..---G...G----..............................C

seq\_6 ..........A..---G...C----..............................T

seq\_7 ..........A..---G....----..............................A

;

END;

[The Trait block is specific to Popart. The number in the matrix are the numbers of samples associated with each trait. The order of the columns should match the order of TraitLabels. Separator can be Comma, Space, or Tab.]

**BEGIN TRAITS**;

[Ntraits = number of geographical locations you have;**]**

Dimensions NTRAITS=5;

Format labels=yes missing=? separator=Comma;

**[**Optionally if you include TraitLatitude and TraitLongitude they will be used to place trait groups on the map.]

TraitLatitude 53 43.6811 5.4 -25.61 -0;

TraitLongitude 16.75 87.3311 26.5 134.355 -76;

[TraitLabels are the name of locations **without commas**;]

TraitLabels Europe Asia Africa Australia America;

Matrix

seq\_2 10,5,0,6,0 [i.e. seq\_2 has **10** samples from Europe, **5** from Asia, **0** from Africa, **6** from Australia and **0** from America**]**

seq\_7 0,0,5,0,0

seq\_5 4,0,10,0,0

seq\_4 0,0,0,4,2

seq\_3 0,0,0,3,5

seq\_1 0,0,0,3,3

seq\_6 0,0,0,7,3

;

END;

**Running Popart (Method1):**

* After creating the nexus file as shown above, Open the Popart software.
* Import the Nexus file from file menu. The data view lets you look at your alignment or traits data.
* Then choose the type of network you want to draw. PopART includes several network inference methods like Minimum Spanning, Median Joining, TCS, Ancestral Parsimony, Integer Neighbor-Joining, and Tight Span Walker.
* Once the haplotype network is drawn, user can change the color, move the nodes around, or edit node sizes and label fonts.
* Popart only allows 10 colors for traits. However, user can add different color by right click on keys to change the colors associated with different traits.
* Once the network is finalized, export the network as a graphical file.

Method2:

In Popart, User can also import alignment (nexus file) and traits (.csv file) separately and create a network as described in method1.

Nexus file is simple DNASP output file. Import .nex file in popart.

Than create a trait file in csv format

Example:

,Africa,SouthEastAsia,PapuaNewGuniea,SouthAmerica

Hap\_1,45,1,1,1

Hap\_2,16,0,0,1

Hap\_3,1,1,18,0

Hap\_4,0,3,4,1

Insert this trait file in popart and build the network as described above.