



Welcome to ShiNyP



This work, primarily authored by Yen-Hsiang Huang from the Department of Agronomy, National Chung Hsing University, Taiwan, is licensed under the [GNU Affero General Public License](#).

Last Updated: Jun 2025

On this page

[Welcome to ShiNyP](#)

[Quickstart](#)

[Overview](#)

[Publication](#)

[Updates and Support](#)

[View source](#)

[Edit this page](#)

Quickstart

[Run ShiNyP via R](#)

[Run ShiNyP via Docker](#)

[ShiNyP Online Version – Trial Platform](#)

Overview

ShiNyP is a platform designed for real-time processing, analysis, and visualization of SNP datasets.

Input: Genome-wide biallelic SNP in Variant Call Format (VCF) file.

Analysis: Data QC, population genetics analysis, core collection, and more.

Output: Publication-ready figures, tables, R objects, and free AI-driven report.

The screenshot shows the ShiNyP home page. At the top is a navigation bar with links: Home, Data Input, Data QC, Data Transform, Population Structure, Genetic Diversity, Selection Sweep, Core Collection, and AI Report. Below the navigation bar is a section titled "ShiNyP: SNP Analysis and Visualization Platform". It includes a brief introduction: "Yen-Hsiang Huang, National Chung-Hsing University (NCHU), Taiwan. For any inquiries, please email us at: teddyhuangyh@gmail.com". There are four main sections: "Key Features" (Real-time Processing, Analysis, and Visualization of SNP Datasets; Input: VCF file; Analysis: Data QC, population genetics analysis, core collection, and more; Output: Publication-ready figures, tables, R objects, and AI-driven report); "Quickstart" (Brief instructions to begin with Data Input); "Publication" (Citation: Huang et al. (upcoming 2025) ShiNyP: An Interactive Shiny-Based Platform for Genome-Wide SNP Analysis and Visualization); and "Support" (Feedback form: <https://forms.gle/GPCqgS05czvNLf0B7> (Google Form) or email: teddyhuangyh@gmail.com).

Publication

If you use *ShiNyP* in your research, please cite:

Huang, Y.-H., Chen, L.-Y., Septiningsih E. M., Kao, P.-H., Kao, C.-F. (2025) *ShiNyP: Unlocking SNP-Based Population Genetics—AI-Assisted Platform for Rapid and Interactive Visual Exploration*. *Molecular Biology and Evolution*, 42(6), msaf117. <https://doi.org/10.1093/molbev/msaf117>

In addition, please acknowledge the R packages utilized in your analysis. The relevant citations and descriptions for each module are detailed in the *ShiNyP User Guide*.

Updates and Support

Aug 2024: alpha version

Oct 2024: v0.1.0

Feb 2025: v0.1.1

Apr 2025: v0.2.0

- Enhanced AI report functionality with new options and models.
- Improved readability of preliminary results.
- Added more methods for constructing core SNP set.
- Added the Docker-based installation.

May 2025: v1.0.0 on GitHub.

- Introduced the **new ShiNyP AI chatbot**.
- Enhanced AI report features and deprecated older AI models.
- Added publication details.
- Made UI/UX improvements.
- Fixed several bugs.

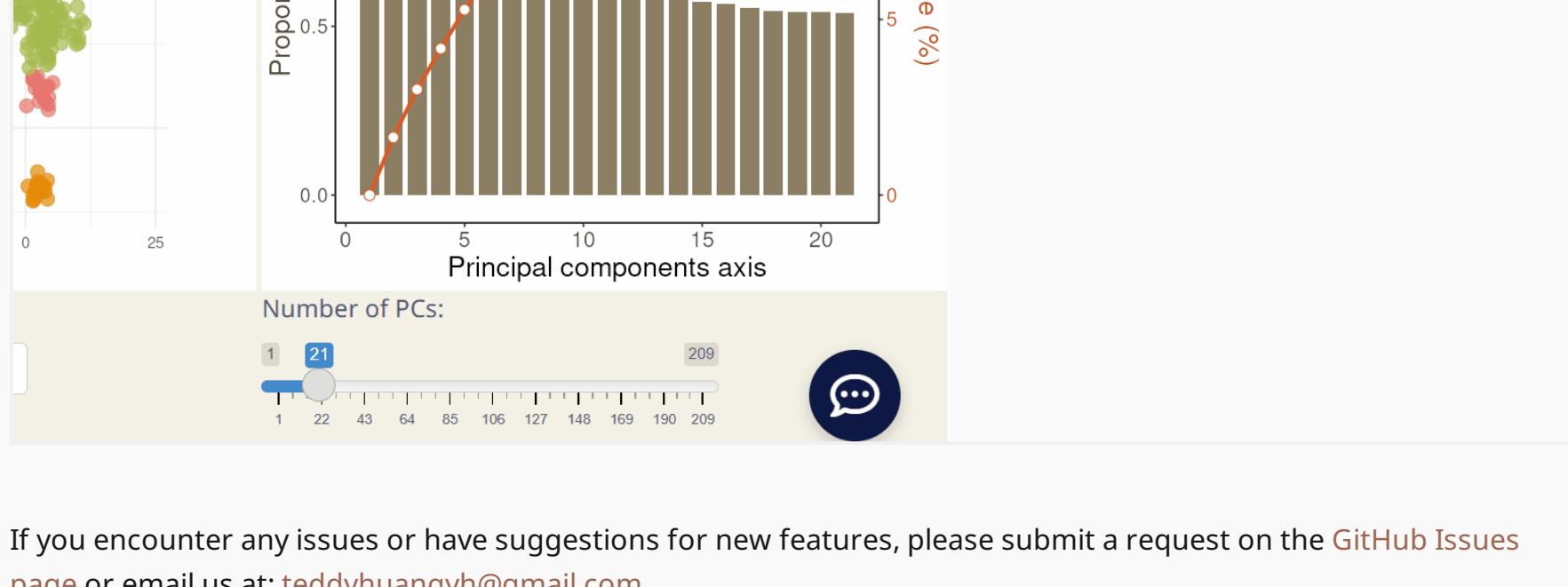
Jun 2025: v1.1.0 on GitHub.

- Enhanced AI report features.
- Enhanced *ShiNyP AI* chatbot.
- Made UI/UX improvements.
- Fixed several bugs.

Aug 2025: v1.2.0.

- Adding new features.

ShiNyP AI Chatbot!



If you encounter any issues or have suggestions for new features, please submit a request on the [GitHub Issues page](#) or email us at: teddyhuangyh@gmail.com

"User Guide for ShiNyP: SNP Analysis and Visualization Platform" was written by Yen-Hsiang Huang, Chung-Feng Kao. It was last built on Jun 2025.

This book was built by the [bookdown](#) R package.

On this page

[Get Started](#)[Run ShiNyP via R](#)[Run ShiNyP via Docker](#)[Main Features](#)[View source](#)[Edit this page](#)

Get Started

There are two easy ways to install and run *ShiNyP*:

1. Using R:

This method is suitable if you already have R installed or prefer working within the R environment. You'll need to install some R packages and then launch *ShiNyP* directly from R environment. ► [R/RStudio/Rtools Tutorial for Beginners! \(#Developing\)](#)

2. Using Docker:

This is the simpler option if you'd rather skip installing R or any packages. With Docker, you can run *ShiNyP* in a ready-to-use setup with just one command. ► [Docker Tutorial for Beginners! \(#Developing\)](#)

Run *ShiNyP* via R

✓ Prerequisites

Before installing *ShiNyP*, ensure your system meets the following requirements:

- **R:** Version \geq 4.4.

Check your current version in R: `getRversion()`

- **Bioconductor:** Version \geq 3.20.

Match your Bioconductor version with your R version (e.g., use Bioconductor 3.21 if R = 4.5).

1 Install Required Package

```
install.packages("BiocManager")
BiocManager::install(version = "3.21") # Use the version that matches your R
BiocManager::install(c("qvalue", "SNPRelate", "ggtree", "snpStats"), force = TRUE)
```

2 Install the *ShiNyP* Package

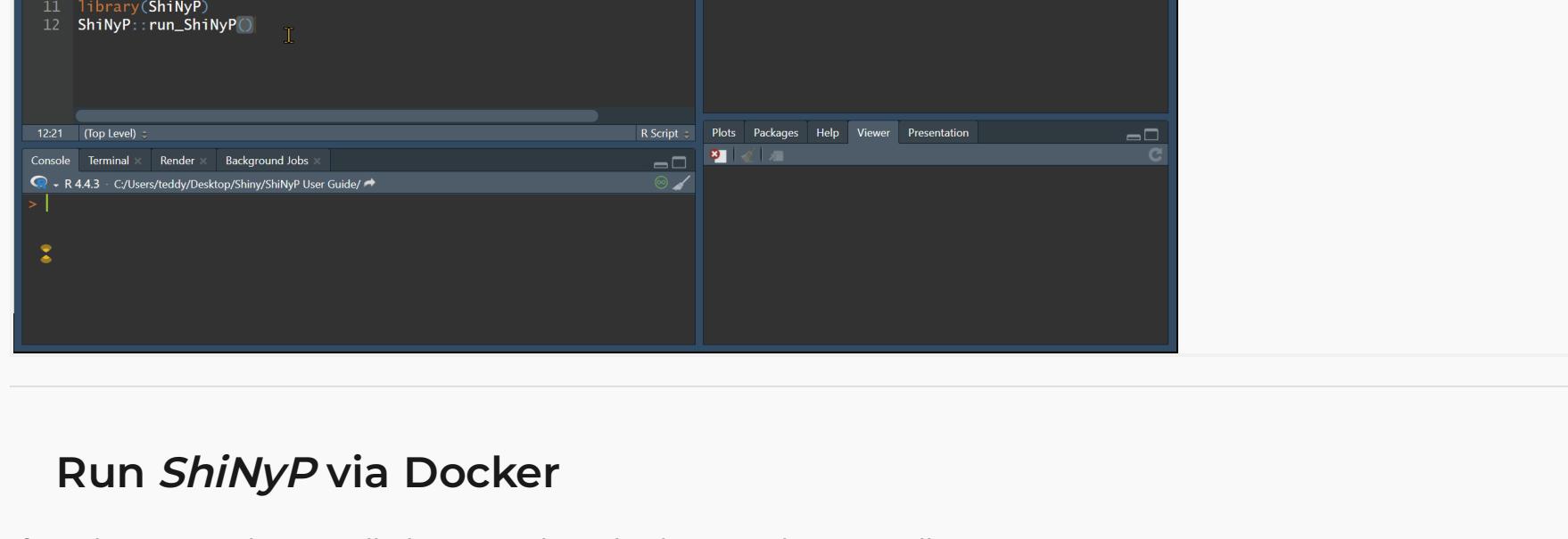
```
install.packages("remotes")
remotes::install_github("TeddyYenn/ShiNyP", force = TRUE)
```

3 Start the *ShiNyP* Platform

```
library(ShiNyP)
ShiNyP::run_ShinyP()
```

4 Run Analysis on *ShiNyP*

Input your SNP dataset in VCF, or try the built-in [Demo Data](#).



Run *ShiNyP* via Docker

If you have Docker installed, you can launch *ShiNyP* without installing R.

✓ Prerequisite

- **Docker.**

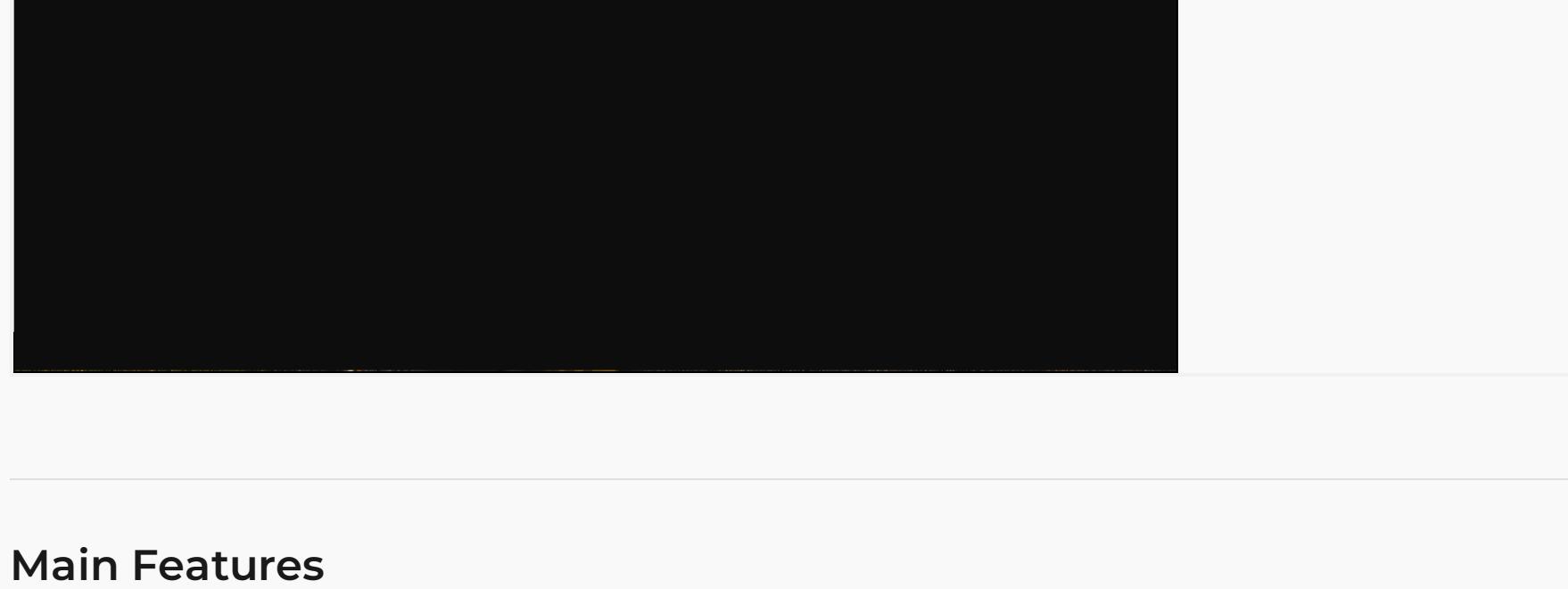
Verify your Docker installation in Terminal: `docker --version`

1 Pull the Docker Image

```
docker run -d -p 3838:3838 teddyenn/shinyp-platform
```

2 Start the *ShiNyP* Platform

Open your browser and visit <http://localhost:3838>



Main Features



Overview of the *ShiNyP* Platform Workflow for SNP Analysis.

► Data Input & Processing:

The workflow begins with Variant Call Format (VCF) [Data Input](#), followed by essential steps such as [Data Quality Control \(QC\)](#) and [Data Transformation](#) to prepare the data for analysis.

► Modular Analysis & Output:

Analytical functions are organized into distinct modules—each accessible as a separate page within the platform. These include: [Population Structure](#), [Genetic Diversity](#), [Selection Sweep](#), and [Core Collection](#). Each module contains multiple subpages offering specialized tools for detailed analysis.

► Customizable Output:

ShiNyP delivers publication-ready visualizations and [AI Report](#) that summarize analytical results in clear, structured narratives. Users can tailor output formats to fit specific research needs.

[« Welcome to ShiNyP](#)

[1 Data Input »](#)

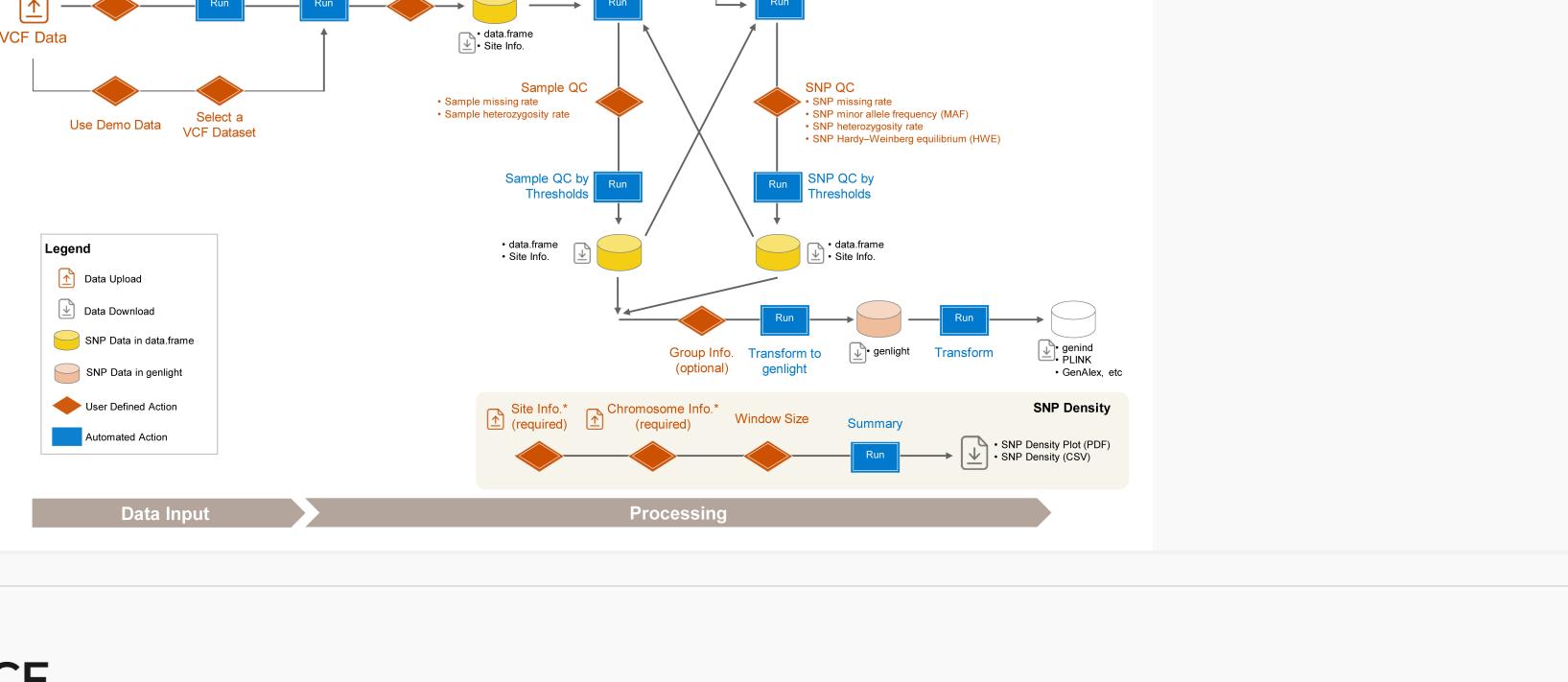
"User Guide for ShiNyP: SNP Analysis and Visualization Platform" was written by Yen-Hsiang Huang, Chung-Feng Kao. It was last built on Jun 2025.

This book was built by the [bookdown R](#) package.

On this page

[1 Data Input](#)[1.1 VCF](#)[1.2 data.frame/genlight](#)[View source](#)[Edit this page](#)

Data Input, QC, and Transform Pages



1.1 VCF

Required File:

ShiNyP accepts VCF files in the following formats:

- VCF file from [PLINK](#)
- VCF or gzipped VCF (vcf.gz) file from [VCFTools](#)
- VCF file in RDS format previously saved by *ShiNyP*

Note: The VCF file should contain chromosome and position information in the first two columns (#CHROM and POS), along with sample names and their genotypic information. For some whole genome sequencing (WGS) data, where SNP marker ID information is missing, *ShiNyP* will auto-generate the SNP ID names as #CHROM:POS, such as 2:12500, indicating chromosome 2, position 12500.

Step 1: Upload VCF File

1. Click Browse to select and upload your VCF file.
2. If your file was generated using VCFTools, make sure to check the "VCF file from VCFTools" option.
3. After the progress bar shows 'Upload complete', click [Input VCF File](#) to proceed.

Or Use Demo Data

1. Click [Use Demo Data](#) and choose a species from the list.
2. For details about the demo datasets, visit: https://github.com/TeddYenn/ShiNyP/tree/main/inst/demo_data.

Note: By default, the interactive table will display genotype data for 5 samples and 10 SNPs as a preview.

Step 2: Transform to data.frame

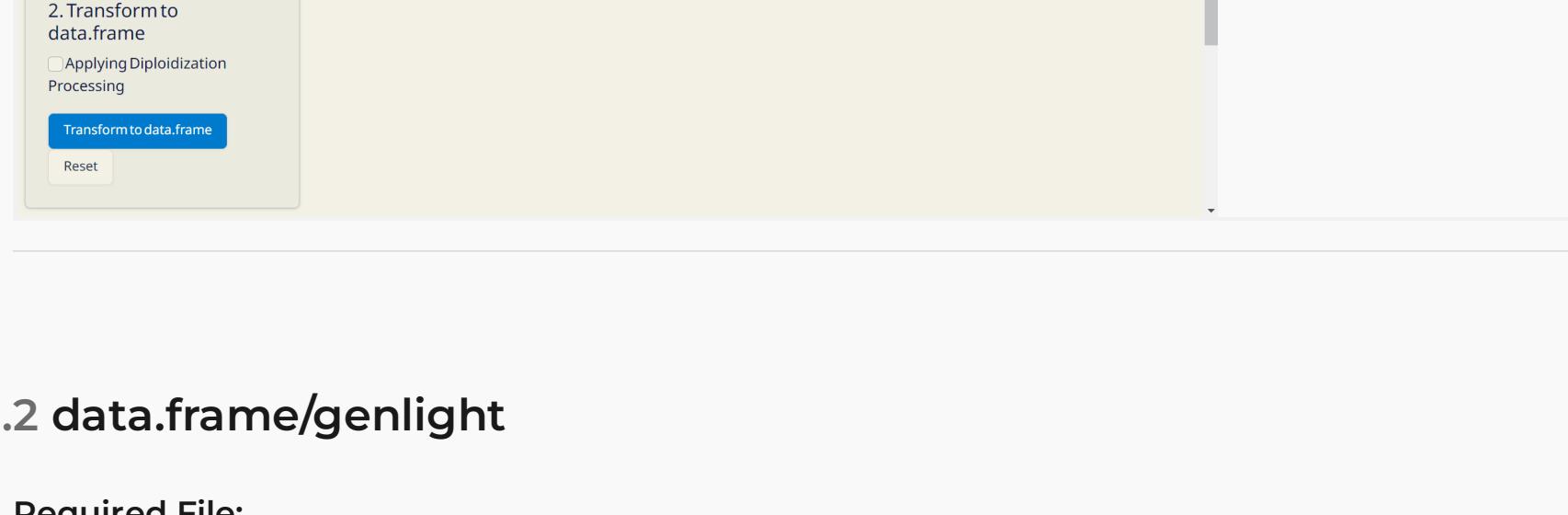
1. After uploading your VCF file, click [Transform to data.frame](#).
2. Download the generated *data.frame* and Site Info (both in RDS format) to skip VCF upload next time by directly importing these files.

Note: *ShiNyP* is optimized for genome-wide SNP analysis in diploid species. For haploid or polyploid data, please the check "Applying diploidization processing" option. *This approach simplifies genotype data and does not account for allelic dosage effects.*

Outputs:

- VCF Data (RDS): Raw VCF data in RDS format, readable in R.
- *data.frame* (RDS): Contains genotypic data — required for downstream analysis.
- Site Info. (RDS): Contains SNP site information — required for downstream analysis.

Note: For large datasets (>1GB), processing may take time. *ShiNyP* handles one task at a time — please wait for each step to complete before proceeding.



1.2 data.frame/genlight

Required File:

- *data.frame* in RDS format
- *genlight* in RDS format
- *genind* in RDS format

Note: *data.frame* available on subpages such as [VCF](#), [Sample QC](#), [SNP QC](#), [Core Sample Set](#), or [Core SNP Set](#). *genlight* and *genind* available on [Data Transform](#) page.

Step 1: Upload File

Browse and select your file, then click [Input](#) to upload.

VCF data.frame/genlight

Input data.frame File

Browse... No file selected

Input Reset

Input genlight File

Browse... No file selected

Input Reset

You can input  data.frame or  genlight files (in RDS) that have already been transformed.

« Get Started

2 Data QC »

"User Guide for ShiNyP: SNP Analysis and Visualization Platform" was written by Yen-Hsiang Huang, Chung-Feng Kao. It was last built on Jun 2025.

This book was built by the [bookdown](#) R package.

On this page

[2 Data QC](#)
[2.1 Sample QC](#)
[2.2 SNP QC](#)
[2.3 SNP Density](#)

2 Data QC

→ This section includes three subpages: [Sample QC](#), [SNP QC](#), and [SNP Density](#), allowing you to assess the quality of samples and SNPs in *data.frame*, as well as visualize SNP density across the genome.

2.1 Sample QC

Required File:

- *data.frame* file from the [Data Input](#) page or
- SNP post-QC *data.frame* file from the [Data QC/SNP QC](#) subpage .

Step 1: Get Summary

First, obtain the sample summary statistics (missing rate and heterozygosity rate) by clicking both [Summary](#) buttons and you will see the results.

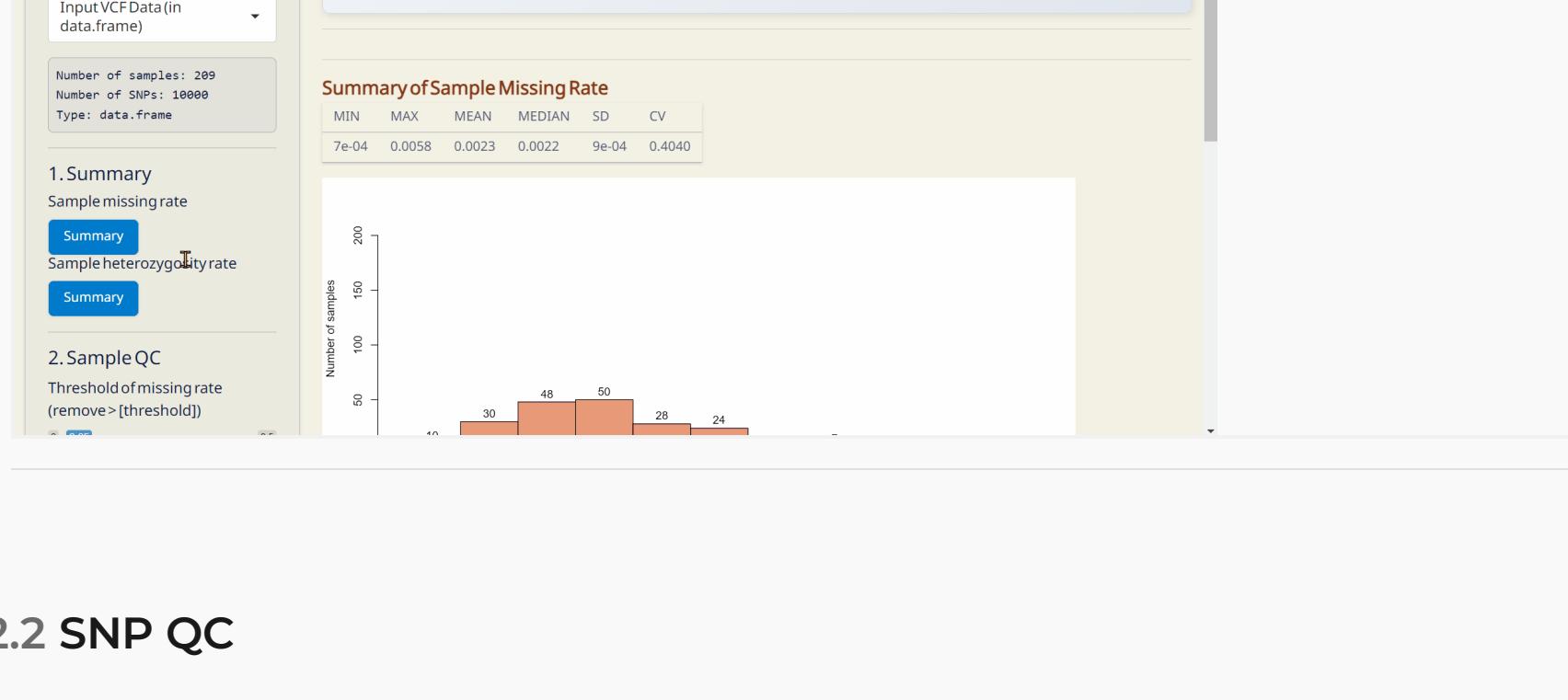
Step 2: Sample QC

Adjust the thresholds and click [Sample QC by Thresholds](#). This will generate the Post-QC *data.frame* file.

Note: If you prefer not to perform sample QC by sample missing rate or heterozygosity rate, please set the threshold to 1.

Outputs:

- *data.frame* (RDS): Updated *data.frame* file — required for downstream analysis.
- Site Info. (RDS): Updated SNP site information file — required for downstream analysis.



2.2 SNP QC

Required File:

- *data.frame* file from the [Data Input](#) page or
- Sample post-QC *data.frame* file from the [Data QC/Sample QC](#) subpage.

Step 1: Get Summary

First, obtain the SNP summary statistics [missing rate, minor allele frequency (MAF), heterozygosity rate, and Hardy-Weinberg equilibrium (HWE)] by clicking all [Summary](#) buttons and you will see the results.

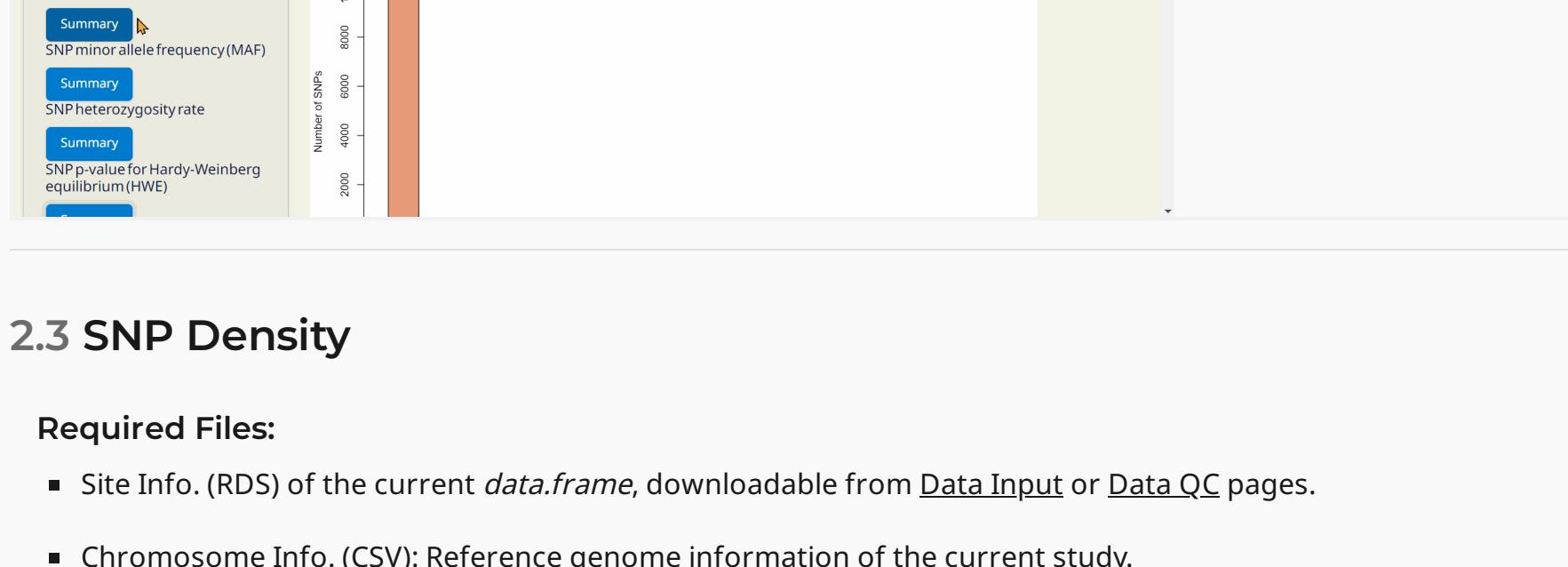
Step 2: Sample QC

Adjust the thresholds and click [SNP QC by Thresholds](#). This will generate the Post-QC *data.frame* file.

Note: If you prefer not to perform QC based on SNP missing rate or heterozygosity rate, set the missing rate threshold to 1, the MAF to 0, and the heterozygosity rate to 0 and 1. Also, leave the 'Do SNP QC by HWE' checkbox unticked to skip QC based on SNP HWE.

Outputs:

- *data.frame* (RDS): Updated *data.frame* file — required for downstream analysis.
- Site Info. (RDS): Updated SNP site information file — required for downstream analysis.



2.3 SNP Density

Required Files:

- Site Info. (RDS) of the current *data.frame*, downloadable from [Data Input](#) or [Data QC](#) pages.
- Chromosome Info. (CSV): Reference genome information of the current study.

[Download an example of Chromosome Info. \(CSV\).](#)

This file should contain three columns: "Chr", "Start", and "End".

- "Chr" column should specify the chromosome names (as characters, e.g., "Chr01", "Chr11")

- "End" column should indicate the length of each chromosome (numeric)

- "Start" column can be set to 0 or 1 for each chromosome.

Steps:

1. Upload Site Info. (RDS) and Chromosome Info. (CSV).

2. Choose a window size in kilobases (kb).

3. Click **Summary**.

Outputs:

- SNP Density Plot (PDF): An ideogram visualizing SNP density across the genome within a defined window size. A gradient color palette is used to represent varying SNP densities: *green* for lower densities, *yellow* for medium densities, and *red* for higher densities, with *grey* indicating regions with zero SNP.
- SNP Density (CSV): A table detailing SNP density across each chromosome. *bp_over_SNPs*: The total base pairs (bp) per SNP in each window, representing the average spacing between SNPs. *SNPs_over_1000bp*: The number of SNPs per 1,000 base pairs, providing a normalized measure of SNP density across the genome.

The screenshot shows the ShiNyP web application interface. At the top, there's a navigation bar with links like Home, Data Input, Data QC, Data Transform, Population Structure, Genetic Diversity, Selection Sweep, Core Collection, and AI Report. Below the navigation bar, the main content area has tabs for Sample QC, SNP QC, and SNP Density. The SNP Density tab is active. On the left, there are two file upload fields: 'Site Info.*(required)' and 'Chromosome Info.*(required)', both currently showing 'No file selected'. Below these is a 'Window size (kb)' slider with a scale from 0 to 1,000, with a blue marker set at 500. At the bottom of the form are two buttons: 'Summary' (highlighted in blue) and 'Reset'. A message box in the center of the page says: 'Need to upload the Site Info file (in RDS format) and Chromosome Info file (in CSV format). Please select the optimal window size and step, then click the "Summary" button.'

« 1 Data Input

3 Data Transform »

"User Guide for ShiNyP: SNP Analysis and Visualization Platform" was written by Yen-Hsiang Huang, Chung-Feng Kao. It was last built on Jun 2025.

This book was built by the [bookdown](#) R package.

3 Data Transform

→ This section allow you to convert your SNP data in *data.frame* into multiple formats, including *genind*, *genind* and *PLINK*.

Required File:

- Input VCF Data (*data.frame* file) from the [Data Input](#) page or
- Post-QC Data (*data.frame* file) from the [Data QC](#) page.

Step 1: Transform *data.frame* to *genlight*

1. Click [Transform to genlight](#).
2. Download the generated *genlight* (in RDS format) to skip VCF or *data.frame* upload next time by directly importing this file.

Output:

- *genlight* (RDS): *genlight* file — required for downstream analysis.

Step 2: Transform *genlight* to others

Select the desired data format to export from *genlight* and click [Transform](#).

Outputs:

- *genlight* (RDS): *genlight* file with Group Info. — required for downstream analysis.
- *genind* (RDS): One of the input format for *ShiNyP DAPC* subpage, optimized for DAPC analysis to reduce computation time.

[Download an example of Group Info. \(CSV\).](#)

This file should have at least two columns: "ID" and "Group".

Note: You can obtain a template after completing a DAPC analysis (Section 4.2) first to generate the initial file ("DAPC_Group_Info.csv"). If your SNP data is large, you can create and use a core SNP set as input to obtain DAPC results more efficiently. You can then expand this file based on your own metadata (e.g., origin, type).

The following transformed files will be generated at the specified path you provide.

- PLINK (PED & MAP): Input format for [PLINK](#) program, designed to perform a range of basic and large-scale SNP analyses.
- GenAIEx (CSV): Input format for [GenAIEx](#) program, offers a wide range of population genetic analysis in Excel.
- LEA (GENO & LFMM): Input format for [LEA](#) R package, designed for population genomics, landscape genomics and genotype-environment association tests.
- GDS (GDS): Input format for [SNPRelate](#) R package, designed for efficient SNP data analysis.
- STRUCTURE (STR): Input format for [STRUCTURE](#) program, used for inferring population structure.
- fastStructure (STR): Input format for [fastStructure](#) program, used for inferring population structure from large SNP data.
- PHYLIP (TXT): Input format for [PHYLIP](#) program, used for phylogenetic tree reconstruction and evolutionary analysis.
- Treemix (GZ): Input format for [Treemix](#) program, designed for modeling population splits and migration events.
- BayeScan (TXT): Input format for [BayeScan](#), used for detecting loci under selection.

The screenshot shows the ShiNyP Data Transform interface. On the left, there's a sidebar with navigation links: Home, Data Input, Data QC, Data Transform (which is active), Population Structure, Genetic Diversity, Selection Sweep, Core Collection, and AI Report. The main area has two main sections: '1. data.frame to genlight' and '2. genlight to ...'. In '1. data.frame to genlight', the user has selected 'Input VCFData (in data.frame)' and 'data.frame'. The output is a 'genlight' file with 209 samples and 10000 SNPs. A 'Download genlightFile' button is available. In '2. genlight to ...', the user has selected 'genind (RDS)' as the output type. The output is a 'genind' file with 209 samples and 10000 SNPs. A 'Download genindFile' button is available. Both sections show a summary of the transformation parameters and the resulting file details.

« 2 Data QC

4 Population Structure »



4 Population Structure

→ This section includes seven subpages: PCA, DAPC, UPGMA Tree, NJ Tree, Kinship, Scatter Plot^{Plus}, and Tree Plot^{Plus}, allowing you to conduct various population structure analyses and customize your plot.

On this page

4 Population Structure

4.1 PCA (Principal Component Analysis)

4.2 DAPC (Discriminant Analysis of Principal Components)

4.3 UPGMA (Unweighted Pair Group Method with Arithmetic mean) Tree

4.4 NJ (Neighbor-Joining) Tree

4.5 Kinship Analysis

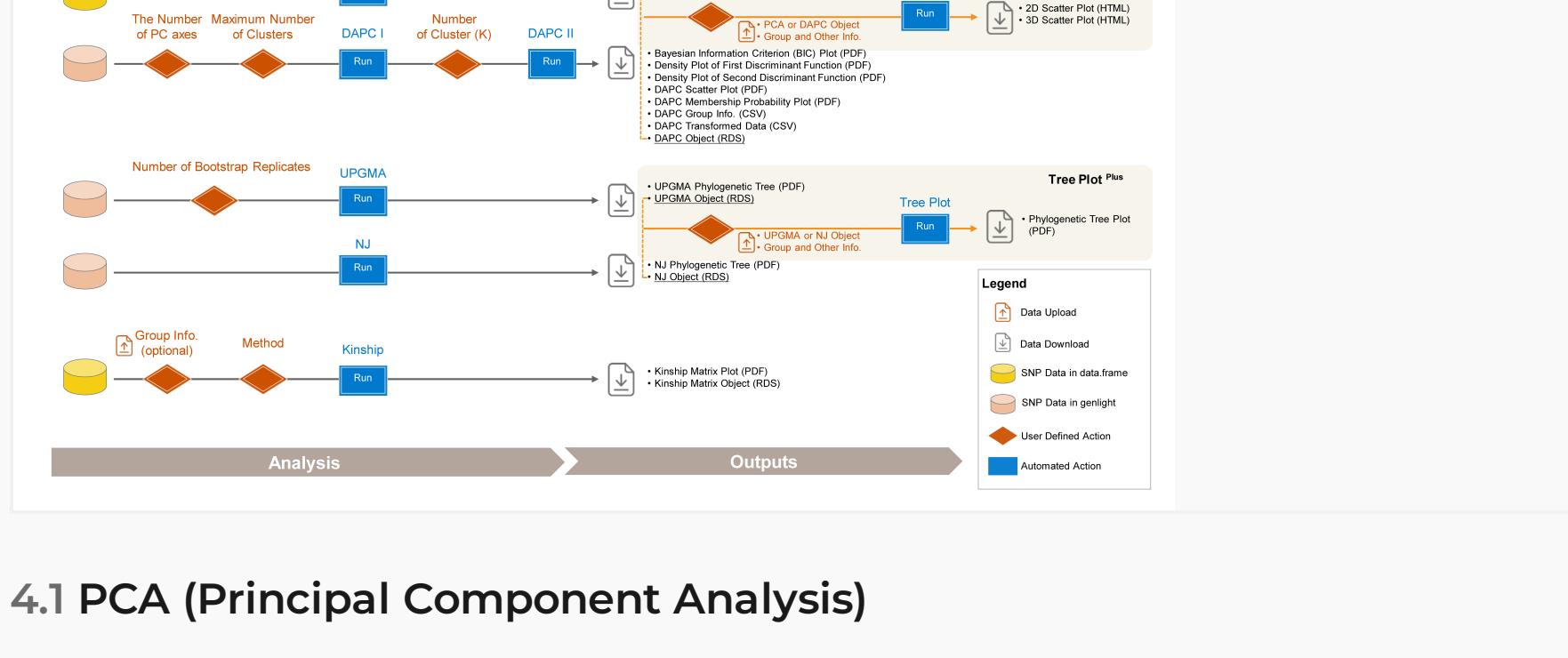
4.6 Scatter Plot Plus

4.7 Tree Plot Plus

View source

Edit this page

Population Structure Page



4.1 PCA (Principal Component Analysis)

A widely used method to uncover underlying population structure by reducing the dimensionality of genetic data.

Required File:

- *data.frame*

One Step:

1. Click **Run PCA** to generate PCA plots and the following downloadable files.

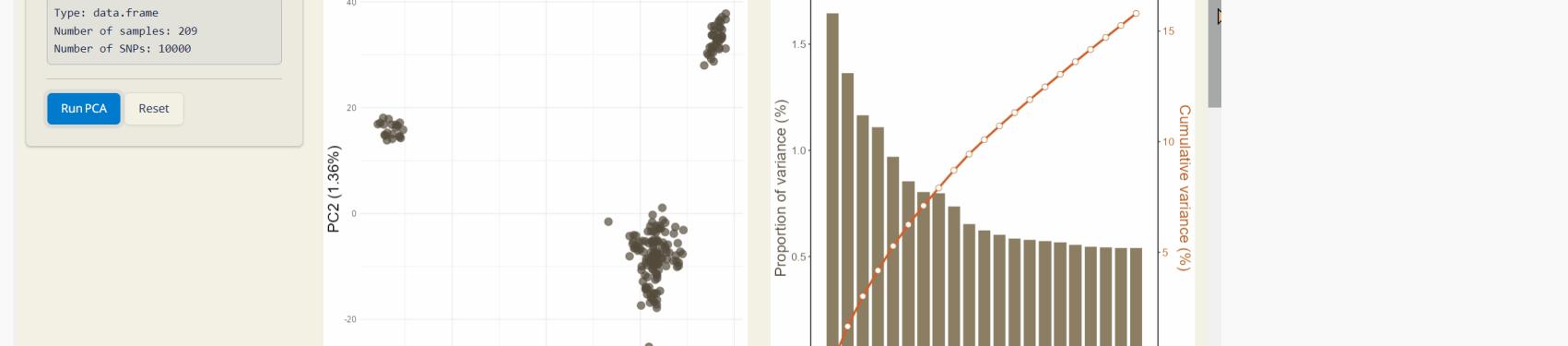
Outputs:

- PCA Scatter Plot (PDF): A scatter plot showing the distribution of samples based on principal components, with each dot representing an individual.

Note: You can upload the Group Info. (from [Population Structure/DAPC](#)) or Core Sample Info. (from [Core Collection/Core Sample Set](#)) to classify individuals and color them in the PCA Scatter Plot.

[Download an example of Group Info. \(CSV\).](#)

- PC Explained Variance Plot (PDF): Visualizes the variance explained by each principal component.
- Explained Variance (CSV): Contains the explained variance of each principal component.
- PCA Transformed Data (CSV): Dataset transformed into principal components, with samples as rows and principal components as columns.
- PCA Object (RDS): Contains all PCA results for future use and reproducibility, and can be used as input data in the [Population Structure/Scatter Plot^{Plus}](#) subpage.



4.2 DAPC (Discriminant Analysis of Principal Components)

A multivariate method for identifying and visualizing genetic clusters by combining PCA and Linear Discriminant Analysis (LDA) (Jombart, Devillard, and Balloux 2010). For more information, visit <https://adegenet.r-forge.r-project.org/files/tutorial-dapc.pdf>.

Required File:

- *genlight* or
- *genind* (faster)

Step 1: Cluster Identification

1. Click **Run DAPC I** to determine the optimal number of clusters (the lowest BIC value indicates the optimal number of clusters).

Note: The default number of PC axes for cluster identification is set to retain PCs that capture up to 80% of the total variance. You can refer the "PC Explained Variance Plot" in the [Population Structure/PCA](#) subpage.

Step 2: DAPC Analysis

1. Choose the number of cluster (K) based on the "Bayesian Information Criterion (BIC) Plot".

2. Click **Run DAPC II** to generate DAPC plots and the following downloadable files.

Note: You can download the "DAPC Object" and upload it on [Population Structure/Scatter Plot^{Plus}](#) subpage to customize your 2D and 3D scatter plots.

Outputs:

- Bayesian Information Criterion (BIC) Plot (PDF): Visual representation of the BIC for model selection.

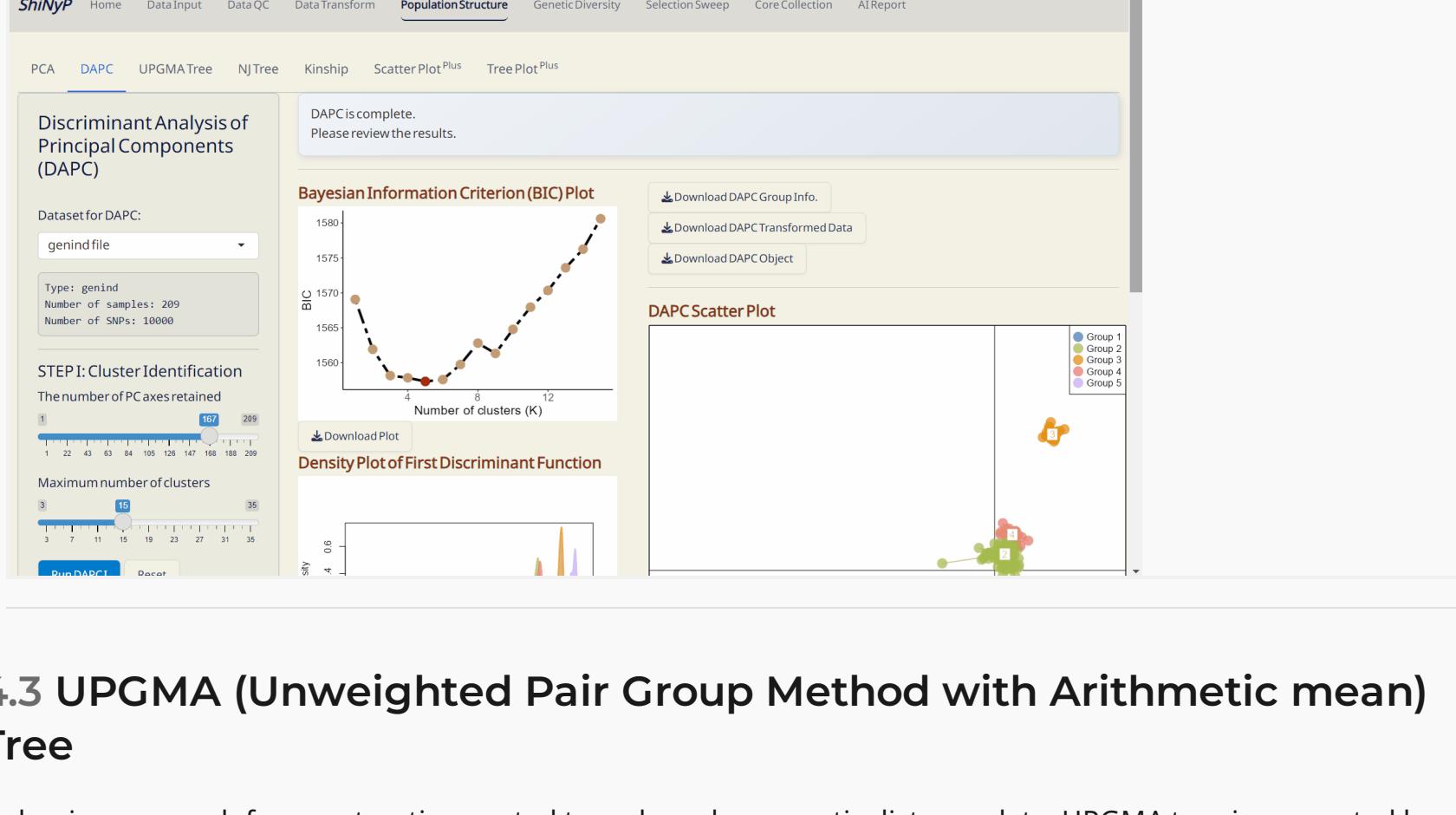
- Density Plot of First & Second Discriminant Function (PDF): Displays the density of the first and second discriminant functions, with each row bar representing an individual.

- DAPC Scatter Plot (PDF): A scatter plot showing the distribution of samples based on discriminant functions (x-axis: first discriminant function; y-axis: second discriminant function), with each dot representing an individual.

- DAPC Membership Probability Plot (PDF): Visualizes membership probabilities of individuals in different groups, with each row bar representing an individual.

- DAPC Group Info. (CSV): Contains the group assignments for each individual based on DAPC. This file used in various subpages.

- DAPC Transformed Data (CSV): Dataset transformed into discriminant functions with samples as rows and discriminant functions as columns.
- DAPC Object (RDS): Contains all results from the DAPC analysis for future reproducibility. It can be used as input data in the [Population Structure/Scatter Plot^{Plus}](#) and [Core Collection/Core SNP Set](#) subpages.



4.3 UPGMA (Unweighted Pair Group Method with Arithmetic mean) Tree

A classic approach for constructing rooted trees based on genetic distance data. UPGMA tree is generated by *poppr* and *ggtree* packages ([Yu et al. 2016](#); [Kamvar, Tabima, and Grünwald 2014](#)).

Required File:

- genlight*

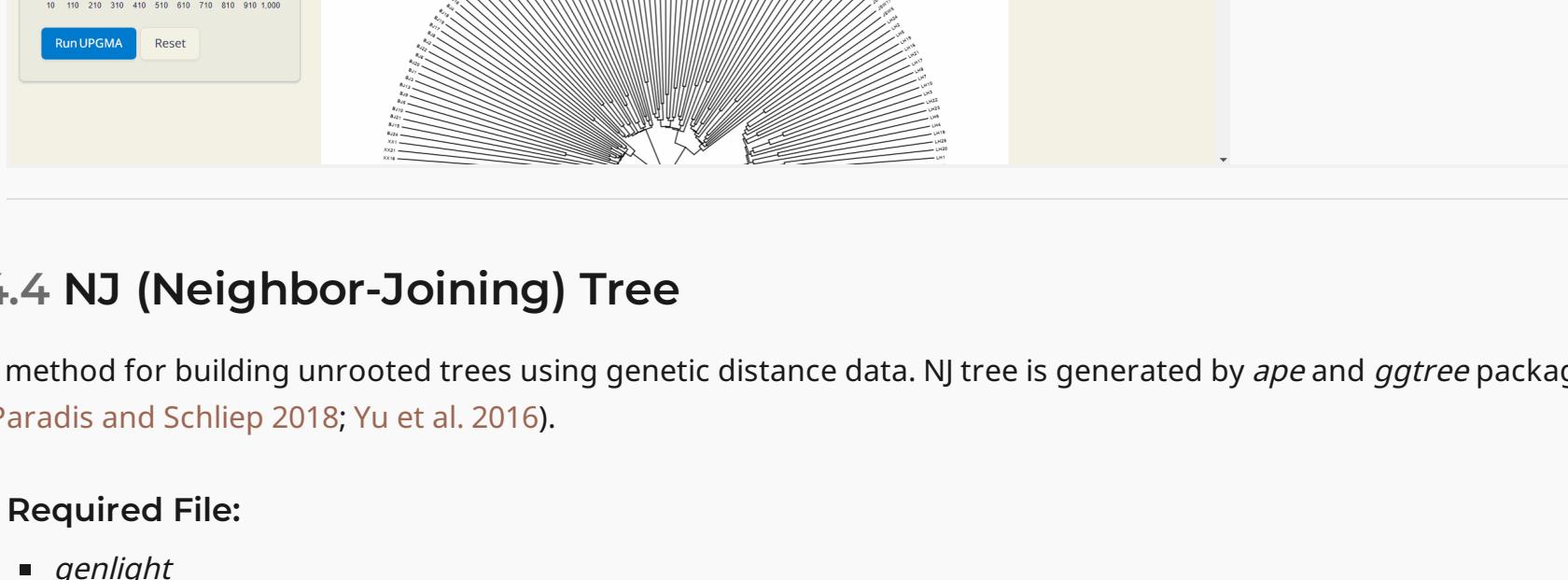
Steps:

- Choose the number of bootstrap replicates, which will be used for assessing the confidence of the branching structure.
- Click **Run UPGMA** to generate tree plot.

Note: You can download the "UPGMA Object" and upload it on [Population Structure/Tree Plot^{Plus}](#) subpage to customize your phylogenetic tree.

Outputs:

- UPGMA Phylogenetic Tree (PDF): A UPGMA rooted tree with a user-defined layout style.
- UPGMA Object (RDS): Contains all information of the UPGMA tree, and can be used as input data in the [Population Structure/Tree Plot^{Plus}](#) subpage.



4.4 NJ (Neighbor-Joining) Tree

A method for building unrooted trees using genetic distance data. NJ tree is generated by *ape* and *ggtree* packages ([Paradis and Schliep 2018](#); [Yu et al. 2016](#)).

Required File:

- genlight*

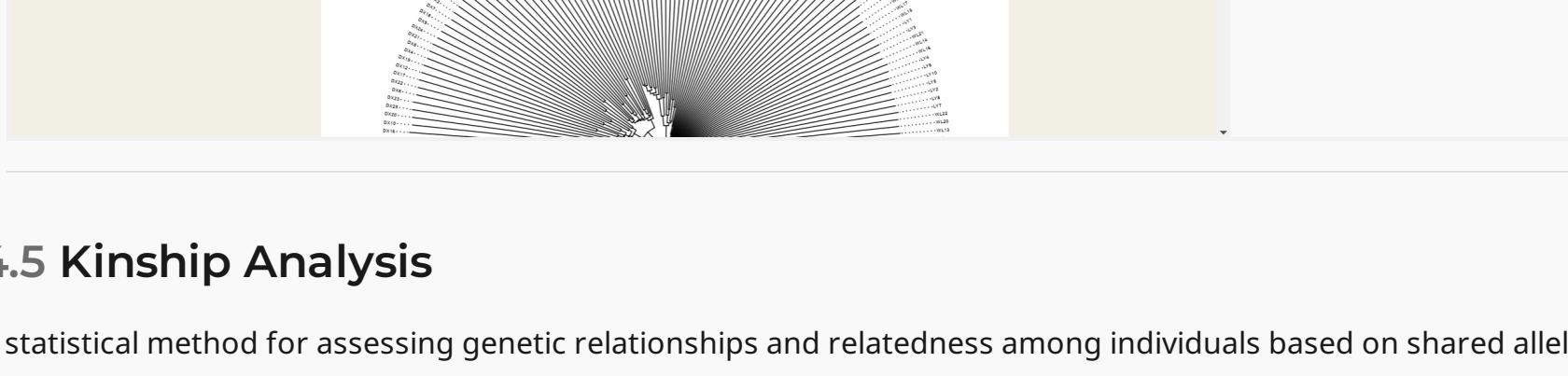
One Step:

- Click **Run NJ** to generate tree plot.

Note: You can download the "NJ Object" and upload it on [Population Structure/Tree Plot^{Plus}](#) subpage to customize your phylogenetic tree.

Outputs:

- NJ Phylogenetic Tree (PDF): A NJ unrooted tree with a user-defined layout style.
- NJ Object (RDS): Contains all information of the NJ tree, and can be used as input data in the [Population Structure/Tree Plot^{Plus}](#) subpage.



4.5 Kinship Analysis

A statistical method for assessing genetic relationships and relatedness among individuals based on shared alleles ([Kang et al. 2010](#)). Kinship matrix is generated by *statgenGWAS* package. For more information, visit <https://rdrr.io/cran/statgenGWAS/man/kinship.html>.

Required File:

- data.frame*

Steps:

- Upload Group Info. from [Population Structure/DAPC](#) subpage (optional). If uploaded, the order of samples will follow the group assignment; otherwise, it will follow the order of the original VCF data.

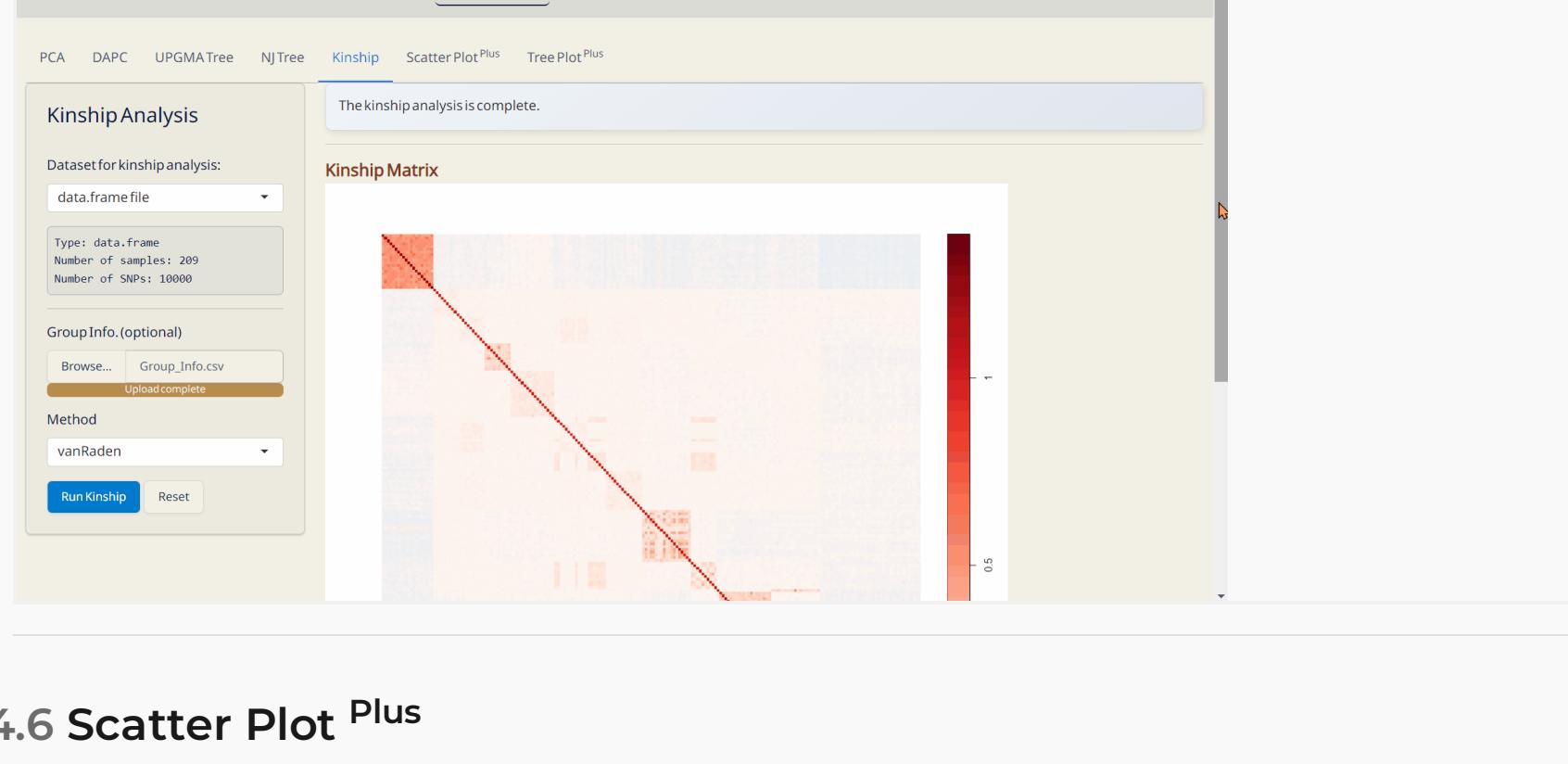
2. Choose a method to run kinship analysis.

3. Click the **Run Kinship** button to generate the kinship matrix.

Outputs:

- Kinship Matrix Plot (PDF): A visual representation of the kinship matrix.
- Kinship Matrix (RDS): Contains the kinship matrix data.

Note: This kinship matrix can be directly used as input for *GAPIT* package in genome-wide association studies (GWAS), helping to control for confounding effects.



4.6 Scatter Plot Plus

Customize your scatter plot based on the results from Population Structure/PCA or Population Structure/DAPC.

Required Files:

- PCA Object (PCA_prcmp_Object.rds file) or DAPC Object (DAPC_dapc_Object.rds file)
- Group and Other Info. (i.e. metadata, modifiable from DAPC_Group_Info.csv)

[Download an example of Group Info. \(CSV\).](#)

Note: If you are using a DAPC object as input, ensure that there are at least 4 groups. Since N groups have N-1 LD axes, you need more than 3 axes to generate a 3D scatter plot.

Note: You can add metadata (more information about samples) by adding new variables to the Group Info. file. Ensure that the sample order remains unchanged.

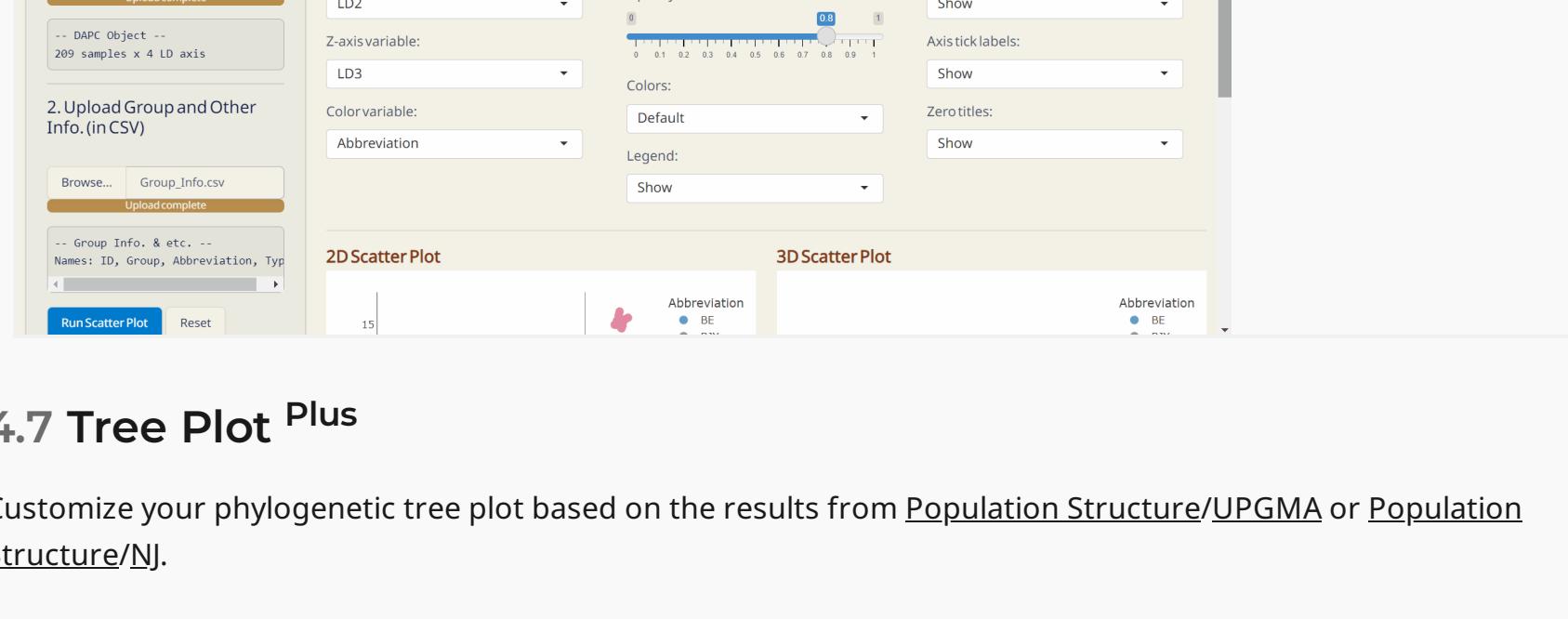
Steps:

1. Upload PCA or DAPC Object (RDS).
2. Upload Group and Other Info. (CSV).
3. Click **Run Scatter Plot** to generate the 2D and 3D interactive scatter plots.
4. Customize the scatter plot and click **Run Scatter Plot** again.

Note: The scatter plots are downloaded as HTML files and can be opened with browsers like Chrome or Edge.

Outputs:

- 2D Scatter Plot (HTML): Two-dimensional interactive scatter plot with user-defined attributes.
- 3D Scatter Plot (HTML): Three-dimensional interactive scatter plot with user-defined attributes.



4.7 Tree Plot Plus

Customize your phylogenetic tree plot based on the results from Population Structure/UPGMA or Population Structure/NJ.

Required Files:

- UPGMA Object (UPGMA_phylo_Object.rds file) or NJ Object (NJ_phylo_Object.rds file)
- Group and Other Info. (modifiable from DAPC_Group_Info.csv)

[Download an example of Group Info. \(CSV\).](#)

Note: You can add more information about samples by adding new variables to the Group Info. file. Ensure that the sample order remains unchanged.

Steps:

1. Upload UPGMA or NJ Object (RDS).
2. Upload Group and Other Info. (CSV).
3. Click **Run Tree Plot** to generate the tree plot.
4. Customize the tree plot and click **Run Tree Plot** again.

Outputs:

- Phylogenetic Tree Plot (PDF): A phylogenetic tree plot with user-defined layout style and attributes.

http://127.0.0.1:3878 | Open in Browser | Republish

ShiNyP Home Data Input Data QC Data Transform Population Structure Genetic Diversity Selection Sweep Core Collection AI Report

PCA DAPC UPGMA Tree NJ Tree Kinship Scatter Plot Plus Tree Plot Plus

You can customize the tree plot and then click the Run Tree Plot button again.

Tree Plot Plus

1. Upload UPGMA or NJ Object (in RDS)

Browse... NL_phyl0_Object.rds Upload complete
-- NJ Tree Object -- 209 samples

2. Upload Group and Other Info.(in CSV)

Browse... Group_Info.csv Upload complete
-- Group Info. & etc. -- Names: ID, Group, Abbreviation, Type

Run Tree Plot Reset

Phylogenetic Tree Plot

Layout: Taxa label: Symbol: Treescale:
Circular Show Show NULL

Line color variable: Text color variable: Symbol color variable: Bootstrap values:
Type Group Group NULL

Line colors: Text colors: Symbol colors: Legend:
Default Grey-single color Default Top

Line size: Text size: Symbol shape variable:
0 0.5 1 1.5 2 2.5 3 3.5 4 4.5 5 Group

Align label: Symbols size:
TRUE 0 0.5 1 1.5 2 2.5 3 3.5 4 4.5 5

« 3 Data Transform

5 Genetic Diversity »

"User Guide for ShiNyP: SNP Analysis and Visualization Platform" was written by Yen-Hsiang Huang, Chung-Feng Kao. It was last built on Jun 2025.

This book was built by the [bookdown](#) R package.

5.3 Genetic Distance

Pairwise genetic distance between populations is computed using *hierfstat* package. For more information, visit <https://rdrr.io/cran/hierfstat/man/genet.dist.html>.

Required File:

- *data.frame*
- Group Info. from [Population Structure/DAPC](#) subpage (required).

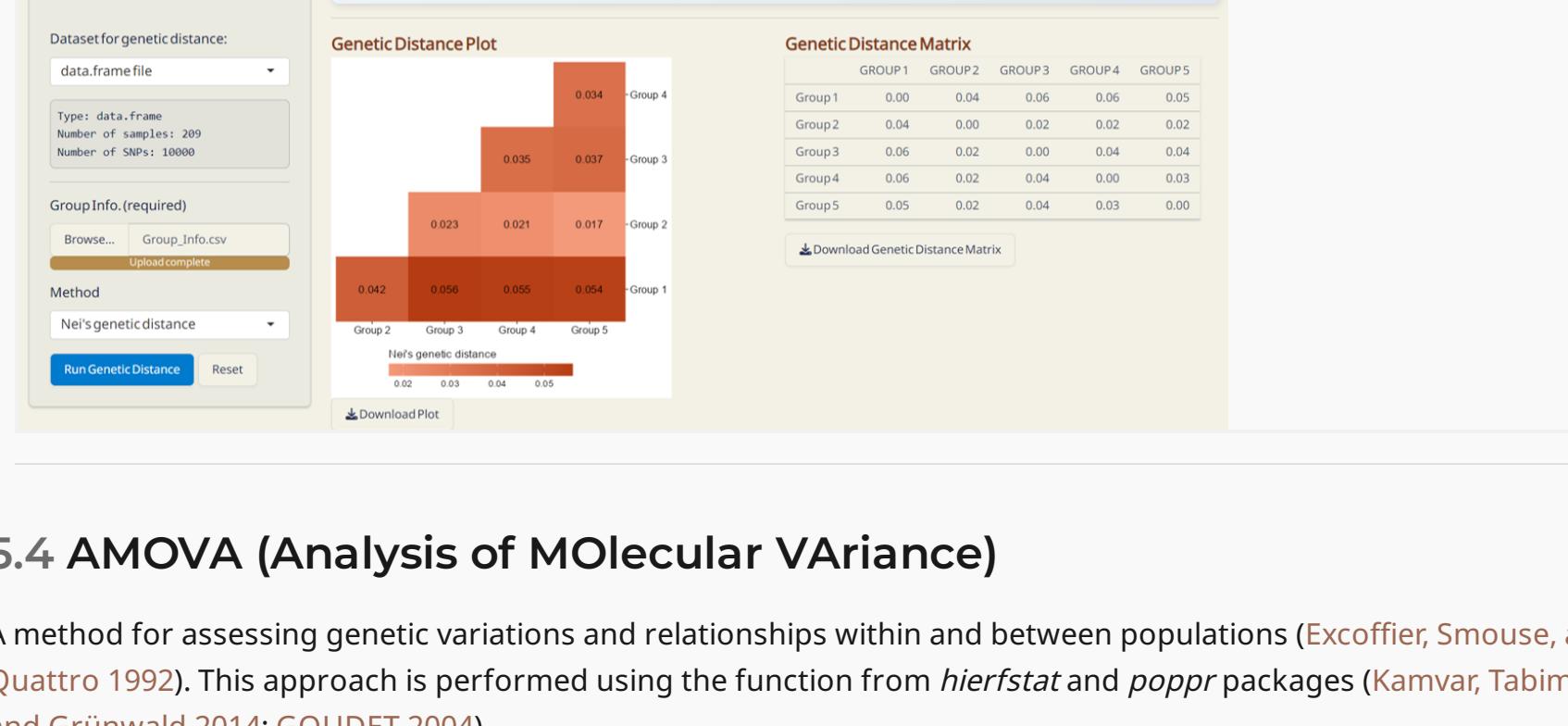
[Download an example of Group Info. \(CSV\).](#)

Steps:

1. Upload Group Info.
2. Select a method.
3. Click **Run Genetic Distance** to generate the pairwise genetic distance.

Outputs:

- Genetic Distance Plot (PDF): A plot of the pairwise genetic distance matrix based on the user-selected method.
- Genetic Distance Matrix (CSV): A pairwise genetic distance matrix based on the user-selected method.



5.4 AMOVA (Analysis of MOlecular VAriance)

A method for assessing genetic variations and relationships within and between populations ([Excoffier, Smouse, and Quattro 1992](#)). This approach is performed using the function from *hierfstat* and *poppr* packages ([Kamvar, Tabima, and Grünwald 2014](#); [GOUDET 2004](#)).

Required File:

- *genlight* with 'Group Info.', downloadable from [Data Transform](#) page after you have Group Info.

Step 1: Run AMOVA

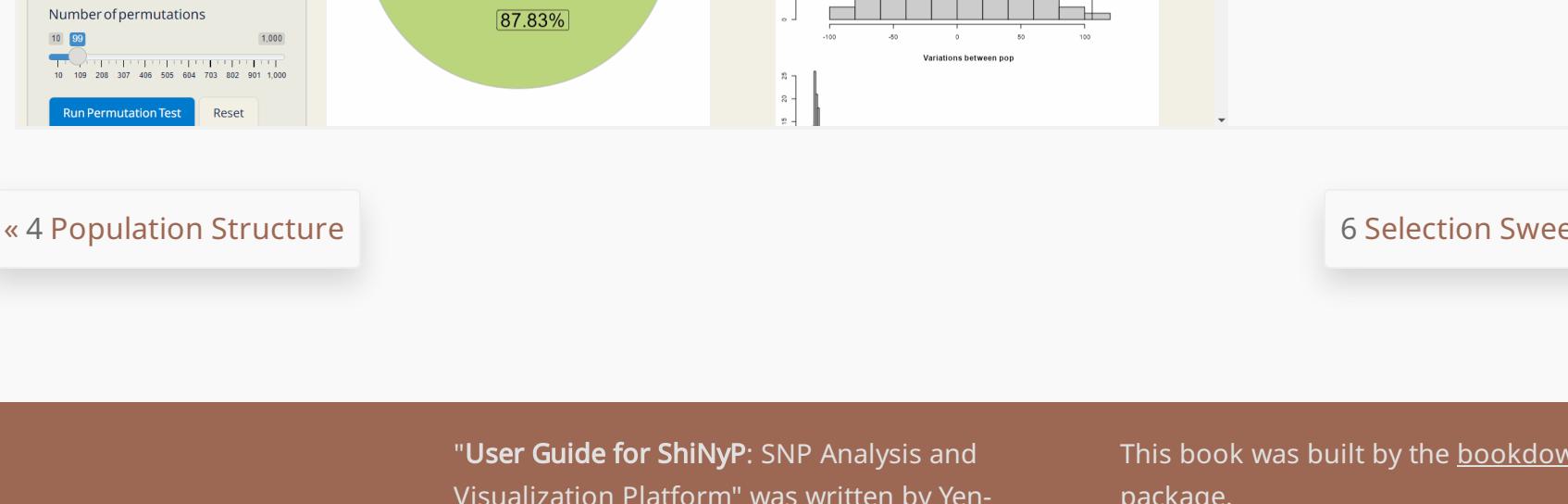
1. Click **Run AMOVA** to partition genetic variation among and within populations.

Step 2: Run Permutation Test

1. Choose the number of randomizations for the permutation test to detect the significance of three hierarchical levels. We recommend using 9, 99 (default), 199, 499, 799, or 999 permutations for more classical p-values.
2. Click **Run Permutation Test** to perform the statistical test.

Outputs:

- AMOVA Variance Plot (PDF): A pie chart showing the explained genetic variance of population strata among defined groups.
- AMOVA Variance Test (PDF): A plot showing the significance test of population strata among defined groups. The histograms depict randomized strata distributions, with the black line representing genetic variance components.
- AMOVA Table (CSV): A table with detailed AMOVA results.



[« 4 Population Structure](#)

[6 Selection Sweep »](#)

"User Guide for ShiNyP: SNP Analysis and Visualization Platform" was written by Yen-Hsiang Huang, Chung-Feng Kao. It was last built on Jun 2025.

This book was built by the [bookdown R](#) package.



On this page

6 Selection Sweep

6.1 pcadapt

6.2 OutFLANK

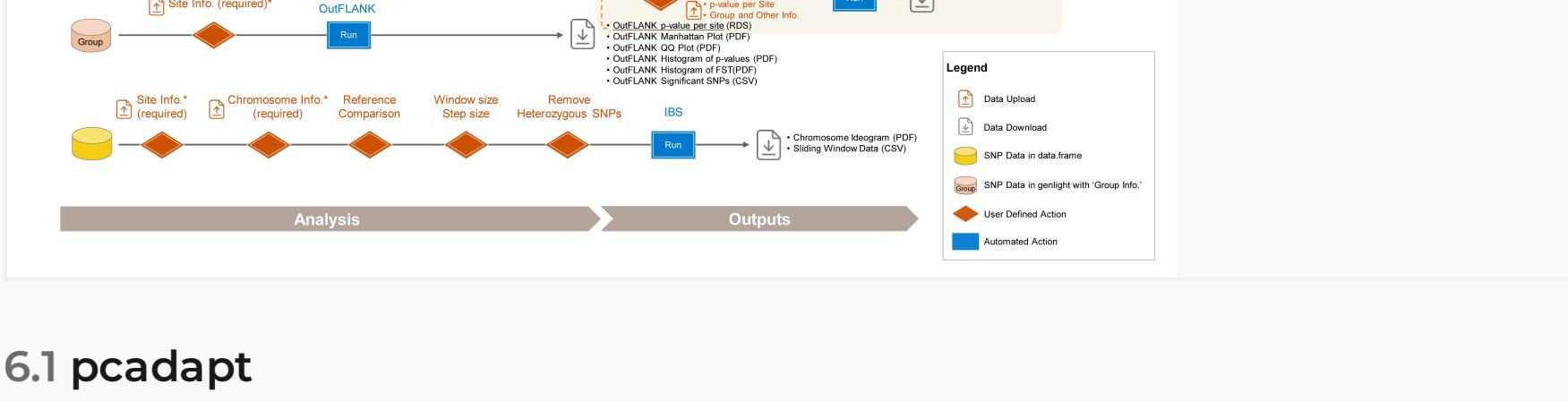
6.3 IBS (Identity By State)

6.4 Manhattan Plot Plus

View source

Edit this page

Selection Sweep Page



6.1 pcadapt

A PCA-based approach identifies selective outliers relative to population structure (Luu, Bazin, and Blum 2016).

Required Files:

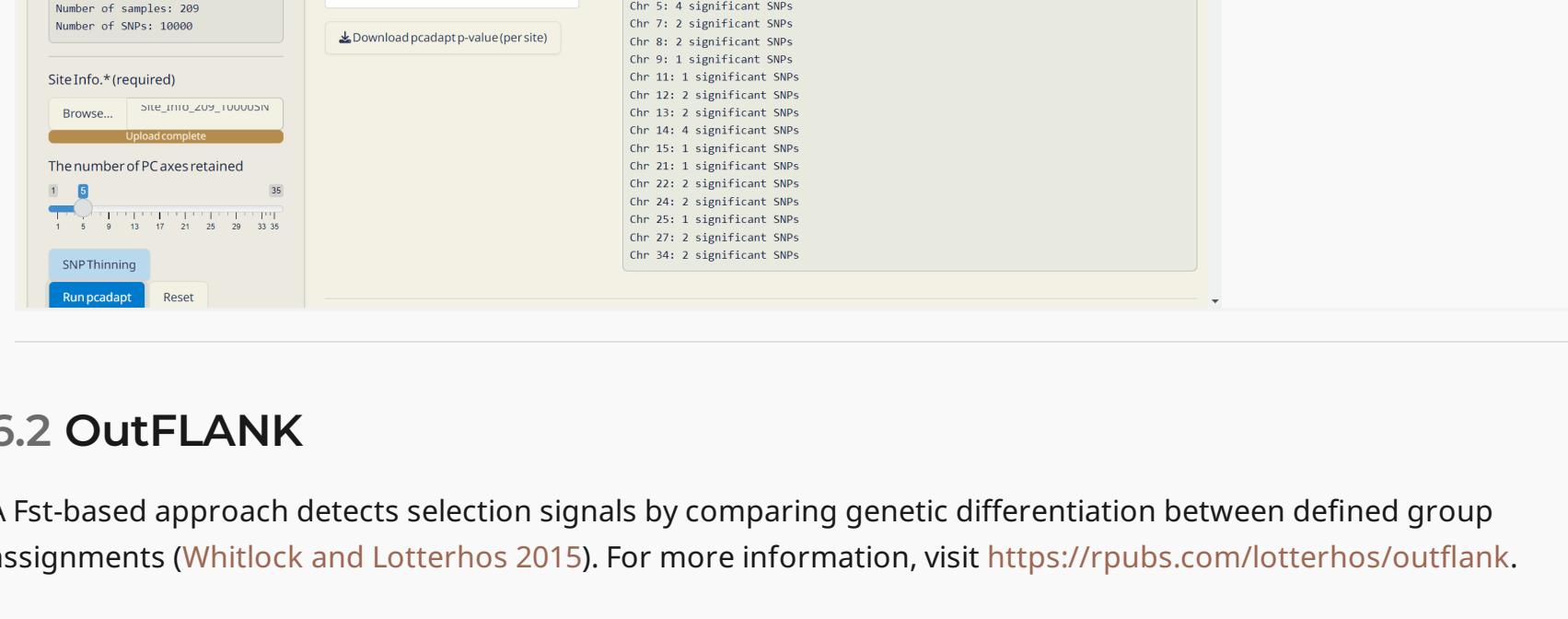
- *data.frame*
- Site Info. (RDS) of the current *data.frame*, downloadable from [Data Input](#) or [Data QC](#) pages.

Steps:

1. Upload Site Info. (required).
2. Click **SNP Thinning** (optional) and choose window size (number of SNPs) and r^2 threshold. For more information, visit <https://bcm-uga.github.io/pcadapt/articles/pcadapt.html>.
3. Click **Run pcadapt** to perform genome scan for selection.

Outputs:

- pcadapt p-value per site (RDS): A dataset containing p-values and adjusted p-values for each site.
- pcadapt Manhattan Plot (PDF): A Manhattan plot visualizing the p-values per site across the genome. Significant SNPs are highlighted in red.
- pcadapt QQ Plot (PDF): A QQ plot comparing the distribution of observed p-values to the expected distribution under the null hypothesis.
- pcadapt Histogram of p-values (PDF): A histogram showing the distribution of p-values across all sites.
- pcadapt Histogram of Test Statistics (PDF): A histogram showing the distribution of test statistics across all sites.
- pcadapt Significant SNPs (CSV): A table listing SNPs identified as significant by pcadapt, including their site info., p-values, and adjusted p-values.



6.2 OutFLANK

A Fst-based approach detects selection signals by comparing genetic differentiation between defined group assignments (Whitlock and Lotterhos 2015). For more information, visit <https://rpubs.com/lotterhos/outflank>.

Required Files:

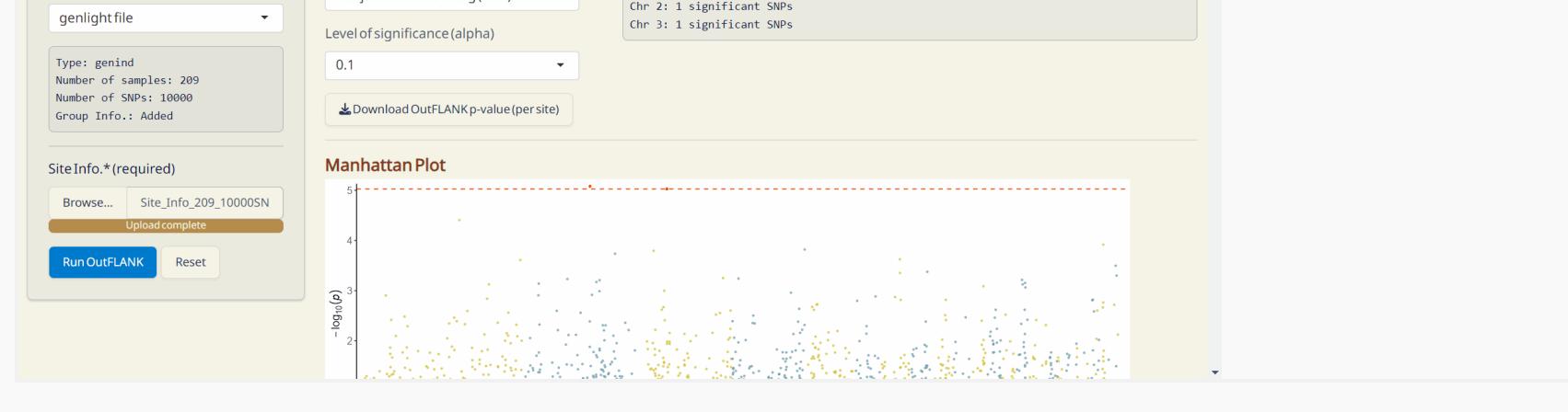
- *genlight* with 'Group Info.', downloadable from [Data Transform](#) page after you have Group Info.
- Site Info. (RDS) of the current *genlight* downloadable from [Data Input](#) or [Data QC](#) pages.

Steps:

1. Upload Site Info. (required).
2. Click **Run OutFLANK** to perform genome scan for selection.

Outputs:

- OutFLANK p-value per site (RDS): A dataset containing p-values and adjusted p-values for each site.
- OutFLANK Manhattan Plot (PDF): A Manhattan plot visualizing the p-values per site across the genome. Significant SNPs are highlighted in red.
- OutFLANK QQ Plot (PDF): A QQ plot comparing the distribution of observed p-values to the expected distribution under the null hypothesis.
- OutFLANK Histogram of p-values (PDF): A histogram showing the distribution of p-values across all sites.
- OutFLANK Histogram of Fst (PDF): A histogram showing the distribution of Fst values across all sites.
- OutFLANK Significant SNPs (CSV): A table listing SNPs identified as significant by OutFLANK, including their site info., Fst values, and p-values.



6.3 IBS (Identity By State)

An approach to detect differences in genomic regions between pairs of individuals, useful for identifying pedigree relationships.

Required Files:

- *data.frame*
- Site Info. (RDS) of the current *data.frame*, downloadable from [Data Input](#) or [Data QC](#) pages.
- Chromosome Info. (CSV): Reference genome information of the current study.

For more details about this file, refer to Section 2.3 (SNP Density).

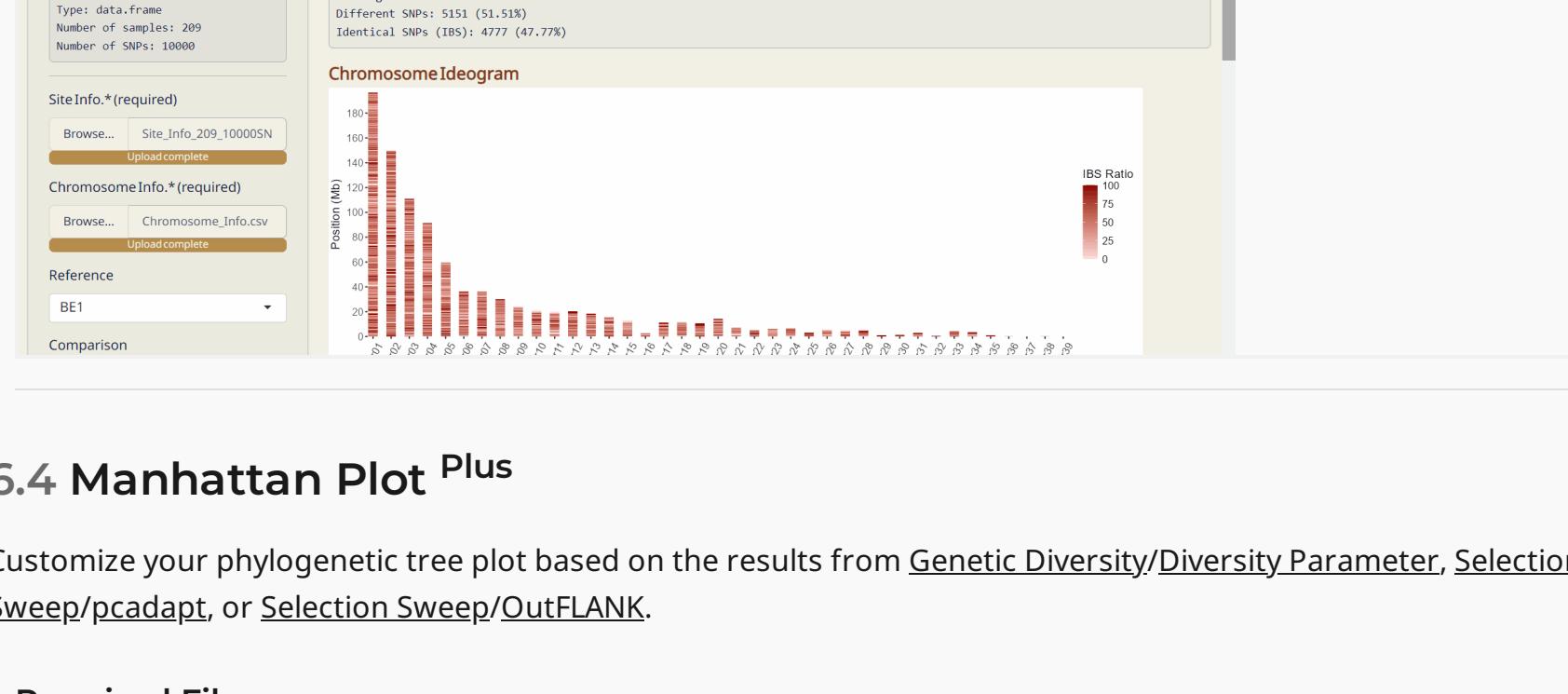
[Download an example of Chromosome Info. \(CSV\).](#)

Steps:

1. Upload Site Info. (required).
2. Upload Chromosome Info. (CSV) (required).
3. Choose the reference and comparison samples.
4. Select window size (kb) and step size (kp).
5. To remove heterozygous SNPs from the reference sample, click [Remove heterozygous SNPs](#) checkbox (optional).
6. Click [Run IBS](#) to perform IBS analysis.

Outputs:

- Chromosome Ideogram (PDF): An ideogram visualizing the IBS results, using a gradient palette to represent the differences across chromosomes.
- Sliding Window Data (CSV): A sliding window dataset with IBS results, including SNP count, different SNPs, and the ratio of different SNPs per window.



6.4 Manhattan Plot Plus

Customize your phylogenetic tree plot based on the results from [Genetic Diversity/Diversity Parameter](#), [Selection Sweep/pcadapt](#), or [Selection Sweep/OutFLANK](#).

Required Files:

- Genetic Diversity per Site (Genetic_Diversity_per_Site.rds), pcadapt p-value per Site (pcadapt_p-value_per_site.rds), or OutFLANK p-value per Site (OutFLANK_p-value_per_site.rds).
- Chromosome Info. (CSV): Reference genome information of the current study.

For more details about this file, refer to Section 2.3 (SNP Density).

[Download an example of Chromosome Info. \(CSV\).](#)

Steps:

1. Upload genetic_diversity/pcadapt_pvalue/OutFLANK_pvalue per site object (RDS).
2. Upload Chromosome Info. (CSV).
3. Click the [Run Manhattan Plot](#) to generate the Manhattan plot.
4. Customize the Manhattan plot and click the [Run Manhattan Plot](#) again.

Outputs:

- Manhattan Plot (PDF): A Manhattan plot with user-defined layout style and attributes.

Note: If generating a plot for p-values, make sure to use '-log10' transformation for the Y axis.



[« 5 Genetic Diversity](#)

[7 Core Collection »](#)

"User Guide for ShiNyP: SNP Analysis and Visualization Platform" was written by Yen-Hsiang Huang, Chung-Feng Kao. It was last built on Jun 2025.

This book was built by the [bookdown](#) R package.



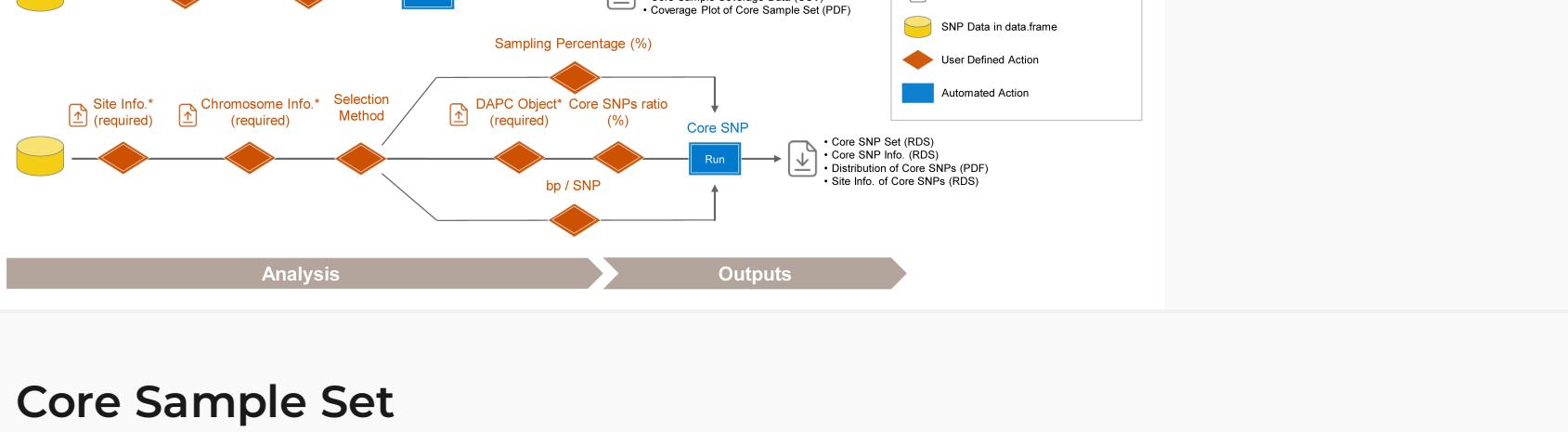
7 Core Collection

→ This section includes two subpages: [Core Sample Set](#), and [Core SNP Set](#), allowing you to capture the key samples and SNPs.

On this page

[7 Core Collection](#)[7.1 Core Sample Set](#)[7.2 Core SNP Set](#)[View source](#)[Edit this page](#)

Core Collection Page



7.1 Core Sample Set

Establish a core collection that represents the genetic variation of the entire population. This approach is modified function from GenoCore ([Jeong et al. 2017](#)).

Required File:

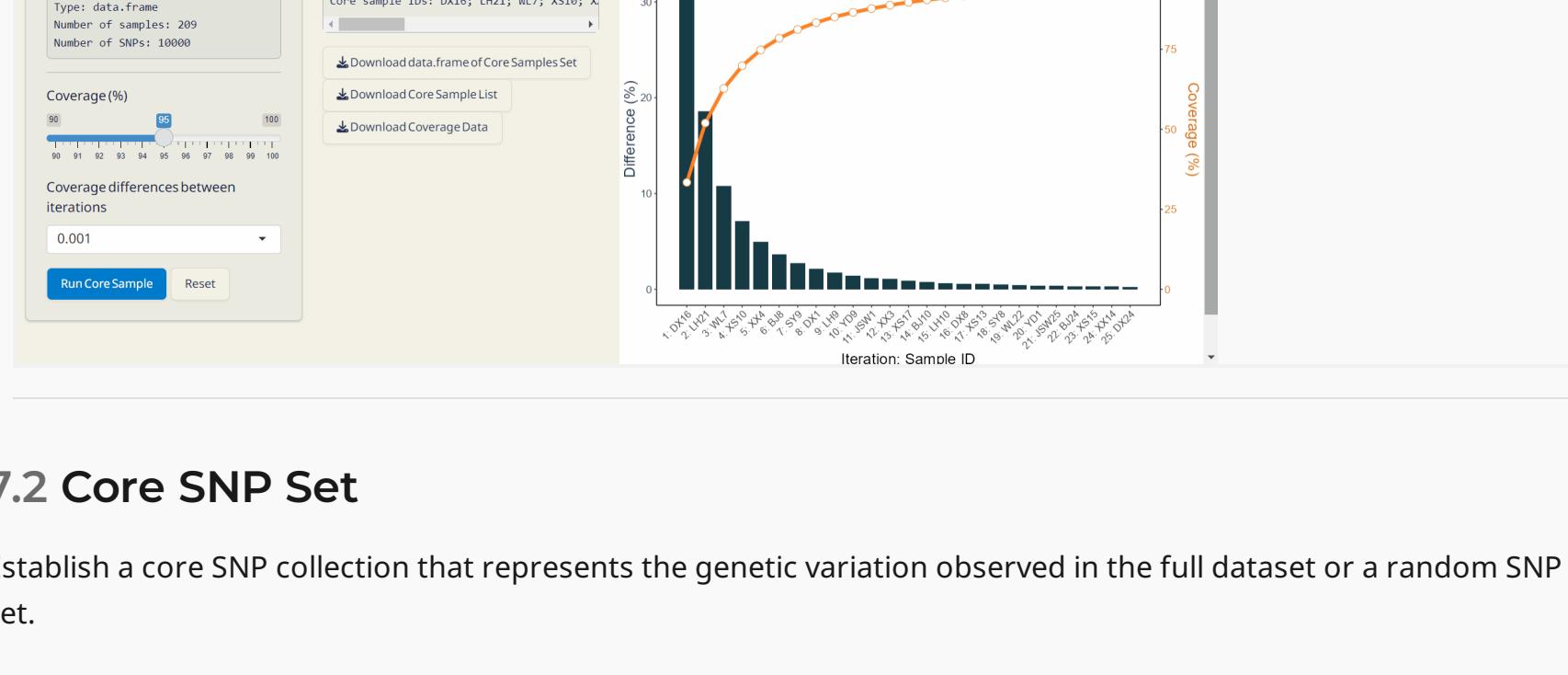
- data.frame*

Steps:

- Choose the minimum genetic coverage (%).
- Choose the minimum genetic coverage differences between iterations.
- Click [Run Core Sample](#) to perform core collection.

Outputs:

- Core Sample Coverage Data (CSV): A table listing the coverage (%) of each iteration and coverage differences between iterations.
- Core Sample Set (RDS): A *data.frame* of core samples and their genotypic information.
- Core Sample Info. (CSV): A table listing whether each sample is included in the core collection or not, and can be used as input data in the [Population Structure/PCA](#) subpage.
- Coverage Plot of Core Sample Set (PDF): Visualizes the sample coverage by each iteration.



7.2 Core SNP Set

Establish a core SNP collection that represents the genetic variation observed in the full dataset or a random SNP set.

Required Files:

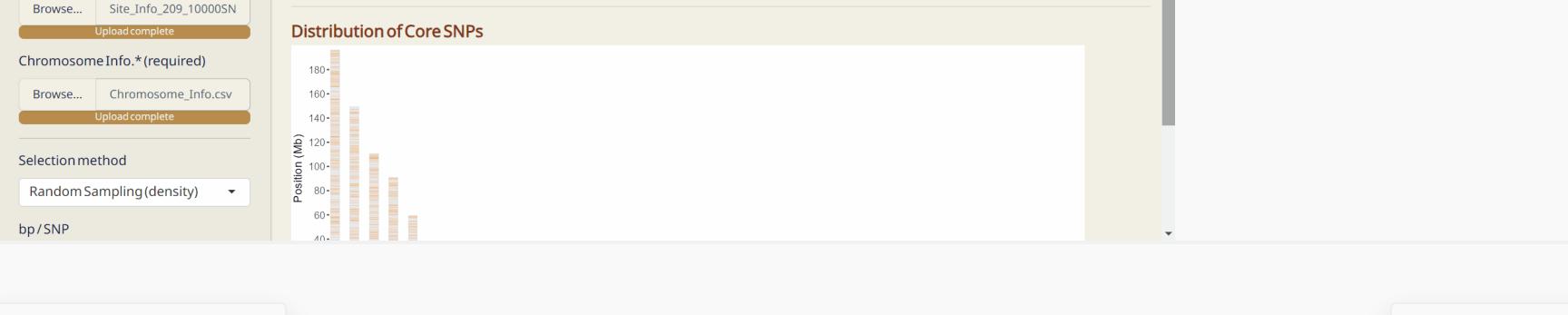
- data.frame*
- Site Info. (RDS) of the current *data.frame*, downloadable from [Data Input](#) or [Data QC](#) pages
- Chromosome Info. (CSV): Reference genome information of the current study. For more details about this file, refer to Section 2.3 (SNP Density).
- [Download an example of Chromosome Info. \(CSV\)](#).
- DAPC Object (DAPC_dapc_Object.rds), downloadable from [Population Structure/DAPC](#) subpage.

Steps:

- Upload required datasets: Site Info. (RDS) and Chromosome Info. (CSV).
- Choose the method and Upload DAPC Object (RDS) or set the parameter.
- Click the [Run Core SNP](#) to perform core collection.

Outputs:

- Core SNP Set (RDS): A *data.frame* of core SNPs and their genotypic information.
- Core SNP Info. (RDS): A table listing whether each SNP is included in the core collection or not.
- Distribution of Core SNPs (PDF): An ideogram labeling the core SNPs.
- Site Info. of Core SNPs (RDS): Core SNPs site information file.

[« 6 Selection Sweep](#)[8 AI Report »](#)

"Visualization Platform" was written by Yen-Hsiang Huang, Chung-Feng Kao. It was last built on Jun 2025.
package.



8 AI Report

→ This page is designed to generate the preliminary results from prior analysis, select an AI model, input the API key, and get an AI-driven report. Powered by the *ellmer* package.

On this page

[8 AI Report](#)

Your AI report is almost ready!

AI models

[Tasks for AI](#)

How to get the API key

[Encountered an error? Let's fix it!](#)

[View source](#)

[Edit this page](#)

Your AI report is almost ready!

Step 1: Preliminary Results

1. Enter the species name for your current analysis.
2. Click **Auto-generate** button to obtain Preliminary Results.

Note: You can download the Preliminary Results as a `.txt` file, edit it as needed, and upload it again for 'AI-driven Report' use.

This screenshot shows the ShiNyP AI Report interface. At the top, there's a navigation bar with links like Home, Data Input, Data QC, Data Transform, Population Structure, Genetic Diversity, Selection Sweep, Core Collection, and AI Report. The AI Report tab is active. Below the navigation, there are two main sections: '1. Preliminary Results' and '2. AI-Driven Report'. In '1. Preliminary Results', there's a dropdown for 'Specify the species for your SNP data' set to 'Ex: Wild rice (*Oryza rufipogon*)', and a 'Auto-generate' button. In '2. AI-Driven Report', there are dropdowns for 'AI model' (set to 'Gemini 2.0 Flash'), 'AI task' (set to 'Data Interpretation'), 'Conversation' (set to 'Single-Turn'), and an 'API keyfile' section where a file named 'APIKey-Gemini.txt' is selected.

Step 2: AI-Driven Report

1. Select an AI model.
2. Define the task for the AI model.
3. Select the conversation mode (*Single-Turn*: one request; *Multi-Turn*: up to five requests).
4. Select the report language.
5. Upload your AI API key: Provide a `.txt` file containing your API key (e.g., `sk-...`).
Model: "Gemini 2.5 Flash (API Free)" — no API required!

▼ Example of API key file.

A screenshot of a text editor window titled 'API Key Example.txt'. The content of the file is a long string of characters: 'sk-QzQlf_ShiNyPLeDPPt1MBNU19cjShiNyP3XazNgWgvT4B1ShiNyPFZB_1bS5xAShiNyPD_aoGZGShiNyPExampleAPIKEY'. The text editor has standard controls like File, Edit, View, and a status bar at the bottom indicating 'Ln 1, Col 99 | 98 characters | 100% | Windows (CRLF) | UTF-8'.

6. Click **Get Report** to obtain AI-driven Report.

7. Download the `.docx` (Word) file with a fully styled, professional layout.

This screenshot shows the '2. AI-Driven Report' section of the ShiNyP interface. It includes dropdowns for 'AI model' (set to 'Gemini 2.0 Flash-Lite'), 'AI task' (set to 'Data Interpretation'), 'Conversation' (set to 'Single-Turn'), 'Language' (set to 'English'), and an 'API keyfile' section where 'APIKey-Gemini.txt' is selected. A 'Download' button is visible above the form. Above the form, there's a preview window showing a snippet of the generated report content.

AI models

NEW For ShiNyP v1.1.0, we support 11 AI models:

Models	Developer	Price ^a
DeepSeek-V3	DeepSeek	\$0.07/\$1.10
Gemini 2.0 Flash	Google	Free or \$0.10/\$0.40
NEW Gemini 2.5 Flash	Google	Free or \$0.30/\$2.50
NEW Gemini 2.5 Flash-Lite	Google	Free or \$0.10/\$0.40
o4-mini	OpenAI	\$1.10/\$4.40
o3-mini	OpenAI	\$1.10/\$4.40
GPT-4.1	OpenAI	\$2.00/\$8.00
GPT-4.1 mini	OpenAI	\$0.40/\$1.60
GPT-4.1 nano	OpenAI	\$0.10/\$0.40
GPT-4o	OpenAI	\$2.50/\$10.00
GPT-4o mini	OpenAI	\$0.15/\$0.60

^a Price (USD) per 1M Tokens (Input / Output)

→ We recommend using Google **Gemini 2.5 Flash**—the flagship model—for generating AI reports. *It is free to get API key & use* (see below).

If you have paid API access from OpenAI, we recommend using **GPT-4.1** for stable performance. Please note that **o4-mini** and **o3-mini** are more suitable for "Idea Expansion" task within single-turn conversation.

Tasks for AI

- **Summary Request:** Provide a clear, structured overview of key SNP analysis results.

- **Data Interpretation:** Clarify findings and highlight key insights into population genetics.

- **Report Structuring:** Create a framework for presenting SNP analysis results in a scientific report.

- **Idea Expansion:** Propose future research directions informed by SNP findings.
- **NEW Custom Template:** You can upload a template (.txt file) tailored to specific project, and AI will automatically fill in the analysis results.

This reports fully comply with designated standards, enabling truly automated and standardized report generation for your lab. *Only single-turn conversation is supported.*

[Download an example of Custom Template \(TXT\).](#)

How to get the API key

1 Sign Up or Log In:

- Google: <https://aistudio.google.com/>
- OpenAI: <https://auth.openai.com/create-account>
- DeepSeek: https://platform.deepseek.com/sign_in

2 Generate Your API Key:

- Google: Click "Create API Key" bottom <https://aistudio.google.com/apikey>
- OpenAI: <https://platform.openai.com/api-keys>
- DeepSeek: https://platform.deepseek.com/api_keys

3 Save Your API Key:

- Copy the generated key and paste it into a Notepad file.
- Save the file as .txt for *ShiNyP* use.

4 Manage Billing & Payments:

- OpenAI: <https://platform.openai.com/settings/organization/billing/overview>
- DeepSeek: https://platform.deepseek.com/top_up

Note: Costs vary depending on the model, conversation mode, and the AI task requested.

✗ Encountered an error? Let's fix it!

⚠ Error: HTTP 401 Unauthorized

This error message indicates that your authentication credentials are invalid. This could happen for several reasons, such as:

1. You are using a revoked API key.
2. You are using a different API key than one under the requesting organization.
3. You are using an API key that does not have the required permissions for the endpoint you are calling.

To resolve this error, first check that you are using the correct API key and organization ID in your request header.

⚠ Error: Failed to Perform HTTP Request

This error may indicate that the request timed out, possibly due to an excessive input token count, which prevents the AI model from completing the task within the allotted time.

[« 7 Core Collection](#)

[9 FAQ »](#)

"User Guide for ShiNyP: SNP Analysis and Visualization Platform" was written by Yen-Hsiang Huang, Chung-Feng Kao. It was last built on Jun 2025.

This book was built by the [bookdown](#) R package.

9 FAQ

If you encounter any issues or have suggestions for new features, please submit a request on the [GitHub Issues page](#) or email us at: teddyhuangyh@gmail.com

General Usage

1. Can I use the online version of *ShiNyP* platform?

Yes, a trial version of *ShiNyP* is available online at: https://teddyhuang.shinyapps.io/ShiNyP_Demo/

This web-based DEMO is hosted on Shinyapps.io and is intended for trial purposes. Please note that, due to server memory limitations (1GB RAM), this version is not suitable for the analysis of large-scale SNP datasets. For complete functionality and to analyze larger datasets, we strongly recommend downloading *ShiNyP* from GitHub.

For more information: [Get Started](#)

2. Do I need programming skills to use *ShiNyP*?

No programming experience is required. *ShiNyP* provides a user-friendly graphical interface that allows you to perform all analyses interactively without coding.

3. How do I install and start the *ShiNyP* platform?

You can install *ShiNyP* in R via:

```
remotes::install_github("TeddyYenn/ShiNyP")
library(ShiNyP)
ShiNyP::run_ShnyP()
```

Alternatively, use the Docker image for a one-step deployment:

```
docker run -d -p 3838:3838 teddyenn/shinyp-platfor
```

For more information: [Get Started](#)

4. Is my data secure when using *ShiNyP*?

All analyses are conducted locally on your machine or server. No user data is transmitted to third-party servers unless you explicitly use AI features (*ShiNyP AI* chatbot and reporting system).

5. Does *ShiNyP* require internet access to function?

ShiNyP runs locally and does not require internet access for core analyses. Internet is only needed when installing the packages and using AI features (*ShiNyP AI* chatbot and reporting system).

Analysis & Features

1. What makes *ShiNyP* different from other SNP analysis tools?

ShiNyP uniquely integrates a modular Graphical User Interface (GUI), *ShiNyP AI*(real-time chatbot), cross-species compatibility, AI-based interpretation, customizable visualizations, and open-source accessibility—all in one platform, making it a versatile tool for researchers in genomics, breeding, and evolutionary biology.

2. What types of input data are supported by *ShiNyP*?

ShiNyP supports genome-wide biallelic SNP datasets in Variant Call Format (VCF). It is also compatible with data.frame and genlight files, covering both diploid and polyploid species. NOTE: *The diploidization processing simplifies genotype data and does not account for allelic dosage effects*.

For more information: [Data Input](#)

3. What kind of output does the platform provide?

The platform generates publication-ready figures (PDF, PNG, or JPEG) and tables, reusable R data objects, and AI-assisted reports that summarize analytical results in natural language.

4. Can I customize analysis parameters?

Yes, users can adjust analysis thresholds, models, and filtering criteria directly through the interface to suit their research needs.

5. Can I analyze large SNP datasets with *ShiNyP*?

Yes, *ShiNyP* is optimized for both moderate and large-scale SNP datasets. However, performance may depend on your system's hardware specifications.

6. Can I analyze datasets from different size/species in one session?

Yes, as long as your data are properly formatted, *ShiNyP* supports datasets from different size/species. However, it is recommended to analyze each dataset separately for clearer results.

7. Can *ShiNyP* export data to formats used by other genetics software?

Yes, results and transformed data can be exported in formats compatible with tools such as R, STRUCTURE, PLINK, GenAIEx, and others.

For more information: [Data Transform](#)

Support

1. Is *ShiNyP* open-source? Is it possible to extend *ShiNyP* with custom modules or scripts?

Yes, *ShiNyP* is released under an open-source license ([GNU Affero General Public License](#)). The source code is available on GitHub for transparency and community contributions.

2. Where can I report bugs or request new features?

If you encounter any issues or have suggestions for new features, please submit a request on the [GitHub Issues page](#) or email us at: teddyhuangyh@gmail.com. The project is actively maintained and welcomes community feedback and collaboration.

3. Future?

Installation Issues

If you encounter any issues while installing *ShiNyP*, please don't hesitate to let us know. The issue may not be unique to you, and by reporting it, you help improve the entire community. Below are some common installation issues:

■ R and Bioconductor version mismatch

The installation specifies Bioconductor version 3.21, which requires R version ≥ 4.5 . If your R version is below 4.5, update it from [CRAN](#).



■ Permission denied: curl.dll

Please close all R/RStudio programs and terminal windows to ensure no sessions are using the package, then restart and try installing again.

■ Installation of dependencies

Installing packages like *shiny* and *dartR* may require additional developer tools. You might encounter error messages such as:

ERROR: dependencies 'shiny', 'dartR' are not available for package 'ShiNyP'



- **For Windows:** Download Rtools from [CRAN Rtools](#).

- **For macOS:** Open Terminal and run the installation command:

```
xcode-select --install
```



■ GitHub installation issues

Verify that your system is connected to the internet. Check for any firewall or proxy settings that might block GitHub access.

■ Package loading issues

Errors occur when loading *ShiNyP* or its dependencies, possibly due to outdated packages or conflicts. [Update all installed packages](#).

Index

- AI Report 8
- AMOVA (Analysis of MOlecular VAriance) 5.4
- API key 8
- Bayesian Information Criterion (BIC) 4.2
- Chromosome Info. 2.3 5.2 6.3 6.4 7.2
- Circos Plot 5.2
- Core Sample Set 7.1
- Core SNP Set 7.2
- DAPC (Discriminant Analysis of Principal Components) 4.2
- data.frame 1.1
- Demo Data 1.1
- Diversity Parameter 5.1
- Genetic Distance 5.3
- genind 3
- genlight 5
- Group Info. 3 4.5 4.6 5.1 6.2
- Hardy-Weinberg equilibrium (HWE) 2.2
- Heterozygosity rate 2.1
- IBS (Identity By State) 6.3
- Kinship Analysis 4.5
- Manhattan Plot 6.4
- Minor allele frequency (MAF) 2.2
- Missing rate 2.1
- NJ (Neighbor-Joining) Tree 4.4
- OutFLANK 6.2
- PCA (Principal Component Analysis) 4.1
- pcadapt 6.1
- Permutation Test 5.4
- Sample QC 2.1
- Scatter Plot 4.6
- *ShiNyP* ShiNyP
- Site Info. 1.1 2.1 2.2 2.3 5.1 6.1 6.2 6.3 7.2
- SNP Density 2.3
- SNP QC 2.2
- Tree Plot 4.7
- UPGMA (Unweighted Pair Group Method with Arithmetic mean) Tree 4.3
- VCF 1.1

Bibliography

- Excoffier, L, P E Smouse, and J M Quattro. 1992. "Analysis of Molecular Variance Inferred from Metric Distances Among DNA Haplotypes: Application to Human Mitochondrial DNA Restriction Data." *Genetics* 131 (2): 479–91. <https://doi.org/10.1093/genetics/131.2.479>.
- GOUDET, JÉRÔME. 2004. "Hierfstat, a Package for r to Compute and Test Hierarchical F-Statistics." *Molecular Ecology Notes* 5 (1): 184–86. <https://doi.org/10.1111/j.1471-8286.2004.00828.x>.
- Granato, Italo S. C., Giovanni Galli, Evelyn Giselly de Oliveira Couto, Massaine Bandeira e Souza, Leandro Freitas Mendonça, and Roberto Fritzsche-Neto. 2018. "snpReady: A Tool to Assist Breeders in Genomic Analysis." *Molecular Breeding* 38 (8). <https://doi.org/10.1007/s11032-018-0844-8>.
- Gu, Zuguang, Lei Gu, Roland Eils, Matthias Schlesner, and Benedikt Brors. 2014. "Circlize implements and Enhances Circular Visualization in R." *Bioinformatics* 30 (19): 2811–12. <https://doi.org/10.1093/bioinformatics/btu393>.
- Jeong, Seongmun, Jae-Yoon Kim, Soon-Chun Jeong, Sung-Taeg Kang, Jung-Kyung Moon, and Namshin Kim. 2017. "GenoCore: A Simple and Fast Algorithm for Core Subset Selection from Large Genotype Datasets." Edited by Alexandre G. de Brevern. *PLOS ONE* 12 (7): e0181420. <https://doi.org/10.1371/journal.pone.0181420>.
- Jombart, Thibaut, Sébastien Devillard, and François Balloux. 2010. "Discriminant Analysis of Principal Components: A New Method for the Analysis of Genetically Structured Populations." *BMC Genetics* 11 (1): 94. <https://doi.org/10.1186/1471-2156-11-94>.
- Kamvar, Zhian N., Javier F. Tabima, and Niklaus J. Grünwald. 2014. "Poppr: An R Package for Genetic Analysis of Populations with Clonal, Partially Clonal, and/or Sexual Reproduction." *PeerJ* 2 (March): e281. <https://doi.org/10.7717/peerj.281>.
- Kang, Hyun Min, Jae Hoon Sul, Susan K Service, Noah A Zaitlen, Sit-ye Kong, Nelson B Freimer, Chiara Sabatti, and Eleazar Eskin. 2010. "Variance Component Model to Account for Sample Structure in Genome-Wide Association Studies." *Nature Genetics* 42 (4): 348–54. <https://doi.org/10.1038/ng.548>.
- Luu, Keurcien, Eric Bazin, and Michael G. B. Blum. 2016. "Pcadapt: An R Package to Perform Genome Scans for Selection Based on Principal Component Analysis." *Molecular Ecology Resources* 17 (1): 67–77. <https://doi.org/10.1111/1755-0998.12592>.
- Paradis, Emmanuel, and Klaus Schliep. 2018. "Ape 5.0: An Environment for Modern Phylogenetics and Evolutionary Analyses in R." Edited by Russell Schwartz. *Bioinformatics* 35 (3): 526–28. <https://doi.org/10.1093/bioinformatics/bty633>.

Whitlock, Michael C., and Katie E. Lotterhos. 2015. "Reliable Detection of Loci Responsible for Local Adaptation: Inference of a Null Model Through Trimming the Distribution of F_{ST} ." *The American Naturalist* 186 (S1): S24–36.

<https://doi.org/10.1086/682949>.

Yu, Guangchuang, David K. Smith, Huachen Zhu, Yi Guan, and Tommy Tsan-Yuk Lam. 2016. "Ggtree: An r Package for Visualization and Annotation of Phylogenetic Trees with Their Covariates and Other Associated Data." Edited by Greg McInerny. *Methods in Ecology and Evolution* 8 (1): 28–36. <https://doi.org/10.1111/2041-210x.12628>.

« 8 AI Report

"User Guide for ShiNyP: SNP Analysis and Visualization Platform" was written by Yen-Hsiang Huang, Chung-Feng Kao. It was last built on Jun 2025.

This book was built by the [bookdown](#) R package.