atresult g value (GEBV)ic estimated breeding value (GEBV)

**User Manual for**

****

**Intelligent Prediction and Association Tool**

**(Version 1.3)**

**Last updated on September 12, 2017**



**Disclaimer**: While extensive testing has been performed by Zhiwu Zhang Lab at Washington State University, results are, in general, reliable, correct or appropriate. However, results are not guaranteed for any specific set of data. We strongly recommend that users validate iPat results with other original software packages, such as GAPIT, PLINK, rrBLUP and BGLR.

**Support documents**: Extensive support documents, including this user manual, source code, demonstration scripts, data, and results, are available at iPat website Zhiwu Zhang Laboratory: <http://zzlab.net/iPat>

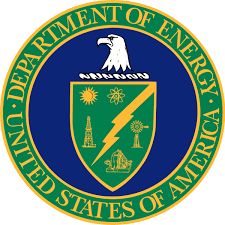
**Questions and comments**: Users and developers are recommended to post questions and comments at iPat forum: <https://github.com/Poissonfish/iPat/issues>. Answers from other users and developers are appreciated. The iPat team members will periodically go through these questions and comments and address them accordingly.

The iPat project is partially under supports from USDA-ARS, DOE, NSF, the Agricultural Research Center at Washington State University, and Washington Grain Commission.

**Citation**: James Chen and Zhiwu Zhang, User manual for Intelligent Prediction and Association Tool, version 1.3, <http://zzlab.net/iPat>, accessed on MM/DD/YYYY.







**Table of Contents**

[Why iPat? 6](#_Toc492991113)

[1. Getting start 6](#_Toc492991114)

[1.1 Operation environment 6](#_Toc492991115)

[1.2 Set up R environment 6](#_Toc492991116)

[1.3 Windows users 7](#_Toc492991117)

[1.4 Mac OS users 7](#_Toc492991118)

[2. Interface 8](#_Toc492991119)

[2.1 Import files 8](#_Toc492991120)

[2.2 Create a project 8](#_Toc492991121)

[2.3 Covariates and kinship 9](#_Toc492991122)

[2.4 Define Your Analysis 9](#_Toc492991123)

[2.5 Run an analysis 11](#_Toc492991124)

[2.6 Inspect the result 12](#_Toc492991125)

[2.7 Remove files from iPat 12](#_Toc492991126)

[3. File formats 13](#_Toc492991127)

[3.1 Phenotype 13](#_Toc492991128)

[3.2 Genotype 14](#_Toc492991129)

[3.2.1 Hapmap 14](#_Toc492991130)

[3.2.2 Numeric 14](#_Toc492991131)

[3.2.3 VCF 14](#_Toc492991132)

[3.2.4 PLINK (the header](#_Toc492991133) **[should](#_Toc492991133)** [be removed) 14](#_Toc492991133)

[3.2.5 Binary PLINK (the header](#_Toc492991134) **[should](#_Toc492991134)** [be removed) 15](#_Toc492991134)

[3.3 Covariates 15](#_Toc492991135)

[3.4 Kinship 15](#_Toc492991136)

[4. Incorporated packages 16](#_Toc492991137)

[4.1 GAPIT 16](#_Toc492991138)

[4.2 FarmCPU 16](#_Toc492991139)

[4.3 PLINK 16](#_Toc492991140)

[4.4 rrBLUP 17](#_Toc492991141)

[4.5 BGLR 17](#_Toc492991142)

[5. Output files 18](#_Toc492991143)

[5.1 Phenotype 18](#_Toc492991144)

[5.1.1 Overview 18](#_Toc492991145)

[5.2 Population structure 18](#_Toc492991146)

[5.2.1 Heterozygosity 18](#_Toc492991147)

[5.3 GWAS 19](#_Toc492991148)

[5.3.1 Genomewise Manhattan plot 19](#_Toc492991149)

[5.3.2 Q-Q plot 19](#_Toc492991150)

[5.3.3 GWAS result 19](#_Toc492991151)

[5.4 GS 20](#_Toc492991152)

[5.4.1 Genomic estimated breeding value (GEBV) 20](#_Toc492991153)

[5.4.2 Standard deviation of GEBV 20](#_Toc492991154)

[5.4.3 Histogram of GEBV 21](#_Toc492991155)

[5.4.4 Prediction result 21](#_Toc492991156)

[6. Tutorial 22](#_Toc492991157)

[6.1 Perform GWAS in iPat 22](#_Toc492991158)

[6.2 Perform GS and add user-define covariates in iPat 23](#_Toc492991161)

[6.3 Perform GWAS-assist GS in iPat 24](#_Toc492991164)

[7. References 27](#_Toc492991167)

# Why iPat?

Genome Wide Association Study (GWAS) and Genomic Prediction/Selection (GS) are two types of analyses in genomic research. Numerous software packages have been developed for the analyses with different models on different format of data. Most of these software packages were executed through Command Line Interface (CLI), including PLINK(Purcell *et al.* 2007), GAPIT(Lipka *et al.* 2012; Tang *et al.* 2016), FarmCPU(Liu *et al.* 2016), rrBLUP(Endelman 2011) and BGLR(Pérez and De Los Campos 2014). Researchers are hindered by factors such as the programming requirements, data format incompatibilities, and zero tolerance on typo of commands. Intelligent Prediction and Association Tool (iPat) is a software package with a user-friendly graphical user interface (GUI) to conduct GWAS and GS with multiple available CLI packages such as the ones listed above. Researchers can simply drag and/or click on graphical elements to specify input data files, select models, and choose define parameters. Multiple data formats are acceptable and converted automatically to the required format. Furthermore, a uniform and comprehensive presentation of results is provided to enhance interpretation of data analyses.

# 1. Getting start

## 1.1 Operation environment

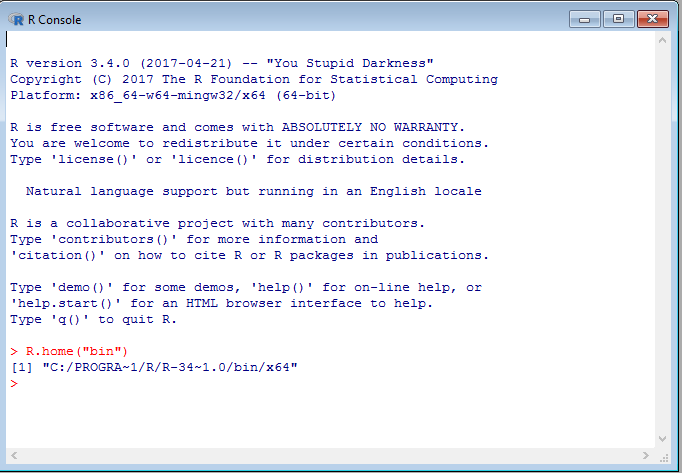
* The operation environment need to meet the following requirement:
* Operation System: Windows or Mac OS X.
* [Java Runtime Environment (JRE)](http://www.oracle.com/technetwork/java/javase/downloads/index.html): Version 8 or later.
* [R](https://www.r-project.org): Version 3.4.1 or later.

## 1.2 Set up R environment

Open R software and run,

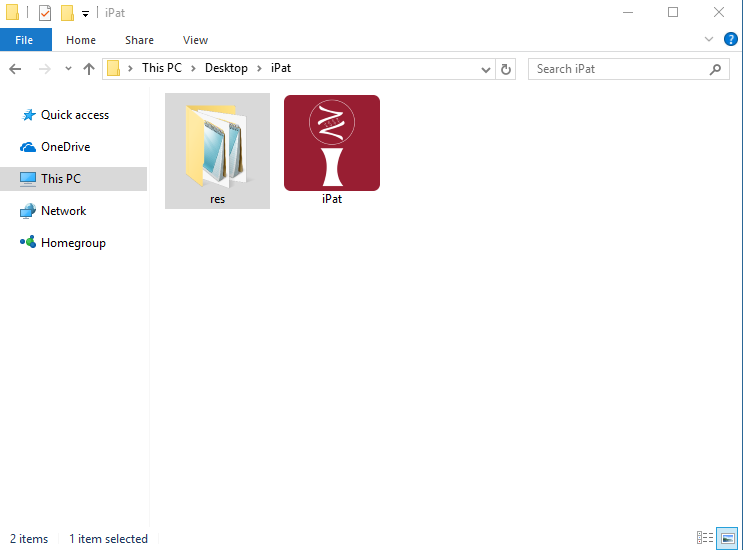


This command will install all the required r packages automatically



## 1.3 Windows users

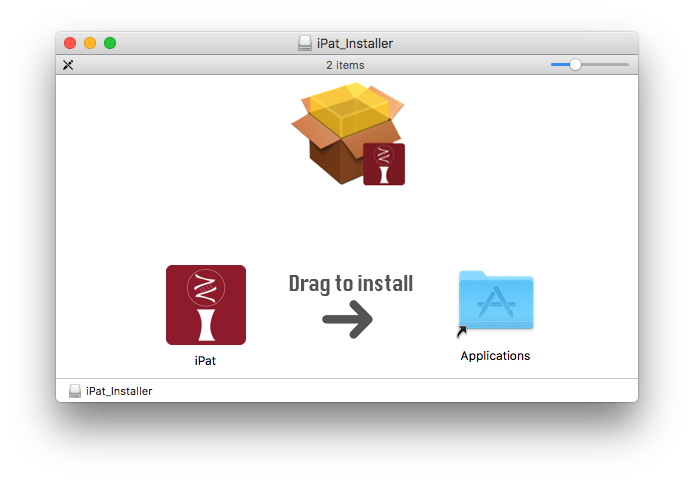
* Download [iPat.zip](http://zzlab.net/iPat/iPat.zip) and decompress it. You will then get a folder named "iPat", which contains a executable file "iPat.exe" and a folder "libs".
* It's noted that users are always required to place "iPat.exe" and the folder "res" in the same folder (directory) so that iPat can function normally.



* Double click 'iPat.exe' to launch iPat.

## 1.4 Mac OS users

* Download [iPat\_Installer.dmg](http://zzlab.net/iPat/iPat_Installer.dmg) and mount it on Mac.
* Follow the instruction to install iPat.

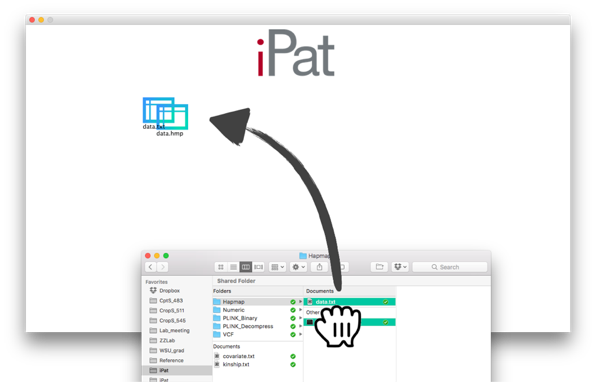


* Double click 'iPat.app' to launch iPat.

# 2. Interface

## 2.1 Import files

* At beginning, iPat will show nothing but an icon "iPat" at the top of screen.
* Users can import files simply by dragging and dropping.



## 2.2 Create a project

* After importing the files, double clicking on anywhere in iPat to create a new project (a gear icon).
* Build a project by dragging a files over the project icon. A dash line will be shown between the file and project, which mean this file has been already included in this project. The below two are examples for a valid project.



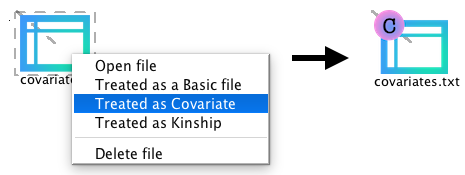
* **[IMPORTANT]** A valid project must include a certain number of required files, **no less, no more.** Otherwise iPat won't work and will return an error message. Valid datasets for each format can be found from the table below (See section 3 for details):

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Format** | **File 1 (required)** | **File 2 (required)** | **File 3 (required)** | **File 4 (required)** |
| Hapmap | **Genotype** (.hmp) | **Phenotype** (.txt) | None | None |
| Numeric | **Genotype** (.dat) | **Phenotype** (.txt) | **Map information** (.map) | None |
| VCF | **Genotype** (.vcf) | **Phenotype** (.txt) | None | None |
| PLINK | **Genotype** (.ped) | **Phenotype** (.txt) | **Map information** (.map) | None |
| PLINK  (binary) | **Genotype** (.bed) | **Phenotype** (.txt) | **Map information** (.bim) | **Individual information** (.fam) |

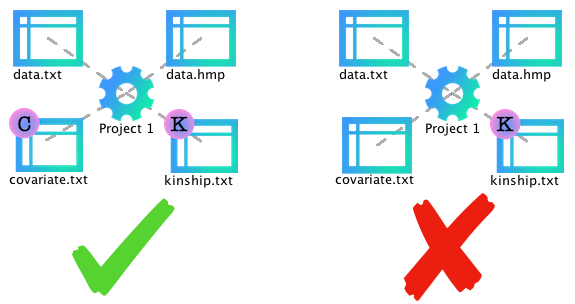
* Map information in numerical format is only required for GWAS
* In PLINK, phenotype file is only required for multiple traits analysis

## 2.3 Covariates and kinship

* Covariates provided by users will be treated as **fixed effect** in the selected model except in BGLR.
* It's **optional** that users can add **user-define** covariates or kinship into the project. Right clicking on the file to tell iPat what type of file it is. (i.e. covariates, kinship or a basic required file)



* Be aware that apart from the basic required file (i.e. phenotype and genotype), optional files must be properly labeled in a project.



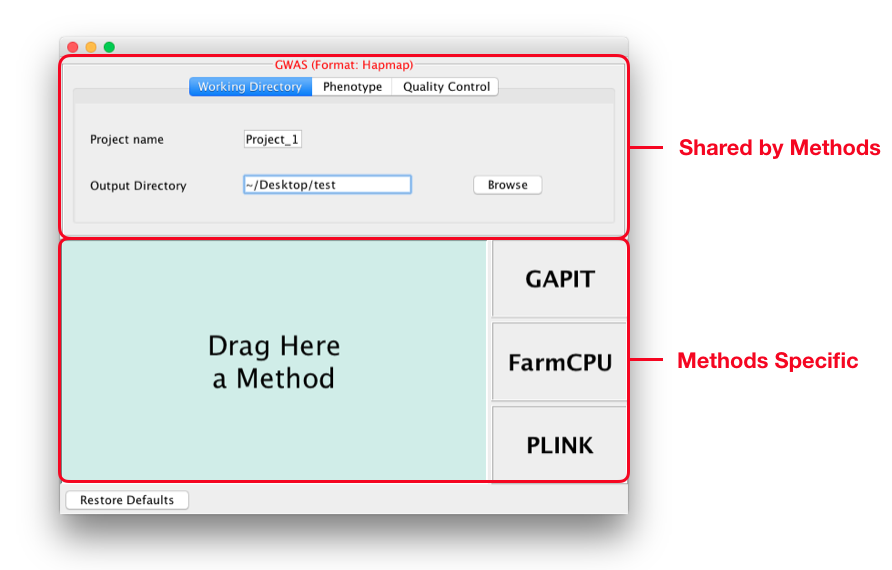
* The file labeled “C” stands for a covariate file, while files labeled “K” is identified as a kinship file by iPat. For the example of a valid project below (Left one), the file "covariate.txt" and "kinship.txt" are treated as covariates and a kinship in this project, respectively. Each project can contain **one** covariate files and **one** single kinship.

## 2.4 Define Your Analysis

* After linking every files needed in the project, right click on the project and choose either GWAS or GS to open a configuration panel.



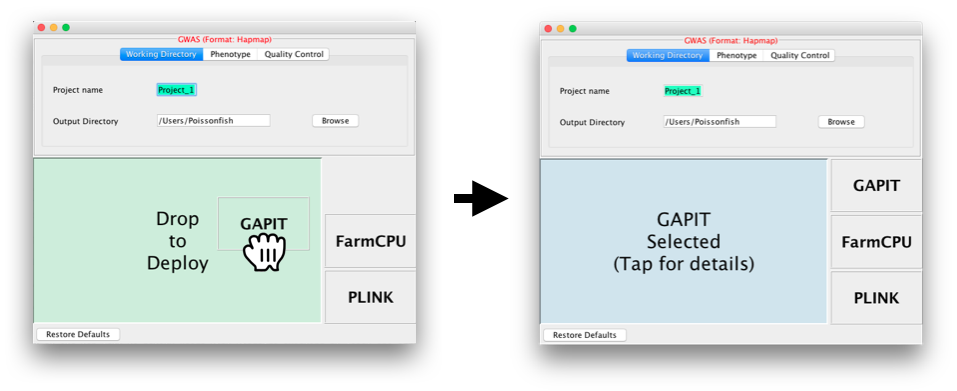
* The panel consist of two sections. The upper one presents a set of input arguments shared by all methods, while users can define method-specific arguments from the lower section.



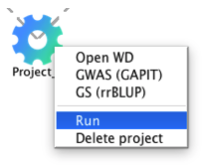
* Available parameters in the upper section:

|  |  |  |  |
| --- | --- | --- | --- |
| **Tab** | **Parameters** | **Definition** | **Default** |
| Working Directory | Project name | Prefix for output files | Project\_x  (x is a number starts from 1) |
| Working Directory | Output Directory | A path where output files will be  generated | Home directory |
| Phenotype | Trait names | Subsetting traits data | All traits are selected |
| Quality Control | By missing rate | Filtering out markers where certain rate  of value is missing | No threshold |
| Quality Control | By MAF | Filtering out markers based on minor  allele frequency | 0.05 |

* To select a method, simply drag a "method block" to the left-side area. And tap on this area for further defining (see section 3 for details)

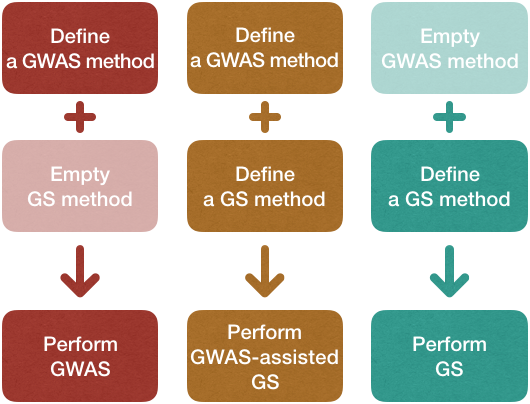


* After defining the analysis, user can start to run the procedure by clicking ‘Run’ at the pop-up menu of the project.



## 2.5 Run an analysis

* In iPat, users are allowed to do genomic studies such as GWAS, GS and GWAS-Assisted GS (Associated SNPs reported by GWAS will be treated as fixed effect in GS). iPat will detect the project configuration and decide which analysis should be implemented afterward.

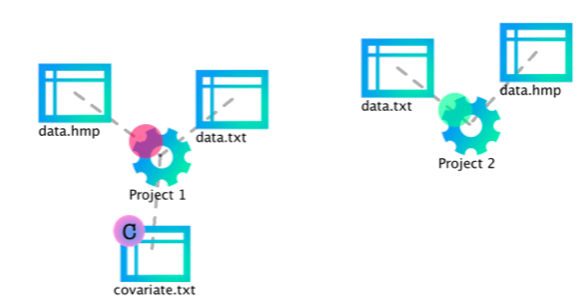


* Each project will generate a console window while running the analysis. User can track the progress of the task from window messages.
* iPat also capable of multitasking. Users can arrange another project even when the previous one have not done yet.

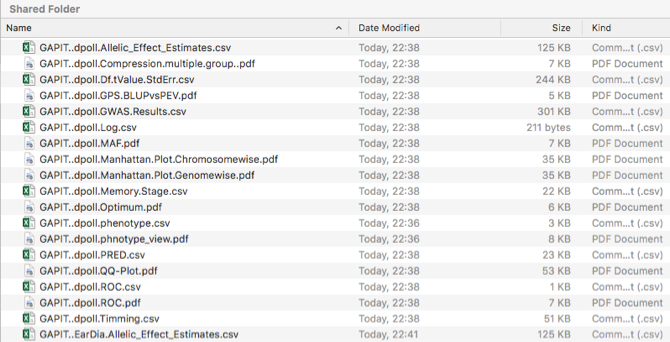


## 2.6 Inspect the result

* When iPat complete a project, the gear icon will show a green dot if the task run successfully without any error occurred. Otherwise it will show a red dot at its top-left to notify users that there’re existing at least one error message during the analysis.



* Users can inspect the results by double clicking on the gear icon, which will direct users to the folder where output files generated (See section 4 for details of output files).



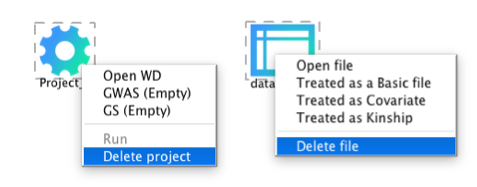
## 2.7 Remove files from iPat

Users are allowed to remove projects, files and linkage from iPat in 3 ways:

* Drag any object to the bottom-left area and release to remove it.



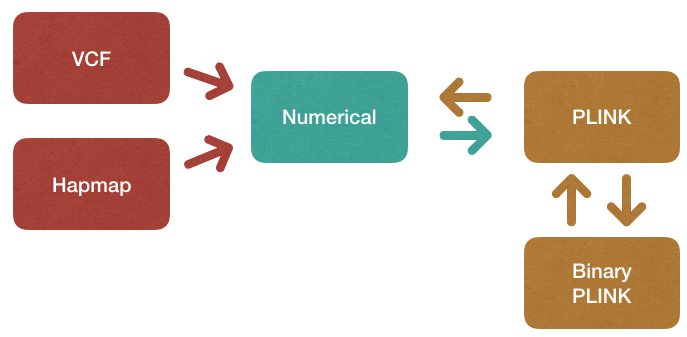
* Right click on an object to open a pop-up menu then delete it



* Or, simply press Backspace (press delete on Mac) after selecting an object.

# 3. File formats

* iPat mainly work with files in numerical format, but it can also work fine with Hapmap, VCF and PLINK format. iPat will recognize the format of input files and do a format conversion automatically if needed.



## 3.1 Phenotype

* Phenotype data for every formats except PLINK must contain **sample names** in the first column and **traits names** as the header:

|  |  |  |
| --- | --- | --- |
| **taxa** | **trait 1** | **trait 2** |
| sample1 |  |  |
| sample2 |  |  |
| sample3 |  |  |

* Phenotype data for PLINK must contain **sample and family names** in the first 2 columns and **traits names** as the header:

|  |  |  |  |
| --- | --- | --- | --- |
| **FID** | **SID** | **trait 1** | **trait 2** |
| family 1 | sample1 |  |  |
| family 2 | sample2 |  |  |
| family 3 | sample3 |  |  |

## 3.2 Genotype

### 3.2.1 Hapmap

* Genotype data, the header is **required** to be provided:

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **rs** | **alleles** | **Chr.** | **pos** | **strand** | **Assembl** | **Cent.** | **protLS** | **assay** | **panel** | **QC** | **Sample 1** | **Sample 2** |
| Marker  1 | A/C | 1 | 157104 | + | AGPv1 | Panzea | NA | NA | maize282 | NA | CC | CC |
| Marker  2 | C/G | 1 | 1947984 | + | AGPv1 | Panzea | NA | NA | maize282 | NA | GG | GG |

### 3.2.2 Numeric

* Genotype data, samples are recorded in rows. The header and sample names can be **omitted**:

|  |  |  |  |
| --- | --- | --- | --- |
| **taxa** | **marker 1** | **marker 2** | **marker 3** |
| sample1 | 0 | 0 | 1 |
| sample2 | 0 | 0 | 0 |
| sample3 | 1 | 0 | 0 |

* Map information, the header is **required** to be provided:

|  |  |  |
| --- | --- | --- |
| **SNP** | **Chromosome** | **Position** |
| marker 1 | 1 | 157104 |
| marker 2 | 1 | 1947984 |
| marker 3 | 1 | 2914066 |

### 3.2.3 VCF

* Genotype data, the header is **required** to be provided:

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Chr** | **POS** | **ID** | **REF** | **ALT** | **QUAL** | **FILTER** | **INFO** | **FORMAT** | **sample 1** | **sample 2** |
| 1 | 157104 | marker 1 | A | C | . | PASS | . | GT | 0/0 | 1/1 |
| 1 | 1947984 | marker 2 | C | G | . | PASS | . | GT | 0/0 | 1/1 |

### 3.2.4 PLINK (the header **should** be removed)

* Genotype data (.ped). Missing value can be filled as "0":

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Family**  **ID** | **Sample**  **ID** | **Paternal**  **ID** | **Maternal**  **ID** | **Sex** | **Affection** | **marker 1** | **marker 2** | **marker 3** |
| FAM1 | NA06985 | 0 | 0 | 1 | 1 | A A | T T | A A |
| FAM1 | NA06991 | 0 | 0 | 1 | 1 | A A | T T | A A |
| 0 | NA06993 | 0 | 0 | 1 | 1 | C T | C C | T T |

* Map information (.map):

|  |  |  |  |
| --- | --- | --- | --- |
| **Chromosome** | **Marker ID** | **Genetic distance** | **Physical Position** |
| 1 | marker 1 | 0 | 157104 |
| 1 | marker 2 | 0 | 1947984 |

### 3.2.5 Binary PLINK (the header **should** be removed)

* Genotype data (.bed):  Please follow the instruction from [here](http://www.cog-genomics.org/plink2/formats" \l "bed)
* FAM file:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Family ID** | **Sample ID** | **Paternal ID** | **Maternal ID** | **Sex** | **Affection** |
| FAM1 | NA06985 | 0 | 0 | 1 | 1 |
| FAM1 | NA06991 | 0 | 0 | 1 | 1 |

* BIM file:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Chromosome** | **Marker ID** | **Genetic**  **distance** | **Physical**  **Position** | **Allele 1** | **Allele 2** |
| 1 | marker 1 | 0 | 157104 | A | C |
| 1 | marker 2 | 0 | 1947984 | A | T |

## 3.3 Covariates

* Demo format for a covariate file. The header is **required** to be provided:

|  |  |  |
| --- | --- | --- |
| **PC1** | **PC2** | **PC3** |
| -1.8942149 | -4.91532916 | 0.8674568 |
| 1.6858820 | -5.08378277 | -0.4069675 |
| 0.2579269 | -6.29547725 | 2.6867939 |

## 3.4 Kinship

* Demo format for a kinship file. Taxa name is **required** while the header can be omitted:

|  |  |  |  |
| --- | --- | --- | --- |
| **taxa** | **sample 1** | **sample 2** | **sample 3** |
| sample 1 | 2.00000000 | 0.22883683 | 0.22932180 |
| sample 2 | 0.22883683 | 2.00000000 | 0.24496455 |
| sample 3 | 0.22932180 | 0.24496455 | 2.00000000 |

* If there is no user-define kinship, a kinship will be generated by the selected package:

|  |  |
| --- | --- |
| **Package** | **Kinship algorithm** |
| GAPIT | VanRaden (*VanRaden, 2008*),  Loiselle (*Loiselle et al., 1995*)  or EMMA (*Kang et al., 2008*) |
| FarmCPU | FARM-CPU (*Liu et al., 2016*) |
| PLINK | Not available |
| rrBLUP | VanRaden (*VanRaden, 2008*) |
| BGLR | User-provided |

# 4. Incorporated packages

Tools implemented in iPat allow users to do genome-wide associate study (GWAS) and genomic selection (GS). Currently GWAS can be performed by GAPIT, FarmCPU and PLINK, and GS can be done by GAPIT, rrBLUP and BGLR in iPat. Tables below are the input arguments available in iPat:

## 4.1 GAPIT

|  |  |  |  |
| --- | --- | --- | --- |
| **Tab** | **Parameters** | **Definitions** | **Default** |
| Covariates | Covariate names | Subsetting covariates data | All covariates are selected |
| GAPIT input | Model | Which linear model to be used in GWAS | GLM |
| GAPIT input | kinship.cluster | Clustering algorithm to group individuals  based on their kinship | average |
| GAPIT input | kinship.group | Method to derive kinship among groups | Mean |
| Advance | SNP.fraction | Fraction of SNPs Sampled to Estimate  Kinship and PCs | 1 |
| Advance | File.fragment | The Fragment Size to Read Each Time within  a File | 512 |
| Advance | Model selection | Conduct Bayesian information criterion  (BIC)-based model selection to find the  optimal number of PCs for inclusion in the  GWAS models | FALSE |

## 4.2 FarmCPU

|  |  |  |  |
| --- | --- | --- | --- |
| **Category** | **Parameters** | **Definitions** | **Default** |
| Covariates | Covariate  names | Subsetting covariates data | All covariates are selected |
| FarmCPU  inpute | method.bin | It uses fixed or optimized of possible QTN window  size and number of possible QTNs selected into  FarmCPU model. | static |
| FarmCPU  inpute | maxLoop | Maximum number of iterations allowed | 10 |

## 4.3 PLINK

|  |  |  |  |
| --- | --- | --- | --- |
| **Category** | **Parameters** | **Definitions** | **Default** |
| Covariates | Covariate names | Subsetting covariates data | All covariates are selected |
| PLINK  input | C.I. | The desired coverage for a confidence interval | 0.95 |
| PLINK  input | Method | Regression methods of the study, available  options are "GLM" and "Logistic regression" | GLM |

## 4.4 rrBLUP

|  |  |  |  |
| --- | --- | --- | --- |
| **Category** | **Parameters** | **Definitions** | **Default** |
| Covariates | Covariate names | Subsetting covariates data | All covariates are selected |
| rrBLUP  input | Shrinkage  estimation | Shrinkage estimation can improve the accuracy of  genome-wide marker-assisted selection,  particularly at low marker density (*Endelman and Jannink 2012*) | TRUE |
| rrBLUP  input | impute.method | Imputation algorithm for missing values in  markers data | mean |

## 4.5 BGLR

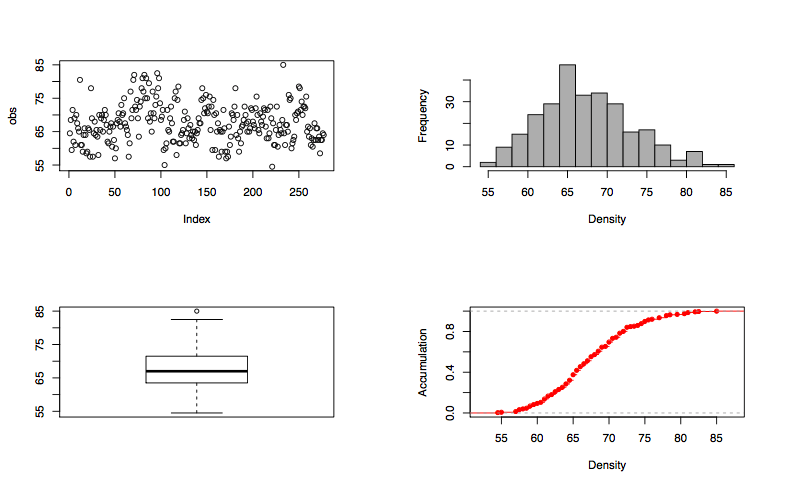
|  |  |  |  |
| --- | --- | --- | --- |
| **Category** | **Parameter** | **Definitions** | **Default** |
| Subset | Subset of traits data | Users can select all or partial of traits to  be analyzed | All traits |
| BGLR | Regression model for predictor (Markers) | The regression type for the markers data | BRR |
| BGLR | response\_type | Data type of the response (y) | gaussian |
| BGLR | nIter | The number of iterations of the sampler | 1200 |
| BGLR | burnIn | The number of samples discarded | 200 |
| BGLR | thin | The number of thinning | 5 |

# 5. Output files

## 5.1 Phenotype

### 5.1.1 Overview

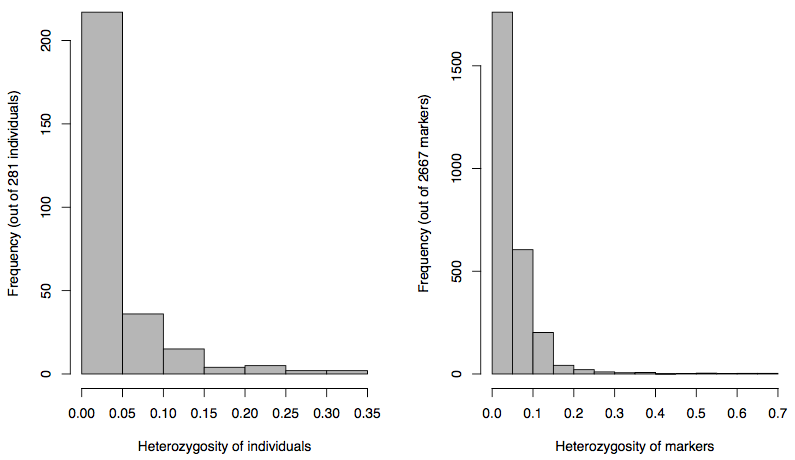
An overview of phenotype, including a scatter plot and histogram of the distribution. (suffix: \_phenotype\_view.pdf)



## 5.2 Population structure

### 5.2.1 Heterozygosity

Histograms that show the heterozygosity by individuals and by markers. (suffix: \_heterozygosity.pdf)



## 5.3 GWAS

### 5.3.1 Genomewise Manhattan plot

(suffix: Manhattan.Plot.Genomewise.pdf)

../../.Trash/GAPIT.Project_1_EarHT.Y...i..Manhattan.Plot.Genomewise.pdf

### 5.3.2 Q-Q plot

(suffix: QQ-Plot.pdf)

../../.Trash/GAPIT.Project_1_EarHT.Y...i..QQ-Plot.pdf

### 5.3.3 GWAS result

A table shows the marker information and its tested P-value. (suffix: \_GWAS.txt)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **SNP** | **Chromosome** | **Position** | **P.value** | **MAF** |
| PZB00859.1 | 1 | 157104 | 0.4636 | 0.2402 |
| PZA01271.1 | 1 | 1947984 | 0.1585 | 0.4893 |

## 5.4 GS

### 5.4.1 Genomic estimated breeding value (GEBV)

A scatter plot shows the correlation between phenotype and genomic estimated breeding value (GEBV) (suffix: \_GEBV\_value.pdf)

../../Desktop/test/tutorial/iPat_Project_1_dpoll_GEBV_value.pdf

### 5.4.2 Standard deviation of GEBV

A scatter plot shows the correlation between phenotype and SD of GEBV (suffix: \_GEBV\_var.pdf)

../../Desktop/test/tutorial/iPat_Project_1_dpoll_GEBV_var.pdf

### 5.4.3 Histogram of GEBV

(suffix: \_GEBV\_hist.pdf)

../../Desktop/test/tutorial/iPat_Project_1_dpoll_GEBV_hist.pdf

### 5.4.4 Prediction result

(suffix: \_EBV.dat)

|  |  |  |
| --- | --- | --- |
| **Taxa** | **Prediction** | **SD of prediction** |
| 33-16 | 36.576341 | 4.0006 |
| 38-11 | 38.401256 | 0.7429 |

# 6. Tutorial

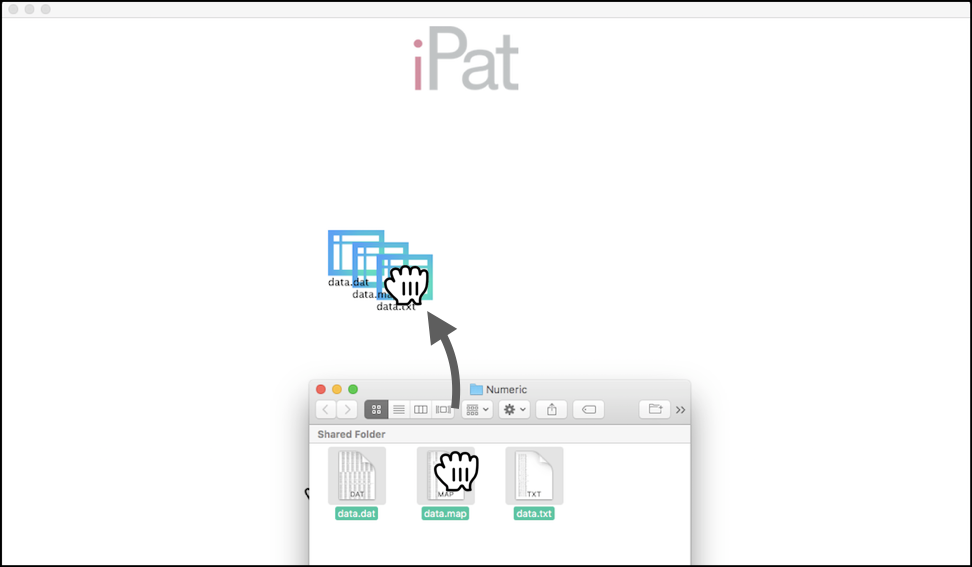
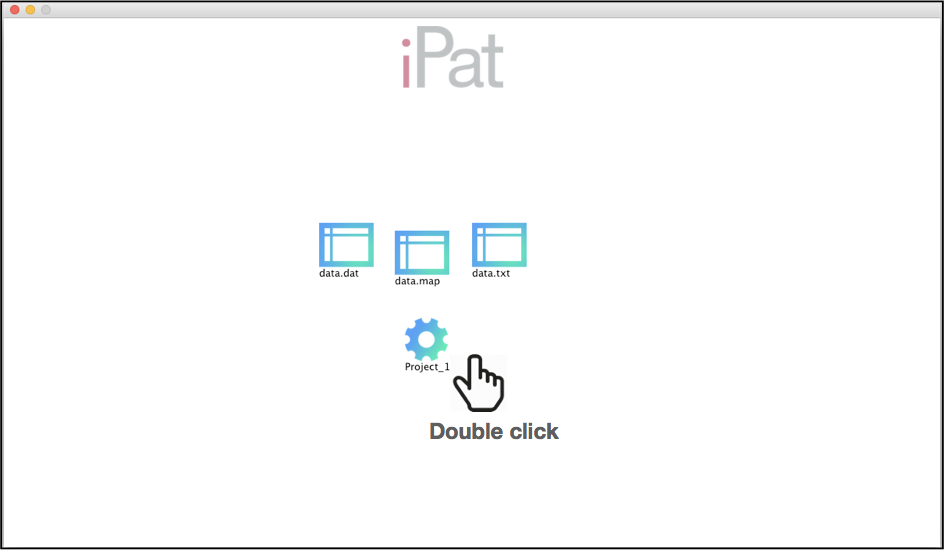
## 6.1 Perform GWAS in iPat

### **Data format** : Numerical (Columns as markers) **Data required** :

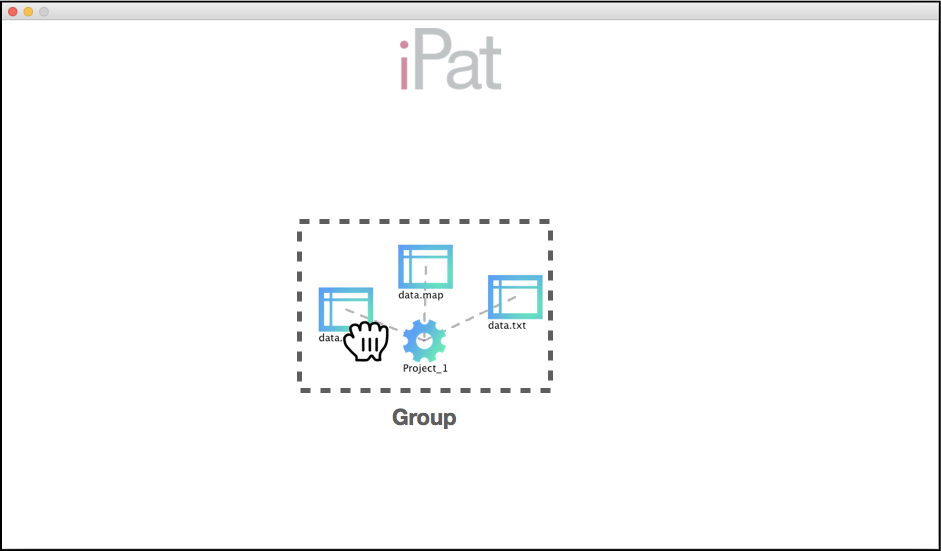
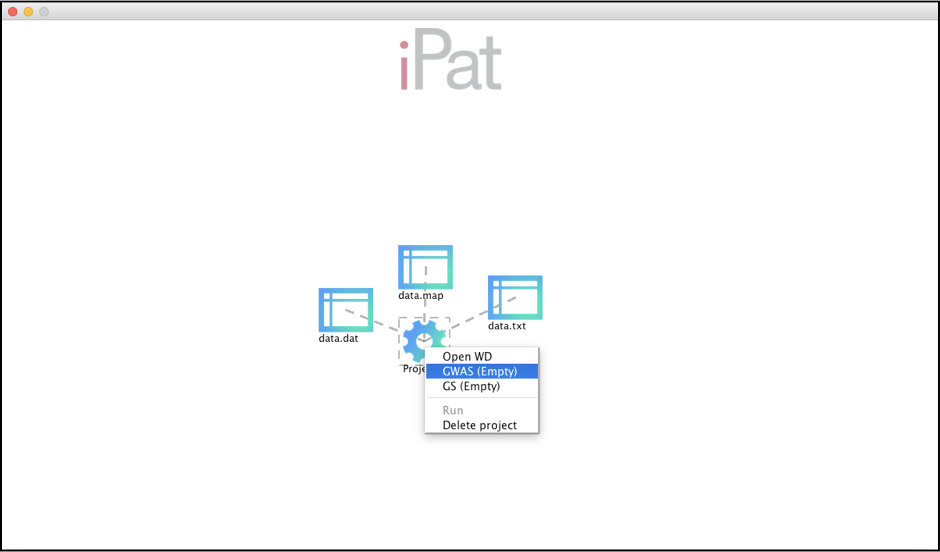
* Phenotype
* Genotype
* Map information

### **Implemented package**: FarmCPU

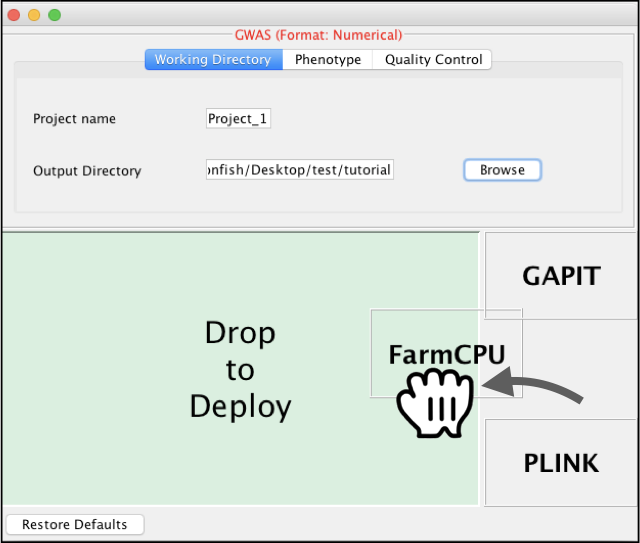
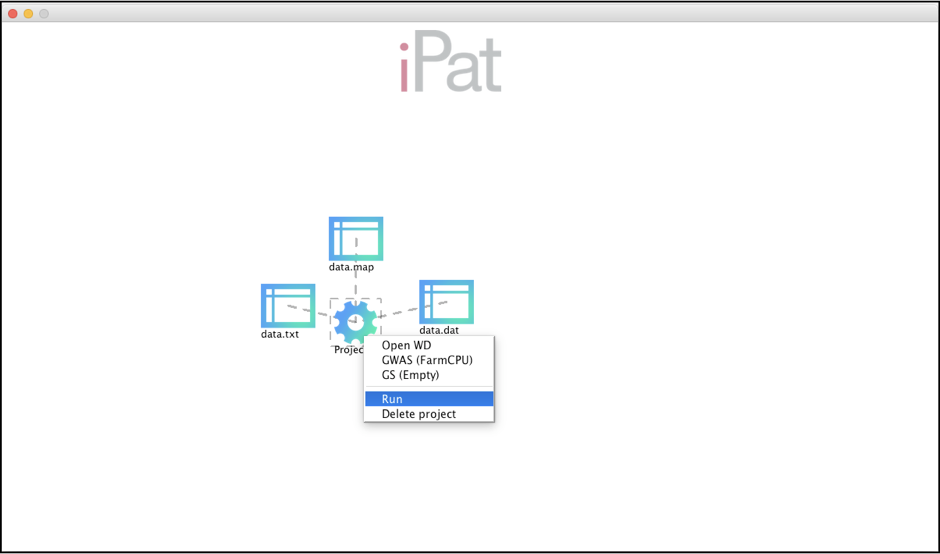
Step 1: Drag all the required files into iPat Step 2: Create a new project for files

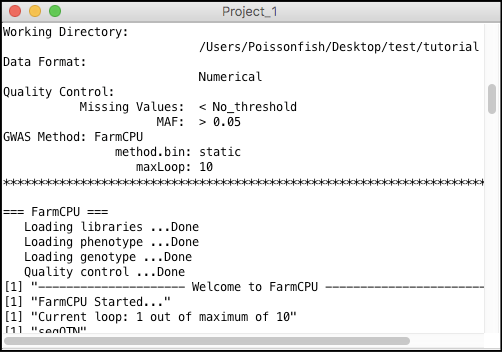
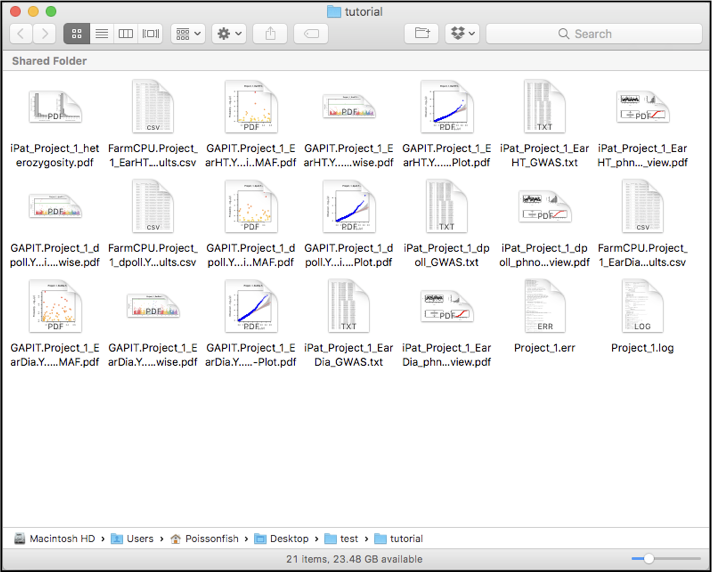
 

Step 3: Drag files and hover over the project to build a group Step 4: Right click on the project and choose "GWAS"

Step 5: Drag "FarmCPU" to the left, then close the window Step 6: Right click on the project and choose “Run”

Step 7: Double click on the project to inspect results after a finished computing 

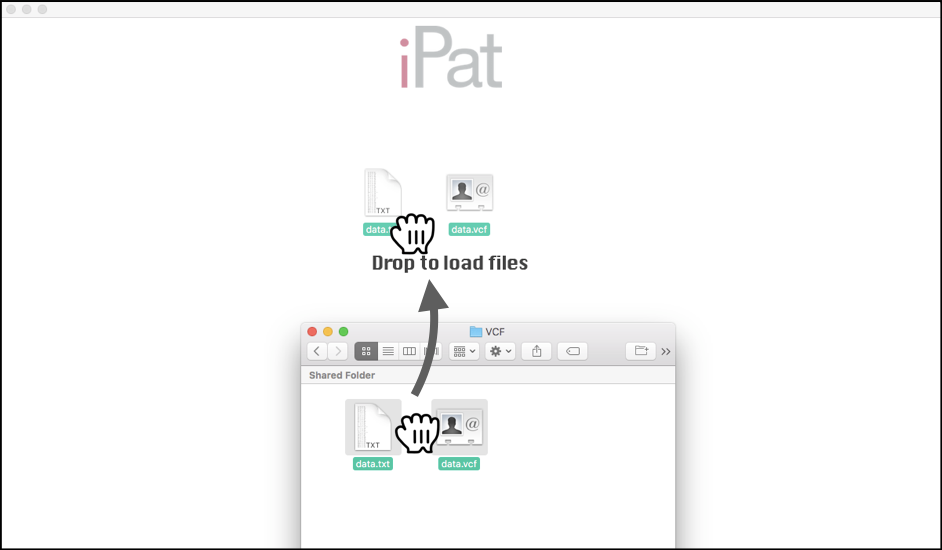
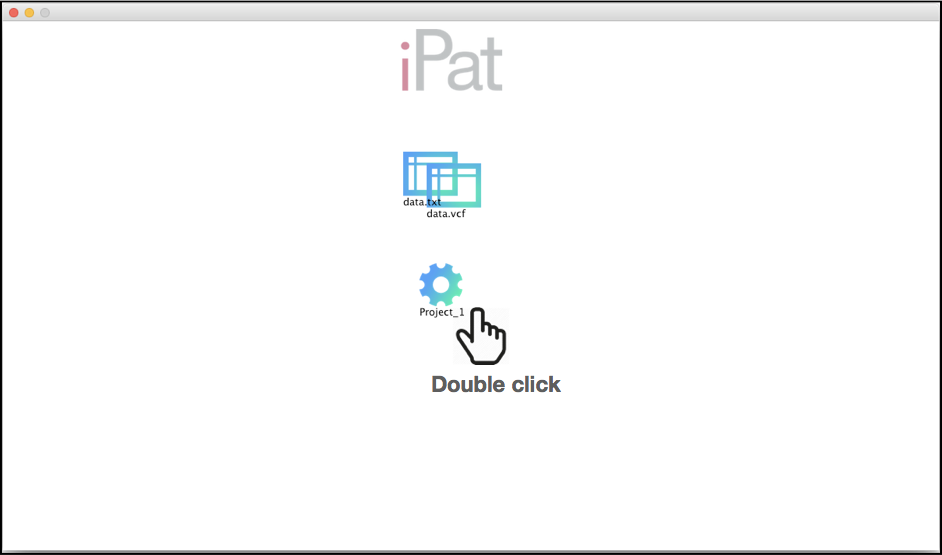
## 6.2 Perform GS and add user-define covariates in iPat

### **Data format** : VCF (Columns as markers) **Data required** :

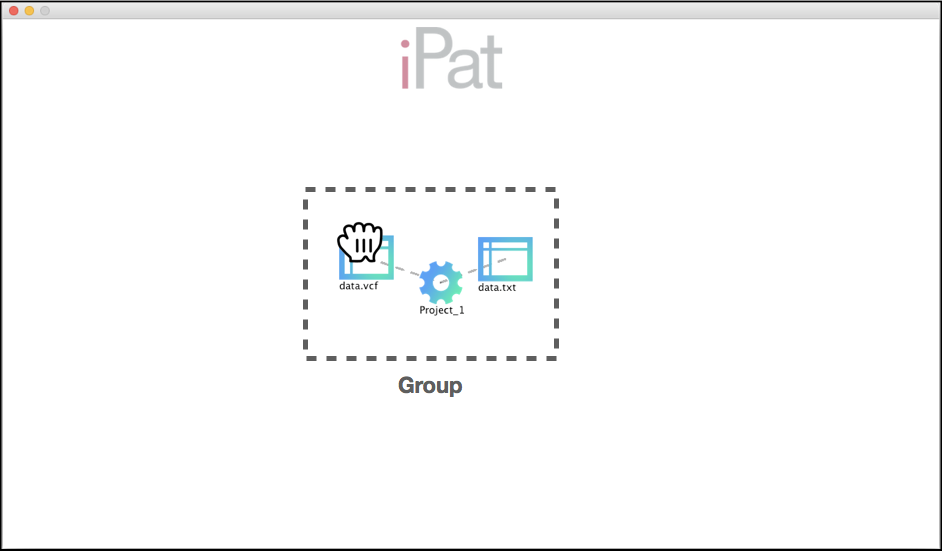
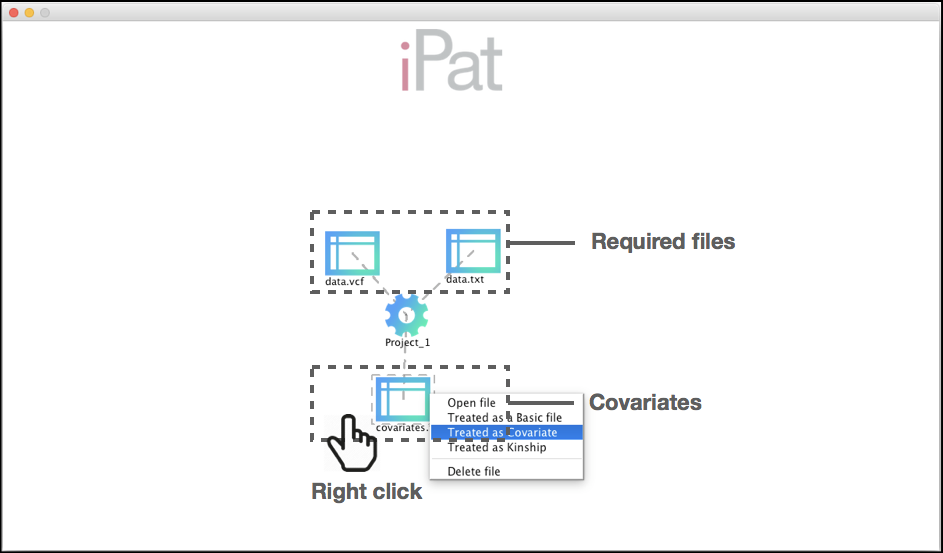
* Phenotype
* Genotype (.vcf)
* Covariates

### **Implemented package**: rrBLUP

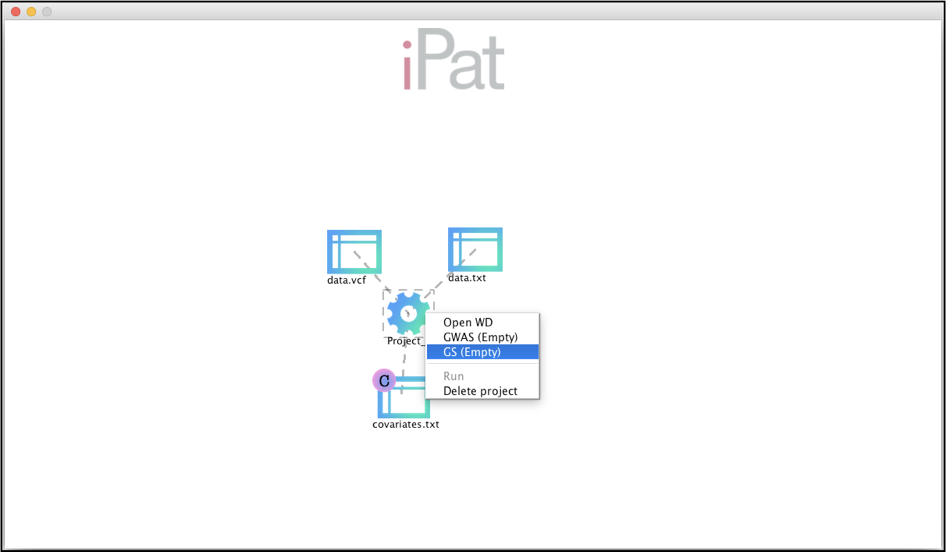
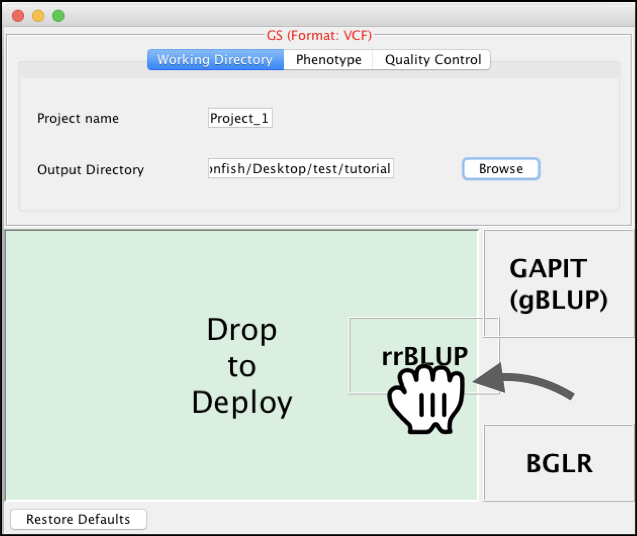
Step 1: Drag all the required files into iPat Step 2: Create a new project for files

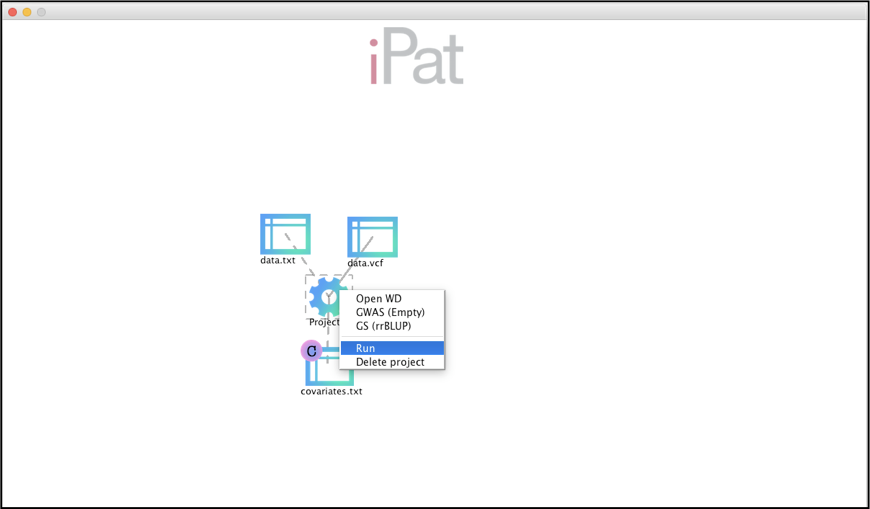
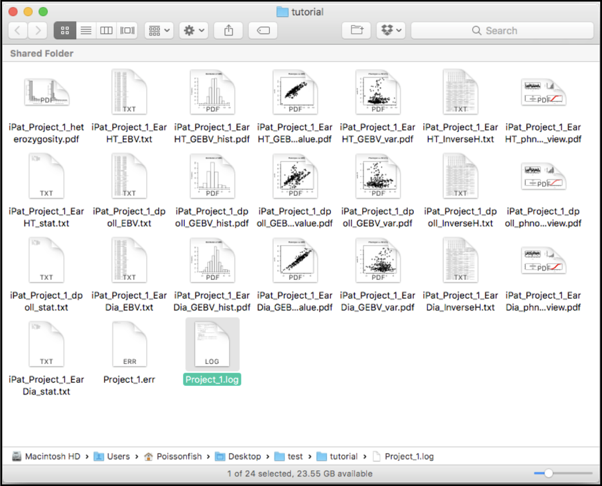
Step 3: Drag files and hover over the project to build a group Step 4: Add a covariates file and assign it as covariates

Step 5: Right click on the project and choose "GS" Step 6: Drag a label "GAPIT" to the left and close window

Step 7: Right click on the project and choose “Run” Step 8: Double click on the project to inspect results

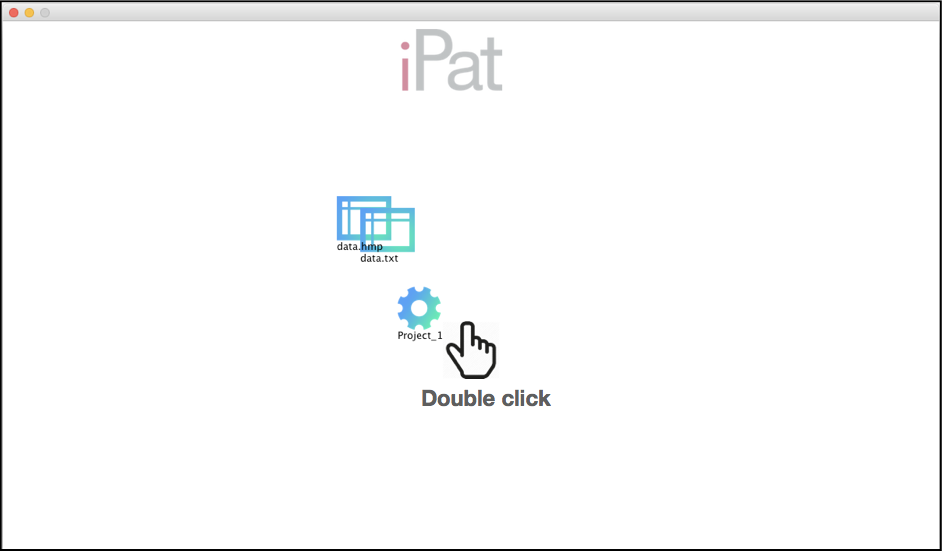
## 6.3 Perform GWAS-assist GS in iPat

### **Data format** : Hapmap **Data required** :

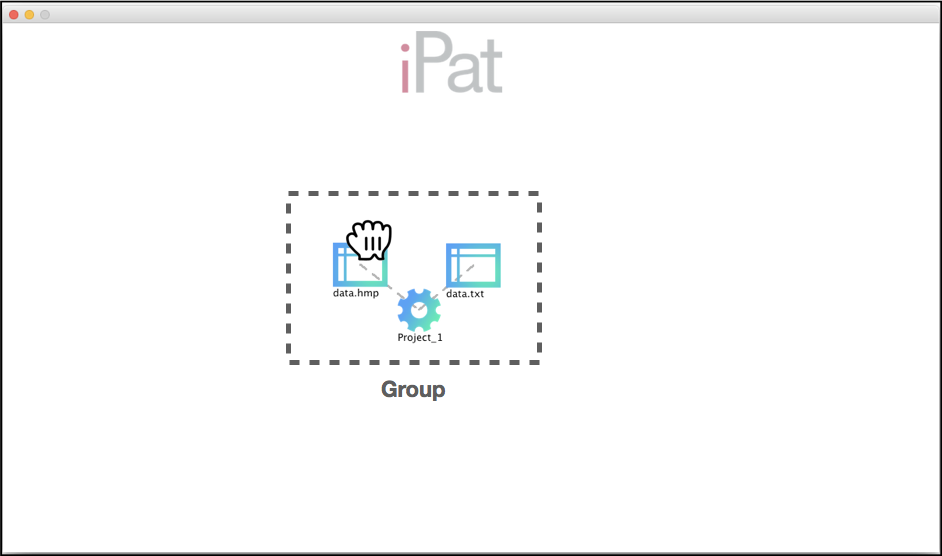
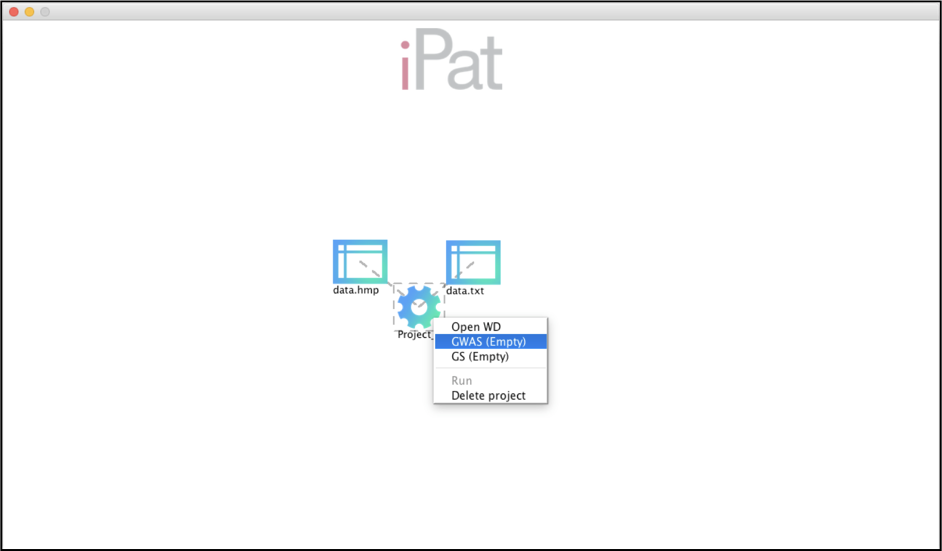
* Phenotype
* Genotype (.hmp)

### **Implemented package**: BGLR

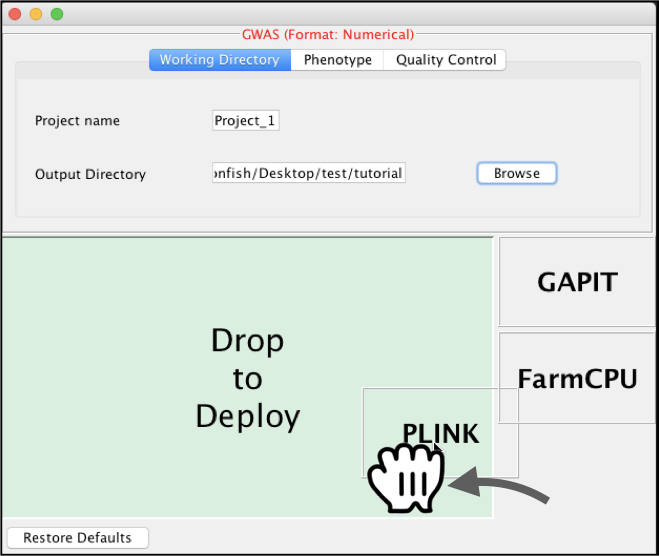
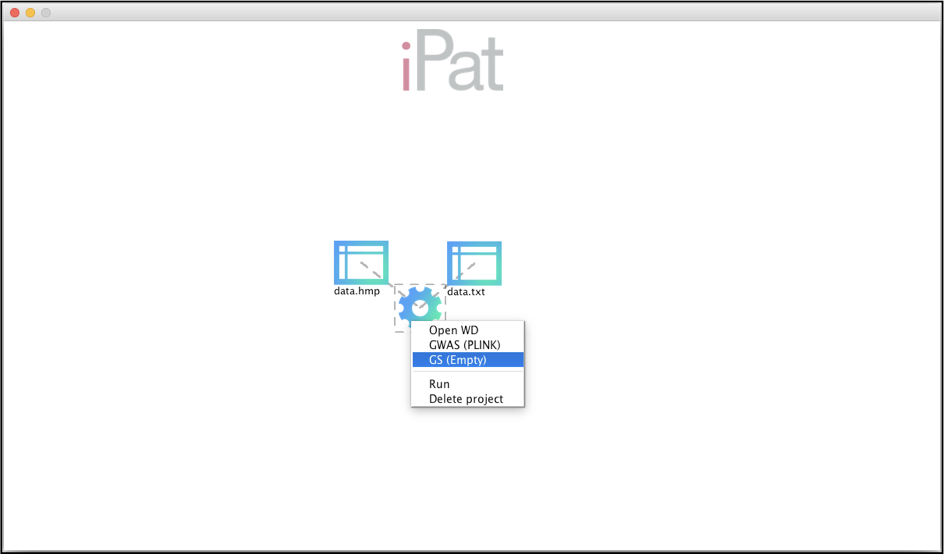
Step 1: Drag all the required files into iPat Step 2: Create a new project for files

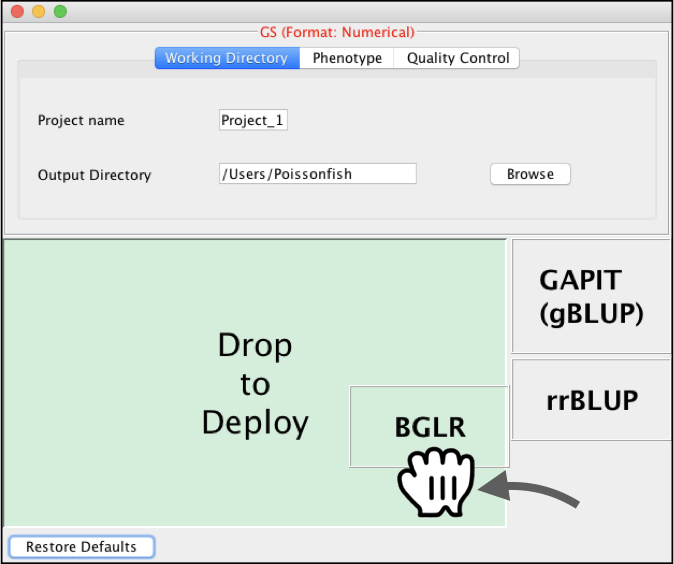
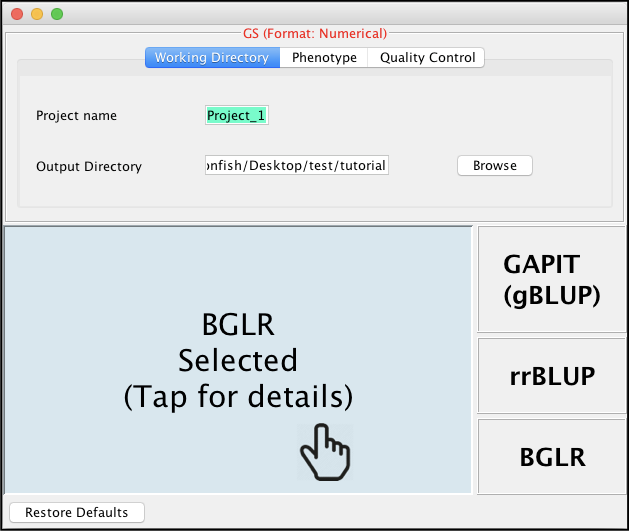
Step 3: Drag files and hover over the project to build a group Step 4: Right click on the project and choose "GWAS"

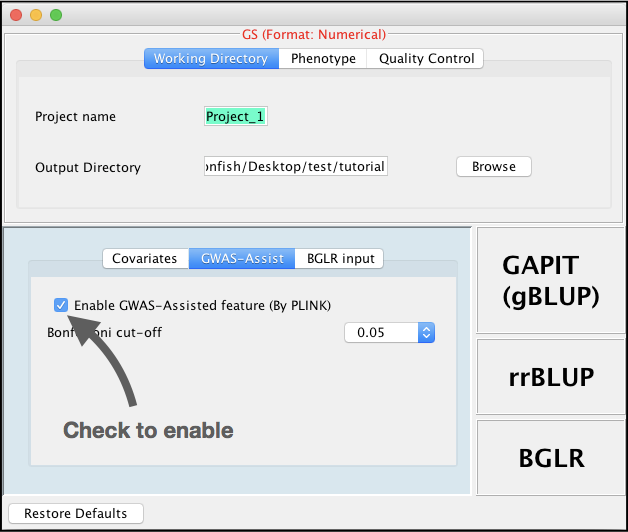
Step 5: Drag "PLINK" to the left, then close the window Step 6: Right click on the project and choose "GS"

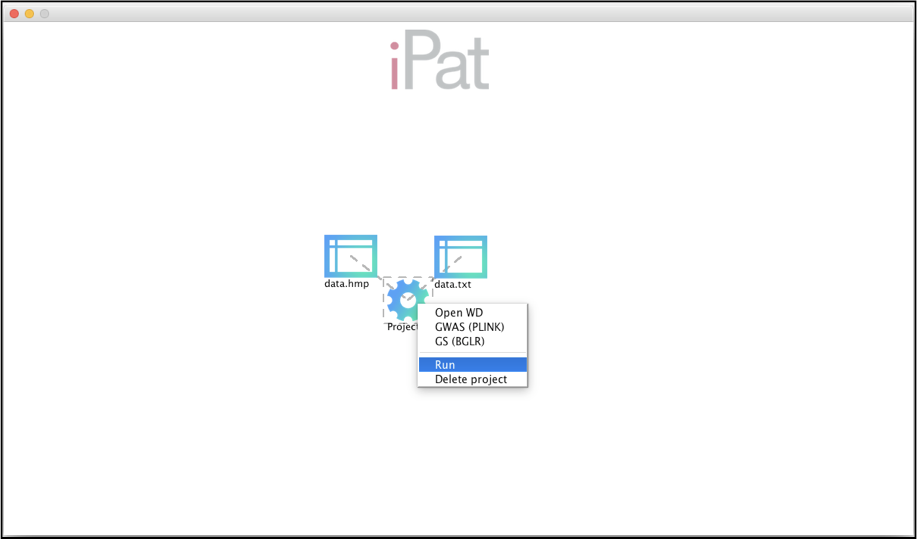
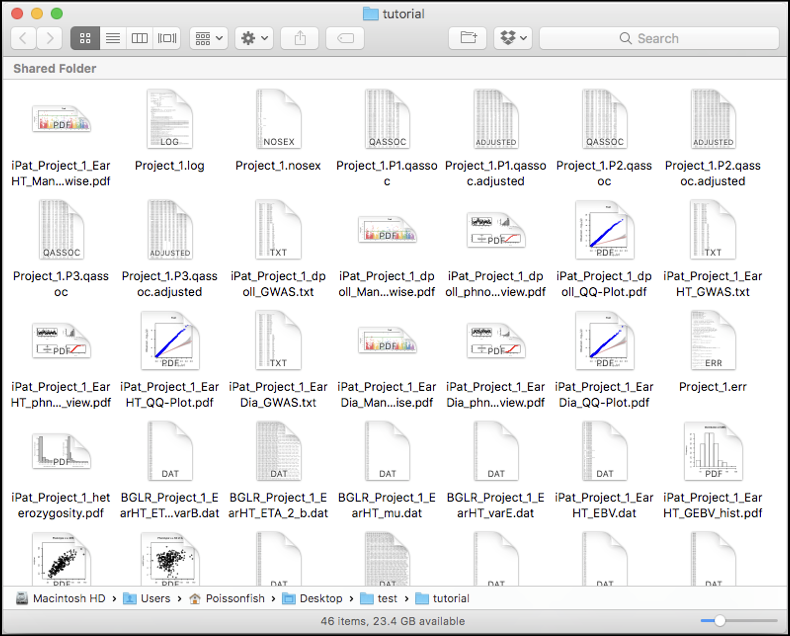
Step 7: Drag "BGLR" to the left, then close the window Step 8: Click on the left to further define GS

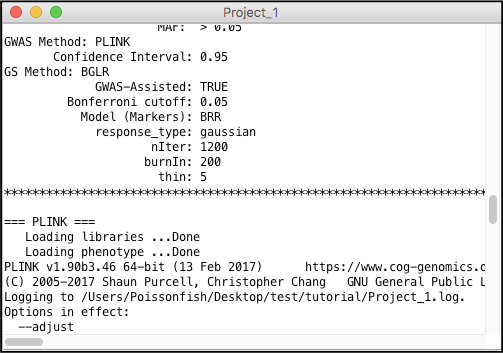
 

Step 9: Select “GWAS-Assist” tab and check “Enable GWAS-Assisted...” to enable GWAS-assisted GS feature



Step 10: Right click on the project and choose “Run” Step 11: Double click on the project to inspect results



# 7. References

Endelman J., 2011 Ridge regression and other kernels for genomic selection in the R package rrBLUP. Plant Genome 4: 250–255.

Lipka A. E., Tian F., Wang Q., Peiffer J., Li M., *et al.*, 2012 GAPIT: genome association and prediction integrated tool. Bioinformatics 28: 2397–2399.

Liu X., Huang M., Fan B., Buckler E. S., Zhang Z., 2016 Iterative Usage of Fixed and Random Effect Models for Powerful and Efficient Genome-Wide Association Studies. PLoS Genet. 12: e1005767.

Pérez P., Los Campos G. De, 2014 Genome-wide regression and prediction with the BGLR statistical package. Genetics 198: 483–495.

Purcell S., Neale B., Todd-Brown K., Thomas L., Ferreira M. A. R., *et al.*, 2007 PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet 81: 559–575.

Tang Y., Liu X., Wang J., Li M., Wang Q., *et al.*, 2016 GAPIT Version 2: An Enhanced Integrated Tool for Genomic Association and Prediction. Plant J. 9.