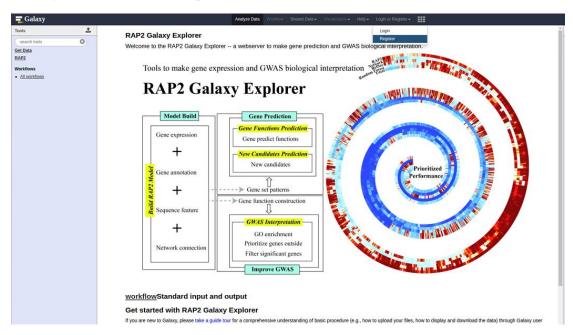
Welcome to RAP2 Galaxy Explorer

This user manual guides you for basic and professional application step by step.

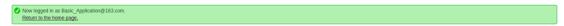
1) Register in RAP2 Galaxy Explorer



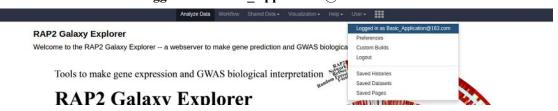
After select Register button, you will go into the register form:



Then click **Return to the home page**:



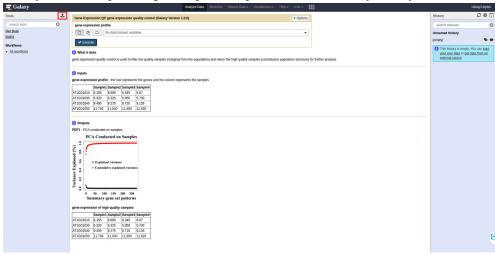
Select the User menu to Logged in as Basic_Application@163.com:



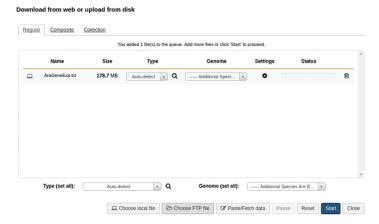
After the registration, you can use the RAP2 Galaxy Explorer smoothly.

- 2) For basic application, we provide four tools: Gene Expression QC, Gene Functions Prediction, New Candidates Prediction, and GWAS Interpretation.
 - a. Select *Gene Expression QC* to filter low quality samples (marginal from the population) and return the high quality samples (consistance population structure) for further analysis.

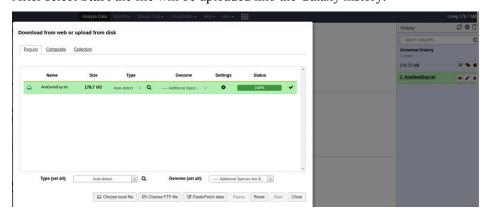
First, you should upload a gene expression profile into the Galaxy history:



After click the **upload button** (red rectangle), you can Choose local file (e.g., AraGeneExp.txt can be download from https://github.com/cma2015/RAP2/tree/master/exampleData):



After select **Start** the file will be uploaded into the Galaxy history:



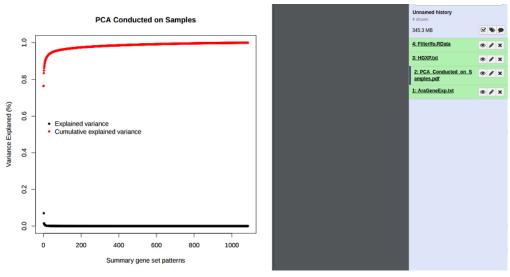
Then, select the file in the **gene expression profile** window from Galaxy history:



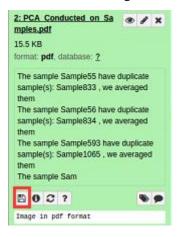
After click **Execute**, the *Gene Expression QC* will be work:



These three files will be output. You can view the output files through the eye button:



And download through the **disk button** (red rectangle):

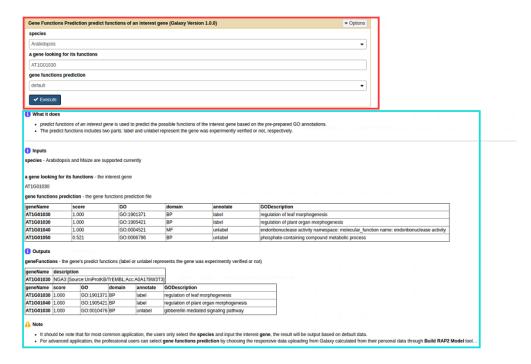


b. Select *Gene Functions Prediction* to predict functions of an interested gene.

First, you should select the **species** (e.g., *Arabidopsis*),

Second, you should input a gene (e.g., AT1G01030),

For basic application, keep the **gene function prediction** by default (generated from our data):



The red rectangle is the main run window, the blue rectangle is the help window. After click **Execute** button, the *Gene Functions Prediction* will be work:

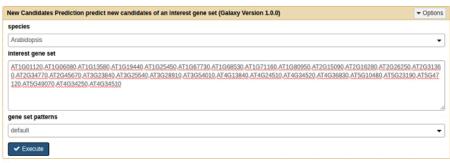


c. Select *New Candidates Prediction* to predict new candidates of an interest gene set.

First, you should select the **species** (e.g., *Arabidopsis*),

Second, you should input a gene set (e.g., AT1G01120,AT1G06080,AT1G13580,AT1G19440,AT1G25450,AT1G67730,AT1G68530,AT1G71160,AT1G80950,AT2G15090,AT2G16280,AT2G26250,AT2G31360,AT2G34770,AT2G45670,AT3G23840,AT3G25540,AT3G28910,AT3G54010,AT4G13840,AT4G24510,AT4G34520,AT4G36830,AT5G10480,AT5G23190,AT5G47120,AT5G49070,AT4G34250,AT4G34510),

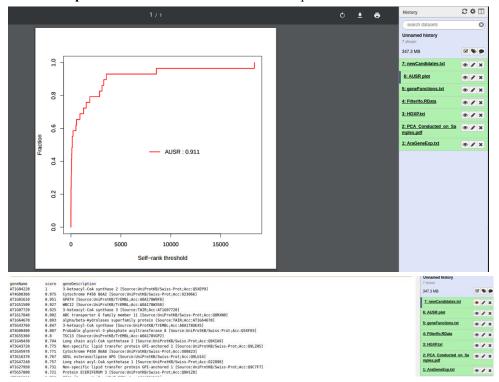
For basic application, keep the **gene set patterns** by default (generated from our data):



After click Execute button, the New Candidates Prediction will be work:



The AUSR plot and newCandidates.txt will be output:



d. Select *GWAS Interpretation* to improve biological interpretation of GWAS.

First, you should upload a gene expression profile into the Galaxy history (e.g., SNPMapGene.txt can be download from https://github.com/cma2015/RAP2/tree/master/exampleData),

Then, select the file in the **GWAS data** window from Galaxy history,

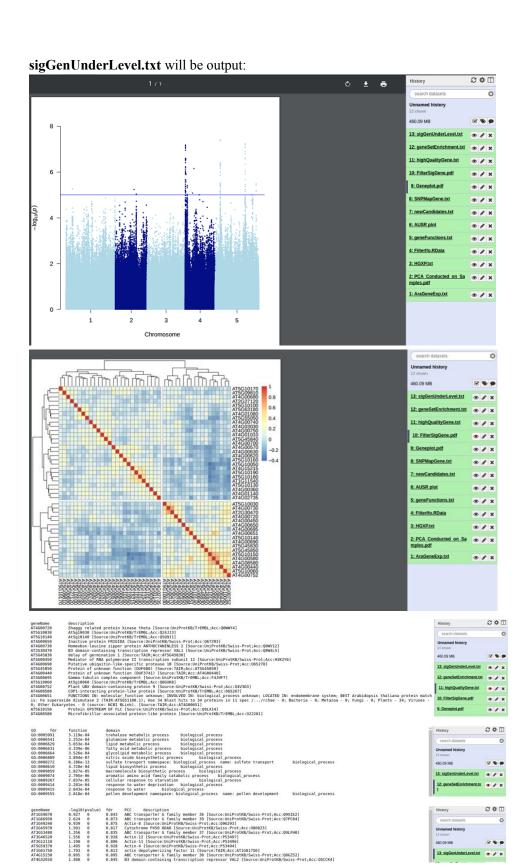
For basic application, keep the **gene function construction** by default (generated from our data):



After click Execute button, the GWAS Interpretation will be work:



The Geneplot.pdf, FilterSigGene.pdf, highQualityGene.txt, geneSetEnrichment.txt,



3) For advanced application, we provide four tools: *Build RAP2 Model*, *Gene Functions Prediction*, *New Candidates Prediction*, and *GWAS Interpretation*. It should be note that, after build RAP2 model based on your personal data to fit for your rigorous demand, the following application is the same with those in the **basic application** by just selecting the corresponding

parameters upload from Galaxy history instead of default.

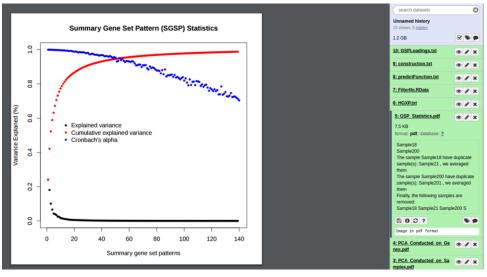
a. Select *Build RAP2 Model* to build RAP2 model based on your personal data. We also provide default data to help you understand what we used to build our RAP2 model. First, upload the gene expression and GO data, then select them in the Galaxy history (we used the Maize as an example, the MaizeGeneEXP.txt and MaiGO.txt can be download from https://github.com/cma2015/RAP2/tree/master/exampleData):



After click Execute, the Build RAP2 Model will be work:



These eight files will be output. You can view the output files through the eye button:

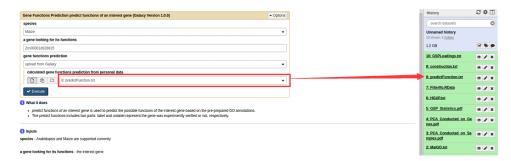


b. Select *Gene Functions Prediction* to predict functions of an interested gene.

First, you should select the species (e.g., Maize),

Second, you should input a gene (e.g., Zm00001d028815),

For advanced application, select the **gene function prediction** from Galaxy history (generated from your personal data):



After click **Execute** button, the *Gene Functions Prediction* will be work:



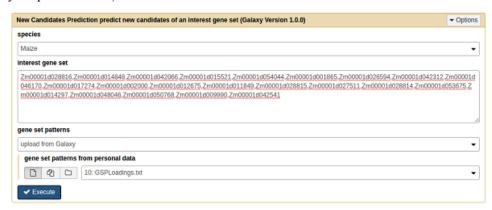
c. Select New Candidates Prediction to predict new candidates of an interest gene set.

ion label label label label label

First, you should select the species (e.g., Maize),

Second, you should input a gene set (e.g., Zm00001d028816,Zm00001d014848,Zm00001d042066,Zm00001d015521,Zm00001d05 4044,Zm00001d001865,Zm00001d026594,Zm00001d042312,Zm00001d046170,Zm000 01d017274,Zm00001d002000,Zm00001d012675,Zm00001d011849,Zm00001d028815,Z m00001d027511,Zm00001d028814,Zm00001d053675,Zm00001d014297,Zm00001d048 046,Zm00001d050768,Zm00001d009990,Zm00001d042541),

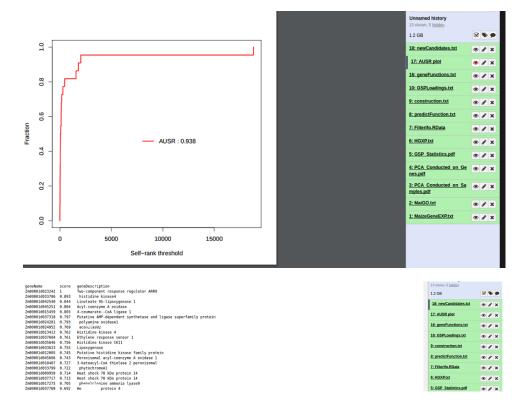
For advanced application, select the **gene set patterns** from Galaxy history (generated from your personal data):



After click Execute button, the New Candidates Prediction will be work:

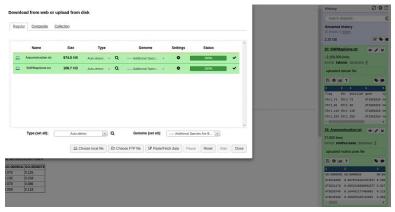


The AUSR plot and newCandidates.txt will be output:



Note:

- (i) It should be note that the maize gene expression data contains 279 samples of 36,580 genes, the maize annotation data contains 97 GO terms of 168 genes (download from https://github.com/cma2015/RAP2/tree/master/exampleData). It takes about 3 hours to build the RAP2 model. While for *Arabidopsis*, gene expression data contains 1,096 samples of 20,356 genes, the annotation data contains 2623 GO terms of 9509 genes. It takes about 3 days to build the RAP2 model. For efficient guide, we only provide how to build RAP2 model on the maize data, it is the same process with *Arabidopsis*.
- (ii) Since the experimentally verified genes are not large enough for improving biological interpretation of GWAS. Currently, we only support Arabidopsis for this tool.
- (iii) For efficient guide, we upload the GWAS data (e.g., SNPMapGene.txt) and the gene function construction (e.g., Araconstruction.txt) generated from our default data (the SNPMapGene.txt and Araconstruction.txt can be download from https://github.com/cma2015/RAP2/tree/master/exampleData).



d. Select *GWAS Interpretation* to improve biological interpretation of GWAS.

First, you should upload a gene expression profile into the Galaxy history (e.g., SNPMapGene.txt can be download from https://github.com/cma2015/RAP2/tree/master/exampleData),

Then, select the file in the GWAS data window from Galaxy history,

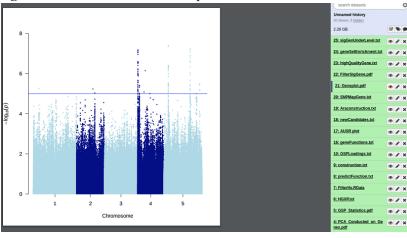
For advanced application, select the **gene function construction** from Galaxy history (generated from your personal data):

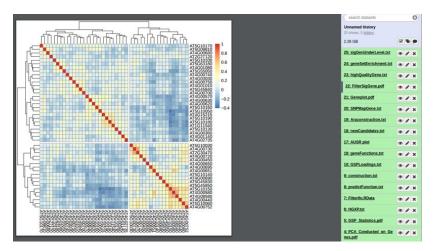


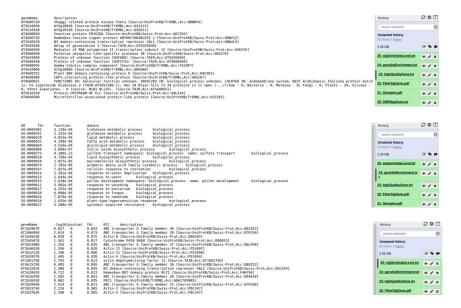
After click **Execute** button, the *GWAS Interpretation* will be work:



The Geneplot.pdf, FilterSigGene.pdf, highQualityGene.txt, geneSetEnrichment.txt, sigGenUnderLevel.txt will be output:







Thanks for reading this user manual for basic and advanced application! Enjoy RAP2 Galaxy Explorer!