



ImmPort Genetic Analysis Tool (IGAT)

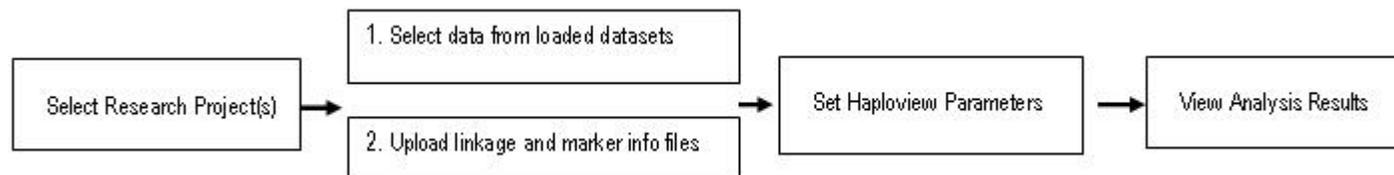
The ImmPort Genetic Analysis Tool (IGAT) allows users to analyze their genotype data. The algorithms implemented in IGAT are adapted from the Haploview algorithm developed by Mark Daly's laboratory at the Broad Institute. More algorithms for genetic and association studies will be added in future releases. This tutorial is an adaptation of the Haploview documentation (<http://www.broad.mit.edu/mpg/haploview/>).

Haploview is designed to simplify and expedite the process of haplotype analysis by providing a common interface to several tasks relating to such analyses. ImmPort currently supports the following functionalities adapted from Haploview:

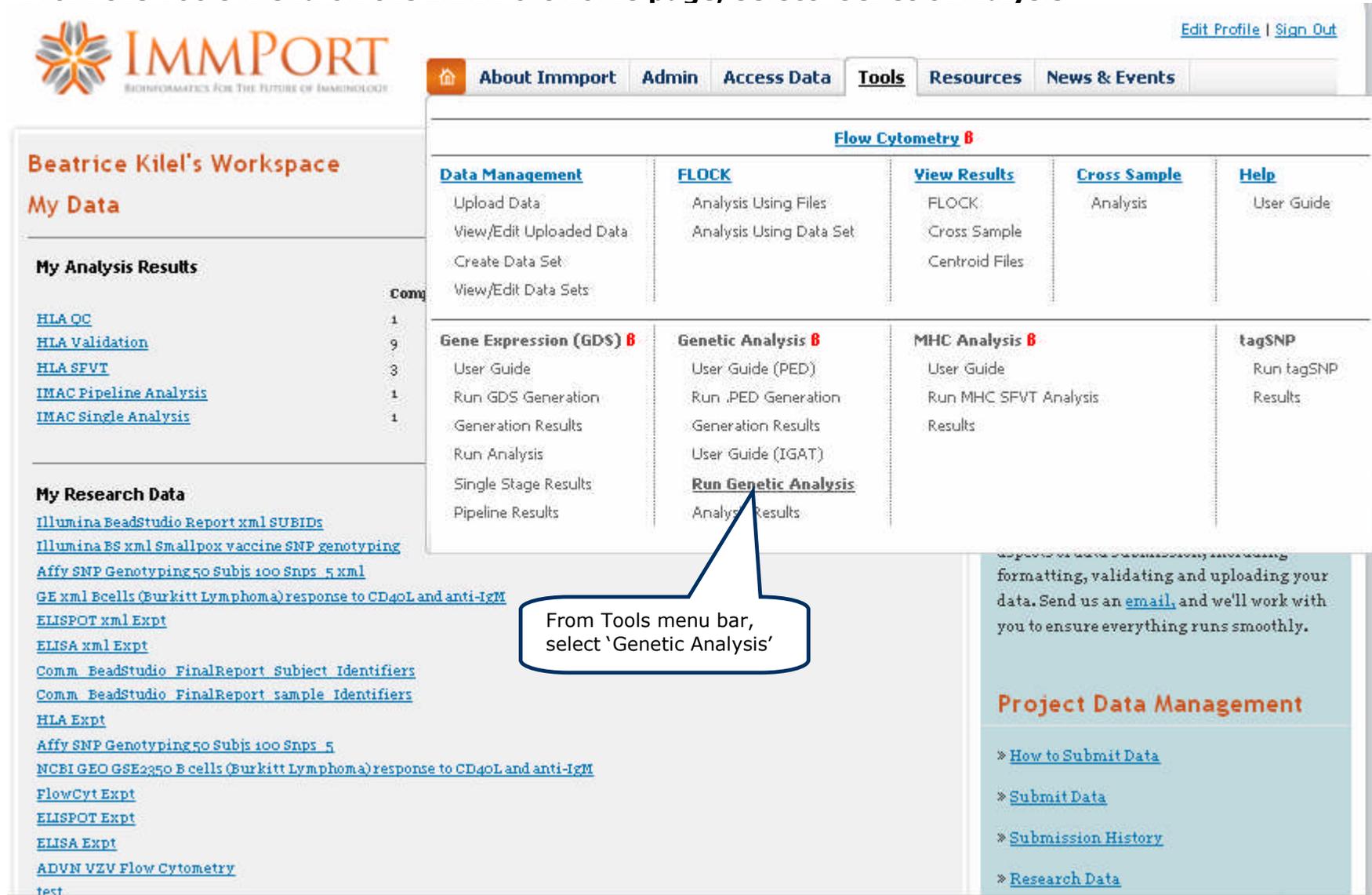
- Linkage Disequilibrium (LD) and haplotype block analysis
- Haplotype population frequency estimation
- Single SNP and haplotype association tests on population-based case/control studies
- Permutation testing for association significance

For any questions or comments using IGAT, contact us at helpdesk@import.org.

The schematic in the next slide indicates the flow of Genetic Analysis activities: -



From the Tools menu on the ImmPort home page, select 'Genetic Analysis'



Beatrice Kilel's Workspace
My Data

My Analysis Results

HLA QC	1
HLA Validation	9
HLA SFVT	3
IMAC Pipeline Analysis	1
IMAC Single Analysis	1

My Research Data

- [Illumina BeadStudio Report xml SUBIDs](#)
- [Illumina BS xml Smallpox vaccine SNP genotyping](#)
- [Affy SNP Genotyping 50 Subs 100 Snps 5 xml](#)
- [GE xml Bcells \(Burkitt Lymphoma\) response to CD40L and anti-IgM](#)
- [ELISPOT xml Expt](#)
- [ELISA xml Expt](#)
- [Comm BeadStudio FinalReport Subject Identifiers](#)
- [Comm BeadStudio FinalReport sample Identifiers](#)
- [HLA Expt](#)
- [Affy SNP Genotyping 50 Subs 100 Snps 5](#)
- [NCBI GEO GSE2350 B cells \(Burkitt Lymphoma\) response to CD40L and anti-IgM](#)
- [FlowCyt Expt](#)
- [ELISPOT Expt](#)
- [ELISA Expt](#)
- [ADVN VZV Flow Cytometry test](#)

Tools Menu:

- Flow Cytometry**
 - Data Management: Upload Data, View/Edit Uploaded Data, Create Data Set, View/Edit Data Sets
 - FLOCK: Analysis Using Files, Analysis Using Data Set
 - View Results: FLOCK, Cross Sample, Centroid Files
 - Cross Sample: Analysis
 - Help: User Guide
- Gene Expression (GDS)**
 - User Guide
 - Run GDS Generation
 - Generation Results
 - Run Analysis
 - Single Stage Results
 - Pipeline Results
- Genetic Analysis**
 - User Guide (PED)
 - Run .PED Generation
 - Generation Results
 - User Guide (IGAT)
 - Run Genetic Analysis**
 - Analysis Results
- MHC Analysis**
 - User Guide
 - Run MHC SFVT Analysis
 - Results
- tagSNP**
 - Run tagSNP Results

Project Data Management

- » [How to Submit Data](#)
- » [Submit Data](#)
- » [Submission History](#)
- » [Research Data](#)



ImmPort Genetic Analysis Tool (IGAT)

Select the Research Project

The first step in using tools in the IGAT is to choose a research project with which the analysis input data and result data can be associated. Once a project is chosen, Haploview can be run using standard linkage format files and associated marker info files. The data can be selected from the custom upload option or from previously loaded datasets.

ImmPort offers the option to parse genotype data loaded into ImmPort and generate standard linkage files and marker info files. This allows you to have the convenience of running Haploview directly from your genotype experiment results.

ImmPort Genetic Analysis Tool (IGAT)

The ImmPort Genetic Analysis Tool analyzes genotype data using the [Haploview](#) software developed by Mark Daly's laboratory at the Broad Institute.

This software is used to perform

- LD & haplotype block analysis
- haplotype population frequency estimation
- single SNP and haplotype association tests
- permutation testing for association significance

You will have the option of either analyzing an existing linkage file or creating linkage file from genotype data stored in Immport.

Select a research project below and then follow directions on the following pages.

Research Project(s) 

Next

Click 'Next' to
run Haploview

Step 1: Choose or Load Datasets

Workflow 1: Select data from previously loaded datasets

ImmPort Genetic Analysis Tool (IGAT)

Research Project: Packages4

The ImmPort implementation of Haploview accepts [linkage files](#) (.ped) and [marker info files](#) (.info) as Input Data Sets. Click [here](#) to view a detailed description of the two formats

Please make a selection from the options below. One option is to use an existing ImmPort dataset for Haploview analysis. Another is to upload a dataset from your computer. You can create Haploview input dataset (linkage and marker info files) from ImmPort geno-typing results by using [PED file generation tool](#). Please note that currently only genotype data in the format of Illumina BeadStudio Final report or Affymetrix genotype .txt file are available for conversion to linkage and marker files.

Workflow Selection:

- Run Haploview using existing Linkage File and Marker Info file stored in ImmPort.
- Upload Linkage File and Marker Info File and Run Haploview analysis.

Existing Input Data Sets

Research Project: Packages4

Select an existing data set and click Next. You may also click on a Data Set Accession to see details about a data set.

3 items found, displaying all items.

	Data Set Accession	Data Source	Data Set Name	Description	Creation Date
<input checked="" type="radio"/>	HDS411	import_created	December 08 dataset ped	December 08 dataset ped 12	2008-12-11
<input type="radio"/>	HDS556	import_created	tt 2245	test 1	2009-02-17
<input type="radio"/>	HDS361	custom_upload	rene again	rene again	2008-07-28

3 items found, displaying all items.

Click 'Next' to display the Existing Input Data Sets

Click 'Next' to view dataset details

Download and view details of downloaded data sets

Data Set Details

Data Set Accession No: HDS411
Data Set Name: December 08 dataset ped
Description: December 08 dataset ped 12
Research Project: Packages4
Source: Import created dataset

To view the linkage or marker files click download. To continue with the analysis submission, click 'Next' to set the Haploview parameters

The files associated with the dataset are:

Linkage file: December 08 dataset ped .ped [Download](#)
Marker file: December 08 dataset ped .info [Download](#)
Missing Markers: MissingMarkers.info [Download](#)

Markers that did not satisfy all criteria for .ped file generation. Please see user guide for further details.

- Continue on to Haploview Analysis
- Edit the linkage file by uploading subject affection status. Your edit will be based on the latest affection status if the linkage file has been updated previously.

Click "Next" to continue with the analysis submission.

Click 'Next' to set Haploview parameters



ImmPort Genetic Analysis Tool (IGAT)

Setting the Haploview Parameters

Further options are presented on the Parameter Setting page.

Haploview saves time by only computing pairwise LD statistics for markers within a certain distance of each other. The default is 500KB. Enter a value of zero to force all pairwise computations.

Haploview excludes individuals with less than 50% complete genotypes. This threshold can be adjusted.

The program needs to know which algorithm(s) to use to calculate haplotype blocks, as well as cutoff parameters associated with each algorithm.

Please choose a color scheme(s) for LD plot display.

Choose allele display mode(s) for haplotype display.

In order to perform association tests, you **MUST** indicate whether you want to perform single marker association tests or haplotype association tests.

You **MUST** also specify how to perform permutation tests and how many permutations are needed for the program to run.

Set the Haploview Parameters

Set parameters from the given Haploview list. Then view the results. If no parameters are selected, a pop-up window is displayed asking the user to select the minimum parameters: Block definitions, LD plot color scheme and Allele display mode.

ImmPort Genetic Analysis Tool (IGAT)

Select Research Project | Choose Workflow | Choose Data Set | Edit Affection Status | **Select Haploview Parameters**

Select Haploview Parameters

Input File

Ignore pair wise comparison of markers >: 500 Kb apart
 Exclude individuals with >: 50 % missing genotypes

Blocks Definitions

Confidence Intervals (Gabriel et al.)
 Four Gamete Rule
 Solid Spine of LD

LD Plot Color Scheme

4th Gamete
 Confidence Bounds
 R Squared
 Standard D' / LOD
 All D' / LOD
 GOLD

Alleles Display Mode

Alleles as colored squares
 Alleles as letters
 Alleles as numbers

Association Tests

Association Tests (Single Marker)
 Association Tests (Haplotypes)

Permutation Tests

Perform Permutations for Association Tests Single Marker
 Perform Permutations for Association Tests Single Markers and Haplotypes in Blocks
 Number of permutation tests: 1000 (a large number may take a long time to run)

Blocks

Block Definition

Gabriel et al.

Upper confidence interval value: 0.98
 Lower confidence interval value: 0.7
 Upper confidence interval max for strong recombination value: 0.9
 Min fraction of strong ld in info comparison: 0.95
 Exclude marks below: 0.05 MAF

4th gamete rule

Fourth gamete frequency min: 0.01

Strong SD Spine

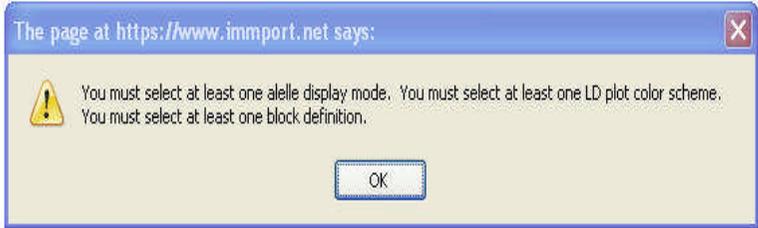
Extend spine id D prime: 0.8

Block Display

Lines Display

Thick line threshold: 10 %
 Thin line threshold: 1 %

Submit | Cancel



Set the Haploview parameters of interest and click submit. Ref: <http://www.broad.mit.edu/mpg/haploview/>

View Analysis Results from the selected parameters

Select Research Project
Choose Workflow
Choose Data Set
Edit Affection Status
Select Haploview Parameters

View Analysis Results

Your Genotyping Analysis request was submitted successfully. Your task id is **4587**.
Click the "Show Analysis History" button to proceed to the IGAT Analysis History page.

Show Analysis History

View Analysis History and the Task Id for genotyping analysis detailed information

IGAT Analysis History

This page is a list of the tasks you have submitted to the ImmPort Genetic Analysis Toolset. The history list can show up to 50 tasks. If the number exceeds 50, the oldest tasks will be dropped off the list. Click on a task ID to view more information about the task or to see the results when the task has completed execution.

126 items found, displaying 1 to 20. [First/Prev] 1, 2, 3, 4, 5, 6, 7 [Next/Last]

Task Id	Start Date	End Date	Task Type	Data Set Accession/Name	Dataset Source	Algorithm	Status
4587	05/04/2009 10:34:10	05/04/2009 10:34:21	Run_Haploview	HDS411 / December 08 dataset ped	import_created	Haploview	Completed with error
4584	05/01/2009 23:07:54	05/01/2009 23:07:55	Create_PED_File	HDS620 / ZP test PED 2PF	import_created	n/a	Completed
4239	05/01/2009 11:46:21	05/01/2009 11:46:42	Run_Haploview	HDS617 / New subject affection status	import_created	Haploview	Completed
4235	05/01/2009 09:09:11	05/01/2009 09:09:19	Create_PED_File	HDS619 / 2.5 test	import_created	n/a	Completed

LD Color Scheme Display

The following tables provide details on the color schemes in the parameter settings page.

Table 1.1: Standard D'/LOD Color Scheme

	D' < 1	D' = 1
LOD < 2	White	blue
LOD >= 2	shades of pink/red	bright red

Table 1.2: Confidence Bounds Color Scheme

Strong Evidence of LD	dark grey
Uninformative	light grey
Strong Evidence of Recombination	white

Table 1.3: r² Color Scheme

r ² = 0	white
0 < r ² < 1	shades of grey
r ² = 1	black

Table 1.4: Alternate D'/LOD Color Scheme

	Low D'	High D'
Low LOD	white	Shades of pink
High LOD	white	grey

(r² and Alt D'/LOD courtesy of Will Fitzhugh)

Table 1.5: Four Gamete Color Scheme

4 distinct 2-marker haplotypes	white
< 4 distinct 2-marker haplotypes	black

Step 3: View Analysis Results

A job status will change from "In Process" to "Completed" once the job is finished. An email notification is sent when the task is complete.

For **Haplotypes**, *Note*: There is no text download feature for the haplotype result so the limit is the same as the display limit.

For **Association Tests**, *Note*: The limit in IGAT is about 100,000 markers.

Haploview is not intended to be the only way of testing association results, but to provide a straightforward way to do simple association tests. It is important to try out multiple approaches to analyze the data this way (Barrett et al., 2005).

IGAT Analysis Result Detail

Task Id	Start Date	End Date	Status
2619	03/18/2008 13:55:19	03/18/2008 13:55:35	Completed

Algorithm: Haploview
 Data Set Accession: HDS300
 Data Set Name: ZP modified jie set
 Data Set Description: ZP modified jie set: replace the affection status from 0 to 1

Result: Displayed below. Note that the size of the analysis result could be very large and may take several minutes to load.

Haploview Result

[Expand All Nodes](#)
[Collapse All Nodes](#)

- Parameters
 - LD Plot
 - Confidence Intervals (Gabriel et al.)
 - GOLD
- [\[Download Image\]](#) [\[Download Text\]](#) [\[Browse\]](#)
- Haplotypes
 - Alleles as colored squares
 - Confidence Intervals (Gabriel et al.)
 - [\[Download\]](#) [\[Browse\]](#)
- Association Tests
 - Single Marker

#	Name	Assoc Allele	Case,Control Ratio	Counts	Case,Control Frequencies	Chi square	P value
1	IGR1118a_1	G	12:0, 119:25	1,000, 0.826	2,481	0.1152	
2	IGR1119a_1	G	12:0, 117:23	1,000, 0.836	2,323	0.1275	
3	IGR1143a_1	A	12:0, 120:28	1,000, 0.811	2,752	0.0971	
4	IGR1144a_1	C	12:0, 122:26	1,000, 0.824	2,517	0.1126	
5	IGR1169a_2	A	12:0, 117:21	1,000, 0.848	2,123	0.1451	
6	IGR1218a_2	A	12:0, 114:26	1,000, 0.814	2,688	0.1011	
7	IGR1219a_2	C	12:0, 113:25	1,000, 0.819	2,609	0.1063	
8	IGR1286a_1	C	12:0, 114:22	1,000, 0.838	2,28	0.131	
9	TSC0101718	T	1:11, 9:129	0.083, 0.065	0.058	0.8093	
10	IGR1373a_1	T	12:0, 122:26	1,000, 0.824	2,517	0.1126	
11	IGR1371a_1	T	12:0, 112:24	1,000, 0.824	2,528	0.1119	
12	IGR1369a_2	A	12:0, 101:25	1,000, 0.802	2,908	0.0882	
13	IGR1369a_1	C	12:0, 123:25	1,000, 0.831	2,402	0.1211	
14	IGR1367a_1	G	12:0, 123:25	1,000, 0.831	2,402	0.1211	
15	IGR2008a_2	T	5:5, 43:93	0.500, 0.316	1,426	0.2324	
16	IGR2008a_1	G	4:6, 33:89	0.400, 0.270	0.768	0.3807	
17	IGR2010a_3	A	5:5, 38:106	0.500, 0.264	2.59	0.1075	
18	IGR2011b_1	C	5:1, 57:57	0.833, 0.500	2.536	0.1113	
19	IGR2016a_1	T	5:7, 41:99	0.417, 0.293	0.803	0.3703	
20	IGR2020a_1	G	10:0, 135:3	1,000, 0.978	0.222	0.6376	

[\[Download\]](#)

Drag the image below by either using your mouse to click and drag or by using the arrow keys

Workflow 2: Upload a Linkage File and a Marker info File

A second method that Haploview input dataset can be created is by uploading a linkage file (.ped) and a marker info file (.info) by formatting genotype data stored in ImmPort. Only genotype data in the format of Illumina BeadStudio Final Report or Affymetrix genotype.txt file are available for conversion to linkage and marker files.

ImmPort Genetic Analysis Tool (IGAT)

Research
Project: Packages3

The ImmPort implementation of Haploview accepts [linkage files](#) (.ped) and [marker info files](#) (.info) as Input Data Sets. Click [here](#) to view a detailed description of the two formats

Please make a selection from the options below. One option is to use an existing ImmPort dataset for Haploview analysis. Another is to upload a dataset from your computer. You can create Haploview input dataset (linkage and marker info files) from ImmPort geno-typing results by using [PED file generation tool](#). Please note that currently only genotype data in the format of Illumina BeadStudio Final report or Affymetrix genotype .txt file are available for conversion to linkage and marker files.

Workflow Selection:

- Run Haploview using existing Linkage File and Marker Info file stored in ImmPort.
- Upload Linkage File and Marker Info File and Run Haploview analysis.

Select workflow and
click 'Next' to continue

Create a new input dataset

ImmPort Genetic Analysis Tool (IGAT) 

Create a new Input Data Set

Research Project: Packages3

Fields marked with an asterisk * are required.

Data Set Name*

Description

In order to create a new input data set, you will need to upload a linkage file (.ped) and its associated marker info file (.info).

Genotype linkage file (.ped)*

Genotype marker file (.info)*

NOTE: Clicking "Next" will upload the above files into the private project workspace of your chosen project (Packages3) and create the data set.

Create a new Input Data Set

Research Project: Packages3

The following files have been uploaded and the data set has been created:

Data Set ID: HDS432
Data Set Name: Bea's dataset
Description: Bea's dataset 12
Linkage file: ZP.ped
Marker file: ZP.info

Click "Next" to continue with the analysis submission.

Create a new input dataset providing a name and a description and then click next to see the dataset that has been uploaded.

Set Haploview Parameters

ImmPort Genetic Analysis Tool (IGAT)

Select Research Project | Choose Workflow | Upload Linkage and Marker File | **Select Haploview Parameters**

Select Haploview Parameters

Input File

Ignore pair wise comparison of markers >: 500 kb apart
 Exclude individuals with >: 50 % missing genotypes

Blocks Definitions

Confidence Intervals (Gabriel et al.)
 Four Gamete Rule
 Solid Spine of LD

LD Plot Color Scheme

4th Gamete
 Confidence Bounds
 R Squared
 Standard D' / LOD
 Alt D' / LOD
 GOLD

Alleles Display Mode

Alleles as colored squares
 Alleles as letters
 Alleles as numbers

Association Tests

Association Tests (Single Marker)
 Association Tests (Haplotypes)

Permutation Tests

Perform Permutations for Association Tests Single Marker
 Perform Permutations for Association Tests Single Markers and Haplotypes in Blocks
 Number of permutation tests: 1000 (a large number may take a long time to run)

Blocks

Block Definition
 Gabriel et al.

Upper confidence interval value: 0.98
 Lower confidence interval value: 0.7
 Upper confidence interval max for strong recombination value: 0.9
 Min fraction of strong ld in info comparison: 0.95
 Exclude marks below: 0.05 MAF

4th gamete rule

Fourth gamete frequency min: 0.01

Strong SD Spine

Extend spine id D prime: 0.8

Block Display

Lines Display

Thick line threshold: 10 %
 Thin line threshold: 1 %

Submit | Cancel

The page at <https://www.immport.net> says:

You must select at least one allele display mode. You must select at least one LD plot color scheme. You must select at least one block definition.

OK

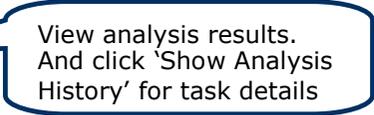
Set parameters from the given Haploview list. Then view the results. If no parameters are selected, a pop-up window is displayed asking the user to select the minimum parameters: Block definitions, LD plot color scheme and Allele display mode.

Click submit to view analysis results

View Analysis Results

ImmPort Genetic Analysis Tool (IGAT) 

[Select Research Project](#)
[Choose Workflow](#)
[Upload Linkage and Marker File](#)
[Select Haploview Parameters](#)

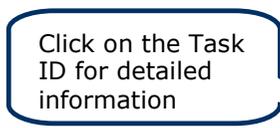
View Analysis Results  View analysis results. And click 'Show Analysis History' for task details

Your Genotyping Analysis request was submitted successfully. Your task id is **3448**.
Click the "Show Analysis History" button to proceed to the IGAT Analysis History page.

[Show Analysis History](#)

IGAT Analysis History

This page is a list of the tasks you have submitted to the ImmPort Genetic Analysis Toolset. The history list can show up to 50 tasks. If the number exceeds 50, the oldest tasks will be dropped off the list. Click on a task ID to view more information about the task or to see the results when the task has completed execution.

 Click on the Task ID for detailed information

117 items found, displaying 1 to 20. [First/Prev] [1](#), [2](#), [3](#), [4](#), [5](#), [6](#) [Next/Last]

Task Id	Start Date	End Date	Task Type	Data Set Accession/Name	Dataset Source	Algorithm	Status
3448	12/31/2008 11:59:11	12/31/2008 13:22:21	Run_Haploview	HDS433 / Bea's dataset	custom_upload	Haploview	Completed
3430	12/18/2008 16:44:36	12/18/2008 16:46:13	Run_Haploview	HDS427 / ZP test PED 2.4.1 on Immport SNP list	import_created	Haploview	Completed
3429	12/18/2008 16:28:24	12/18/2008 16:30:24	Create_PED_File	HDS427 / ZP test PED 2.4.1 on Immport SNP list	import_created	n/a	Completed
3423	12/18/2008 11:34:34	12/18/2008 12:33:54	Run_Haploview	HDS426 / ZP test PED 2.4.1 bigset with mid 1K SNPs	import_created	Haploview	Completed
3422	12/18/2008 11:06:39	12/18/2008 11:20:48	Create_PED_File	HDS426 / ZP test PED 2.4.1 bigset with mid 1K SNPs	import_created	n/a	Completed

View Analysis Results Detail

[Back to History](#)

IGAT Analysis Result Detail

Task Id	Start Date	End Date	Status
3448	12/31/2008 11:59:11	12/31/2008 13:22:21	Completed

Algorithm: Haploview
 Data Set Accession: HDS433
 Data Set Name: Bea's dataset
 Data Set Description: Bea's dataset 12

Result: Displayed below. The ImmPort enterprise version 2.3 of Haploview can handle datasets consisting up to about 100,000 markers; The viewable LD plot size limit (browse and png file displays) is ~1000 markers. There is a 4 Gbyte limit for memory allocation to the algorithm.

Haploview Result

[Expand All Nodes](#)
[Collapse All Nodes](#)

- Parameters
- LD Plot
- Haplotypes
- Association Tests

Expand the nodes to get detailed information on the Haploview results. The same repertoire is repeated in **Slides 10 - 18**

View Analysis Results Detail

Haploview Result

[Expand All Nodes](#)
[Collapse All Nodes](#)

<input type="checkbox"/> Parameters	
No Block	false
Assoc Test Type	CC
Exclude Individuals Missing Genotypes	50
Ignore Pairwise Comp Markers Threshold	500
Number Of Permutation Tests	1000

Expand all the nodes to view the details of the analysis results. NO image was available for more than 1000 markers

- LD Plot
 - Confidence Intervals (Gabriel et al.)
 - 4th Gamete

[\[Download Image\]](#) [\[Download Text\]](#) [\[Browse\]](#)

- Haplotypes
 - Alleles as colored squares
 - Confidence Intervals (Gabriel et al.)

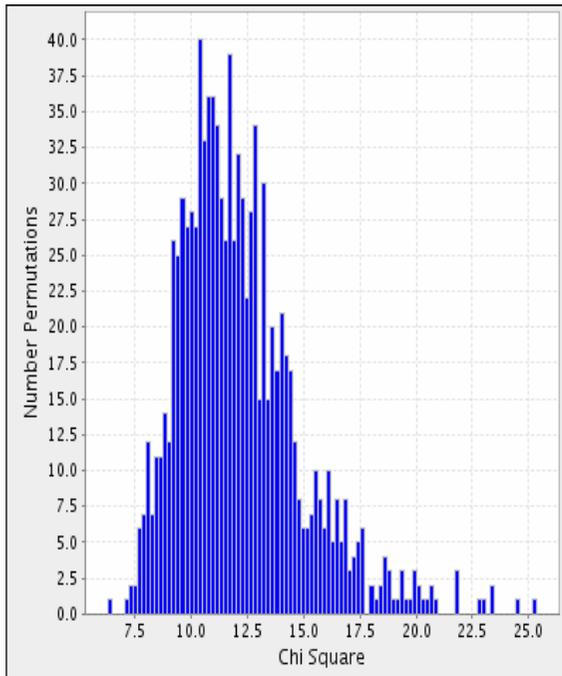
[\[Download Image\]](#) [\[Download Text\]](#) [\[Browse\]](#)

- Association Tests
 - Single Marker

#	Name	Assoc Allele	Case,Control Ratio Counts	Case,Control Frequencies	Chi square	P-value
2051	rs1035798	T	583:983, 129:647	0.372, 0.166	104.115	1.909E-24
2059	rs2071286	A	503:995, 124:622	0.336, 0.166	71.113	3.3734E-17
2032	rs7766862	A	622:950, 178:606	0.396, 0.227	66.333	3.8076E-16
1317	rs6905389	C	1461:83, 663:119	0.946, 0.848	63.402	1.6856E-15

- Permutation Tests
 - Single Markers Only

# 1000 permutations performed.		
Name	Chi Square	Permutation p-value
rs3016013	25.42	0.0000E0
rs9468209	25.446	0.0000E0
rs370520	25.464	0.0000E0
rs2894325	25.556	0.0000E0





ImmPort Genetic Analysis Tool (IGAT)

Blocks generated by Haploview based on Algorithms

- Confidence Intervals
- Four Gamete Rule
- Solid Spine of LD

Haplotypes

- Display
- Display Controls

Association Tests

Permutation Testing

- Single Marker only
- Single Marker and Haplotypes in blocks

File Type Formats

a). Linkage Format

Pedigree name
Individual ID
Father's ID
Mother's ID
Sex
Affected status
Marker genotypes

b). Marker Information File

The marker info file has two columns; marker name and position. The positions can be either absolute chromosomal coordinates or relative positions.

c). Output File

Single Marker Association Text Output File
Haplotype Association Text Output
Permutation Text Output File