

# Package ‘BIGDAWG’

July 21, 2025

**Type** Package

**Title** Case-Control Analysis of Multi-Allelic Loci

**Version** 3.0.3

**Date** 2021-11-01

**Author** Derek Pappas <djpappas75@gmail.com>, Steve Mack <Steven.Mack@ucsf.edu>, Jill Hollenbach <Jill.Hollenbach@ucsf.edu>

**Maintainer** Steve Mack <Steven.Mack@ucsf.edu>

**URL** <http://tools.immunogenomics.org/>,  
<https://github.com/IgDAWG/BIGDAWG>

**BugReports** <https://github.com/IgDAWG/BIGDAWG/issues>

**Description** Data sets and functions for chi-squared Hardy-Weinberg and case-control association tests of highly polymorphic genetic data [e.g., human leukocyte antigen (HLA) data]. Performs association tests at multiple levels of polymorphism (haplotype, locus and HLA amino-acids) as described in Pappas DJ, Marin W, Hollenbach JA, Mack SJ (2016) <[doi:10.1016/j.humimm.2015.12.006](https://doi.org/10.1016/j.humimm.2015.12.006)>. Combines rare variants to a common class to account for sparse cells in tables as described by Hollenbach JA, Mack SJ, Thomson G, Gourraud PA (2012) <[doi:10.1007/978-1-61779-842-9\\_14](https://doi.org/10.1007/978-1-61779-842-9_14)>.

**License** GPL (>= 3)

**Depends** R (>= 3.5.0)

**Imports** XML, httr, haplo.stats, parallel

**Suggests** knitr, rmarkdown

**VignetteBuilder** knitr

**RoxygenNote** 7.1.2

**Encoding** UTF-8

**NeedsCompilation** no

**Repository** CRAN

**Date/Publication** 2021-11-17 11:50:14 UTC

## Contents

A	3
A.wrapper	4
AA.df.check	5
AA.df.cs	5
AAtable.builder	6
AlignmentFilter	6
AlignObj.Create	7
AlignObj.Update	7
Append.System	8
BIGDAWG	8
Build.Matrix	10
buildHAPnames	10
buildHAPsets	11
cci	11
cci.pval	12
cci.pval.list	13
Check.Cores	13
Check.Data	14
Check.Params	14
Check.Params.GLS	15
CheckAlleles	16
CheckHLA	17
CheckLoci	17
CheckRelease	18
CheckString.Allele	18
CheckString.Locus	19
Condense.EPL	19
Create.Null.Table	20
DRB345.Check.Wrapper	20
DRB345.Check.Zygoty	21
DRB345.Exp	21
DRB345.parser	22
Err.Log	22
EVSremoval	23
Exon.Filter	23
ExonPtnAlign.Create	24
ExonPtnList	24
Filler	25
Format.Allele	25
Format.Tab	26
getAllele.Count	26
getCS.Mat	27
getCS.stat	27
GetField	28
getFileName	28
GetFiles	29

getHap . . . . .	29
getObsFreq . . . . .	30
GL2Tab.Loci . . . . .	30
GL2Tab.Sub . . . . .	31
GL2Tab.wrapper . . . . .	31
GLSconvert . . . . .	32
H.MC . . . . .	32
H.MC.wrapper . . . . .	33
HLA_data . . . . .	34
HWE . . . . .	34
HWE.ChiSq . . . . .	35
HWE.wrapper . . . . .	35
L . . . . .	36
L.wrapper . . . . .	36
make2x2 . . . . .	37
makeComb . . . . .	37
MergeData_Output . . . . .	38
PgrpExtract . . . . .	38
PgrpFormat . . . . .	39
PreCheck . . . . .	39
prepData . . . . .	40
rmABstrings . . . . .	40
RunChiSq . . . . .	41
RunChiSq_c . . . . .	41
Stripper . . . . .	42
summaryGeno.2 . . . . .	42
Tab2GL.Loci . . . . .	43
Tab2GL.Sub . . . . .	43
Tab2GL.wrapper . . . . .	44
TableMaker . . . . .	44
UpdateRelease . . . . .	45
<b>Index</b>	<b>46</b>

**Description**

This is the workhorse function for the amino acid analysis.

**Usage**

A(Locus, loci.ColNames, genos, grp, Strict.Bin, ExonAlign, Cores)

**Arguments**

Locus	Locus being analyzed.
loci.ColNames	The column names of the loci being analyzed.
genos	Genotype table.
grp	Case/Control or Phenotype groupings.
Strict.Bin	Logical specify if strict rare cell binning should be used in ChiSq test.
ExonAlign	Exon protein alignment filtered for locus.
Cores	Number of cores to use for analysis.

**Note**

This function is for internal BIGDAWG use only.

---

A.wrapper

*Amino Acid Wrapper*


---

**Description**

Wrapper function for amino acid analysis.

**Usage**

```
A.wrapper(
  loci,
  loci.ColNames,
  genos,
  grp,
  Exon,
  EPL,
  Cores,
  Strict.Bin,
  Output,
  Verbose
)
```

**Arguments**

loci	Loci being analyzed.
loci.ColNames	The column names of the loci being analyzed.
genos	Genotype table.
grp	Case/Control or Phenotype groupings.
Exon	Exon(s)for targeted analysis.
EPL	Protein Alignment List.

Cores	Number of cores to use for analysis.
Strict.Bin	Logical specify if strict rare cell binning should be used in ChiSq test
Output	Data return carryover from main BIGDAWG function
Verbose	Summary display carryover from main BIGDAWG function

**Note**

This function is for internal BIGDAWG use only.

---

AA.df.check	<i>Contingency Table Check</i>
-------------	--------------------------------

---

**Description**

Checks amino acid contingency table data frame to ensure required variation exists.

**Usage**

AA.df.check(x, Strict.Bin)

**Arguments**

x	contingency table.
Strict.Bin	Logical specify if strict rare cell binning should be used in ChiSq test.

**Note**

This function is for internal BIGDAWG use only.

---

AA.df.cs	<i>Contingency Table Amino Acid ChiSq Testing</i>
----------	---

---

**Description**

Runs ChiSq test on amino acid contingency table data frames.

**Usage**

AA.df.cs(x, Strict.Bin)

**Arguments**

x	contingency table.
Strict.Bin	Logical specify if strict rare cell binning should be used in ChiSq test.

**Note**

This function is for internal BIGDAWG use only.

AAtable.builder      *Amino Acid Contingency Table Build*

---

**Description**

Build Contingency Tables for Amino Acid Analysis.

**Usage**

```
AAtable.builder(x, y)
```

**Arguments**

x	Filtered alignment list element.
y	Phenotype groupings.

**Note**

This function is for internal BIGDAWG use only.

---

AlignmentFilter      *Alignment Filter*

---

**Description**

Filter Protein Exon Alignment File for Specific Alleles.

**Usage**

```
AlignmentFilter(Align, Alleles, Locus)
```

**Arguments**

Align	Protein Alignment Object.
Alleles	to be pulled.
Locus	Locus to be filtered against.

**Note**

This function is for internal BIGDAWG use only.

---

AlignObj.Create      *Alignment Object Creator*

---

**Description**

Create Object for Exon Protein Alignments.

**Usage**

```
AlignObj.Create(Loci, Release, RefTab)
```

**Arguments**

Loci	Loci to be bundled.
Release	IMGT/HLA database release version.
RefTab	Data of reference exons used for protein alignment creation.

**Note**

This function is for internal BIGDAWG use only.

---

AlignObj.Update      *Updated Alignment Object Creator*

---

**Description**

Synthesize Object for Exon Protein Alignments.

**Usage**

```
AlignObj.Update(Loci, Release, RefTab)
```

**Arguments**

Loci	Loci to be bundled.
Release	IMGT/HLA database release version.
RefTab	Data of reference exons used for protein alignment creation.

**Note**

This function is for internal BIGDAWG use only.

Append.System

*Append Genetic System Locus Designation to Allele String*

---

**Description**

Adds genetic system (HLA/KIR) to each allele name

**Usage**

```
Append.System(x, df.name)
```

**Arguments**

x	Vector Column genotypes to append
df.name	String SystemLocus name for each allele.

**Note**

This function is for internal use only.

---

BIGDAWG

*BIGDAWG Main Wrapper Function*

---

**Description**

This is the main wrapper function for each analysis.

**Usage**

```
BIGDAWG(  
  Data,  
  HLA = TRUE,  
  Run.Tests,  
  Loci.Set,  
  Exon,  
  All.Pairwise = FALSE,  
  Trim = FALSE,  
  Res = 2,  
  EVS.rm = FALSE,  
  Missing = 2,  
  Strict.Bin = FALSE,  
  Cores.Lim = 1L,  
  Results.Dir,  
  Return = FALSE,  
  Output = TRUE,
```

```

Merge.Output = FALSE,
Verbose = TRUE
)

```

### Arguments

Data	Name of the genotype data file.
HLA	Logical Indicating whether data is HLA class I/II genotyping data only.
Run.Tests	Specifics which tests to run.
Loci.Set	Input list defining which loci to use for analyses (combinations permitted).
Exon	Numeric Exon(s) for targeted amino acid analysis.
All.Pairwise	Logical indicating whether all pairwise loci should be analyzed in haplotype analysis.
Trim	Logical indicating if HLA alleles should be trimmed to a set resolution.
Res	Numeric setting what desired resolution to trim HLA alleles.
EVS.rm	Logical indicating if expression variant suffixes should be removed.
Missing	Numeric setting allowable missing data for running analysis (may use "ignore").
Strict.Bin	Logical specify if strict rare cell binning should be used in ChiSq test.
Cores.Lim	Integer setting the number of cores accessible to BIGDAWG (Windows limit is 1 core).
Results.Dir	Optional, string of full path directory name for BIGDAWG output.
Return	Logical Should analysis results be returned as list.
Output	Logical Should analysis results be written to output directory.
Merge.Output	Logical Should analysis results be merged into a single file for easy access.
Verbose	Logical Should a summary of each analysis be displayed in console.

### Examples

```

## Not run:
### The following examples use the synthetic data set bundled with BIGDAWG

# Haplotype analysis with no missing genotypes for two loci sets
# Significant haplotype association with phenotype
# BIGDAWG(Data="HLA_data", Run.Tests="H", Missing=0, Loci.Set=list(c("DRB1", "DQB1")))

# Hardy-Weinberg and Locus analysis ignoring missing data
# Significant locus associations with phenotype at all but DQB1
# BIGDAWG(Data="HLA_data", Run.Tests="L", Missing="ignore")

# Hardy-Weinberg analysis trimming data to 2-Field resolution with no output to files (console only)
# Significant locus deviation at DQB1
BIGDAWG(Data="HLA_data", Run.Tests="HWE", Trim=TRUE, Res=2, Output=FALSE)

## End(Not run)

```

**Build.Matrix***Build Output Matrix for GL2Tab Conversion*

---

**Description**

Initializes output matrix format for GL2Tab conversion

**Usage**

```
Build.Matrix(System, Loci)
```

**Arguments**

System	Character Genetic system HLA- or KIR
Loci	The loci for header names

**Note**

This function is for internal use only.

---

**buildHAPnames***Haplotype Name Builder*

---

**Description**

Builds table of names for HAPsets

**Usage**

```
buildHAPnames(Combn, loci)
```

**Arguments**

Combn	Combination of loci to extraction from genos
loci	Character vector of unique loci being analyzed.

**Note**

This function is for internal BIGDAWG use only.

---

buildHAPsets	<i>Haplotype List Builder</i>
--------------	-------------------------------

---

**Description**

Builds table of haplotypes from combinations

**Usage**

```
buildHAPsets(Combn, genos, loci, loci.ColNames)
```

**Arguments**

Combn	Combination of loci to extraction from genos
genos	The genotype columns of the loci set being analyzed.
loci	Character vector of unique loci being analyzed.
loci.ColNames	Character vector of genos column names.

**Note**

This function is for internal BIGDAWG use only.

---

cci	<i>Case-Control Odds ratio calculation and graphing</i>
-----	---

---

**Description**

cci function port epicalc version 2.15.1.0 (Virasakdi Chongsuvivatwong, 2012)

**Usage**

```
cci(  
  caseexp,  
  controlex,  
  casenonex,  
  controlnonex,  
  cctable = NULL,  
  graph = TRUE,  
  design = "cohort",  
  main,  
  xlab,  
  ylab,  
  xaxis,  
  yaxis,  
  alpha = 0.05,
```

```

    fisher.or = FALSE,
    exact.ci.or = TRUE,
    decimal = 2
  )

```

### Arguments

caseexp	Number of cases exposed
controlex	Number of controls exposed
casenonex	Number of cases not exposed
controlnonex	Number of controls not exposed
cctable	A 2-by-2 table. If specified, will supercede the outcome and exposure variables
graph	If TRUE (default), produces an odds ratio plot
design	Specification for graph; can be "case control", "case-control", "cohort" or "prospective"
main	main title of the graph
xlab	label on X axis
ylab	label on Y axis
xaxis	two categories of exposure in graph
yaxis	two categories of outcome in graph
alpha	level of significance
fisher.or	whether odds ratio should be computed by the exact method
exact.ci.or	whether confidence limite of the odds ratio should be computed by the exact method
decimal	number of decimal places displayed

### Note

This function is for internal BIGDAWG use only.

---

cci.pval

*Case Control Odds Ratio Calculation from Epicalc*

---

### Description

Calculates odds ratio and pvalues from 2x2 table

### Usage

```
cci.pval(x)
```

### Arguments

x List of 2x2 matrices for calculation, output of TableMaker.

**Note**

This function is for internal BIGDAWG use only.

---

cci.pval.list	<i>Case Control Odds Ratio Calculation from Epicalc list variation</i>
---------------	--

---

**Description**

Variation of the cci.pvalue function

**Usage**

```
cci.pval.list(x)
```

**Arguments**

x	List of 2x2 matrices to apply the cci.pvalue function. List output of TableMaker.
---	---

**Note**

This function is for internal BIGDAWG use only.

---

Check.Cores	<i>Check Cores Parameters</i>
-------------	-------------------------------

---

**Description**

Check cores limitation for OS compatibility

**Usage**

```
Check.Cores(Cores.Lim, Output)
```

**Arguments**

Cores.Lim	Integer How many cores can be used.
Output	Logical Should analysis results be written to output directory.

Check.Data

*Check Data Structure*

---

**Description**

Check data structure for successful conversion.

**Usage**

```
Check.Data(Data, System, Convert)
```

**Arguments**

Data	String Type of output.
System	Character Genetic system HLA or KIR
Convert	String Direction for conversion.

**Note**

This function is for internal use only.

---

Check.Params

*Check Input Parameters*

---

**Description**

Check input parameters for invalid entries.

**Usage**

```
Check.Params(  
  HLA,  
  Loci.Set,  
  Exon,  
  All.Pairwise,  
  Trim,  
  Res,  
  EVS.rm,  
  Missing,  
  Cores.Lim,  
  Return,  
  Output,  
  Merge.Output,  
  Verbose  
)
```

**Arguments**

HLA	Logical indicating whether data is HLA class I/II genotyping data only.
Loci.Set	Input list defining which loci to use for analyses (combinations permitted).
Exon	Numeric Exon(s) for targeted amino acid analysis.
All.Pairwise	Logical indicating whether all pairwise loci should be analyzed in haplotype analysis.
Trim	Logical indicating if HLA alleles should be trimmed to a set resolution.
Res	Numeric setting what desired resolution to trim HLA alleles.
EVS.rm	Logical indicating if expression variant suffixes should be removed.
Missing	Numeric setting allowable missing data for running analysis (may use "ignore").
Cores.Lim	Integer setting the number of cores accessible to BIGDAWG (Windows limit is 1 core).
Return	Logical Should analysis results be returned as list.
Output	Logical Should analysis results be written to output directory.
Merge.Output	Logical Should analysis results be merged into a single file for easy access.
Verbose	Logical Should a summary of each analysis be displayed in console.

**Note**

This function is for internal use only.

---

Check.Params.GLS

*Check Input Parameters for GLS conversion*

---

**Description**

Check input parameters for invalid entries.

**Usage**

```
Check.Params.GLS(
  Convert,
  File.Output,
  System,
  HZY.Red,
  DRB345.Check,
  Cores.Lim
)
```

**Arguments**

Convert	String Direction for conversion.
File.Output	String Type of output.
System	String Genetic system (HLA or KIR) of the data being converted
HZY.Red	Logical Reduction of homozygote genotypes to single allele.
DRB345.Check	Logical Check DR haplotypes for consistency and flag unusual haplotypes.
Cores.Lim	Integer How many cores can be used.

**Note**

This function is for internal use only.

---

CheckAlleles

*HLA Allele Legitimacy Check for Amino Acid Analysis*

---

**Description**

Checks available alleles against data to ensure complete overlap.

**Usage**

CheckAlleles(x, y)

**Arguments**

x	Exon protein list alignment object.
y	Genotypes from data file

**Note**

This function is for internal BIGDAWG use only.

---

CheckHLA	<i>HLA Formatting Check for Amino Acid Analysis</i>
----------	---

---

**Description**

Checks data to see if HLA data is properly formatted .

**Usage**

CheckHLA(x)

**Arguments**

x                    All columns of HLA genotyping data.

**Note**

This function is for internal BIGDAWG use only.

---

CheckLoci	<i>HLA Loci Legitimacy Check for Amino Acid Analysis</i>
-----------	--

---

**Description**

Checks available loci against data to ensure complete overlap.

**Usage**

CheckLoci(x, y)

**Arguments**

x                    Loci available in exon protein list alignment object.  
y                    Unique column names

**Note**

This function is for internal BIGDAWG use only.

---

CheckRelease	<i>Function to Check Release Versions</i>
--------------	---

---

**Description**

This updates the protein alignment used in checking HLA loci and alleles as well as in the amino acid analysis.

**Usage**

```
CheckRelease(Package = T, Alignment = T, Output = F)
```

**Arguments**

Package	Logical to check for BIGDAWG package versions
Alignment	Logical to check the IMGT/HLA database version for the alignment bundled with BIGDAWG.
Output	Should any error be written to a file

**Note**

Requires active internet connection.

---

CheckString.Allele	<i>GL String Allele Check</i>
--------------------	-------------------------------

---

**Description**

GL String check for allele ambiguity formatting

**Usage**

```
CheckString.Allele(x)
```

**Arguments**

x	GL String to check against
---	----------------------------

**Note**

This function is for internal use only.

---

CheckString.Locus	<i>GL String Locus Check</i>
-------------------	------------------------------

---

**Description**

Check GL string for loci appearing in multiple gene fields.

**Usage**

```
CheckString.Locus(x, Loci)
```

**Arguments**

x	GL String to check against
Loci	Loci to check

**Note**

This function is for internal use only.

---

Condense.EPL	<i>Condensing Exon Specific Alignments to Single Dataframe</i>
--------------	--

---

**Description**

Combines multiple Exon Specific Alignments into a single Alignment object

**Usage**

```
Condense.EPL(EPL.Exon)
```

**Arguments**

EPL.Exon	Exon-Locus Specific Amino Acid Alignment.
----------	---

**Note**

This function is for internal BIGDAWG use only.

Create.Null.Table      *Create Empty Table*

---

**Description**

Creates matrix of NA for no result tables.

**Usage**

```
Create.Null.Table(Locus, Names, nr)
```

**Arguments**

Locus	Locus being analyzed.
Names	Column names for final matrix.
nr	Number of rows.

**Note**

This function is for internal BIGDAWG use only.

---

DRB345.Check.Wrapper      *DRB345 haplotype zygosity wrapper*

---

**Description**

Checks DR haplotypes for correct zygosity and flags unanticipated haplotypes

**Usage**

```
DRB345.Check.Wrapper(Genotype, Loci.DR)
```

**Arguments**

Genotype	Row of data set data frame following DRB345 parsing
Loci.DR	DRBx Loci of interest to test for consistency

**Note**

This function is for internal use only.

---

DRB345.Check.Zygoty *DRB345 haplotype zygoty checker single locus*

---

**Description**

Checks DR haplotypes for correct zygoty and flags unanticipated haplotypes for a single DRBx

**Usage**

DRB345.Check.Zygoty(Locus, Genotype)

**Arguments**

Locus	Locus of interest to test for consistency
Genotype	Row of data set data frame following DRB345 parsing

**Note**

This function is for internal use only.

---

DRB345.Exp	<i>DRB345 Expected</i>
------------	------------------------

---

**Description**

Checks DRB1 Genotype and Returns Expected DR345 Loci

**Usage**

DRB345.Exp(DRB1.Genotype)

**Arguments**

DRB1.Genotype	DRB1 Subject Genotypes
---------------	------------------------

**Note**

This function is for internal use only.

---

 DRB345.parser

*DRB345 Column Processing*


---

**Description**

Separates DRB345 column pair into separate columns for each locus

**Usage**

DRB345.parser(Tab)

**Arguments**

Tab                      Data frame of sampleIDs, phenotypes, and genotypes

**Note**

This function is for internal BIGDAWG use only.

---

 Err.Log

*Error Code Display and Logging*


---

**Description**

Displays error codes attributable to data formatting and Locus/Allele naming. Writes to log file.

**Usage**

Err.Log(Output, x, y = NULL, z = NULL)

**Arguments**

Output	Logical indicating if Error logging should be written to a file.
x	Log Code.
y	Misc information relevant to error.
z	Misc information relevant to error.

**Note**

This function is for internal BIGDAWG use only.

---

EVSremoval	<i>Expression Variant Suffix Removal</i>
------------	--

---

**Description**

Removes expression variant suffixes from HLA alleles in the exon protein alignment object.

**Usage**

EVSremoval(Locus, EPList)

**Arguments**

Locus	Locus to be filtered against.
EPList	Exon Protein Alignment Object

**Note**

This function is for internal BIGDAWG use only.

---

Exon.Filter	<i>Filter Exon Specific Alignment Sections</i>
-------------	--

---

**Description**

Filters the ExonPtnAlign object by locus and exon.

**Usage**

Exon.Filter(Locus, Exon, EPL.Locus, RefExons, E.Ptn.Starts)

**Arguments**

Locus	Locus being analyzed.
Exon	Exon being analyzed.
EPL.Locus	ExonPtnAlign object filtered by Locus
RefExons	Reference Exon Table
E.Ptn.Starts	Exon Protein Overlay Map

**Note**

This function is for internal BIGDAWG use only.

---

ExonPtnAlign.Create     *Protein Exon Alignment Formatter*

---

**Description**

Dynamically creates an alignment of Allele exons for Analysis.

**Usage**

ExonPtnAlign.Create(Locus, RefTab)

**Arguments**

Locus	Locus alignment to be formatted.
RefTab	Reference exon protein information for alignment formatting.

**Note**

This function is for internal BIGDAWG use only.

---

ExonPtnList     *Exon protein alignments.*

---

**Description**

Alignment object for use in the amino acid analysis.

**Usage**

ExonPtnList

**Format**

A list where each element is an alignment dataframe for a single locus.

---

Filler	<i>Replace or Fill 00:00 allele strings</i>
--------	---

---

**Description**

Replaces or Fills absent allele strings.

**Usage**

Filler(x, Locus = NULL, Type)

**Arguments**

x	Genotype
Locus	Locus column to adjust.
Type	String specifying whether to pad ('Fill') or leave blank ('Remove') absent calls

**Note**

This function is for internal use only.

---

Format.Allele	<i>Ambiguous Alleles Locus Name Formatting</i>
---------------	--

---

**Description**

Remove or Append Locus name from/to allele in an ambiguous allele string

**Usage**

Format.Allele(x, Type)

**Arguments**

x	Allele String
Type	String specifying whether to strip ('off') or append ('on') locus prefix

**Note**

This function is for internal use only.

---

`Format.Tab`*Tabular Data Locus Format Tool*

---

**Description**

Correctly orders the expanded GL string

**Usage**

```
Format.Tab(x, Order)
```

**Arguments**

<code>x</code>	Single row of converted GL string
<code>Order</code>	Single row data frame for mapping converted GL strings

**Note**

This function is for internal use only.

---

`getAllele.Count`*Recompute number of alleles*

---

**Description**

Using `Freq.Final`, recompute number of alleles

**Usage**

```
getAllele.Count(x)
```

**Arguments**

<code>x</code>	Locus specific contingency matrix <code>getCS.Mat</code> output.
----------------	--

**Note**

This function is for internal BIGDAWG use only.

---

`getCS.Mat`*Chi square matrices*

---

**Description**

Chi Square contingency matrix builder with rare cell binning

**Usage**

```
getCS.Mat(Locus, genos.sub, Allele.Freq, Allele.Combn)
```

**Arguments**

Locus	Locus of interest.
genos.sub	Genotypes for locus of interest.
Allele.Freq	Allele frequencies.
Allele.Combn	Allele combinations.

**Note**

This function is for internal BIGDAWG use only.

---

`getCS.stat`*Chi square test statistic*

---

**Description**

Calculate chi square test statistic

**Usage**

```
getCS.stat(Locus, Freq.Final)
```

**Arguments**

Locus	Locus of interest.
Freq.Final	Contingency Matrix getCS.Mat output.

**Note**

This function is for internal BIGDAWG use only.

---

GetField	<i>HLA trimming function</i>
----------	------------------------------

---

**Description**

Trim a properly formatted HLA allele to desired number of fields.

**Usage**

```
GetField(x, Res)
```

**Arguments**

x	HLA allele.
Res	Resolution desired.

**Note**

This function is for internal BIGDAWG use only.

---

getFileName	<i>File Name Extraction</i>
-------------	-----------------------------

---

**Description**

Function to extract file path.

**Usage**

```
getFileName(x)
```

**Arguments**

x	File name.
---	------------

**Note**

This function is for internal use only.

---

`GetFiles`*File Fetcher*

---

**Description**

Download Protein Alignment and Accessory Files

**Usage**

```
GetFiles(Loci)
```

**Arguments**

Loci                    HLA Loci to be fetched. Limited Loci available.

**Note**

This function is for internal BIGDAWG use only.

---

`getHap`*Haplotype Table Maker*

---

**Description**

Builds table of haplotypes

**Usage**

```
getHap(SID, HaploEM)
```

**Arguments**

SID                    Index number (i.e., row number) of sample ID from genotype matrix.

HaploEM                Haplotype output object from haplo.stat::haplo.em function.

**Note**

This function is for internal BIGDAWG use only.

---

getObsFreq	<i>Observed Frequency</i>
------------	---------------------------

---

**Description**

Get observed frequency of genotypes

**Usage**

```
getObsFreq(x, genos.locus)
```

**Arguments**

x	Single genotype.
genos.locus	Locus genotypes.

**Note**

This function is for internal BIGDAWG use only.

---

GL2Tab.Loci	<i>Locus Ordering for GL2Tab</i>
-------------	----------------------------------

---

**Description**

Orders Locus Calls

**Usage**

```
GL2Tab.Loci(Locus, Genotype, System)
```

**Arguments**

Locus	Locus to condense
Genotype	Row of loci to condense
System	Character Genetic system HLA or KIR

**Note**

This function is for internal use only.

---

GL2Tab.Sub	<i>Genotype List String Expander</i>
------------	--------------------------------------

---

**Description**

Expands GL string into a table of adjacent loci

**Usage**

GL2Tab.Sub(x, System)

**Arguments**

x	Character GL string to expand
System	Character Genetic system HLA or KIR

**Note**

This function is for internal use only.

---

GL2Tab.wrapper	<i>Genotype List String to Tabular Data Conversion</i>
----------------	--

---

**Description**

Expands GL strings to columns of adjacent locus pairs.

**Usage**

GL2Tab.wrapper(df, System, Strip.Prefix, Abs.Fill, Cores)

**Arguments**

df	Data frame containing GL strings
System	Character Genetic system HLA or KIR
Strip.Prefix	Logical Should System/Locus prefixes be stripped from table data.
Abs.Fill	Logical Should absent loci special designations be used.
Cores	Integer How many cores can be used

**Note**

This function is for internal use only

---

GLSconvert	<i>Genotype List String Conversion</i>
------------	--

---

**Description**

Main Workhorse wrapper for cross converting columnar table to GL string representaion.

**Usage**

```
GLSconvert(
  Data,
  Convert,
  File.Output = "txt",
  System = "HLA",
  HZY.Red = FALSE,
  DRB345.Check = FALSE,
  Strip.Prefix = TRUE,
  Abs.Fill = FALSE,
  Cores.Lim = 1L
)
```

**Arguments**

Data	String File name or R Data Frame.
Convert	String Direction for conversion.
File.Output	String Type of File.Output.
System	String Genetic system (HLA or KIR) of the data being converted
HZY.Red	Logical Reduction of homozygote genotypes to single allele.
DRB345.Check	Logical Check DR haplotypes for consistency and flag unusual haplotypes.
Strip.Prefix	Logical Should System/Locus prefixes be stripped from table data.
Abs.Fill	Logical Should absent loci special designations be used.
Cores.Lim	Integer How many cores can be used.

---

H.MC	<i>Haplotype Analysis Function for Multicore</i>
------	--

---

**Description**

This is the workhorse function for the haplotype analysis.

**Usage**

```
H.MC(genos.sub, grp, Strict.Bin, Verbose)
```

**Arguments**

genos.sub	The genotype columns of the loci(locus) set being analyzed.
grp	Case/Control or Phenotype groupings.
Strict.Bin	Logical specify if strict rare cell binning should be used in ChiSq test
Verbose	Summary display carryover from main BIGDAWG function

**Note**

This function is for internal BIGDAWG use only.

---

H.MC.wrapper

*Haplotype Wrapper for Multicore*


---

**Description**

Wrapper for main H function

**Usage**

```
H.MC.wrapper(
  SID,
  Tabsub,
  loci,
  loci.ColNames,
  genos,
  grp,
  All.Pairwise,
  Strict.Bin,
  Output,
  Verbose,
  Cores
)
```

**Arguments**

SID	Character vector of subject IDs.
Tabsub	Data frame of genotype calls for set being analyzed.
loci	Character vector of unique loci being analyzed.
loci.ColNames	Character vector of genos column names.
genos	The genotype columns of the loci set being analyzed.
grp	Case/Control or Phenotype groupings.
All.Pairwise	Haplotype argument carryover from main BIGDAWG function
Strict.Bin	Logical specify if strict rare cell binning should be used in ChiSq test
Output	Data return carryover from main BIGDAWG function
Verbose	Summary display carryover from main BIGDAWG function
Cores	Cores carryover from main BIGDAWG function

**Note**

This function is for internal BIGDAWG use only.

---

HLA\_data

*Example HLA Dataset*

---

**Description**

A synthetic dataset of HLA genotypes for using bigdawg.

**Usage**

HLA\_data

**Format**

A data frame with 2000 rows and 14 variables

---

HWE

*Hardy Weinberg Equilibrium Function*

---

**Description**

This is the main function for the HWE analysis.

**Usage**

HWE(Tab)

**Arguments**

Tab                    data frame of genotype files post processing.

**Note**

This function is for internal BIGDAWG use only.

---

HWE.ChiSq	<i>Hardy Weinberg Equilibrium Function</i>
-----------	--

---

**Description**

This is the workhorse function for each group analysis.

**Usage**

```
HWE.ChiSq(genos.sub, loci, nloci)
```

**Arguments**

gnos.sub	data frame of genotype files post processing.
loci	list of loci.
nloci	number of loci in list

**Note**

This function is for internal BIGDAWG use only.

---

HWE.wrapper	<i>Hardy-Weinberg Wrapper</i>
-------------	-------------------------------

---

**Description**

Wrapper for main HWE function

**Usage**

```
HWE.wrapper(Tab, Output, Verbose)
```

**Arguments**

Tab	Data frame of genotype files post processing.
Output	Data return carryover from main BIGDAWG function
Verbose	Summary display carryover from main BIGDAWG function

**Note**

This function is for internal BIGDAWG use only.

---

L *Locus Analysis Function*

---

**Description**

This is the workhorse function for the locus level analysis.

**Usage**

```
L(loci.ColNames, Locus, genos, grp, Strict.Bin)
```

**Arguments**

loci.ColNames	The column names of the loci being analyzed.
Locus	Locus being analyzed.
genos	Genotype table
grp	Case/Control or Phenotype groupings.
Strict.Bin	Logical specify if strict rare cell binning should be used in ChiSq test

**Note**

This function is for internal BIGDAWG use only.

---

L.wrapper *Locus Wrapper*

---

**Description**

Wrapper for main L function

**Usage**

```
L.wrapper(nloci, loci, loci.ColNames, genos, grp, Strict.Bin, Output, Verbose)
```

**Arguments**

nloci	Number of loci being analyzed.
loci	Loci being analyzed.
loci.ColNames	The column names of the loci being analyzed.
genos	Genotype table
grp	Case/Control or Phenotype groupings.
Strict.Bin	Logical specify if strict rare cell binning should be used in ChiSq test
Output	Data return carryover from main BIGDAWG function
Verbose	Summary display carryover from main BIGDAWG function

**Note**

This function is for internal BIGDAWG use only.

---

make2x2	<i>Creation of a 2x2 table using the indicated orientation.</i>
---------	---

---

**Description**

make2x2 function port epicalc version 2.15.1.0 (Virasakdi Chongsuvivatwong, 2012)

**Usage**

```
make2x2(caseexp, controlex, casenonex, controlnonex)
```

**Arguments**

caseexp	Number of cases exposed
controlex	Number of controls exposed
casenonex	Number of cases not exposed
controlnonex	Number of controls not exposed

**Note**

This function is for internal BIGDAWG use only.

---

makeComb	<i>Genotype Combination Maker</i>
----------	-----------------------------------

---

**Description**

Make data frame of possible genotype combinations

**Usage**

```
makeComb(x)
```

**Arguments**

x	Number of alleles.
---	--------------------

**Note**

This function is for internal BIGDAWG use only.

MergeData\_Output      *Data Object Merge and Output*

---

**Description**

Whole data set table construction of per haplotype for odds ratio, confidence intervals, and pvalues

**Usage**

```
MergeData_Output(BD.out, Run, OutDir)
```

**Arguments**

BD.out	Output of analysis as list.
Run	Tests that are to be run as defined by Run.Tests.
OutDir	Output directory defined by Results.Dir or default.

**Note**

This function is for internal BIGDAWG use only.

---

PgrpExtract      *HLA P group Finder*

---

**Description**

Identify P group for a given allele if exists.

**Usage**

```
PgrpExtract(x, y)
```

**Arguments**

x	Allele of interest.
y	Formatted P groups.

**Note**

This function is for internal BIGDAWG use only.

---

PgrpFormat	<i>HLA P group File Formatter</i>
------------	-----------------------------------

---

**Description**

Format the hla\_nom\_p.txt read table object for a specific locus.

**Usage**

PgrpFormat(x, Locus)

**Arguments**

x	P group object from read.table command.
Locus	Locus to be filtered on.

**Note**

This function is for internal BIGDAWG use only.

---

PreCheck	<i>Data Summary Function</i>
----------	------------------------------

---

**Description**

Summary function for sample population within data file.

**Usage**

PreCheck(Tab, All.ColNames, rescall, HLA, Verbose, Output)

**Arguments**

Tab	Loci available in exon protein list alignment object.
All.ColNames	Column names from genotype data.
rescall	HLA resolution set for analysis.
HLA	HLA BIGDAWG argument passed to function
Verbose	Summary display carryover from BIGDAWG function.
Output	Data output carryover form BIGDAWG function

**Note**

This function is for internal BIGDAWG use only.

prepData                      *Prepare imported data*

---

**Description**

Prepare imported data for processing, checks, and analysis.

**Usage**

```
prepData(Tab)
```

**Arguments**

Tab                      Genotypes dataframe.

**Note**

This function is for internal BIGDAWG use only.

---

rmABstrings                      *Replace absent allele strings*

---

**Description**

Replaces allowable absent allele strings with ^ symbol.

**Usage**

```
rmABstrings(df)
```

**Arguments**

df                      Genotypes dataframe.

**Note**

This function is for internal BIGDAWG use only.

---

`RunChiSq`*Strict Chi-squared Contingency Table Test*

---

**Description**

Calculates chi-squared contingency table tests and bins all rare cells.

**Usage**`RunChiSq(x)`**Arguments**

`x` Contingency table.

**Note**

This function is for internal BIGDAWG use only.

---

`RunChiSq_c`*Contextual Binning Chi-squared Contingency Table Test*

---

**Description**

Calculates chi-squared contingency table tests and bins rare cells at 20

**Usage**`RunChiSq_c(x)`**Arguments**

`x` Contingency table.

**Note**

This function is for internal BIGDAWG use only.

---

Stripper	<i>Removes System and Locus from Alleles</i>
----------	--

---

**Description**

Removes the System and Locus designations for alleles calls in GL2Tab

**Usage**

```
Stripper(x)
```

**Arguments**

x	Allele
---	--------

**Note**

This function is for internal use only.

---

summaryGeno.2	<i>Haplotype missing Allele summary function</i>
---------------	--

---

**Description**

Summary function for identifying missing alleles in a matrix of genotypes.

**Usage**

```
summaryGeno.2(geno, miss.val = 0)
```

**Arguments**

geno	Matrix of genotypes.
miss.val	Vector of codes for allele missing values.

**Note**

This function is for internal BIGDAWG use only and is ported from haplo.stats.

---

Tab2GL.Loci	<i>Locus Condenser for Tab2GL</i>
-------------	-----------------------------------

---

**Description**

Condenses alleles calls of a single locus string using "+"

**Usage**

Tab2GL.Loci(Locus, Genotype, System, HZY.Red)

**Arguments**

Locus	Locus to condense
Genotype	Row of loci to condense
System	Character Genetic system HLA or KIR
HZY.Red	Logical Should homozygote genotypes be a single allele for non-DRB345.

**Note**

This function is for internal use only.

---

Tab2GL.Sub	<i>Genotype List String Condenser</i>
------------	---------------------------------------

---

**Description**

Condenses column of loci into a GL string using "^"

**Usage**

Tab2GL.Sub(x, System, HZY.Red)

**Arguments**

x	Row of loci to condense
System	Character Genetic system HLA or KIR
HZY.Red	Logical Should homozygote genotypes be a single allele for non-DRB345.

**Note**

This function is for internal use only.

---

Tab2GL.wrapper	<i>Genotype List String to Tabular Data Conversion</i>
----------------	--

---

**Description**

Expands GL strings to columns of adjacent locus pairs.

**Usage**

```
Tab2GL.wrapper(df, System, HZY.Red, Abs.Fill, Cores)
```

**Arguments**

df	Data frame containing GL strings
System	Character Genetic system HLA or KIR
HZY.Red	Logical Should homozygote genotypes be a single allele for non-DRB345.
Abs.Fill	Logical Should absent loci special designations be used
Cores	Integer How many cores can be used.

**Note**

This function is for internal use only.

---

TableMaker	<i>Table Maker</i>
------------	--------------------

---

**Description**

Table construction of per haplotype for odds ratio, confidence intervals, and pvalues

**Usage**

```
TableMaker(x)
```

**Arguments**

x	Contingency table with binned rare cells.
---	---

**Note**

This function is for internal BIGDAWG use only.

---

UpdateRelease	<i>Update function for protein alignment upon new IMGT HLA data release</i>
---------------	---

---

**Description**

This updates the protein alignment used in checking HLA loci and alleles as well as in the amino acid analysis.

**Usage**

```
UpdateRelease(Force = F, Restore = F, Output = F)
```

**Arguments**

Force	Logical specifying if update should be forced.
Restore	Logical specifying if the original alignment file be restored.
Output	Logical indicating if error reporting should be written to file.

# Index

## \* datasets

- ExonPtnList, 24
- HLA\_data, 34

A, 3

A.wrapper, 4

AA.df.check, 5

AA.df.cs, 5

AAtable.builder, 6

AlignmentFilter, 6

AlignObj.Create, 7

AlignObj.Update, 7

Append.System, 8

BIGDAWG, 8

Build.Matrix, 10

buildHAPnames, 10

buildHAPsets, 11

cci, 11

cci.pval, 12

cci.pval.list, 13

Check.Cores, 13

Check.Data, 14

Check.Params, 14

Check.Params.GLS, 15

CheckAlleles, 16

CheckHLA, 17

CheckLoci, 17

CheckRelease, 18

CheckString.Allele, 18

CheckString.Locus, 19

Condense.EPL, 19

Create.Null.Table, 20

DRB345.Check.Wrapper, 20

DRB345.Check.Zygotity, 21

DRB345.Exp, 21

DRB345.parser, 22

Err.Log, 22

EVSremoval, 23

Exon.Filter, 23

ExonPtnAlign.Create, 24

ExonPtnList, 24

Filler, 25

Format.Allele, 25

Format.Tab, 26

getAllele.Count, 26

getCS.Mat, 27

getCS.stat, 27

GetField, 28

getFileName, 28

GetFiles, 29

getHap, 29

getObsFreq, 30

GL2Tab.Loci, 30

GL2Tab.Sub, 31

GL2Tab.wrapper, 31

GLSconvert, 32

H.MC, 32

H.MC.wrapper, 33

HLA\_data, 34

HWE, 34

HWE.ChiSq, 35

HWE.wrapper, 35

L, 36

L.wrapper, 36

make2x2, 37

makeComb, 37

MergeData\_Output, 38

PgrpExtract, 38

PgrpFormat, 39

PreCheck, 39

prepData, 40

rmABstrings, [40](#)  
RunChiSq, [41](#)  
RunChiSq\_c, [41](#)  
  
Stripper, [42](#)  
summaryGeno.2, [42](#)  
  
Tab2GL.Loci, [43](#)  
Tab2GL.Sub, [43](#)  
Tab2GL.wrapper, [44](#)  
TableMaker, [44](#)  
  
UpdateRelease, [45](#)