

nprcgeneKeepr Application for Genetic Management of NHP Colonies: Colony Manager Tutorial

R. Mark Sharp

Introduction

This tutorial demonstrates the major functions used within the **nprcgenekeepr** provided by the **nprcgenekeepr** package. This is a brief tutorial that illustrates a typical workflow and does not explore all possible workflows.

Please provide any comments, questions, or bug reports through the GitHub issue tracker at <https://github.com/rmsharp/nprcgenekeepr/issues>.

Installation and Help

Downloading R

Go to <https://cran.r-project.org/mirrors.html> and click on one of the links within the US (ex. <https://ftp.osuosl.org/pub/cran/>). Then proceed to download the latest version of **R** for your desired operating system. Note: IT may have to install if you do not have administrative permissions on the computer.

Installing nprcgenekeepr

To get the most recent version you can install **nprcgenekeepr** from GitHub with the following code (Mac: command + return; Windows/Linux: ctrl + enter).

```
install.packages("devtools")
devtools::install_github("rmsharp/nprcgenekeepr")
```

All missing packages should be automatically installed.

You can start the **nprcgenekeepr** from the **R** console with

```
library(nprcgenekeepr)
runGenekeepr()
```

This will result in the opening screen where you tell the application how to find the pedigree you will be using.

Uploading a Pedigree File

Pedigree File Structure

Most of the screen is filled with information about formatting a text or Excel worksheet pedigree file.

http://127.0.0.1:6012 | Open in Browser |  Genetic Management Tools - Version 0.5.30 (20190829) | Publish

Input Pedigree Browser Pedigree Age Plot Genetic Value Analysis Summary Statistics and Plots Breeding Group Formation Genetic Value Analysis and Breeding Group Description

NATIONAL PRIMATE RESEARCH CENTERS

Causes | Preventions | Treatments | Cures

Select how you are submitting data.

File Type

- Excel
- Text

File Content

- Pedigree(s) file only; genotypes not provided
- Pedigree(s) and genotypes in one file
- Pedigree(s) and genotypes in separate files
- Breeders only; pedigree obtained from database

Select Pedigree File

Browse... No file selected

Minimum Parent Age (years)

4.0

Read and Check Pedigree

Debug on

Input File Handling:

- This software requires an input file for each breeding colony to be analyzed. Multiple separate pedigrees within the same breeding colony can be accommodated within the same input file.
- If an animal is missing information on one parent, that cell should be blank. For animals with no information on either parent, both cells under "Sire ID" and "Dam ID" should be blank. For animals with information on one parent a unique "placeholder" ID (designated as UnknownID and UnkID in Pedigree Brower tab) will be assigned by the software to represent the unknown parent. Please be aware that animals with no parents will be treated as founders in these calculations, i.e., sources of new genetic variation in the colony.
- Genotypic information for a single locus can be included in the genetic analyses. See [Pedigree File Format](#) and [Genotype File Format](#) below for file descriptions. Observed genotypic information can be supplied for any members of the pedigree. Designation of two alleles is required as currently there is no accommodation for partial genetic information for an individual. Please contact [R. Mark Sharp](#) if you need to be able to enter partial genetic information.

Pedigree File Format Details:

The following column names are reserved and will be recognized:
Ego ID, Sire ID, Dam ID, and Sex are required columns.
Either Age or Birth must also be provided.

- Use the Allowable Names listed in the table below. Allowable Name case and order are not significant. The first Allowable Name listed for each column is the name used internally and is the name used in examples and output.
- If the Age column is provided, the program will use the user-specified age.
- Otherwise, the program will calculate age based on the Birth column and the current date, or the Death, Departure, or Exit columns, if provided.
- Exit dates, if provided, will supercede Death and Departure dates. If Exit is not provided, Death and Departure will be combined based on which date is chronologically first. (This functions primarily to catch clerical errors or historical changes in record-keeping practices).
- Dates should be in the format YYYY-MM-DD (this date option can be found under the "English-U.K." locale option under the "Format

Scrolling down to the middle of the opening screen exposes a table that describes a pedigree file and further instructions.

http://127.0.0.1:6012 | Open in Browser |

No file selected

Minimum Parent Age (years)
4.0

Debug on

Pedigree File Format Details:

The following column names are reserved and will be recognized:
Ego ID, *Sire ID*, *Dam ID*, and *Sex* are required columns.
 Either *Age* or *Birth* must also be provided.

- Use the Allowable Names listed in the table below. Allowable Name case and order are not significant. The first Allowable Name listed for each column is the name used internally and is the name used in examples and output.
- If the *Age* column is provided, the program will use the user-specified age.
- Otherwise, the program will calculate age based on the *Birth* column and the current date, or the *Death*, *Departure*, or *Exit* columns, if provided.
- Exit* dates, if provided, will supersede *Death* and *Departure* dates. If *Exit* is not provided, *Death* and *Departure* will be combined based on which date is chronologically first. (This functions primarily to catch clerical errors or historical changes in record-keeping practices).
- Dates should be in the format YYYY-MM-DD (this date option can be found under the "English-U.K." locale option under the "Format Cells -> Number -> Date" menu in Excel 2010).
- Genotype data may be supplied within the pedigree file or in a separate genotype file. Only two additional columns (*first* and *second*) are required when the genotypes are provided within the pedigree file. The columns are described below.

Allowable Name	Description of Information	Data Format
<i>id</i> , <i>egoid</i> , <i>ego_id</i>	Ego ID: Unique animal identifier	Alphanumeric characters (no symbols)
<i>sire</i> , <i>sireid</i> , <i>sire_id</i>	Sire ID: Unique identifier of the ego's father	Alphanumeric characters (no symbols)
<i>dam</i> , <i>damid</i> , <i>dam_id</i>	Dam ID: Unique identifier of the ego's mother	Alphanumeric characters (no symbols)
<i>sex</i>	Sex: Ego's sex	Sex can be indicated by any of the following: Male: ("male", "m", "1"), Female: ("female", "f", "2"), Unknown: ("unknown", "u", "3"), Hermaphrodite: ("hermaphrodite", "h", "4")
<i>age</i>	Age: Age at exit, or current age of the ego.	Age in decimal years
<i>birth</i> , <i>birthdate</i> , <i>birth_date</i>	Birth: Ego's date of birth	YYYY-MM-DD Format
<i>death</i>	Death: Ego's date of death	YYYY-MM-DD Format
<i>departure</i>	Departure: Ego's date of sale	YYYY-MM-DD Format
<i>exit</i>	Exit: Ego's date of exit from the colony. (Supersedes Death or Departure information)	YYYY-MM-DD Format
<i>allele_1</i>	Allele 1: Alphanumeric representation of first allele	Currently limited to alphanumeric values, underscores, spaces and dashes. Other characters have not been tested.
<i>allele_2</i>	Allele 2: Alphanumeric representation of second allele	Currently limited to alphanumeric values, underscores, spaces and dashes. Other characters have not been tested.

Scrolling down to the bottom of the opening screen exposes more pedigree file instructions, a table that describes a genotype file and instructions regarding use of a genotype file.

http://127.0.0.1:6012 | Open in Browser |  Representations in second screen | Characters have not been tested. | Publish

- You will need to save your input file as a plain text file (either .txt or .csv), in order to upload it correctly. To do this, you may use the "Save As" function in Excel to save a spreadsheet in a "tab-delimited" or "comma-delimited" file format.
- Select the correct file type, and then click on "Select Input File" to upload the file.
- The pedigree information displayed in the Pedigree Browser will not appear exactly the same as the input file. The pedigree shown in the Pedigree Browser has been built using input file contents that may be used to produce other columns of information displayed. Input file processing will include:
 - A check for the 4 required fields: Ego ID, Sire ID, Dam ID, and Sex.
 - A check for either Age or Birth information.
 - A new row entry will be added for any Sire or Dam that do not already have their own row as an Ego.
 - Animals will be checked to ensure that their sex is consistent throughout the file.
 - Duplicate rows will be removed.
 - If Death or Departure date columns have been included, an "Exit" column will be built. Otherwise, Exit date assignment will be skipped if an Exit date column was provided in the input file.
 - If Birth dates have been included, age will be calculated as (Birth Date - Exit Date) / 365.25. If no Death, Departure, or Exit dates are provided, it is assumed that the ego is still alive, and Age will be calculated as (Birth Date - Today's Date) / 365.25. Age calculation will be skipped if an Age information column is included in the input file.
 - Generation numbers will be added to each ego.
 - Families (Pedigrees) will be numbered and this information will be added to each ego.

Genotype File Format Details:

The following columns names are reserved and will be recognized:

Columns `id`, `first_name`, and `second_name` are required columns.

Columns `first` and `second` will be automatically generated and cannot be present.

- Use the Allowable Names listed in the table below. Allowable Name case and order are not significant. The Allowable Name listed for each column is the name used internally and is the name used in examples and output.
- Genotype data can be supplied within the pedigree file or in this genotype file. Only two additional columns (`first` and `second`) are required when the genotypes are provided within the pedigree file. See table above for a description. The genotype file columns are described below.

Allowable Name	Description of Information	Data Format
<code>id</code>	ID: Unique animal identifier	Alphanumeric characters (no symbols)
<code>allele_1</code>	Allele 1: Alphanumeric representation of first allele	Currently limited to alphanumeric values, underscores, spaces and dashes.
<code>allele_2</code>	Allele 2: Alphanumeric representation of second allele	Currently limited to alphanumeric values, underscores, spaces and dashes. Other characters have not been tested.

- You will need to save your genotype file as a plain text file (either .txt or .csv), in order to upload it correctly. To do this, you may use the "Save As" function in Excel to save a spreadsheet in a "tab-delimited" or "comma-delimited" file format.
- Select the correct file type, and then click on "Select Genotype File" to upload the file.

The following is an example of the pedigree file format.

Without genotypes:

```
Error in readPNG("./shiny_app_use/examplePedigreeTutorial.png") : unable to open
./shiny_app_use/examplePedigreeTutorial.png Calls: <Anonymous> ... withCallingHandlers ->
withVisible -> eval -> eval -> readPNG Execution halted
```

With genotypes:

```
✖ Line 128 Error in readPNG("./shiny_app_use/examplePedigreeTutorial_with_alleles.png") : unable to
open ./shiny_app_use/examplePedigreeTutorial_with_alleles.png Calls: <Anonymous> ...
withCallingHandlers -> withVisible -> eval -> eval -> readPNG Execution halted
```

Uploading the Pedigree

In this introductory tutorial, we will use an Excel file containing a hypothetical pedigree of macaques. We will work with the gray box on the left at the top of the screen.

http://127.0.0.1:6012 | Open in Browser |  NATIONAL PRIMATE RESEARCH CENTERS

Genetic Management Tools - Version 0.5.30 (20190829)

Input Pedigree Browser Pedigree Age Plot Genetic Value Analysis Summary Statistics and Plots Breeding Group Formation Genetic Value Analysis and Breeding Group Description

Genetic Management Tools

Causes | Preventions | Treatments | Cures

Select how you are submitting data.

File Type

- Excel
- Text

File Content

- Pedigree(s) file only; genotypes not provided
- Pedigree(s) and genotypes in one file
- Pedigree(s) and genotypes in separate files
- Breeders only; pedigree obtained from database

Select Pedigree File

Browse... No file selected

Minimum Parent Age (years)

4.0

Read and Check Pedigree

Debug on

Input File Handling:

- This software requires an input file for each breeding colony to be analyzed. Multiple separate pedigrees within the same breeding colony can be accommodated within the same input file.
- If an animal is missing information on one parent, that cell should be blank. For animals with no information on either parent, both cells under "Sire ID" and "Dam ID" should be blank. For animals with information on one parent a unique "placeholder" ID (designated as UnknownID and UnkID in Pedigree Brower tab) will be assigned by the software to represent the unknown parent. Please be aware that animals with no parents will be treated as founders in these calculations, i.e., sources of new genetic variation in the colony.
- Genotypic information for a single locus can be included in the genetic analyses. See [Pedigree File Format](#) and [Genotype File Format](#) below for file descriptions. Observed genotypic information can be supplied for any members of the pedigree. Designation of two alleles is required as currently there is no accommodation for partial genetic information for an individual. Please contact [R. Mark Sharp](#) if you need to be able to enter partial genetic information.

Pedigree File Format Details:

The following column names are reserved and will be recognized:
Ego ID, Sire ID, Dam ID, and Sex are required columns.
Either Age or Birth must also be provided.

- Use the Allowable Names listed in the table below. Allowable Name case and order are not significant. The first Allowable Name listed for each column is the name used internally and is the name used in examples and output.
- If the Age column is provided, the program will use the user-specified age.
- Otherwise, the program will calculate age based on the Birth column and the current date, or the Death, Departure, or Exit columns, if provided.
- Exit dates, if provided, will supersede Death and Departure dates. If Exit is not provided, Death and Departure will be combined based on which date is chronologically first. (This functions primarily to catch clerical errors or historical changes in record-keeping practices).
- Dates should be in the format YYYY-MM-DD (this date option can be found under the "English-U.K." locale option under the "Format

A Microsoft Excel workbook with a single worksheet is the default file type; though comma (.csv), semi-colon (.txt), and tab (.txt) separated value files are all acceptable formats.

The *Example_Pedigree.xlsx* file we are using is from a CSV file created as shown below and then saved in an Excel format. Alternatively, you may download this example pedigree from the nprc consortium web page.

[makeExamplePedigreeFile\(\)](#)

Select the **Browse** button and select the pedigree file from your file system.

Select how you are submitting data.

File Type

Excel
 Text

File Content

Pedigree(s) file only; genotypes not provided
 Pedigree(s) and genotypes in one file
 Pedigree(s) and genotypes in separate files
 Breeders only; pedigree obtained from database

Select Pedigree File

Browse... Example_Pedigree.xlsx

Upload complete

Minimum Parent Age (years)

4.0

Read and Check Pedigree

Debug on

It is important to make sure the minimum parent age is low enough for the animals in your pedigree. For our example pedigree, we are changing it from 4 years to 2 years of age since these macaques may reproduce as early as two years of age.

Select how you are submitting data.

File Type

- Excel
- Text

File Content

- Pedigree(s) file only; genotypes not provided
- Pedigree(s) and genotypes in one file
- Pedigree(s) and genotypes in separate files
- Breeders only; pedigree obtained from database

Select Pedigree File

Example_Pedigree.xlsx

Minimum Parent Age (years)

Debug on

Select how you are submitting data.

File Type

- Excel
- Text

File Content

- Pedigree(s) file only; genotypes not provided
- Pedigree(s) and genotypes in one file
- Pedigree(s) and genotypes in separate files
- Breeders only; pedigree obtained from database

Select Pedigree File

Example_Pedigree.xlsx

Minimum Parent Age (years)

Result: Parents must be at least as old as the minimum parent age at the birthdate of an offspring. If not, the file will not be accepted and a file named lowParentAge.csv containing a list of parents below the minimum age will be written to the user's home directory. Animals without birthdates are not affected by this rule.

Debug on

Select how you are submitting data.

File Type

- Excel
- Text

File Content

- Pedigree(s) file only; genotypes not provided
- Pedigree(s) and genotypes in one file
- Pedigree(s) and genotypes in separate files
- Breeders only; pedigree obtained from database

Select Pedigree File

Example_Pedigree.xlsx

Minimum Parent Age (years)

Debug on

Upload a Pedigree and Testing for Errors

Selected **Read and Check Pedigree** will read in the file and test to see if the pedigree file has all of the columns needed and the pedigree is internally consistent.

Select how you are submitting data.

File Type

- Excel
- Text

File Content

- Pedigree(s) file only; genotypes not provided
- Pedigree(s) and genotypes in one file
- Pedigree(s) and genotypes in separate files
- Breeders only; pedigree obtained from database

Select Pedigree File

Upload complete

Minimum Parent Age (years)

Debug on

Several error types, shown below, are detected by the application.

Error	Definition
failedDatabaseConnection	Database connection failed: configuration or permissions are invalid
missingColumns	Columns that must be within the pedigree file are missing.
invalidDateRows	Values, which are supposed to be dates, cannot be interpreted as a date.
suspiciousParents	Parents were too young on the date of birth of to have been the parent.
femaleSires	Individuals listed as female or hermaphroditic and as a sire.
maleDams	Individuals are listed as male and as a dam.
sireAndDam	Individuals who are listed as both a sire and a dam.
duplicateIds	IDs listed more than once.
fatalError	System Crash. These are unanticipated errors that should be reported to the package maintainer
changedCols	Columns that have been changed to conform to the program (ex: id -> Ego ID).

Pedigree Browser

The **Pedigree Browser** tab defaults to displaying 10 rows of the pedigree at a time, but you can choose to display 10, 25, 50, or 100 rows. You can choose to display UNKNOWN IDs in the rows displayed. UNKNOWN IDs (UIDs) are used to label unknown parents of animals with one known parent. The program calculates additional columns based on the input pedigree.

http://127.0.0.1:6012 | Open in Browser |

Genetic Management Tools - Version 0.5.35 (20191013)

Input Pedigree Browser Pedigree Age Plot Genetic Value Analysis Summary Statistics and Plots Breeding Group Formation Genetic Value Analysis and Breeding Group Description

IDs of selected focal animals may be manually entered here if analysis of all individuals is not needed. IDs may be pasted from Excel or you can browse for and select a file containing the list of focal animals.

Display Unknown IDs Trim pedigree based on focal animals Export

A population must be defined before proceeding to the Genetic Value Analysis.

Choose CSV file with focal animals Browse... No file selected

Update Focal Animals The search field below will search all columns for matches to any text or number entered.

Show 10 entries Search []

Ego ID	Sire ID	Dam ID	Sex	Generation #	Birth Date	Exit Date	Age (in years)	Ancestry	Origin	Status	Original/ Added	NA	Breeding Colony Member	Pedigree #
1 01WY5E			M	0	2005-07-07	2005-08-26	0.1	JAPANESE		DECEASED	original	original	true	1
2 02GZ4L			U	0	2004-07-05	2004-08-15	0.1	JAPANESE		DECEASED	original	original	true	2
3 079ZJK			F	0		1975-01-16		JAPANESE	JAPAN MONKEY CENTER PRIMATE ZOO	DECEASED	original	original	true	3
4 08CF4C			F	0		1966-03-19		JAPANESE	JAPAN MONKEY CENTER PRIMATE ZOO	DECEASED	original	original	true	4
5 093AB5			M	0	2004-09-03	2005-01-29	0.4	JAPANESE		DECEASED	original	original	true	5
6 0CRGND			F	0	1982-04--	1982-04--	0.1	JAPANESE		DECEASED	original	original	true	6

Unknown IDs

I have placed red lines under the UNKNOWN IDs in the partial pedigree list below for clarity. These IDs have no meaning other than they all begin with the letter *U* and are followed with a left alphanumeric string of five places.

3684	I04JZV	AEP5EG	T3QPW5	M		8	2009-07-20	2010-05-17	0.8
3685	IJT6B6	633UQI	7NE2UT	M		8	2014-04-17		5.2
3686	RLJC25	UH5ZMC	46ZHKN	M		8	2014-03-20	2014-04-01	0
3687	S63QDN	4JARCS	Y0TCYX	M		8	2011-06-27		8
3688	W761EE	U5Q4MV	FL170P	F		8	2014-04-02		5.2
3689	X994RC	UBX18G	FL170P	F		8	2008-06-29	2009-08-25	1.2
3690	Y9BAP6	UZ51CU	46ZHKN	M		8	2013-04-09	2013-04-30	0.1

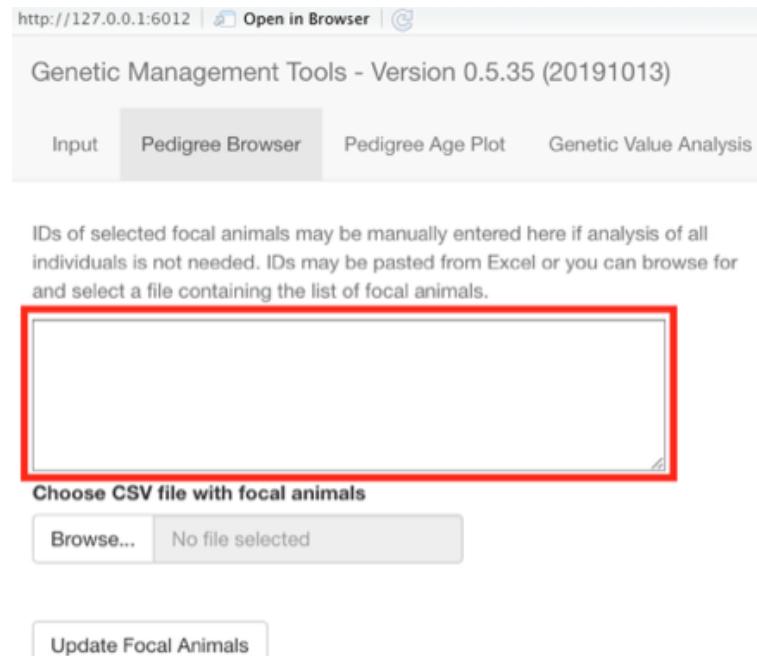
In this example pedigree, when you deselect the **Display Unknown IDs** checkbox. The number of rows reduces from 3,694 to 2,322, because there were 1,372 UNKNOWN animals generated when constructing the pedigree to provide sire and dam placeholders for all animals.

97	B267C6			M	0		03-12	
98	B2CW2H			M	0	1995-09-20	2001-06-20	
99	B2I259			M	0		1982-05-11	
100	B6V285			M	0	1989-07-24	1998-01-28	

Showing 1 to 100 of 2,322 entries

Selecting a Pedigree Subset — Focal Animals

The **Pedigree Browser** tab displays the full pedigree by default but allows you to select a subset of the pedigree by entering a list of animals of interest (*focal animals*).



http://127.0.0.1:6012 | Open in Browser |

Genetic Management Tools - Version 0.5.35 (20191013)

Input Pedigree Browser **Pedigree Age Plot Genetic Value Analysis**

IDs of selected focal animals may be manually entered here if analysis of all individuals is not needed. IDs may be pasted from Excel or you can browse for and select a file containing the list of focal animals.

Choose CSV file with focal animals

No file selected

You can enter in the animal IDs by typing them into the text box directly as shown below (FJS7RQ, H6T2FF, HEVL3L, I04JZV, S63QDN). Deselect the **Display Unknown IDs** checkbox and select the **Trim pedigree based on focal animals** checkbox. (See top right of image below).

Trimming the pedigree based on focal animals will keep only animals in the pedigree that are related to the focal animals selected.

Select the **Update Focal Animals** button to tell the application to read your list of animals, trim the pedigree based on that list, and display the trimmed pedigree below. You will end up with 54 animals in your pedigree.

IDs of selected focal animals may be manually entered here if analysis of all individuals is not needed. IDs may be pasted from Excel or you can browse for and select a file containing the list of focal animals.

- Display Unknown IDs
 Trim pedigree based on focal animals



A population must be defined before proceeding to the Genetic Value Analysis.

Choose CSV file with focal animals

No file selected

Clear Focal Animals

The search field below will search all columns for matches to any text or number entered.

Show 10 entries

	Ego ID	Sire ID	Dam ID	Sex	Generation #	Birth Date	Exit Date	Age (in years)
47	4Z1F6			F	0	1997-09-21		
50	5D0XXP			F	0	1987-10-04		
149	FJLBZR			F	0	1982-12-24		
182	KKYQP2			F	0	1978-09-23		
202	NSMEH6			F	0	1992-01-06		
204	NU5H65			F	0	1987-08-29		
1623	VF5TA6			F	0	1982-05-17		
1626	VJ08BW			F	0	1971-11-21		
1630	VXTG4H			F	0	1979-11-27		
1692	4Y8JHT		NU5H65	F	1	1979-06-04	2002-03-26	

Showing 1 to 10 of 54 entries

Also, you can import a list of focal animals by selecting the **Browse** button under **Choose CSV file with focal animals**. This file can be constructed by creating a simple text file with commas between animal IDs or by placing individual animal IDs on separate lines.

Focal animals are the list of animals that will be used in the following analysis. In most cases, we recommend using all alive animals in the breeding population. By selecting focal animals, the number of pedigree entrees does not change, but the population membership flag will be set to “true” for the focal animals, and “false” for all other animals.

IDs of selected focal animals may be manually entered here if analysis of all individuals is not needed. IDs may be pasted from Excel or you can browse for and select a file containing the list of focal animals.

Choose CSV file with focal animals

focalAnimalsLargeSet.csv

Upload complete

After entering your list of focal animals, you can select to trim the pedigree so that it will only include relatives of the focal animals you have selected. This will reduce the number of members within the pedigree to all animals required to connect all of the focal animals in the pedigree.

IDs of selected focal animals may be manually entered here if analysis of all individuals is not needed. IDs may be pasted from Excel or you can browse for and select a file containing the list of focal animals.

01QRQ4,CLSVU6,1SPLS8,5IAFMK,HLQ9SY,XFWVVX,6X6BG9,B2
2806,B2CKHA,BCJJKN,DCJJYS,DKIM6U,FB5L3N,GCBYDW,GDX
WJ1,JSAP3H,KZM9RB,MB6NYQ,QWKFBH,RJ4JPC,SH3FB7,TR5
L57,WK8919,XC304E,Z7NBA2,ZATMEE,1E8KD1,1QVS67,50D771,5
ERY5Z,5KFB90,5PW7WT,A6A1M1,AEP5EG,AW400C,BW10CL,CF

Display Unknown IDs
 Trim pedigree based on focal animals

Trim the pedigree to include only relatives of the focal animals provided.

Choose CSV file with focal animals

Browse... focalAnimalsLargeSet.csv

Upload complete

A pedigree trimmed based on focal animals will have only the relatives of those animals remaining. In this instance there are only a total of 85 focal animals and their relatives. *Note: focal animals and their relatives will only be included in the same pedigree when the original pedigree file uploaded indicates a common ancestor for them. Otherwise, focal animals and their relatives will be sorted into separate pedigrees in the output, with each separate pedigree indicated by its own number.*

					♂	♀
182	KKYQP2			F	0	1978-09-23
184	L219W7			F	0	1990-11-05
202	NSMEH6			F	0	1992-01-06

Showing 1 to 10 of 522 entries

You can remove the animals from the list of focal animals by selecting the **Clear Focal Animals** checkbox and selecting the **Update Focal Animals** button. This will read in an empty ID list, clear the box of IDs, and bring back all of the trimmed away IDs.

Genetic Management Tools - Version 0.5.42.9009 (20200402)

Input Pedigree Browser Pedigree Age Plot Genetic Value Analysis

IDs of selected focal animals may be manually entered here if analysis of all individuals is not needed. IDs may be pasted from Excel or you can browse for and select a file containing the list of focal animals.

```
01QRQ4,CLSVU6,1SPLS8,5IAFMK,HLQ9SY,XFWVVX,6X  
6BG9,B228Q6,B2CKHA,BCJJKN,DCJJYS,DKIM6U,FB5L  
3N,GCBYDW,GDXWJ1,JSAP3H,KZM9RB,MB6NYQ,QW  
KFBH,RJ4JPC,SH3FB7,TR5L57,WK8919,XC304E,Z7NBA  
2,ZATMEE,1E8KD1,1QVS67,50D771,5ERY5Z,5KFB90,5P
```

Choose CSV file with focal animals

Browse... focalAnimals.csv Upload complete

Clear Focal Animals

Tip: Focal animal list is to be cleared as well using a file with no animal IDs. (Files are not affected).

Genetic Management Tools - Version 0.5.42.9009 (20200401)

Input Pedigree Browser Pedigree Age Plot Genetic Value Analysis

IDs of selected focal animals may be manually entered here if analysis of all individuals is not needed. IDs may be pasted from Excel or you can browse for and select a file containing the list of focal animals.

Choose CSV file with focal animals

Browse... focalAnimals.csv Upload complete

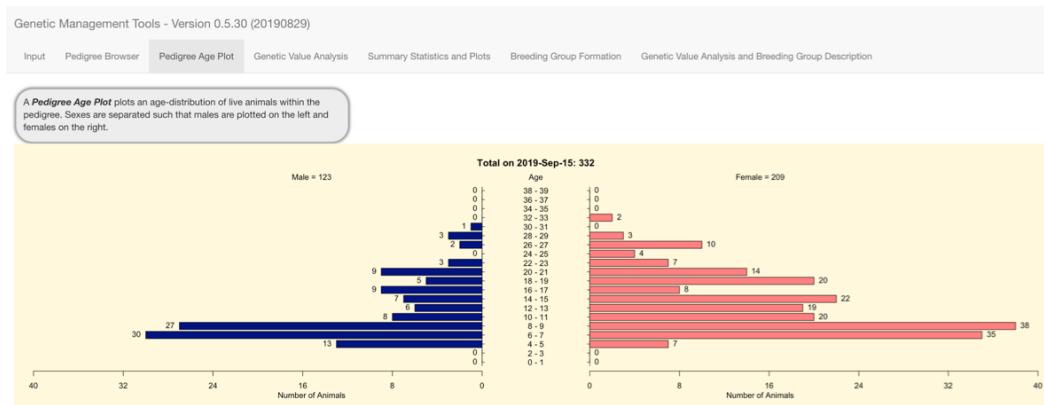
Clear Focal Animals

Showing 1 to 10 of 2,322 entries

Deselect the **Clear Focal Animals** checkbox and reselect the **Update Focal Animals** button before continuing with the tutorial so that we will be working with the trimmed pedigree.

Pedigree Age Plot

The **Pedigree Age Plot** tab displays a standard pyramid plot for the pedigree as selected in the **Pedigree Browser** tab. This is showing 332 living animals from the entire example pedigree.



Genetic Value Analysis

Select the **Genetic Value Analysis** tab and enter the number of simulations and genome uniqueness threshold desired. See the **Genetic Value Analysis and Breeding Group Description** tab for a breakdown of the calculation.

We recommend trying multiple numbers of simulations to arrive at an ideal number that produces consistent results (i.e., 1,000). Genome uniqueness values are calculated using a gene-drop simulation according to MacCluer et al. (1986) and Ballou & Lacy (1995), by assigning unique alleles to all pedigree founders, and simulating their segregation throughout the pedigree according to Mendelian rules. Genome uniqueness is a measure of the probability that an animal possesses founder alleles that are present in at most x other animals (usually 0-3), and thus are rare and at risk of being lost from the population. A range of 2 to 100,000 simulations may be selected. A minimum of 1,000 simulations is recommended. A genome uniqueness threshold value between 0-3 should also be selected, as desired.

Select the **Begin Analysis** button to start the gene dropping process, which you can monitor with the progress meter in the lower right corner of the display.

Genetic Management Tools - Version 0.5.30 (20190829)

Input Pedigree Browser Pedigree Age Plot Genetic Value Analysis Summary Statistics and Plots Breeding Group Formation Genetic Value Analysis and Breeding Group Description

A genetic value analysis ranks animals according to their genetic value to the breeding colony, based on an algorithm that incorporates measures of mean kinship to the rest of the colony, as well as the probability that an animal contains rare founder alleles that should be conserved in the colony.

For a full description of the algorithm, see the **Genetic Value Analysis and Breeding Group Description** tab.

Enter the number of simulations for the gene-drop analysis:

Enter the genome uniqueness threshold:

Analysis may take a significant amount of time (>20 min)

Begin Analysis

Results:
Enter IDs of animals to be viewed (separate with comma, semicolon, tab or return):
(Leave blank to view all)

Filter View

Export All **Export Current Subset**

message Calculating Genome Uniqueness

As soon as the calculations are completed, a table showing the results of the analysis is displayed in 10 rows. The calculations for 1000 iterations of the gene dropping algorithm took 1 minute 38 seconds with the example pedigree of 3,691 animals using a MacBook Pro (Mid 2014), 2.8 GHz Intel Core i7 with 16 GB of 1600 MHz DDR3 memory.

Again you can select how many rows to display at once by changing the values in the **Show entries** selection tool, or you may view results for a specific set of animals by entering their IDs in the box above. The default is to display the analysis for the list of focal animals.

Show 10 entries Search:

	Ego ID	Sex	Age (in years)	Birth Date	Exit Date	Breeding Colony Member	Origin	Individual Mean Kinship	Z-score (Mean Kinship)	Genome Uniqueness (%)	Total Offspring	Living Offspring	Value Designation	Rank
1	CLSVU6	F	23.9	1995-08-02		true		0.00267	-1.97	100	1	1	High Value	1
2	KZM9RB	M	30.1	1989-05-03		true		0.00329	-1.9	90.1	0	0	High Value	2
3	1SPLS8	F	7.9	2011-07-26		true		0.00373	-1.83	83.55	0	0	High Value	3
4	G2GYST	M	19.1	2000-05-16		true		0.00488	-1.64	83.05	2	2	High Value	4
5	01QRQ4	F	18.2	2001-04-04		true		0.00373	-1.83	82.85	0	0	High Value	5
6	Q8U9LB	F	18.1	2001-05-24		true		0.00566	-1.52	82	1	1	High Value	6
7	3MMZD4	M	12.2	2007-03-24		true		0.00536	-1.57	80.85	0	0	High Value	7
8	50D771	F	20.1	1999-05-08		true		0.00451	-1.71	79.3	0	0	High Value	8
9	WK8919	F	21.1	1998-05-26		true		0.00582	-1.49	79.65	0	0	High Value	9
10	IZDV8K	M	7.7	2011-09-29		true		0.0048	-1.66	78.5	0	0	High Value	10

Showing 1 to 10 of 327 entries

Previous 1 2 3 4 5 ... 33 Next

Searching down the table of results in the **Value Designation** column you can see starting at row 268 the values change from *High Value* to *Low Value*. Though not shown here, the value of *Undetermined* in the **Value Designation** column means the animal did not have parentage information. Infants or very young animals without assigned parents are given an “Undetermined” designation. Founders also do not have parentage information but are high value by definition.

267	WRY2SZ	F	10	2009-07-05		true		0.01505	0.01	7.55	0	0	High Value	267
268	1CZM30	F	13.3	2006-03-02		true		0.01547	0.08	9	2	2	High Value	268
269	MQT080	M	19	2000-06-10		true		0.01682	0.3	5.55	12	12	Low Value	269
270	YFCIHJ	F	10.4	2009-01-20		true		0.01715	0.35	7.85	1	1	Low Value	270

Showing 261 to 270 of 327 entries

Previous 1 ... 26 27 28 ... 33 Next

Summary Statistics

The **Summary Statistics and Plots** tab used results from the **Genetic Value Analysis** tab. Definitions of genome uniqueness and kinship are located in the **Genetic Value Analysis and Breeding Group Description** tab. Additionally, definitions of founder equivalents and founder genome equivalents are located at the bottom of the **Summary Statistics and Plots** tab.

Summary Statistics

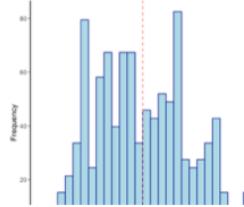
After running a Genetic Value Analysis, this page will contain summary statistics, histograms, and box plots for the population under consideration.

Note: Founder equivalents and founder genome equivalents will be inflated due to partial parentage information. Calculating these statistics necessitates adding placeholder IDs to fill out partial parentage information. Each placeholder ID used is treated as a founder. Calculations of f_u and f_g may be changed at some point in the future to try to correct for this situation. Export Kinship Matrix Export First-Order Relationships

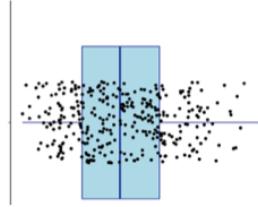
Known Founders	Known Female Founders	Known Male Founders	Founder Equivalents	Founder Genome Equivalents
20	17	3	109.67	47.79

	Min	1st Quartile	Mean	Median	3rd Quartile	Max
Mean Kinship	0.0029	0.0101	0.015	0.0147	0.0195	0.0318
Genome Uniqueness	0.05	12.05	30.2723	24.75	46.3	100

Distribution of Individual Mean Kinship Coefficients



Boxplot of Individual Mean Kinship Coefficients



The **Export Kinship Matrix** button creates a CSV file that has a row and column for each individual in the genetic analysis plus a first row and first column each containing the IDs.

The first few rows of such a file are shown below.

	Kinship.csv													
~./Downloads/Kinship.csv ~														
1	,"", "01QRQ4", "CLSVU6", "1SPLS8", "5IAFMK", "HLQ9SY", "XFWVVX", "6X6BG9", "B228Q6", "B2CKHA", "BCJJKN", "DCJJYS", "DKIM6U", "FB!													
2	"01QRQ4", 0.5,0,													
3	"CLSVU6", 0,0,0.5,0,													
4	"1SPLS8", 0,0,0.5,0,													
5	"5IAFMK", 0,0,0,0.5,0,0,0,0.0625,0,0,0,0.25,0,0,0,0,0.0625,0,0,0,0.0625,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0.25,0,0,0,													
6	"HLQ9SY", 0,0,0,0,0.5,0,0,0.015625,0,0.015625,0,0.015625,0.015625,0,0,0.25,0,0,0,0,0,0,0.015625,0,0,0,0,0,0,0,0,0,0,0.03125,													
7	"XFWVVX", 0,0,0,0,0.5,0,0,0.015625,0,0.015625,0.015625,0,0.015625,0,0,0.25,0,0,0,0,0,0,0.015625,0,0,0,0,0,0,0,0,0,0.03125,													
8	"6X6BG9", 0,0,0,0,0.5,0,0,0.015625,0,0.015625,0,0.015625,0,0.015625,0,0,0.25,0,0,0,0,0,0,0.015625,0,0,0,0,0,0,0,0,0,0.0625,													
9	"B228Q6", 0,0,0,0,0.0625,0,0,0.5,0,0,0,0,0.03125,0,0,0,0,0.03125,0,0,0,0.03125,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0.03125,													
10	"B2CKHA", 0,0,0,0,0.015625,0,0.125,0,0.5,0,0.03125,0.125,0,0,0.0078125,0,0,0,0.0078125,0,0,0,0.0078125,0,0,0,0.03125,0.03125,0,0,0,0,													
11	"BCJJKN", 0,0,0,0,0,0,0,0.5,0,0,0,0,0,0.0078125,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0.00390625,0,0,0,0.00390625,0.00390625,													
12	"DCJJYS", 0,0,0,0,0.015625,0,0.03125,0,0.03125,0,0.03125,0,0.03125,0,0.03125,0,0.03125,0,0,0.0078125,0,0,0,0,0,0,0,0,0.125,0,0,0,0,0,0,0,0.0625,													
13	"DKIM6U", 0,0,0,0,0.015625,0,0.125,0,0.03125,0.125,0,0.03125,0.5,0,0,0.0078125,0,0,0,0,0,0.03125,0,0,0,0,0.0078125,0,0,0,0,													

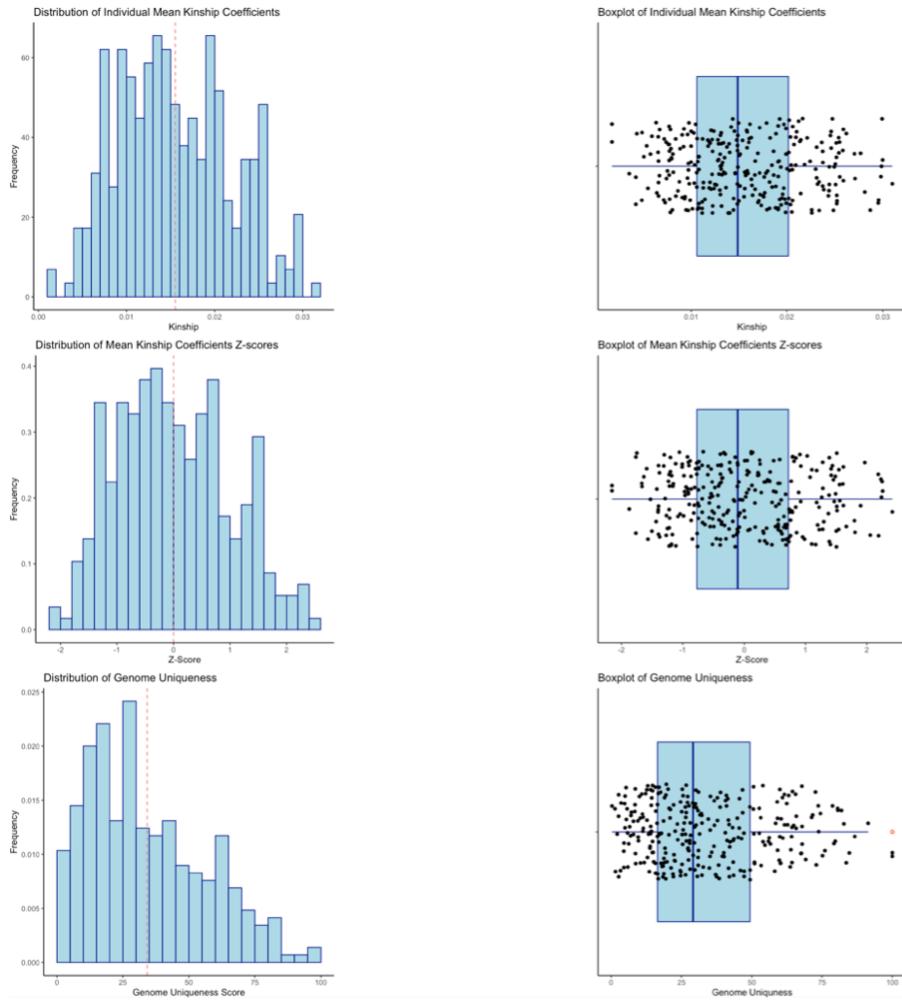
The **First-Order Relationships** button creates a CSV file that has the following columns defined: an unnamed column for row number, *id*, *parents*, *offspring*, *siblings*, and *total*. Counts are based off known relationships. The first few rows of such a file are shown below.

	FirstOrder.csv					
	~/Downloads/FirstOrder.csv					
1	"", "id", "parents", "offspring", "siblings", "total"					
2	"1", "01QRQ4", 0, 0, 0, 0					
3	"2", "CLSVU6", 0, 1, 0, 1					
4	"3", "1SPLS8", 0, 0, 0, 0					
5	"4", "5IAFMK", 0, 2, 0, 2					
6	"5", "HLQ9SY", 0, 3, 0, 3					
7	"6", "XFWVVX", 0, 4, 0, 4					
8	"7", "6X6BG9", 0, 0, 0, 0					
9	"8", "B228Q6", 0, 2, 0, 2					
10	"9", "B2CKHA", 0, 5, 0, 5					
11	"10", "BCJJKN", 0, 1, 0, 1					
12	"11", "DCJJYS", 0, 3, 0, 3					
13	"12", "DKIM6U", 0, 4, 0, 4					
14	"13", "FB5L3N", 1, 3, 0, 4					
15	"14", "GCBYDW", 0, 6, 0, 6					
16	"15", "GDXWJ1", 2, 0, 0, 2					
17	"16", "JSAP3H", 0, 3, 0, 3					
18	"17", "KZM9RB", 0, 0, 0, 0					
19	"18", "MB6NYQ", 0, 0, 0, 0					
20	"19", "QWKFBH", 0, 0, 0, 0					
21	"20", "RJ4JPC", 0, 3, 0, 3					
22	"21", "SH3FB7", 0, 1, 0, 1					
23	"22", "TR5L57", 0, 1, 0, 1					

The **Export Female Founders** and **Export Male Founders** buttons creates a CSV file that has the following columns defined: an unnamed column for row number, *id*, *sires*, *dam*, *sex*, *gen*, *birth*, *exit*, *age*, *recordStatus*, *population*, and *pedNum*. The first few rows of such a file are shown below.

```
✖ Line 487 Error in readPNG("./shiny_app_use/ss_female_founders.png") : unable to open
  ./shiny_app_use/ss_female_founders.png Calls: <Anonymous> ... withCallingHandlers ->
  withVisible -> eval -> eval -> readPNG Execution halted
```

The six plots provide histograms and boxplots for the kinship coefficients, the Z-scores of the kinship coefficients, and the genome uniqueness scores. These plots may also be downloaded as desired.



Breeding Group Formation

Selecting the **Breeding Group Formation** tab brings forward the screen shown below. In this screen you can form breeding groups using one of three workflows based on your source of animals selected under **Choose one group formation workflow:**

Further you must specify how you want to construct the breeding groups with regard to the groups' sex ratios. The third of the three options (*User specified sex ratio of breeders*) causes the appearance of the field where you can fill in the sex ratio (F/M) that you want to have in the formed breeding groups. The sex ratio algorithm will form a group as nearly to the selected ratio as possible given the size of the group. Limits in the availability of either sex will restrict the size of the groups formed.

Choose one group formation workflow:

- Randomly select from only high-value animals in genetic value analysis
- Randomly select from all animals in genetic value analysis
- Use candidate animals entered below to form groups

Sex of animals in groups:

- Ignore sex when forming groups
- Form harems (one male)
- User specified sex ratio of breeders

Optional: Seed Groups with Specific Animals

Number of Groups Desired:

Animals will be grouped with the mother below age:

Animals with kinship above this value will be excluded:

Ignore females at or above the minimum parent age:

Include kinship in display of groups

Number of simulations:

Enter the group to view:

 Export Current Group

 Export Current Group Kinship Matrix

The **Make Groups** button appears once you select the source of animals you are going to use. However, you probably will be making additional selections using other controls on the screen.

The most common source of animals will be the high-value animals found by the genetic analysis.

You can either type in the number of groups that you want to form or select the number of groups using the arrows on the right edge of the **Number of Groups Desired** field, which is outlined in blue in the image below.

Choose one group formation workflow:

- Randomly select from only high-value animals in genetic value analysis
- Randomly select from all animals in genetic value analysis
- Use candidate animals entered below to form groups

Sex of animals in groups:

- Ignore sex when forming groups
- Form harems (one male)
- User specified sex ratio of breeders

Optional: Seed Groups with Specific Animals

Number of Groups Desired:
1

Animals will be grouped with the mother below the minimum parent age of 2.

Animals will be grouped with the mother below age:
1

Animals with kinship above this value will be excluded:
0.015625 (second cousins)

Ignore females at or above the minimum parent age:
Yes

Include kinship in display of groups

Number of simulations:
10

Make Groups

Enter the group to view:
1

There are often behavioral constraints, such as preexisting social groups, that dictate the need to have some animals maintained together. This need is readily accommodated by pre-seeding groups with those social groups. You may select the **Optional: Seed Groups with Specific Animals** field if you decide to place some animals together within the groups because you know them to be compatible with each other.

This has been done in the example below using six groups with differing numbers of seed animals. Note the selection of having animals below the minimum parent age of two being grouped with their mother.

Choose one group formation workflow:

- Randomly select from only high-value animals in genetic value analysis
- Randomly select from all animals in genetic value analysis
- Use candidate animals entered below to form groups

Sex of animals in groups:

- Ignore sex when forming groups
- Form harems (one male)
- User specified sex ratio of breeders

Optional: Seed Groups with Specific Animals

Number of Groups Desired: 6	Seed animals 1 TEACA3 7NE2UT NK802Y GPG4SG	Seed animals 2 N4NV8B 4SUIQ0 H9HL2V	Seed animals 3 AIH8Z DHNQ1W	Seed animals 4 WI38KZ 1CZM30 TTJ0J 0IIAEN 6KLWVC DGZLV3	Seed animals 5 3YHBC1 KEA4QG KX0RJ3	Seed animals 6 CHSCFG GM371F WKY2SZ
---------------------------------------	---	---	--	--	---	---

Animals with kinship above this value will be excluded:
0.015625 (second cousins)

Ignore females at or above the minimum parent age:
Yes

Include kinship in display of groups

Number of simulations:
10

Make Groups

Each group has all of the seed animals that were assigned to it plus additional animals that could be added while satisfying the requirements imposed by the selected settings. I have indicated the seed animals for the first group with red rectangles.

Enter the group to view:
1

Show 25 entries

Search:

	Ego ID	Sex	Age in Years
1OVS67	1OVS67	F	18.3
50D77I	50D77I	F	20.5
7NE2UT	7NE2UT	F	16
8YP6PA	8YP6PA	M	5.4
GPG4SG	GPG4SG	M	5.6
MK4J7G	MK4J7G	F	19.5
NK802Y	NK802Y	F	12.8
R6HV9A	R6HV9A	M	22.5
SH3FB7	SH3FB7	F	28.2
TEACA3	TEACA3	F	18.3
TYEWF1	TYEWF1	F	25.5

Showing 1 to 11 of 11 entries

Previous 1 Next

Display of kinship values requires that the **Include kinship in display of groups** checkbox be selected prior to group formation.

A group of ten animals was formed in the next run after choosing to include kinship and selecting the **Make Groups** button.

Export Current Group Kinship Matrix

Show 10 entries Search:

	Ego ID	Sex	Age in Years
1042	1042	M	28.2
1415	1415	M	19.4
1520	1520	F	17.4
1562	1562	M	16.4
1607	1607	F	15.4
1701	1701	F	14.5
1705	1705	F	15
1759	1759	F	13.4
1846	1846	F	12.4
1954	1954	F	10.1

Showing 1 to 10 of 13 entries Previous 1 2 Next

Show 10 entries Search:

	1042	1415	1520	1562	1607	2087	1701	1705	1759	1846	1954	2315	2396
1042	0.5	0	0	0	0	0.003906	0	0.003906	0	0	0.001953	0	0
1415	0	0.5	0	0	0	0	0	0.001953	0	0	0	0.002441	0
1520	0	0	0.5	0.125	0	0	0	0	0.007812	0.125	0	0.000488	0
1562	0	0	0.125	0.5	0.03125	0	0	0	0.007812	0	0	0	0.007812
1607	0	0	0	0.03125	0.5	0	0	0	0.001953	0	0.000732	0	0.062988
2087	0.003906	0	0	0	0	0.5	0.015625	0.017578	0.007812	0.03125	0.015625	0	0
1701	0	0	0	0	0	0.015625	0.5	0.125	0.003906	0.015625	0	0	0
1705	0.003906	0.001953	0	0	0	0.017578	0.125	0.5	0.003906	0.015625	0.000977	0.000977	0
1759	0	0	0.007812	0.007812	0.001953	0.007812	0.003906	0.003906	0.5	0.007812	0	0	0.001465
1846	0	0	0.125	0	0	0.03125	0.015625	0.015625	0.007812	0.5	0	0.000488	0.000977

Showing 1 to 10 of 13 entries Previous 1 2 Next

The option to select a desired sex ratio allows you to select any ratio desired. However, the ratio obtained is limited by the availability of animals that meet all criteria you have set.

Choose one group formation workflow:

- Randomly select from only high-value animals in genetic value analysis
- Randomly select from all animals in genetic value analysis
- Use candidate animals entered below to form groups

Sex of animals in groups:

- Ignore sex when forming groups
- Form harems (one male)
- User specified sex ratio of breeders

Sex Ratio (F/M):

2.5



Selecting a sex ratio of 2.5 with 6 groups as is illustrated resulted in 5 groups of 20 with a ratio of 14:6 (2.3) and 1 group of 23 with a ratio of 16:7 (2.3).

Groups can be individually exported into file names and locations of your choosing. The corresponding kinship matrix for each group can also be exported.