Shiny Application for Genetic Management of Colonies

R. Mark Sharp

Introduction

This tutorial demonstrates the major functions used within the Shiny application provided by the **nprcmanager** 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/nprcmanager/issues.

Installation and Help

To get the most recent version you can install nprcmanager from GitHub with the following code.

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

All missing dependencies should be automatically installed.

You can start the Shiny application from the ${\bf R}$ console with

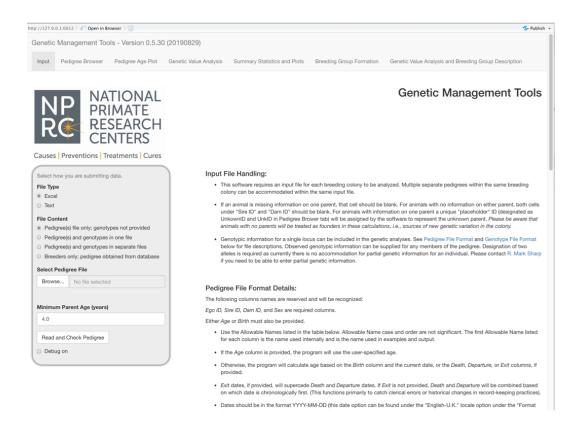
```
library(nprcmanager)
#runManager()
getwd()
```

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

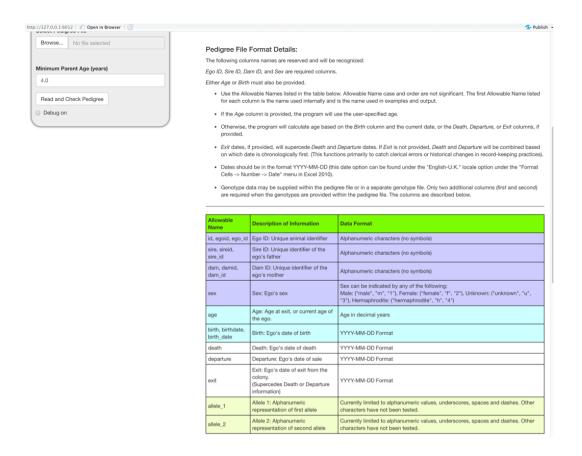
Input

Pedigree File Structure

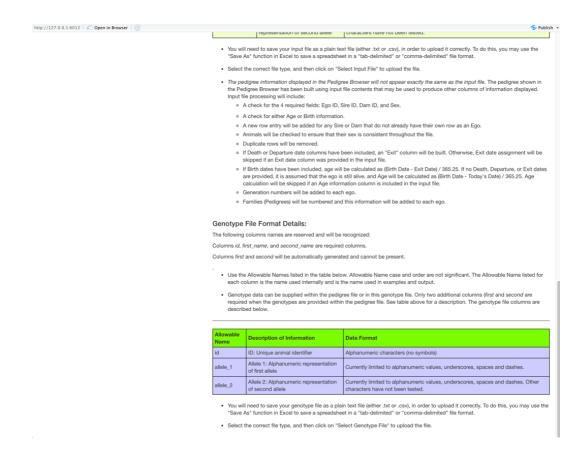
Most of the screen is filled with information about formatting a text or Excel worksheet pedigree file.



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

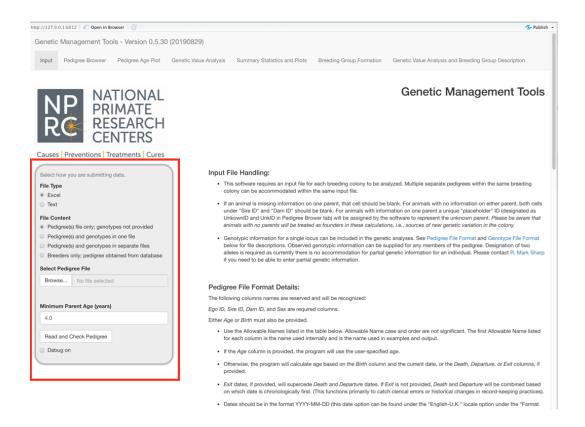


Scrolling down to the bottom of the opening screen exposes more pedigree file instructions, a table that describes a genotype file and instructions regarding use of a genotype file. This tutorial will not include instructions on using a genotype file.



Reading in 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.

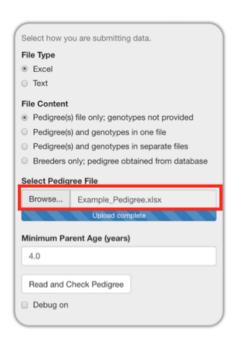


A Microsoft Excel workbook with a single worksheet is the default file type; though comma (.csv), semi-colon (.txt), and tab (.txt) separated value files are all acceptable formats.

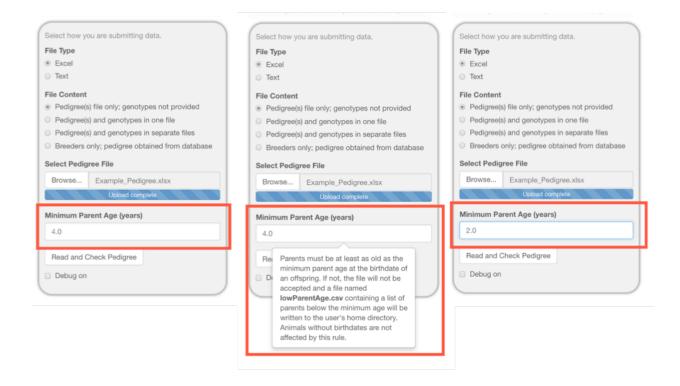
The *Example_Pedigree.xlsx* file we are using is from a CSV file created as shown below and then saved in an Excel format.

makeExamplePedigreeFile()

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

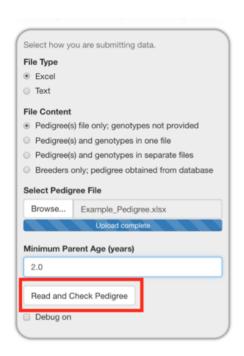


It is important to make sure the minimum parent age is low enough for the animals in your pedigree. For our example pedigree, we are changing it from 4 years to 2 years of age since these macaques may reproduce as early as two years of age. This is shown below in three progressive images with the center image demonstrating how the hovertext provides an explanation of how this value is used.



Reading in 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.

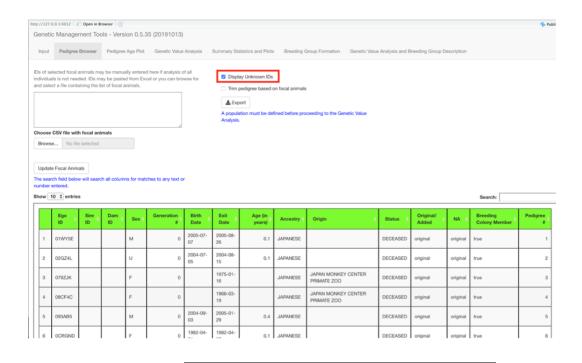


Several error types, shown below, are detected the application.

Error	Definition
failedDatabaseConnection	Database connection failed: configuration or permissions are invalid
missingColumns	Columns that must be within the pedigree file are missing.
invalid Date Rows	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	Fatal Errors.
${\rm changedCols}$	Columns that have been changed to conform to internal naming conventions and what they we

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.



A Note About Obfuscated Data

I have place 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 following with a left alphanumeric string of five places. The form of this ID is set by the obfuscateId function.

3684	I04JZV	AEP5EG	T3QPW5	М	8	2009-07-20	2010-05- 17	0.8
3685	IJT6B6	633UQI	7NE2UT	М	8	2014-04-17		5.2
3686	RLJC25	UH5ZMC	46ZHKN	М	8	2014-03-20	2014-04- 01	0
3687	S63QDN	4JARCS	YOTCYX	М	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	М	8	2013-04-09	2013-04- 30	0.1

Unknown IDs

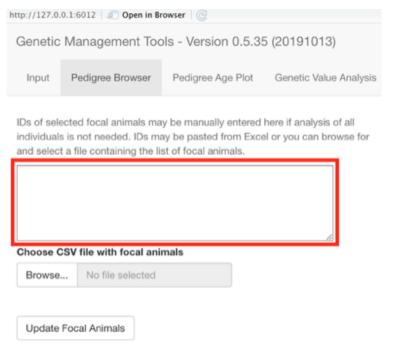
In this example pedigree, when you deselect the **Display Unknown IDs** checkbox. The number of rows reduces from 3,694 to 2,322, because there were 1,372 UNKNOWN animals generated when constructing the pedigree to provide sire and dam placeholders for all animals.

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

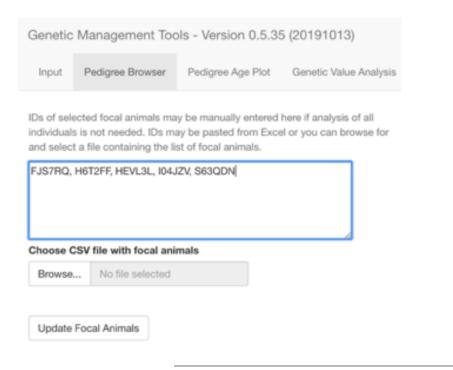
Showing 1 to 100 of 2,322 entries

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*).



You can enter in the animal IDs by typing them into the text box directly as shown below.



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.

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.

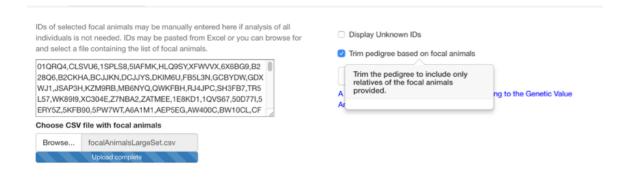
01QRQ4,CLSVU6,1SPLS8,5IAFMK,HLQ9SY,XFWVVX,6X6BG9,B2
28Q6,B2CKHA,BCJJKN,DCJJYS,DKIM6U,FB5L3N,GCBYDW,GDX
WJ1,JSAP3H,KZM9RB,MB6NYQ,QWKFBH,RJ4JPC,SH3FB7,TR5
L57,WK89I9,XC304E,Z7NBA2,ZATMEE,1E8KD1,1QVS67,50D77I,5
ERY5Z,5KFB90,5PW7WT,A6A1M1,AEP5EG,AW400C,BW10CL,CF

Choose CSV file with focal animals

Browse... 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.



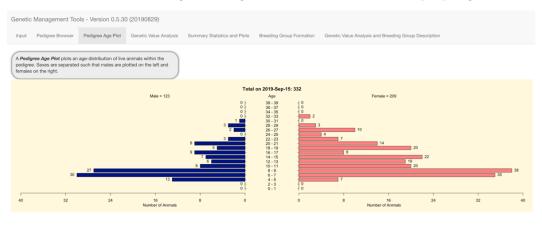
A pedigree trimmed based on focal animals will have only the relatives of those animals remaining. In this instance there are only a total of 85 focal animals and their relatives.

					30	06-23
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

Pedigree Age Pyramid Plot

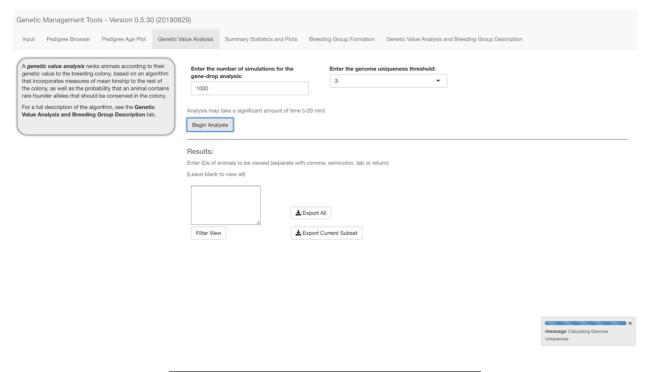
The **Pedigree Age Plot** tab displays a standard pyramid plot for the pedigree as selected in the **Pedigree Browser** tab. This is showing 332 living animals from the entire example pedigree.



Genetic Value Analysis

Selecting the **Genetic Value Analysis** tab and the **Begin Analysis** button will begin the gene dropping process, which you can monitor with the progress meter in the lower right corner of the display.

We are not aware of a systematic study of pedigree structure with this algorithm and have not performed extensive studies with pedigrees of various structures, but 1000 iterations has seemed to provide reproducible results.



As soon as the calculations are completed, a table showing the results of the analysis is displayed in 10 record increments. 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.



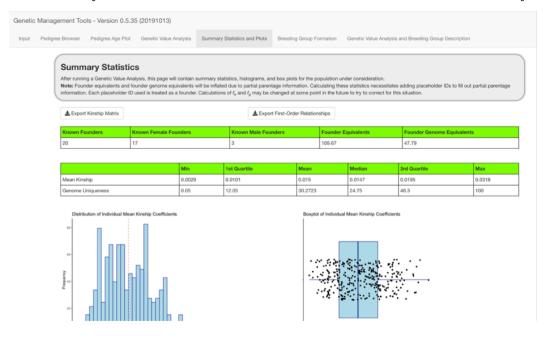
Searching down the table of results in the Value Designation column you can see starting at row 268 the

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.

267	WKY2SZ	F	10	2009-07-		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	М	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	Showing 261 to 270 of 327 entries Previous 1 26 27 28 33 Next												33 Next	

Summary Statistics

The Summary Statistics and Plots tab used results from the Genetic Value Analysis tab.



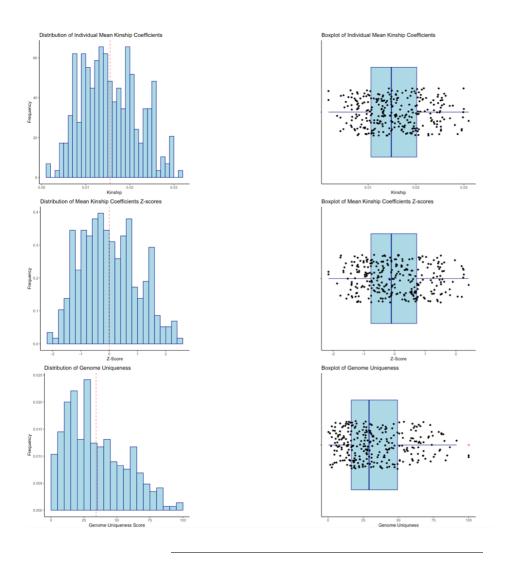
The **Export Kinship Matrix** button creates a CSV file that has a row and column for each individual in the genetic analysis plus a first row and first column each containing the IDs.

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

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

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

The six plots provide histograms and boxplots for the kinship coefficients, the Z-scores of the kinship coefficients, and the genome uniqueness scores.



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 ways based on your source of animals selected under **Choose** one group formation workflow:

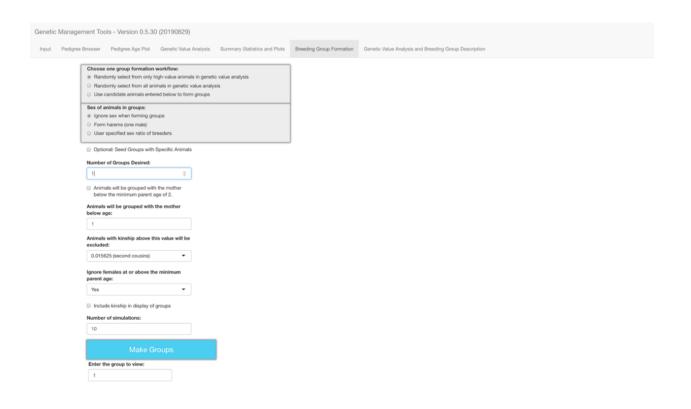
Further you must specify how you want to construct the breeding groups with regard to the groups' sex ratios. The third of the three options ($User\ specified\ sex\ ratio\ of\ breeders$) causes the appearance of the field where you can fill in the sex ratio (F/M) that you want to have in the formed breeding 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.



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.

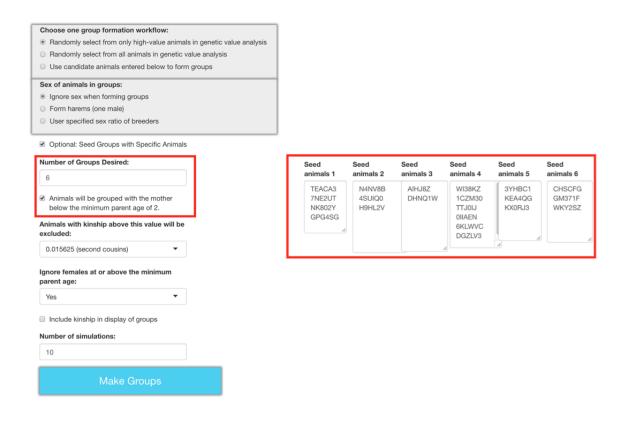
The most common source of animals will be the high-value animals found by the genetic analysis.

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.

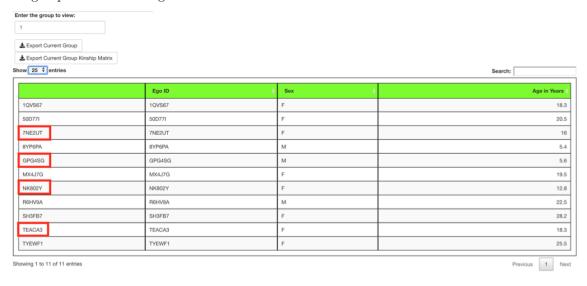


There are often behavioral constraints, such as preexisting social groups, that dictate the need to have some animals maintained together. This need is readily accommodated by pre-seeding groups with those social groups. You may select the **Optional: Seed Groups with Specific Animals** field if you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



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.



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.



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.