

用途

血統情報から近交係数、血縁係数、基礎集団の遺伝的寄与率を R 上で算出し、指定されたグループ (年次等) 毎の平均値等をまとめて出力するプログラムです。オプションで、各個体の近交係数、グループ内ペア毎の血縁係数、遺伝的寄与率の推移なども出力することも可能です。

準備

データセット

血統情報とグループ情報が格納されたデータフレームをあらかじめ準備します。それぞれのデータフレームに格納する情報は以下の通りです。

・血統情報

- 1 列目: 個体番号*
- 2 列目: 父親番号*
- 3 列目: 母親番号*
- 4 列目: 性別 (雄=1、雌=2)
- 5 列目: 基礎世代個体の指定 (基礎=1、非基礎=0)

*番号は整数値とし、英字や記号は含められません。最大 19 桁まで利用可能です。

**始祖世代の両親番号は 0 とする。非始祖個体の両親の番号は既知とする。

***1 列目の個体番号は任意で生年月日順である必要は無いが、個体番号より両親の番号が小さい方が望ましい。

・グループファイル

- 1 列目: 個体番号
- 2 列目: グループ番号 (年次、世代等、必ず整数値)

注: 同一個体が複数のグループ番号を持つ場合、グループ毎に同じ個体番号を入力します。

データフレームの作成方法

ここでは Excel のシートからデータを読み込む方法を例として取り上げます。

必要な環境

このプログラムは次の 2 つの package を利用しています。事前にこれらの package をイン

ストールしておく必要があります。

```
> install.packages("pedigreemm")
> install.packages("dplyr")
```

実行方法

血統データを図1のように入力されているとする。図1のようにデータ全体をコピーし、Rのコンソールで以下のように実行する。

```
> testped <- read.delim("clipboard", header=F)
```

この操作によって、血統データが testped という名前のデータフレームに収録される。同様に、グループデータも図2のように入力されているとし、データ全体をコピーし、Rのコンソールで以下のように実行し、testgroup という名前のデータフレームに収録する。

```
> testgroup <- read.delim("clipboard", header=F)
```

GRC.R とデータファイルがあるディレクトリに移動し、以下のコマンドを実行します。

```
> source("GRC.R")
> GRC(peddata=testped, groupdata=testgroup, name="name",
      GRCmethod=1, RLSmethod=1, conftable=T, inbcoef=T, rlscoef=T, grc=T, Cleandata=T)
```

必須項目

peddata: 血統データのデータフレームを指定

groupdata: グループファイルのデータフレームを指定

任意項目（指定しない場合はデフォルトが選択される）

name: 任意の分析名を指定（デフォルトは分析開始時の時間）

GRCmethod: 平均遺伝的寄与率の計算方法を指定（デフォルトは1）

1: 現集団の雌雄毎に平均遺伝的寄与率を計算し、その2つの平均値を現集団における各基礎個体の平均遺伝的寄与率とする方法

2: 現集団の全個体の平均遺伝的寄与率を計算し、それを各基礎個体の平均遺伝的寄与率とする方法

RLSmethod: 平均血縁係数の計算方法を指定（デフォルトは1）

1: 各グループ内全個体の総当たりの血縁係数の平均値とする方法

2: 各グループ内個体の雌雄間総当たりの血縁係数の平均値とする方法

conftable: 各形質のグループ毎平均などの出力オプション（デフォルトは T:出力）

inbcoef: 各個体の近交係数の出力オプション（デフォルトは F:非出力）

rlscoef: 各グループ内ペアの血縁係数の出力オプション（デフォルトは F:非出力）

grc: 各基礎集団個体の遺伝的変動係数の推移の出力オプション（デフォルトは T:出力）

Cleandata: 重複の削除などの処理後のデータセットの出力オプション（デフォルトは F:非出力）

出力結果

結果はすべて *name* 以下のフォルダーに出力されます。

confable=T の場合

nameconf_data.txt

group: グループ名

aveinb(%): 近交係数 (%) のグループ内平均値

sdinb(%): 近交係数 (%) のグループ内標準偏差

averls(%): グループ内での総当たり*血縁係数 (%) の平均値

sdrls(%): グループ内での総当たり*血縁係数 (%) の標準偏差

GRC: 遺伝的寄与率変動係数

kai2: 遺伝的寄与率変動係数の χ^2 値

p-value: 遺伝的寄与率変動係数の χ^2 値の p 値

Ninb: 近交係数 (%) のグループ内平均値の算出に用いた個体数

Nrls: 血縁係数 (%) のグループ内平均値の算出に用いた個体数

*:RLSmethod のオプションによってかわる。

inbcoef=T の場合

nameInb.txt (各個体の近交係数を出力)

1 列目: 個体番号 (順番は peddata と異なる場合がある)

2 列目: 各個体の近交係数 (%)

注: 血統情報を自動的に整理しているため、血統情報オブジェクトの個体番号の並びと異なる場合があります。

rlscoef=T の場合

namerls_グループ名.txt (雄同士、雌同士のペアを含まない)

1 列目: 雄個体番号

2 列目: 雌個体番号

3 列目: それぞれのペアの血縁係数(%)

nameallrls_グループ名.txt (雄同士、雌同士のペアを含む)

1 列目: 個体 1 番号

2 列目: 個体 2 番号

3 列目: それぞれのペアの血縁係数(%)

grc=T の場合

nameAveGRC1.txt

- 1 列目:各個体の ID
- 2 列目:各個体の性別
- 3 列目以降:各個体のグループ内での平均遺伝的寄与率 (**GRCmethod** で指定した方法に従って計算した値)

nameAveGRC2.txt

- 1 列目:各個体の ID
 - 2 列目:各個体の性別
 - 3 列目以降:各個体のグループ内での平均遺伝的寄与率*
- *雄内、雌内で寄与率の合計が 100%になるように調整。

Cleandata=T の場合

name_Cleanped.txt (修正後の血統情報データフレームを出力)

- 1 列目: 個体番号*
- 2 列目: 父親番号*
- 3 列目: 母親番号*
- 4 列目: 性別 (雄=1、雌=2、その他=0)
- 5 列目: 基礎世代個体の指定 (基礎=1、非基礎=0)

name_Cleangroup.txt (修正後のグループ情報データフレームを出力)

- 1 列目: 個体番号
- 2 列目: グループ番号 (年次、世代等、必ず数字)

注意メッセージ、エラーメッセージについて

血統ファイル、グループファイルに問題がある場合、注意メッセージ (**Warning**) とエラーメッセージ (**Error!!!**) が、*name* フォルダー内に出力されます。

注意メッセージ (**Warning**)

注意メッセージが出た場合でも、分析は継続されます。

****Warning****Some animals were duplicate and automatically removed

血統情報データフレーム内で ID、父親番号、母親番号のすべてが重複している個体が存在する場合には表示されます。重複データは自動的に削除され、*name_RemovedID.txt* に出力されます。

****Warning****Some animals have unknown parent

血統情報データフレーム内で父親番号または母親番号の一方が 0 の個体が存在する場合には表示されます。その個体は *name_UnknownPID.txt* に出力されます。

****Warning****Ordered a pedigree so parents come before offspring

血統情報データフレーム内で各個体の父親番号または母親番号の個体が、その個体より先に出現していない場合に表示されます。先に出現するように血統データを自動的に並び替え、分析を続けます。ID が昇順でない場合にも同様に表示されます。また、プログラムの性質上、取り込んだデータ中の ID の並び順と出力する近交係数等の ID の並び順が異なる場合にも表示されることがあります。

エラーメッセージ (**Error!!!**)

エラーメッセージが出た場合、分析は中断されます。出力ファイルを参考に、血統データ、グループデータを修正し、エラーメッセージが出なくなるまで再実行してください。

****Error!!!****Program terminated unsuccessfully!

Some parents do not have pedigree data.

Check the *test1_NopedinfoID.txt*

父親番号または母親番号が個体番号にない場合、表示されます。分析は中断され、その個体は *name_NopedinfoID.txt* に出力されます。

****Error!!!****Program terminated unsuccessfully!

Some animal having other sire and/or dam.

Check the *name_DuplicatedID.txt*

ID が重複していて、父親番号または母親番号が異なる個体が存在する場合には表示されません。分析は中断され、その個体は *name_DuplicatedID.txt* に出力されます。

****Error!!!****Program terminated unsuccessfully!

Sex of some ancients were unknown.

Check the *name_Improperancients.txt!!*

基礎集団中に性が不明な個体が存在する場合には表示されます。分析は中断され、その個体は *name_Improperancients.txt* に出力されます。

****Error!!!****Program terminated unsuccessfully!

Some animals are both sire and dam.

Check the *test8_bothID.txt*

父親番号と母親番号に同じ番号が存在する場合には表示されます。分析は中断され、その個

体は *name_bothID.txt* に出力されます。

****Error!!!****Program terminated unsuccessfully!

Some animals do not have pedigree information.

Check *name_nopedID.txt* グループデータに存在する個体が、血統データに存在しない場合に表示されます。分析は中断され、その個体は *name_nopedID.txt* に出力されます。

備考

本プログラム中の近交係数および血縁係数の算出には、R のパッケージである”pedigreemm”を利用しています。パッケージの詳細については以下の URL、論文を参照してください。

<https://cran.r-project.org/web/packages/pedigreemm/index.html>

A. I. Vazquez, D. M. Bates, G. J. M. Rosa, D. Gianola, K. A. Weigel, Technical note: An R package for fitting generalized linear mixed models in animal breeding, *Journal of Animal Science*, Volume 88, Issue 2, February 2013, Pages 497–504, <https://doi.org/10.2527/jas.2009-1952>

更新履歴

2020/02/06 v1.0 公開

2020/06/02 v1.1 公開、*nameAveGRC2.txt* の出力を修正、外部導入個体がいる場合に対応

	A	B	C	D	E	F	G	H	I	J
1	1	0	0	1	1					
2	2	0	0	2	1					
3	3	0	0	2	1					
4	4	0	0	1	1					
5	5	0	0	2	1					
6	6	0	0	2	1					
7	11	1	2	1	0					
8	12	1	2	2	0					
9	13	1	3	2	0					
10	14	4	5	1	0					
11	15	4	5	2	0					
12	16	4	6	2	0					
13	21	11	13	1	0					
14	22	11	15	2	0					
15	23	11	15	2	0					
16	24	14	12	1	0					
17	25	14	12	1	0					

図1 血統データの画面

	A	B	C	D	E	F	G	H	I	J
1	1	0								
2	2	0								
3	3	0								
4	4	0								
5	5	0								
6	6	0								
7	11	1								
8	12	1								
9	13	1								
10	14	1								
11	15	1								
12	16	1								
13	21	2								
14	22	2								
15	23	2								
16	24	2								
17	25	2								

図2 グループデータの画面