Fixing 3K (3102) weird PED vs AS Kinships

E. Lavanchy

10/26/2023

```r
library(hierfstat)
library(corrplot)
```

```
## corrplot 0.92 loaded
```

```r
birds = read.csv("~/PhD/Barn_owl/ID_in3Kowls/data/BarnOwls_Legacy_2023.10.23/BarnOwls_Legacy_2023102311

#Read PEDIGREE kinship matrix
BetaPED = as.matrix(read.table("~/PhD/Barn_owl/3Kowls/data/chechinkgGLIMPSu/A_pedigree_kinship_matrix.t
namesBetaPED = read.table("~/PhD/Barn_owl/3Kowls/data/chechinkgGLIMPSu/A_pedigree_kinship_matrix.txt", 
namesBetaPED = as.vector(namesBetaPED[1,], mode = "character")
colnames(BetaPED) = namesBetaPED
rownames(BetaPED) = namesBetaPED

grmAS = readRDS("~/PhD/Barn_owl/ID_in3Kowls/data/checkingINDVsGRMvalues/All3102_AUTOSAUMES_RP502SNPs.RD

kas = grm2kinship(grmAS)

#read pedigree itself
ped = read.table("~/PhD/Barn_owl/ID_in3Kowls/data/pedigree_290623.tab", h = T)
```

Extracting list of individuals with weird kinships AS vs PED

First we need to sub-sample the pedigree matrix so that it only contains the 3102 sequenced individuals. Then we nee to make sure that individuals orders are matching between both matrices (genomic and pedigree kinship) !

```r
#subsample sequenced individuals
BetaPEDsub = BetaPED[rownames(BetaPED) %in% rownames(kas), colnames(BetaPED) %in% colnames(kas)]

#change betaPED order so that individuals' order matches AS kinship matrix order
orderCOL = match(colnames(kas), colnames(BetaPEDsub))
BetaPEDsuborder = BetaPEDsub[orderCOL,orderCOL]

#sanity check
unique(colnames(kas) == colnames(BetaPEDsuborder))
```

```
## [1] TRUE
```

```r
unique(rownames(kas) == rownames(BetaPEDsuborder))
```

```
## [1] TRUE
```

Then we extract the differences between pedigree and genomic matrices

```
#subtract genetic matrix from pedigree matrix
diff_kin = BetaPEDsuborder - kas

#mean difference
mean(diff_kin) #0.001, very low
```

[1] 0.001141341

We can look (and hopefully solve) differences above 0.2 and below -0.2 ! So we'll extract the pairs of individuals which have high differences !

```
#How many problems do we have here ?
length(diff_kin[(diff_kin > .2) | (diff_kin < -.2)]) #1734 (out of 9,622,404)
```

[1] 1734

```
#but each pair actually appears twice (because the matrix is symmetrical)
#so we have 867 weird links !
```

So we have 1734 links which have an absolute difference above 0.2. Since the matrix is symmetrical, all the pairs are represented twice so in the end, we have 867 weird links.

When the difference is below -0.2 it means that the genomic kinship is higher than the pedigree kinship. This probably indicates that our pedigree is missing some links. We have 1610 cases for which genomic kinship is higher than pedigree kinship. We won't fix these problems (at least here) since they are not really problems.

On the contrary, when the difference is above 0.2 it means that the pedigree kinship is higher than the genomic kinship. These are more problematic because it either means that our pedigree is wrong or that we did not sequenced the individual·s we though we sequenced ! We'll dig into these pairs.

Finally, we also need to be wary of very high genomic kinship values ($< .4$) between pairs of individuals which would mean that: individuals are related and inbred (which is good, we want to keep these individuals) OR if close to 0.5 that individuals are twins or the same individual (we mislabeled one of the tubes, which is bad). If the latter, we cannot really solve the problem because we need to be absolutely sure than the phenotypes are linked to the correct individual, so we would usually remove both individuals.

AS kinship higher than 0.4

We'll first dig into the pairs for which the genomic kinship is higher than 0.4!

```
#For this we need to get rid of the diagonal values because they all (should) be above 0.4
kasHIGH = kas

#First set diagonal to NA (because all self-kinship will be above .4)
diag(kasHIGH) = NA
#Then also set upper triangle of the matric to NA (because matrix is symmetrical)
kasHIGH[upper.tri(kasHIGH)] = NA

WeirdgenKin = which(kasHIGH > .4, arr.ind = T)

#Add the pair guy
WeirdgenKinnames = as.data.frame(cbind(rownames(kas)[WeirdgenKin[,1]],
                                       rownames(kas)[WeirdgenKin[,2]],WeirdgenKin))

rownames(WeirdgenKinnames) = NULL
colnames(WeirdgenKinnames)[1:2] = c("RingId1","RindId2")
WeirdgenKinnames
```

```
##    RingId1 RindId2  row  col
## 1 M026477 M026451  928  169
## 2 M031772 M031770  232  230
## 3 M031773 M031770  233  230
## 4 M043615 M038321 2788  320
## 5 M047568 M040628  344  327
## 6 M047568 M043044  344  331
## 7 M043616 M043620 2789  335
## 8 M037798 M032181 1938 1580
## 9 M038425 M038424 2301 2300
```

We only have 9 pairs with a weirdly high kinship! We'll investigate these pairs one by one!

pair 1

The first pair are individuals M026477 and M026451 with a kinship of 0.485.

```r
#what is the genomic kinship
kas[WeirdgenKin[1,][1],WeirdgenKin[1,][2]]
```

```
## [1] 0.4853092
```

```r
#pedigree kinship
BetaPEDsuborder[WeirdgenKin[1,][1],WeirdgenKin[1,][2]]
```
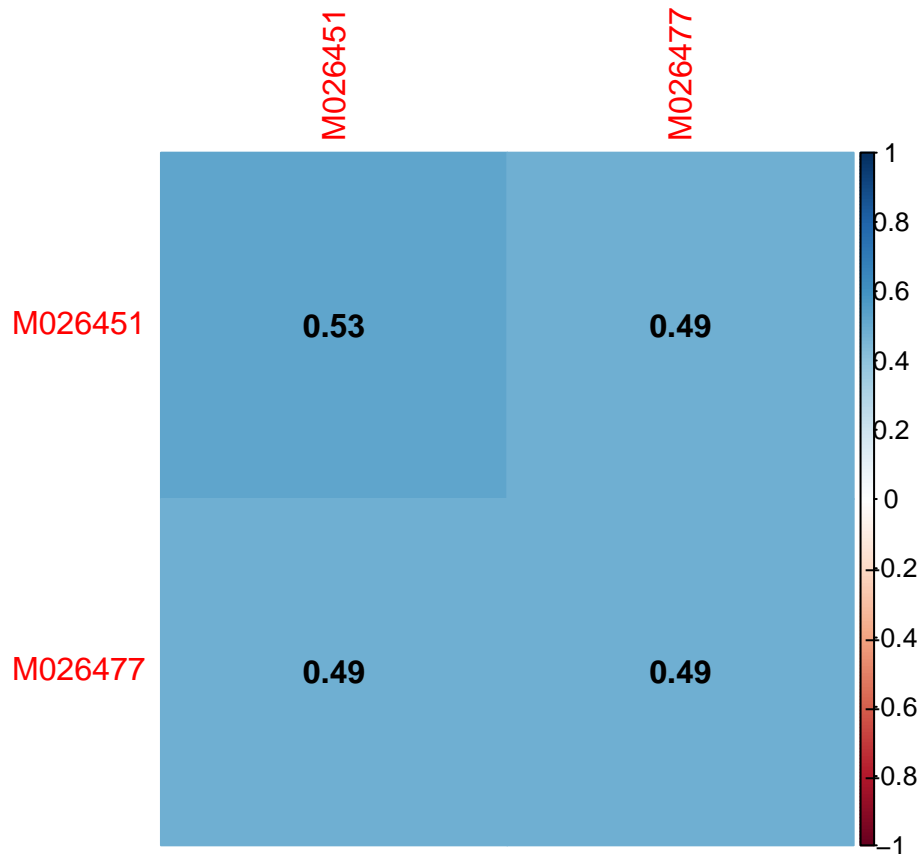
```
## [1] 0.25
```

```r
#subset family matrix
kASweirdFam = kas[rownames(kas) %in% c(WeirdgenKinnames[1,][1],WeirdgenKinnames[1,][2]),
                  colnames(kas) %in% c(WeirdgenKinnames[1,][1],WeirdgenKinnames[1,][2])]

#If we look at this family plot
corrplot(kASweirdFam, method = "color", addCoef.col = "black")
```

```
#parents of both individuals
ped[ped$animal == "M026451",]
```

```
##        animal    sire      dam sex
## 6332 M026451 M022131 M014481   1
```

```
ped[ped$animal == "M026477",]
```

```
##        animal    sire      dam sex
## 6356 M026477 M022131 M014481   1
```

The kinship is close to 0.5 indicating that we have either mislabeled individuals and it's the same individuals we sequenced twice, either that they are twins. The pedigree kinship is 0.25 and they have the same parents. So they might be twins but since we can never be sure, we'll remove both of these individuals from our analyses.

pairs 2 & 3

We'll study the second and third pairs together because it's one individual (M031770) which has high kinship with two of its siblings: M031772 and M031773. Kinship values are 0.402 and 0.443.

```
#what is the genomic kinships
kas[WeirdgenKin[2,][1],WeirdgenKin[2,][2]]
```

```
## [1] 0.4017183
```

```
kas[WeirdgenKin[3,][1],WeirdgenKin[3,][2]]
```

```
## [1] 0.443119
```

```
#pedigree kinships
BetaPEDsuborder[WeirdgenKin[2,][1],WeirdgenKin[2,][2]]
```
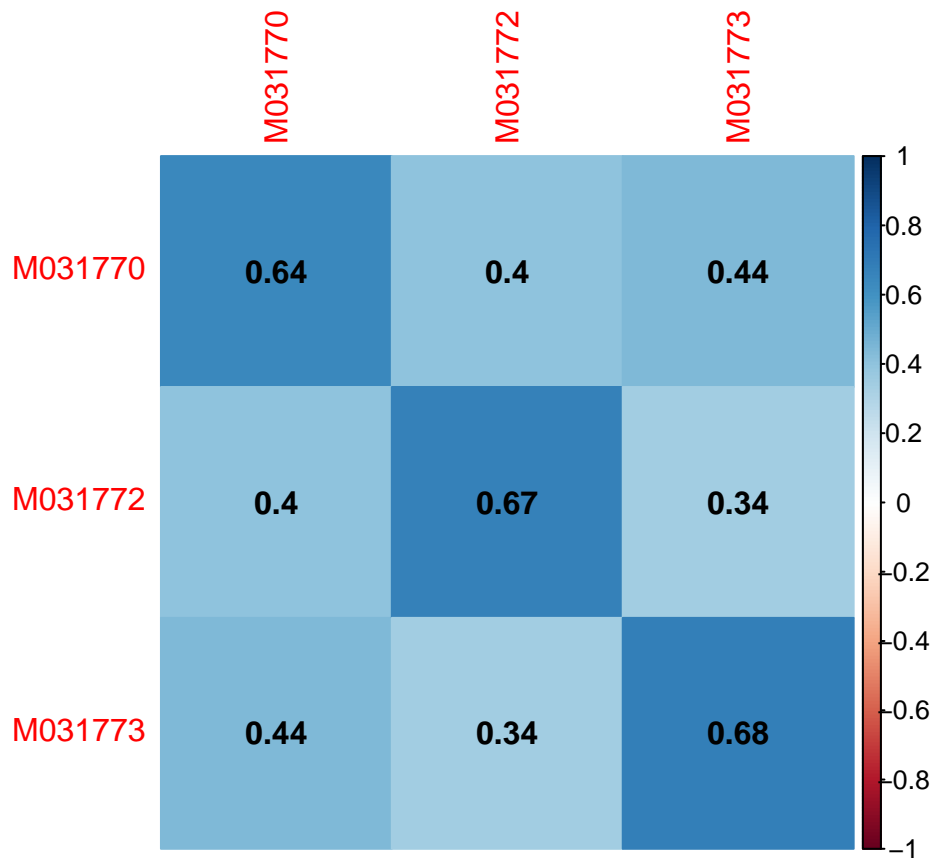
```
## [1] 0.375
```

```
BetaPEDsuborder[WeirdgenKin[3,][1],WeirdgenKin[3,][2]]
```

```
## [1] 0.375
```

```
#subset family matrix
kASweirdFam = kas[rownames(kas) %in% c(WeirdgenKinnames[2,][1], WeirdgenKinnames[2,][2],
                                       WeirdgenKinnames[3,][1]),
                  colnames(kas) %in% c(WeirdgenKinnames[2,][1], WeirdgenKinnames[2,][2],
                                       WeirdgenKinnames[3,][1])]
```

```
#If we look at this family plot
corrplot(kASweirdFam, method = "color", addCoef.col = "black")
```



```
#parents of the three individuals
ped[ped$animal == "M031770",]
```

```
##        animal     sire     dam sex
## 6859 M031770 M032400 M032314    1
```

```
ped[ped$animal == "M031772",]
```

```
##        animal     sire     dam sex
## 6861 M031772 M032400 M032314    1
```

```
ped[ped$animal == "M031773",]
```

```
##      animal    sire     dam sex
## 6862 M031773 M032400 M032314   2
```

The kinship is below (0.5) so it's not twins (or the same individuals). These individuals are siblings (they have the same parents) and the pedigree kinship is above 0.25 and close to the observed genomic value (0.375). This simply indicates that these individuals are siblings and are inbred so we'll keep the three of them !

pair 4

The forth pair are individuals M043615 and M038321 with a kinship of 0.401.

```
#what is the genomic kinship
kas[WeirdgenKin[4,][1],WeirdgenKin[4,][2]]
```

```
## [1] 0.4010053
```

```
#pedigree kinship
BetaPEDsuborder[WeirdgenKin[4,][1],WeirdgenKin[4,][2]]
```

```
## [1] 0.375
```

```
#subset family matrix
kASweirdFam = kas[rownames(kas) %in% c(WeirdgenKinnames[4,][1], WeirdgenKinnames[4,][2]),
                  colnames(kas) %in% c(WeirdgenKinnames[4,][1], WeirdgenKinnames[4,][2])]

#If we look at this family plot
corrplot(kASweirdFam, method = "color", addCoef.col = "black")
```
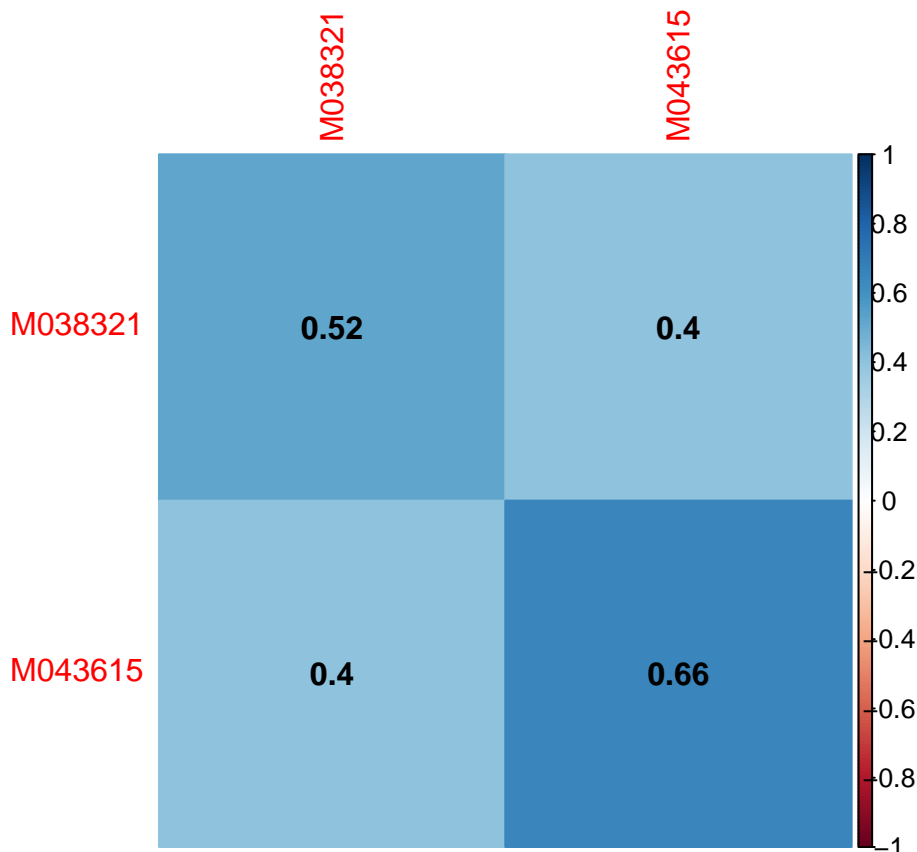
```
#parents of both individuals
ped[ped$animal == "M043615",]
```

```
##        animal    sire     dam sex
## 9037 M043615 M038321 M038012   1
```

```
ped[ped$animal == "M038321",]
```

```
##        animal    sire     dam sex
## 8006 M038321 M042107 M038012   1
```

Just like before, the genomic kinship is close to 0.4 and the matches the pedigree kinship quite well (0.375). These individuals are mother - offspring but the mother here mated with her dad to create this offspring. So the kinship we observe is OK, we'll keep both individuals.

pair 5

The fifth pair are individuals M047568 and M040628 with a kinship of 0.413.

```
#what is the genomic kinship
kas[WeirdgenKin[5,][1],WeirdgenKin[5,][2]]
```

```
## [1] 0.4128979
```

```
#pedigree kinship
BetaPEDsuborder[WeirdgenKin[5,][1],WeirdgenKin[5,][2]]
```

```
## [1] 0.375
```

```
#subset family matrix
kASweirdFam = kas[rownames(kas) %in% c(WeirdgenKinnames[5,][1], WeirdgenKinnames[5,][2]),
                  colnames(kas) %in% c(WeirdgenKinnames[5,][1], WeirdgenKinnames[5,][2])]

#If we look at this family plot
corrplot(kASweirdFam, method = "color", addCoef.col = "black")
```
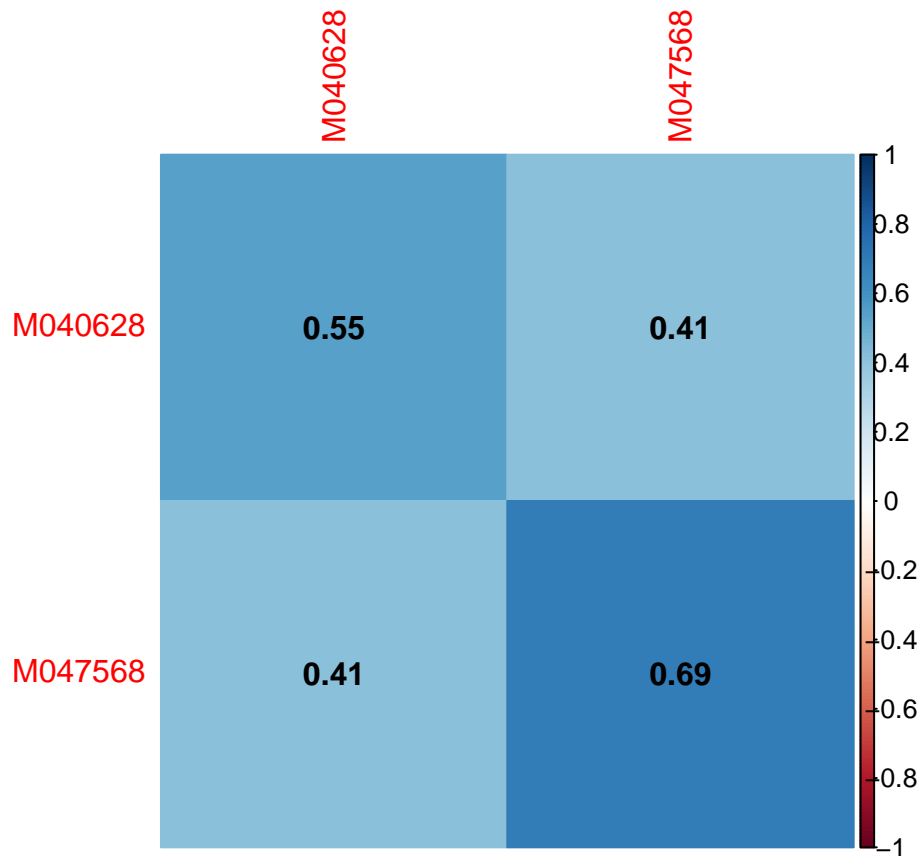
```
#parents of both individuals
ped[ped$animal == "M047568",]
```

```
##        animal    sire     dam sex
## 9939 M047568 M040628 M040627   1
```

```
ped[ped$animal == "M040628",]
```

```
##        animal    sire     dam sex
## 9254 M040628 M032479 M043819   1
```

```
#parents of the dad of M047568
ped[ped$animal == "M040627",]
```

```
##        animal    sire     dam sex
## 9253 M040627 M032479 M043819   2
```

Genomic kinship is close to 0.4 and the matches the pedigree kinship quite well (0.375). These individuals are father - offspring and the father mated with his sibling (they have the same parents) to create this offspring. So the kinship we observe is OK, we'll keep both individuals.

pair 6

The sixth pair are individuals M047568 (same individual as offspring in pair 5) and M043044 with a kinship of 0.406.

```
#what is the genomic kinship
kas[WeirdgenKin[6,][1],WeirdgenKin[6,][2]]
```
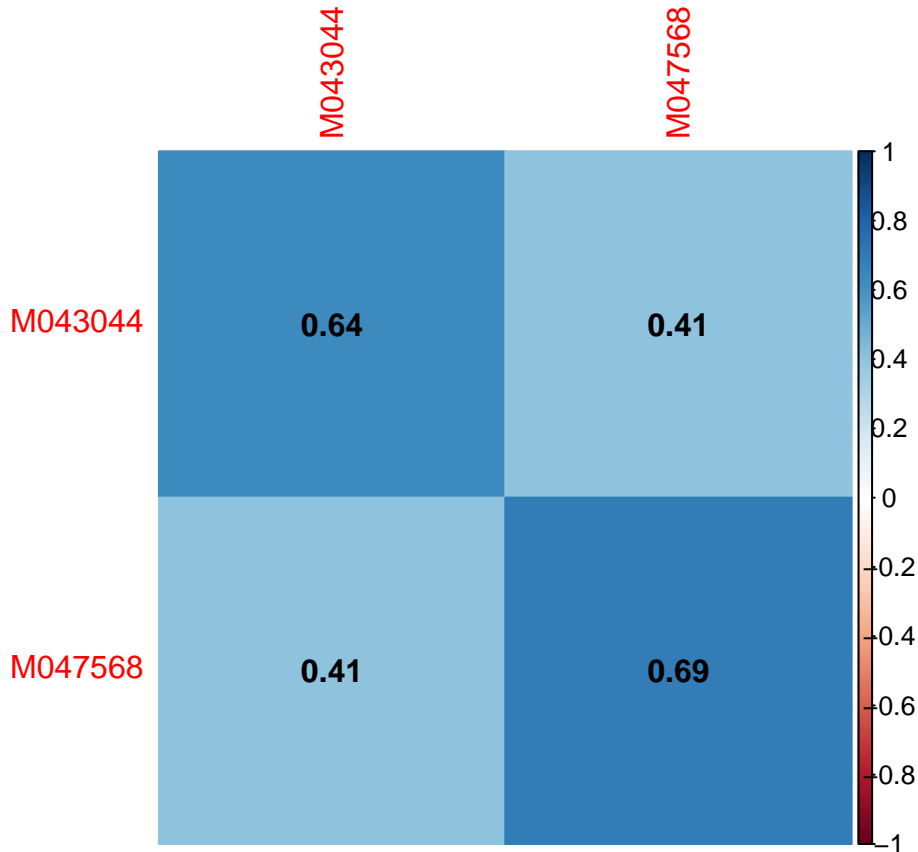
```
## [1] 0.4059577
```

```
#pedigree kinship
BetaPEDsuborder[WeirdgenKin[6,][1],WeirdgenKin[6,][2]]
```

```
## [1] 0.375
```

```
#subset family matrix
kASweirdFam = kas[rownames(kas) %in% c(WeirdgenKinnames[6,][1], WeirdgenKinnames[6,][2]),
                  colnames(kas) %in% c(WeirdgenKinnames[6,][1], WeirdgenKinnames[6,][2])]

#If we look at this family plot
corrplot(kASweirdFam, method = "color", addCoef.col = "black")
```



```
#parents of both individuals
ped[ped$animal == "M047568",]
```

```
##       animal    sire     dam sex
## 9939 M047568 M040628 M040627   1
```

```
ped[ped$animal == "M043044",]
```

```
##       animal    sire     dam sex
## 9418 M043044 M040628 M040627   2
```

Again, the genomic kinship is close to 0.4 and matches the pedigree kinship quite well (0.375). These individuals are siblings and are inbred so they both some from a mating between siblings. We can check the kinship between the second offspring (M047568) and the father we identified before (M040628): 0.397. It is just below 0.4 so we did not detect because of the 0.4 threshold we set, but it is close to 0.4, which makes sense ! We'll keep both individuals !

pair 7

The seventh pair are individuals M043616 and M043620 with a kinship of 0.514.

```
#what is the genomic kinship
kas[WeirdgenKin[7,][1],WeirdgenKin[7,][2]]
```

```
## [1] 0.51378
```

```
#pedigree kinship
BetaPEDsuborder[WeirdgenKin[7,][1],WeirdgenKin[7,][2]]
```

```
## [1] 0
```

```
#subset family matrix
kASweirdFam = kas[rownames(kas) %in% c(WeirdgenKinnames[7,][1], WeirdgenKinnames[7,][2]),
                  colnames(kas) %in% c(WeirdgenKinnames[7,][1], WeirdgenKinnames[7,][2])]

#If we look at this family plot
corrplot(kASweirdFam, method = "color", addCoef.col = "black")
```
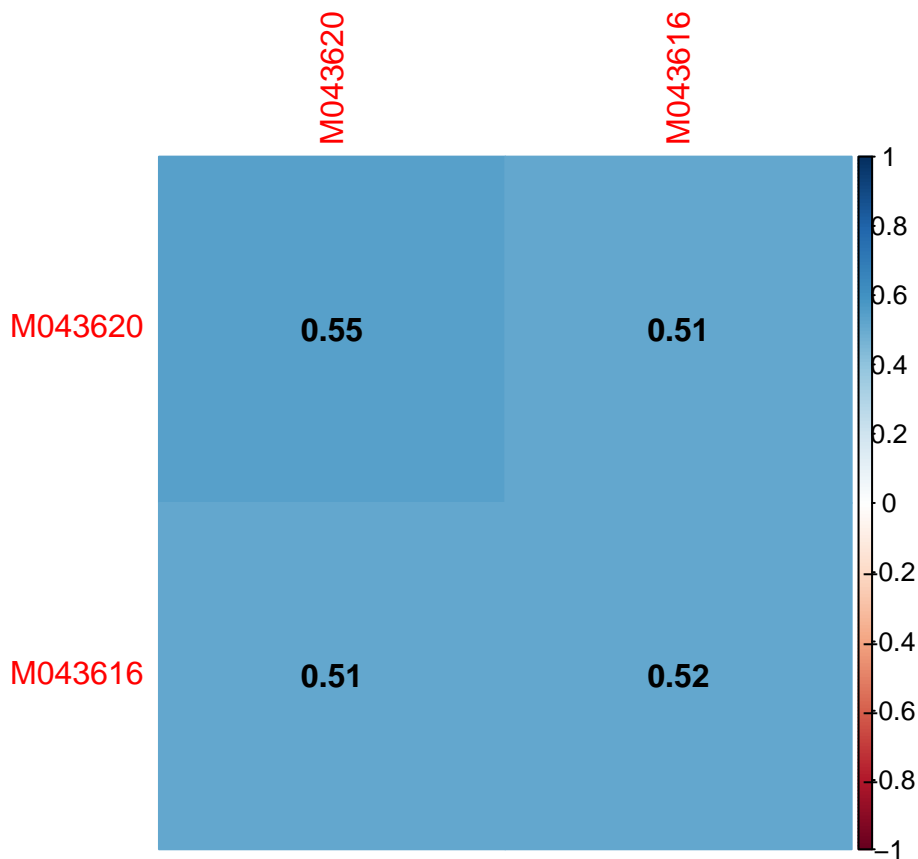


```
#parents of both individuals
ped[ped$animal == "M043616",]
```

```
##        animal    sire     dam sex
## 9038 M043616 M036843 M027838   1
```

```
ped[ped$animal == "M043620",]
```

```
##        animal sire  dam sex
## 9039 M043620 <NA> <NA>   1
```

Here we have a problem. The genomic kinship indicates that these individuals are twins / the same but the pedigree indicates that they should be unrelated, so they are certainly not twins: they are probably the same individual which has been sequenced twice. Now the problem is to know which individual it is (between M043616 & M043620)! Luckily, we know that M043616 has siblings we also sequenced (with which M043620 should be unrelated).

```
#parents of M043616
ped[ped$animal == "M043616",]
```

```
##        animal    sire      dam sex
## 9038 M043616 M036843 M027838   1
```

```
#Look for M043616 siblings
ped$animal[(ped$sire == "M036843") & (ped$dam == "M027838") &
            !(is.na(ped$sire)) & !(is.na(ped$dam))]
```

```
## [1] "M043612" "M043614" "M043616"
```

Did we sequence them ?

```
#Did we sequence them ?
("M043612" %in% colnames(kas))
```

```
## [1] FALSE
```

```
("M043614" %in% colnames(kas))
```

```
## [1] FALSE
```

No we did not sequenced them. However, from the database we know that this clutch (clutch n° 2014) initially contained two more individuals, which were later transferred to the clutch where M043620 comes from (clutch n° 2011). These individuals are individuals M043618 & M043619. We also know that the clutch in which individual M043620 was born (clutch n° 2011) initially contained three other individuals: M043613, M043615 and M043617. Then M043618 and M043619 were raised in clutch 2011 (cross-fostering) and individuals M043613 and M043615 were raised in clutch 2014 (also cross-fostering).

Among these "new" siblings, did we sequence any of them ?

```
#Did we sequence them ?
("M043613" %in% colnames(kas))
```

```
## [1] FALSE
```

```
("M043615" %in% colnames(kas))
```

```
## [1] TRUE
```

```
("M043617" %in% colnames(kas))
```

```
## [1] TRUE
```

```
("M043618" %in% colnames(kas))
```

```
## [1] TRUE
```

```
("M043619" %in% colnames(kas))
```

```
## [1] TRUE
```

We sequenced most of them: M043615, M043617, M043618 and M043619. So now we'll check with which individuals M043616/M043620 are related so we can infer which clutch it came from and finally which individual it is between M043616 or M043620.

```r
#M043615
kas[rownames(kas) == "M043615", colnames(kas) == "M043616"] # not related
```

```
## [1] 0.0136108
```

```r
kas[rownames(kas) == "M043615", colnames(kas) == "M043617"] # not related
```

```
## [1] 0.02542192
```

```r
kas[rownames(kas) == "M043615", colnames(kas) == "M043618"] # not related
```

```
## [1] 0.02055356
```

```r
kas[rownames(kas) == "M043615", colnames(kas) == "M043619"] # not related
```

```
## [1] 0.02223019
```

```r
kas[rownames(kas) == "M043615", colnames(kas) == "M043620"] # not related
```

```
## [1] 0.02041284
```

```r
#M043615 NOT related to anyone, weird

#M043616 (= M043620)
kas[rownames(kas) == "M043616", colnames(kas) == "M043615"] # not related
```

```
## [1] 0.0136108
```

```r
kas[rownames(kas) == "M043616", colnames(kas) == "M043617"] # related
```

```
## [1] 0.2496824
```

```r
kas[rownames(kas) == "M043616", colnames(kas) == "M043618"] # related
```

```
## [1] 0.2325267
```

```r
kas[rownames(kas) == "M043616", colnames(kas) == "M043619"] # related
```

```
## [1] 0.2348672
```

```r
kas[rownames(kas) == "M043616", colnames(kas) == "M043620"] # same individual
```

```
## [1] 0.51378
```

```r
#M043616 related to everyone except 15

#M043617
kas[rownames(kas) == "M043617", colnames(kas) == "M043615"] # not related
```

```
## [1] 0.02542192
```

```r
kas[rownames(kas) == "M043617", colnames(kas) == "M043616"] # related
```

```
## [1] 0.2496824
```

```r
kas[rownames(kas) == "M043617", colnames(kas) == "M043618"] # related
```

```
## [1] 0.2092422
```

```r
kas[rownames(kas) == "M043617", colnames(kas) == "M043619"] # related
```

```
## [1] 0.2313336
```

```r
kas[rownames(kas) == "M043617", colnames(kas) == "M043620"] # related
```

```
## [1] 0.256188
#M043617 related to everyone except 15

#M043618
kas[rownames(kas) == "M043618", colnames(kas) == "M043615"] # not related

## [1] 0.02055356
kas[rownames(kas) == "M043618", colnames(kas) == "M043616"] # related

## [1] 0.2325267
kas[rownames(kas) == "M043618", colnames(kas) == "M043617"] # related

## [1] 0.2092422
kas[rownames(kas) == "M043618", colnames(kas) == "M043619"] # related

## [1] 0.2064427
kas[rownames(kas) == "M043618", colnames(kas) == "M043620"] # related

## [1] 0.2389193
#M043618 related to everyone except 15


#M043619
kas[rownames(kas) == "M043619", colnames(kas) == "M043615"] # not related

## [1] 0.02223019
kas[rownames(kas) == "M043619", colnames(kas) == "M043616"] # related

## [1] 0.2348672
kas[rownames(kas) == "M043619", colnames(kas) == "M043617"] # related

## [1] 0.2313336
kas[rownames(kas) == "M043619", colnames(kas) == "M043618"] # related

## [1] 0.2064427
kas[rownames(kas) == "M043619", colnames(kas) == "M043620"] # related

## [1] 0.2417663
#M043619 related to everyone except 15

#M043620 (= M043616)
kas[rownames(kas) == "M043620", colnames(kas) == "M043615"] # not related

## [1] 0.02041284
kas[rownames(kas) == "M043620", colnames(kas) == "M043616"] # same individual

## [1] 0.51378
kas[rownames(kas) == "M043620", colnames(kas) == "M043617"] # related

## [1] 0.256188
```

```r
kas[rownames(kas) == "M043620", colnames(kas) == "M043618"] # related
```

```
## [1] 0.2389193
```

```r
kas[rownames(kas) == "M043620", colnames(kas) == "M043619"] # related
```

```
## [1] 0.2417663
#M043620 related to everyone except 15
```

Okay so we have a weird pattern here: M043617 is clustering with the wrong clutch ... It should be related to M043615 but it is not. We also sequenced the parents of both clutches: M038321 and M038012 are the parents of clutch 2011 and M027838 and M036843 are the parents of clutch 2014. Let's check the kinship within both clutches:
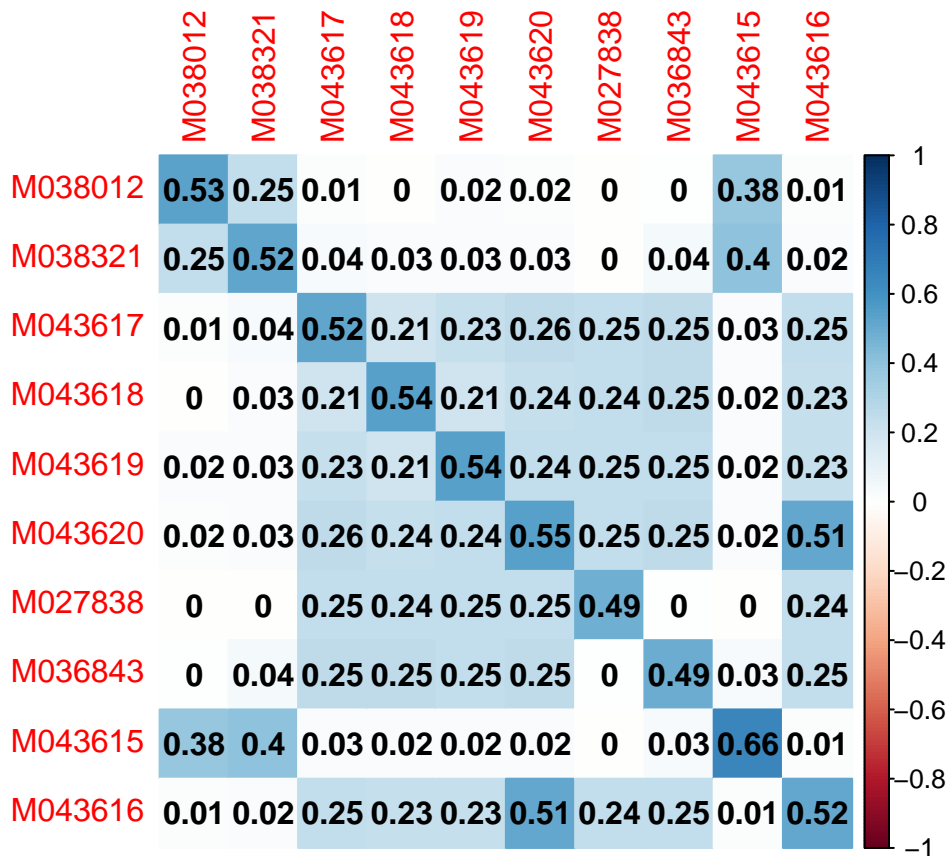
```r
#We also sequenced the parents: "M038321","M038012","M027838","M036843"

#subset family matrix
kASweirdFam = kas[rownames(kas) %in% c("M038321","M038012","M043615","M043617","M027838",
                                       "M036843","M043616","M043618","M043619","M043620"),
                  colnames(kas) %in% c("M038321","M038012","M043615","M043617","M027838",
                                       "M036843","M043616","M043618","M043619","M043620")]

#If we look at this family plot
corrplot(kASweirdFam, method = "color", addCoef.col = "black")
```

	M038012	M038321	M043617	M043618	M043619	M043620	M027838	M036843	M043615	M043616
M038012	0.53	0.25	0.01	0	0.02	0.02	0	0	0.38	0.01
M038321	0.25	0.52	0.04	0.03	0.03	0.03	0	0.04	0.4	0.02
M043617	0.01	0.04	0.52	0.21	0.23	0.26	0.25	0.25	0.03	0.25
M043618	0	0.03	0.21	0.54	0.21	0.24	0.24	0.25	0.02	0.23
M043619	0.02	0.03	0.23	0.21	0.54	0.24	0.25	0.25	0.02	0.23
M043620	0.02	0.03	0.26	0.24	0.24	0.55	0.25	0.25	0.02	0.51
M027838	0	0	0.25	0.24	0.25	0.25	0.49	0	0	0.24
M036843	0	0.04	0.25	0.25	0.25	0.25	0	0.49	0.03	0.25
M043615	0.38	0.4	0.03	0.02	0.02	0.02	0	0.03	0.66	0.01
M043616	0.01	0.02	0.25	0.23	0.23	0.51	0.24	0.25	0.01	0.52

Okay so first, we notice that both parents from clutch 2011: M038321 and M038012 are related (genomic kinship around 0.25). Let's see what the pedigree says about this:

```
ped[ped$animal == "M038321",]
```

```
##      animal    sire      dam sex
## 8006 M038321 M042107 M038012   1
```

```
ped[ped$animal == "M038012",]
```

```
##      animal    sire      dam sex
## 8672 M038012 M032251 M032393   2
```

They are mother and offspring. So the kinship value makes sense it also makes sense that their offspring: M043615 has a high kinship (higher than 0.25) with both of them and a very high self-kinship. So we think we can be confident about these samples being correct: we'll keep the three of them. However, M043617 does not correlate with this family (we saw that with M043615 before but it does not correlate as it should with the supposed parents either) but with clutch 2014 family. We cannot know which of the clutch 2014 kids M043617 really is so we'll just remove this sample.

We also see with the correlation plot that both M043616 and M043620 correlate with the 2014 clutch family suggesting that they both are M043616 (because both their kinship match the expected kinship of M043616). So we'll also remove sample M043620 but we'll keep M043616.

pair 8

The eight pair are individuals M037798 and M032181 with a kinship of 0.403.

```
#what is the genomic kinship
kas[WeirdgenKin[8,][1],WeirdgenKin[8,][2]]
```

```
## [1] 0.4032219
```

```
#pedigree kinship
BetaPEDsuborder[WeirdgenKin[8,][1],WeirdgenKin[8,][2]]
```

```
## [1] 0.25
```

```
#subset family matrix
kASweirdFam = kas[rownames(kas) %in% c(WeirdgenKinnames[8,][1], WeirdgenKinnames[8,][2]),
                  colnames(kas) %in% c(WeirdgenKinnames[8,][1], WeirdgenKinnames[8,][2])]

#If we look at this family plot
corrplot(kASweirdFam, method = "color", addCoef.col = "black")
```
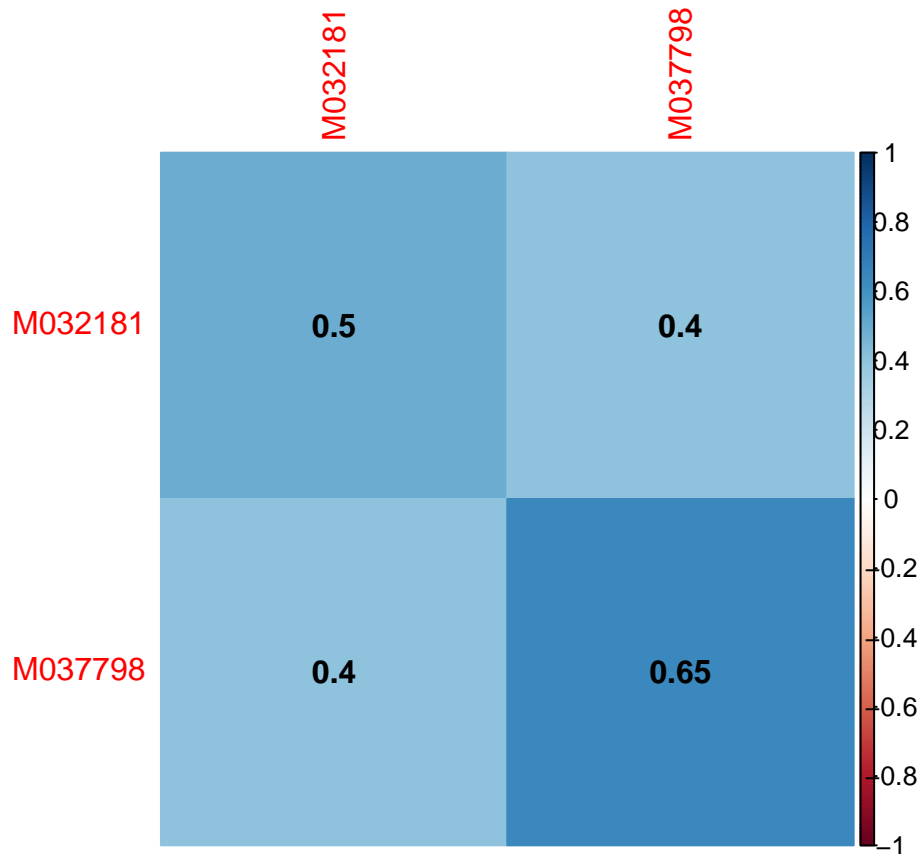
```r
#parents of both individuals
ped[ped$animal == "M037798",]
```

```
##          animal    sire      dam sex
## 8508 M037798 M032181 M022904    2
```

```r
ped[ped$animal == "M032181",]
```

```
##          animal sire   dam sex
## 7007 M032181 <NA> <NA>    1
```

These two individuals are father and offspring and have a genomic kinship which is higher than the expected pedigree kinship ! The observed genomic kinship is around 0.4 which would indicate that there is some inbreeding going on. However we don't know the father's parents so the pedigree kinship is probably missing some links (and the father is somehow related to the mother). We'll keep both individuals assuming that the pedigree is lacking information here !

pair 9

Finally, the ninth pair are individuals M038425 and M038424 with a kinship of 0.4.

```r
#what is the genomic kinship
kas[WeirdgenKin[9,][1],WeirdgenKin[9,][2]]
```

```
## [1] 0.4000155
```

```r
#pedigree kinship
BetaPEDsuborder[WeirdgenKin[9,][1],WeirdgenKin[9,][2]]
```
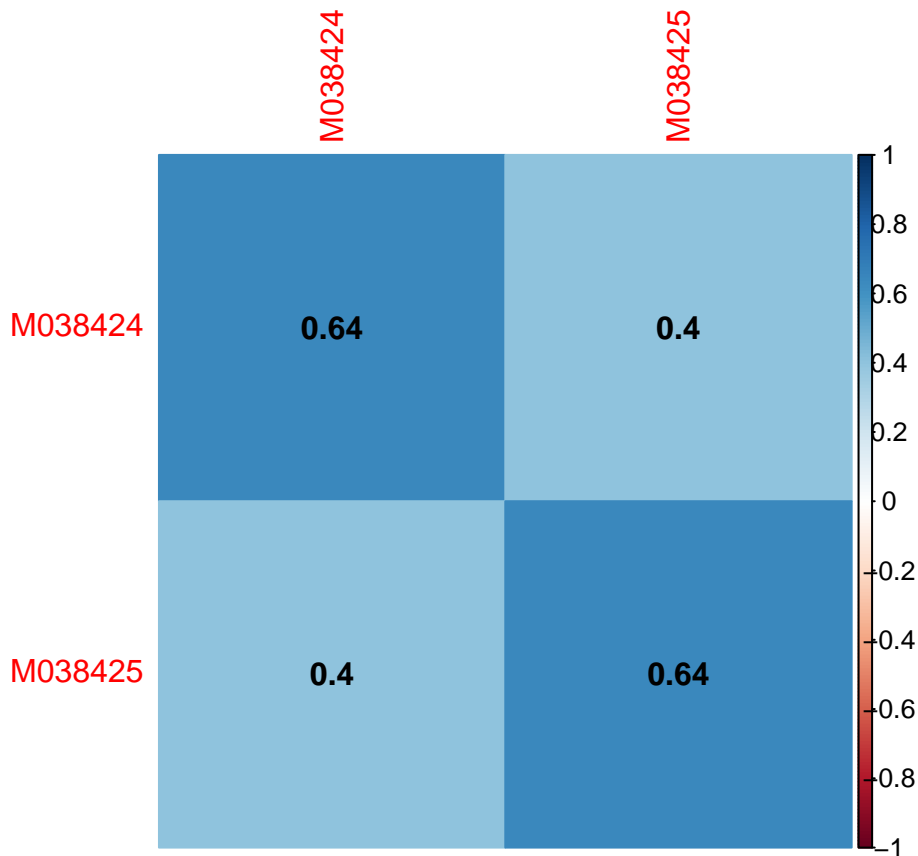
```
## [1] 0.25
```

```
#subset family matrix
kASweirdFam = kas[rownames(kas) %in% c(WeirdgenKinnames[9,][1], WeirdgenKinnames[9,][2]),
                  colnames(kas) %in% c(WeirdgenKinnames[9,][1], WeirdgenKinnames[9,][2])]

#If we look at this family plot
corrplot(kASweirdFam, method = "color", addCoef.col = "black")
```



```
#parents of both individuals
ped[ped$animal == "M038425",]
```

```
##        animal    sire     dam sex
## 8105 M038425 M027966 M027970   1
```

```
ped[ped$animal == "M038424",]
```

```
##        animal    sire     dam sex
## 8104 M038424 M027966 M027970   1
```

These two individuals are siblings and similarly to what we observed with he previous pair, we have a genomic kinship (around 0.4) higher than the pedigree kinship (0.25). We can now check the parents of their parents:

```
#check parents pedigree
ped[ped$animal == "M027966",]
```

```
##        animal sire  dam sex
## 7398 M027966 <NA> <NA>   1
```

```
ped[ped$animal == "M027970",]
```

```
##        animal sire  dam sex
```

17

```
## 7399 M027970 <NA> <NA>    2
```

Since we don't know both parents' parents we can assume that the pedigree is once again missing some link and that the genomic kinship is correct. We'll keep both of these individuals.

Pedigree kinship more than 0.16 higher than genomic kinship

Now we'll investigate the pairs for which the difference between pedigree and genomic kinship is too big ($>$ 0.16).

```r
#pass diagonal and upper tree of difference to NA
diag(diff_kin) = NA
diff_kin[upper.tri(diff_kin)] = NA

#Create df with difference, indv names etc.
high <- as.data.frame(which(diff_kin > 0.16, arr.ind=T))
rownames(high) <- NULL
high$dif<- NA
high$p <- NA
high$b <- NA

#Loop through difference to add indv RingId
for(i in 1:nrow(high)){
high$dif[i] <- diff_kin[high$row[i],high$col[i]]
high$p[i] <- kas[high$row[i],high$col[i]]
high$b[i] <- BetaPEDsuborder[high$row[i],high$col[i]]
}

#Set row.names and col.names
high$RingId1 <- row.names(diff_kin)[high$row]
high$RingId2 <- colnames(diff_kin)[high$col]

#How many individuals are concerned
length(unique(c(high$row,high$col)))
```

```
## [1] 93
```

```r
#high object structure
str(high)
```

```
## 'data.frame':    71 obs. of  7 variables:
##  $ row    : int  1110 1111 1112 591 639 647 1107 1108 1109 1530 ...
##  $ col    : int  88 88 88 116 116 141 245 245 245 245 ...
##  $ dif    : num  0.256 0.266 0.236 0.247 0.247 ...
##  $ p      : num  -0.00633 -0.01581 0.0144 0.00324 0.00325 ...
##  $ b      : num  0.25 0.25 0.25 0.25 0.25 0.25 0.25 0.25 0.25 0.25 ...
##  $ RingId1: chr  "M026819" "M026821" "M026823" "M011181" ...
##  $ RingId2: chr  "M006413" "M006413" "M006413" "M011827" ...
```

group 1

The first three rows are three siblings M026819, M026821 and M026823 and their father: M006413. Their genomic kinship (with the father) are close to 0: -0.006, -0.016 and 0.014 and the pedigree kinship is 0.25.

```r
#what is the genomic kinship
kas[high$row[1],high$col[1]]
```

```
## [1] -0.006331283
```
```
kas[high$row[2],high$col[2]]
```

```
## [1] -0.01581478
```
```
kas[high$row[3],high$col[3]]
```

```
## [1] 0.01440022
```
```
#pedigree kinship
BetaPEDsuborder[high$row[1],high$col[1]]
```

```
## [1] 0.25
```
```
BetaPEDsuborder[high$row[2],high$col[2]]
```

```
## [1] 0.25
```
```
BetaPEDsuborder[high$row[3],high$col[3]]
```

```
## [1] 0.25
```

Let's see who their mom is and check what genomic kinship they have with her:

```
#parents of the three siblings individuals
ped[ped$animal == high$RingId1[1],]
```

```
##          animal     sire      dam sex
## 6610 M026819 M006413 M026255   1
```
```
ped[ped$animal == high$RingId1[2],]
```

```
##          animal     sire      dam sex
## 6612 M026821 M006413 M026255   2
```
```
ped[ped$animal == high$RingId1[3],]
```

```
##          animal     sire      dam sex
## 6614 M026823 M006413 M026255   1
```
```
#kinship with their mom:
kas[colnames(kas) == "M026819", rownames(kas) == "M026255"]
```

```
## [1] 0.2316854
```
```
kas[colnames(kas) == "M026821", rownames(kas) == "M026255"]
```

```
## [1] 0.233123
```
```
kas[colnames(kas) == "M026823", rownames(kas) == "M026255"]
```

```
## [1] 0.2326714
```
```
corrplot(kas[rownames(kas) %in% c(high$RingId1[1], high$RingId1[2], high$RingId1[3],
                                  high$RingId2[1], "M026255"),
         colnames(kas) %in% c(high$RingId1[1], high$RingId1[2], high$RingId1[3],
                                  high$RingId2[1], "M026255")],
        method = "color", addCoef.col = "black")
```

|	M006413	M026255	M026819	M026821	M026823
M006413	0.55	0	−0.01	−0.02	0.01
M026255	0	0.48	0.23	0.23	0.23
M026819	−0.01	0.23	0.5	0.23	0.18
M026821	−0.02	0.23	0.23	0.5	0.19
M026823	0.01	0.23	0.18	0.19	0.5

The kinship with their father is wrong but the kinship with their mom is correct. So we have two possible explanations here: i) the pedigree is wrong and that guy is not their father; ii) we mislabeled the father tube and sequenced someone else. To check this we will 1) see if anyone else among the individuals we sequenced has a genomic kinship with the three siblings close to 0.25 like their father should; 2) we'll see if the father has any siblings which could confirm his identity.

1) who in our data set has a genomic kinship above 0.16 with the kids and could be the father ?

```
#who has a high genomic kinship with one of the kids
row.names(kas)[which(kas[high$row[1],] > 0.16)]
```

```
## [1] "M026255" "M026616" "M026819" "M026821" "M026823"
```

```
#correlation plot
corrplot(kas[rownames(kas) %in% c(high$RingId1[1], high$RingId1[2], high$RingId1[3],
                                  "M026616","M026255"),
         colnames(kas) %in% c(high$RingId1[1], high$RingId1[2], high$RingId1[3],
                              "M026616", "M026255")],
         method = "color", addCoef.col = "black")
```
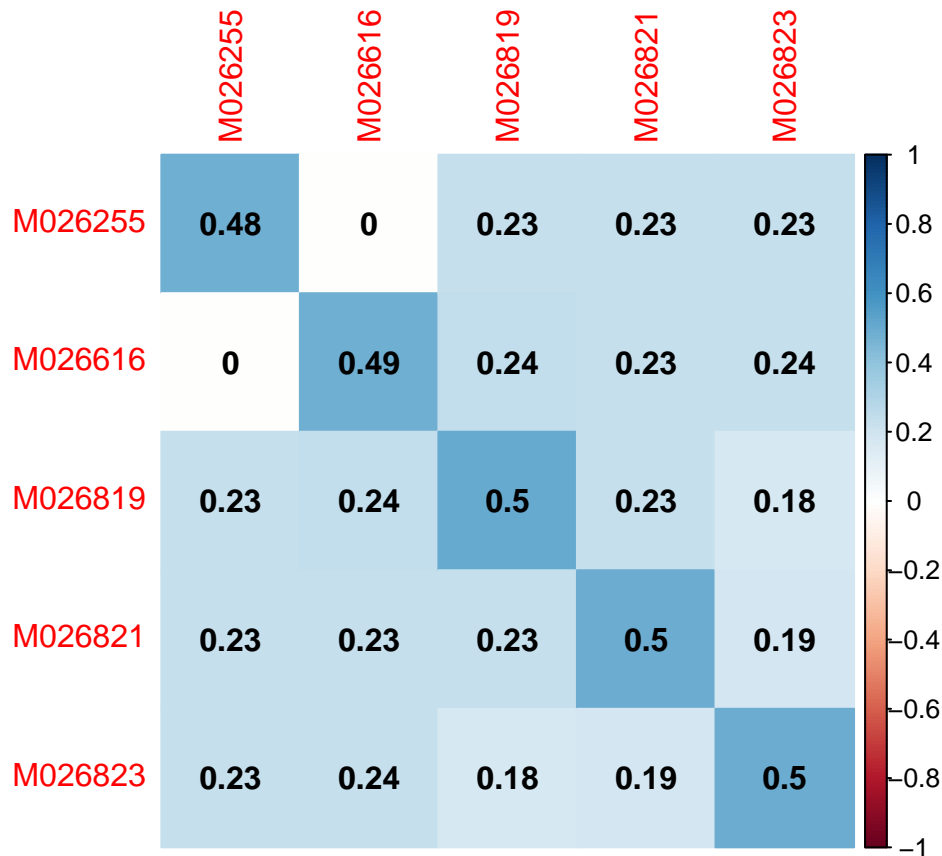
We have the three siblings, the mother and one extra individual which could be the dad (kinship close to 0.25 with the three siblings). According to the database, this potential dad was not seen reproducing with any other females that year.

2) Now about the expected father and it's relatives: we first search for his parents IDs.

```
#Father's parent
ped[ped$animal == high$RingId2[1],]
```

```
##        animal sire  dam sex
## 4917 M006413 <NA> <NA>   1
```

We don't know his parents but in the database, we can see if any other bird comes from the same clutch !

```
#Father's clutch id
birds$BornClutchId[birds$RingId == "M006413"]
```

```
## [1] NA
```

We don't know which clutch it was born in, so we'll see if he has any links with some other birds we sequenced:

```
#Who is related to him
row.names(kas)[which(kas[high$col[2],] > 0.16)] #many links
```

```
##  [1] "M006413" "M017312" "M022861" "M032120" "M022863" "M022864" "M026547"
##  [8] "M026548" "M026790" "M026792" "M026793" "M026848" "M032001" "M032450"
```

```
#genomic kinship with these guys
corrplot(kas[which(kas[high$col[2],] > 0.16),
             which(kas[high$col[2],] > 0.16)],
         method = "color", addCoef.col = "black", number.cex = .7)
```
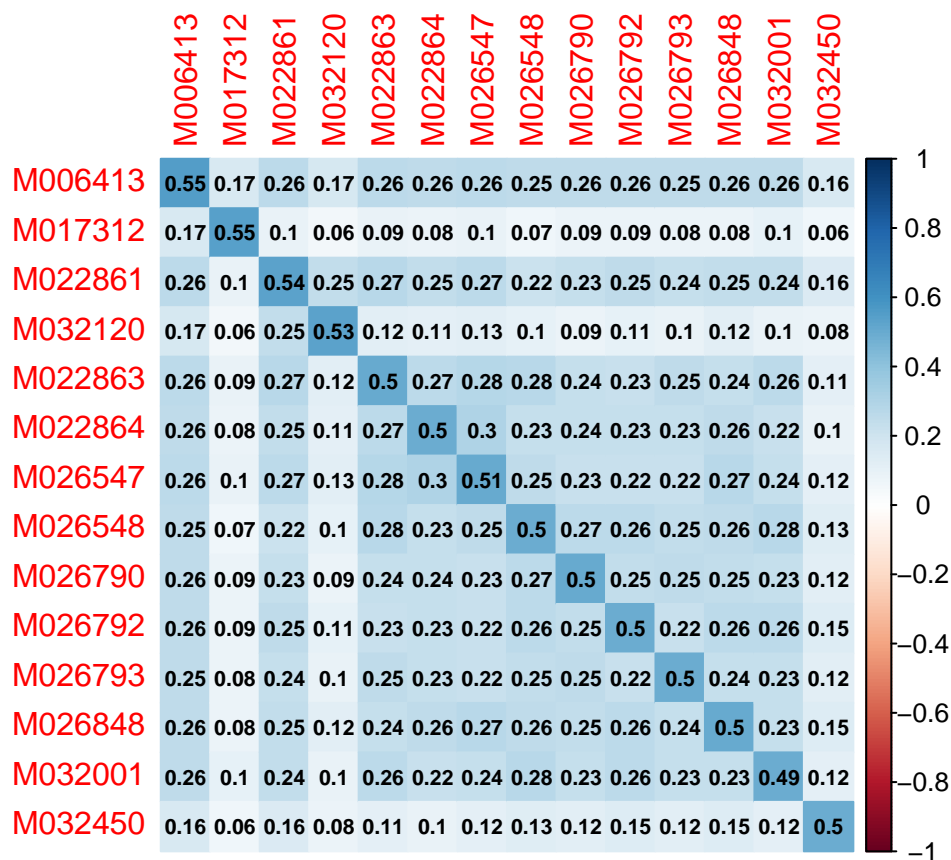
```r
#who are they ?
ped[ped$animal == "M017312",] #we don't know

##        animal sire  dam sex
## 5627 M017312 <NA> <NA>    2

ped[ped$animal == "M022861",] #kid

##        animal    sire     dam sex
## 5972 M022861 M006413 M026459    2

ped[ped$animal == "M032120",] #grand-kid

##        animal    sire     dam sex
## 7750 M032120 M026832 M022861    1

ped[ped$animal == "M022863",] #kid

##        animal    sire     dam sex
## 5974 M022863 M006413 M026459    1

ped[ped$animal == "M022864",] #kid

##        animal    sire     dam sex
## 5975 M022864 M006413 M026459    2

ped[ped$animal == "M026547",] #kid

##        animal    sire     dam sex
## 6424 M026547 M006413 M026459    1
```

```r
ped[ped$animal == "M026548",] #kid
```

```
##       animal    sire     dam sex
## 6425 M026548 M006413 M026459   2
```

```r
ped[ped$animal == "M026790",] #kid
```

```
##       animal    sire     dam sex
## 6585 M026790 M006413 M026459   2
```

```r
ped[ped$animal == "M026792",] #kid
```

```
##       animal    sire     dam sex
## 6587 M026792 M006413 M026459   2
```

```r
ped[ped$animal == "M026793",] #kid
```

```
##       animal    sire     dam sex
## 6588 M026793 M006413 M026459   2
```

```r
ped[ped$animal == "M026848",] #kid
```

```
##       animal    sire     dam sex
## 6639 M026848 M006413 M026459   1
```

```r
ped[ped$animal == "M032001",] #kid
```

```
##       animal    sire     dam sex
## 7678 M032001 M006413 M026459   2
```

```r
ped[ped$animal == "M032450",] #grand-kid
```

```
##       animal    sire     dam sex
## 7784 M032450 M004206 M026849   1
```

He has many links and they match the pedigree: they are either his kids or grand-kids from other clutches. So from this, we know that the individual we sequenced is indeed M006413, he's just not the father of M026819, M026821 and M026823 ! So we'll just keep all the samples and fix the pedigree !

group 2

The next group is composed of one individual M011827 and it's parents: M011181 and M011878 which apparently are not his parents. The parents' genomic kinship with their kid are close to 0: 0.003 and 0.003 but the pedigree kinship is 0.25.

```r
#what is the genomic kinship
kas[high$row[4],high$col[4]]
```

```
## [1] 0.003235013
```

```r
kas[high$row[5],high$col[5]]
```

```
## [1] 0.003253555
```

```r
#pedigree kinship
BetaPEDsuborder[high$row[4],high$col[4]]
```

```
## [1] 0.25
```

```r
BetaPEDsuborder[high$row[5],high$col[5]]
```

```
## [1] 0.25
```

```
#parents of the three siblings individuals
ped[ped$animal == high$RingId2[4],]
```

```
##          animal    sire     dam sex
## 5350 M011827 M011878 M011181   1
```

We'll first look for other kids from the same parents which (in theory) should be siblings of M011827:

```
#Any siblings ?
ped$animal[(ped$sire == "M011878") & (ped$dam == "M011181") &
              (!is.na(ped$sire)) & (!is.na(ped$dam))]
```

```
## [1] "M011257" "M011258" "M011259" "M011260" "M011827"
```

```
#Did we sequence them ?
"M011257" %in% rownames(kas)
```

```
## [1] FALSE
```

```
"M011258" %in% rownames(kas)
```

```
## [1] FALSE
```

```
"M011259" %in% rownames(kas)
```

```
## [1] FALSE
```

```
"M011260" %in% rownames(kas)
```

```
## [1] FALSE
```

We did not sequenced any of his siblings, but does this individual have kids ?

```
#does he has kids ?
ped[ped$sire=='M011827' & !is.na(ped$sire),]
```

```
##          animal    sire     dam sex
## 5246 M011409 M011827 M005415   1
## 5678 M022035 M011827 M005415   1
## 5680 M022037 M011827 M005415   1
## 5682 M022039 M011827 M005415   1
## 5914 M022691 M011827 M022154   2
## 5916 M022694 M011827 M022154   2
## 5922 M022753 M011827 M022154   2
## 6107 M026044 M011827 M022154   2
## 6184 M026301 M011827 M022154   1
## 6185 M026302 M011827 M022154   1
## 6186 M026303 M011827 M022154   1
## 6237 M026354 M011827 M022154   2
## 6238 M026355 M011827 M022154   1
## 6239 M026356 M011827 M022154   2
## 6265 M026383 M011827 M022154   2
## 6266 M026384 M011827 M022154   2
## 6267 M026385 M011827 M022154   2
```

```
#Did we sequence them ?
"M011409" %in% rownames(kas)
```

```
## [1] FALSE
```

```r
"M022035" %in% rownames(kas)
```

```
## [1] FALSE
```

```r
"M022037" %in% rownames(kas)
```

```
## [1] FALSE
```

```r
"M022039" %in% rownames(kas)
```

```
## [1] TRUE
```

```r
"M022691" %in% rownames(kas)
```

```
## [1] FALSE
```

```r
"M022694" %in% rownames(kas)
```

```
## [1] TRUE
```

```r
"M022753" %in% rownames(kas)
```

```
## [1] FALSE
```

```r
"M026044" %in% rownames(kas)
```

```
## [1] FALSE
```

```r
"M026301" %in% rownames(kas)
```

```
## [1] TRUE
```

```r
"M026302" %in% rownames(kas)
```

```
## [1] FALSE
```

```r
"M026303" %in% rownames(kas)
```

```
## [1] FALSE
```

```r
"M026354" %in% rownames(kas)
```

```
## [1] FALSE
```

```r
"M026355" %in% rownames(kas)
```

```
## [1] FALSE
```

```r
"M026356" %in% rownames(kas)
```

```
## [1] FALSE
```

```r
"M026383" %in% rownames(kas)
```

```
## [1] TRUE
```

```r
"M026384" %in% rownames(kas)
```
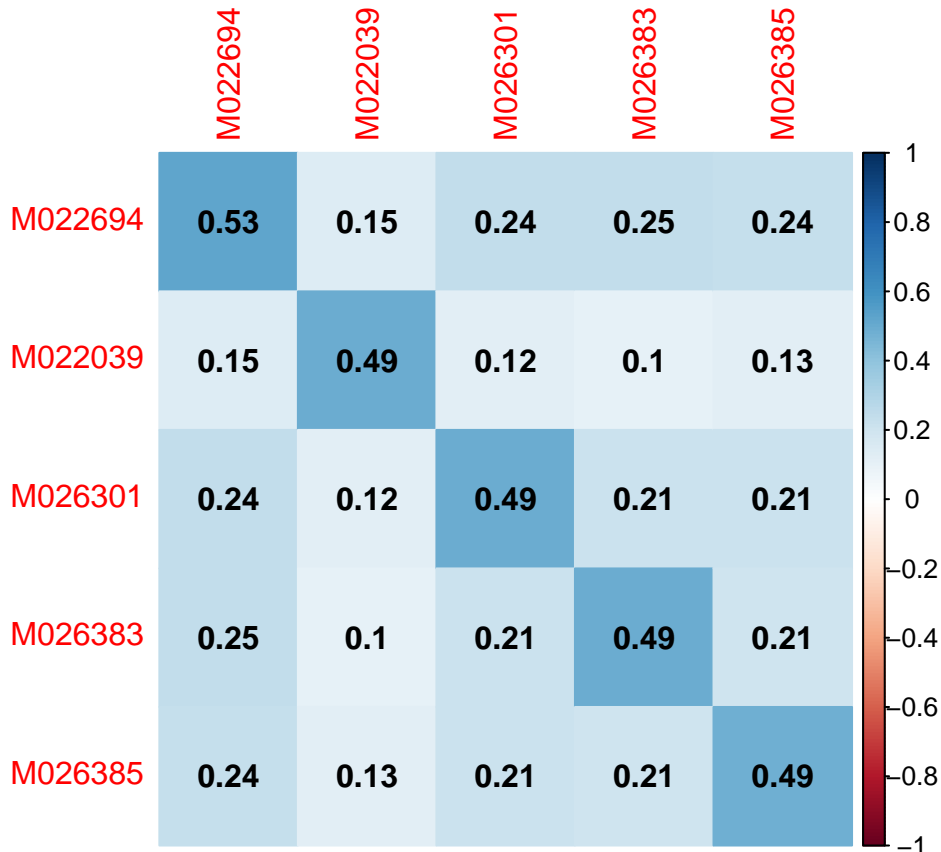
```
## [1] FALSE
```

```r
"M026385" %in% rownames(kas)
```

```
## [1] TRUE
```

We sequenced some of his kids so let's see if their genomic kinship matches the expected pedigree kinship with him. if it does, then we sequenced the correct individual, but the pedigree was wrong about it's parents. If it does not, it means that we did not sequence the correct individual.

```
#genomic kinship with these guys
corrplot(kas[rownames(kas) %in% c("M022039","M022694","M026301",
                                  "M026383","M026385"),
             colnames(kas) %in% c("M022039","M022694","M026301",
                                  "M026383","M026385")],
         method = "color", addCoef.col = "black")
```

	M022694	M022039	M026301	M026383	M026385
M022694	0.53	0.15	0.24	0.25	0.24
M022039	0.15	0.49	0.12	0.1	0.13
M026301	0.24	0.12	0.49	0.21	0.21
M026383	0.25	0.1	0.21	0.49	0.21
M026385	0.24	0.13	0.21	0.21	0.49

That guy really is the father of all these kids (M022039 has lower genomic kinship with the rest of the siblings because he has a different mother). So we can be sure we sequenced the correct individual. We'll keep that sample and set both his parents to NA in the pedigree !

group 3

The next group is composed of two individuals M022383 and M012646 with genomic kinship 0.006 but a pedigree kinship of 0.25. It's supposed to be a father and his son but they are unrelated!

```
#what is the genomic kinship
kas[high$row[6],high$col[6]]
```

```
## [1] 0.005908333
```

```
#pedigree kinship
BetaPEDsuborder[high$row[6],high$col[6]]
```

```
## [1] 0.25
```

We'll first identify who is the mother and whether there are other kids in this clutch we sequenced:

```
#who is the mother ?
ped[ped$animal == high$RingId2[6],]
```

```
##      animal    sire     dam sex
## 5862 M022383 M012646 M011725   1
```
```
#Any siblings ?
ped$animal[(ped$sire == "M012646") & (ped$dam == "M011725") &
           (!is.na(ped$sire)) & (!is.na(ped$dam))]
```
```
## [1] "M022382" "M022383" "M022384" "M022759"
```
```
#Did we sequence them ?
"M022382" %in% rownames(kas)
```
```
## [1] FALSE
```
```
"M022384" %in% rownames(kas)
```
```
## [1] FALSE
```
```
"M022759" %in% rownames(kas)
```
```
## [1] FALSE
```

Unfortunately we did not sequenced any of the other siblings from this clutch. We'll now try to see if individual M022383 is related to anyone in the data set and where these guys come from!

```
#who is related ?
row.names(kas)[which(kas[high$col[6],] > 0.16)]
```
```
## [1] "M022383" "M022385" "M026267"
```
```
#who are these guys ?
ped[ped$animal == "M022385",]
```
```
##      animal sire  dam sex
## 6701 M022385 <NA> <NA>   1
```
```
ped[ped$animal == "M026267",]
```
```
##      animal sire  dam sex
## 6765 M026267 <NA> <NA>   2
```
```
#Which clutch are they coming from ?
birds$BornClutchId[birds$RingId == "M022385"]
```
```
## [1] NA
```
```
birds$BornClutchId[birds$RingId == "M026267"]
```
```
## [1] NA
```

We don't know their parents or which clutch they come from. They might be his true parents because they were seen as adults this year but we really cannot be sure.

Now what about the father, anyone related to him in our data set ?

```
#parents ?
ped[ped$animal == high$RingId1[6],]
```
```
##      animal sire  dam sex
## 5617 M012646 <NA> <NA>   1
```
```
#who is related ?
row.names(kas)[which(kas[high$row[6],] > 0.16)]
```

```
## [1] "M012646" "M022859"
```
```
#who is that guy ?
ped[ped$animal == "M022859",]
```
```
##          animal sire     dam sex
## 5970 M022859 <NA> M026901   1
```

We found one individual but from the pedigree we cannot know what is their relationship ...

It seems to be a complicated case, we'll come back to this later, for now we will remove both guys !

group 4

The next rows are one individual M032002 and all his supposed kids M026816, M026817, M026818, M032035, M032036. But from the genomic kinship we can be sure that he's not the dad.

```
#what is the genomic kinship
corrplot(kas[c(high$col[7], high$row[7:11]),
             c(high$col[7], high$row[7:11])],
         method = "color", addCoef.col = "black")
```
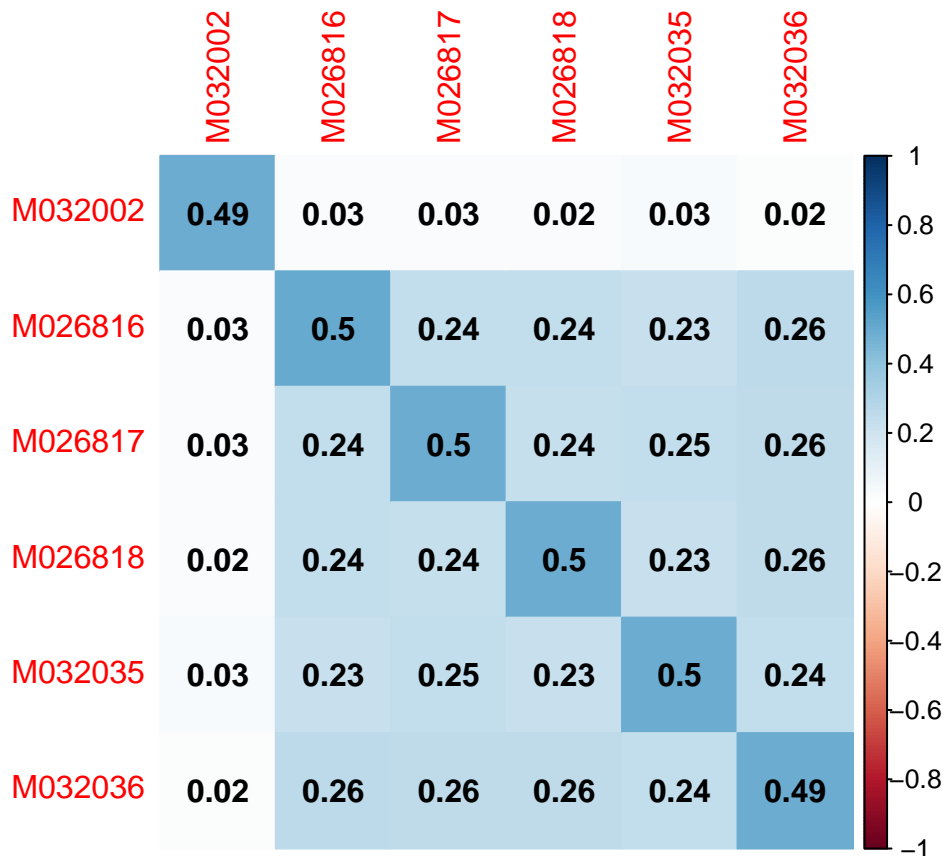


```
#pedigree kinship /what it should look like)
corrplot(BetaPEDsuborder[c(high$col[7],high$row[7:11]),
                         c(high$col[7],high$row[7:11])],
         method = "color", addCoef.col = "black")
```
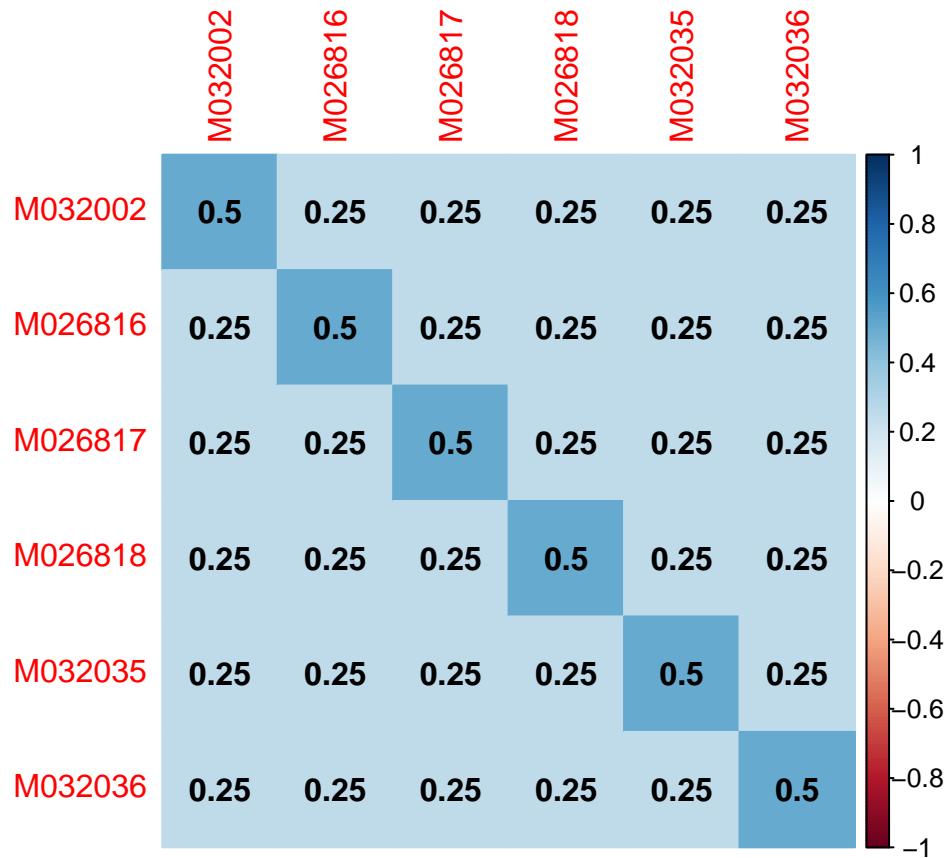
```r
#Who is the mom ?
ped[ped$animal == high$RingId1[7],]
```

```
##        animal    sire      dam sex
## 6607 M026816 M032002 M026985   1
```

```r
#kinship of the kids with the mom:
corrplot(kas[rownames(kas) %in% c("M026985", high$RingId2[7], high$RingId1[7:11]),
             colnames(kas) %in% c("M026985", high$RingId2[7], high$RingId1[7:11])],
         method = "color", addCoef.col = "black")
```

	M026985	M032002	M026816	M026817	M026818	M032035	M032036
M026985	0.49	0	0.24	0.24	0.24	0.24	0.24
M032002	0	0.49	0.03	0.03	0.02	0.03	0.02
M026816	0.24	0.03	0.5	0.24	0.24	0.23	0.26
M026817	0.24	0.03	0.24	0.5	0.24	0.25	0.26
M026818	0.24	0.02	0.24	0.24	0.5	0.23	0.26
M032035	0.24	0.03	0.23	0.25	0.23	0.5	0.24
M032036	0.24	0.02	0.26	0.26	0.26	0.24	0.49
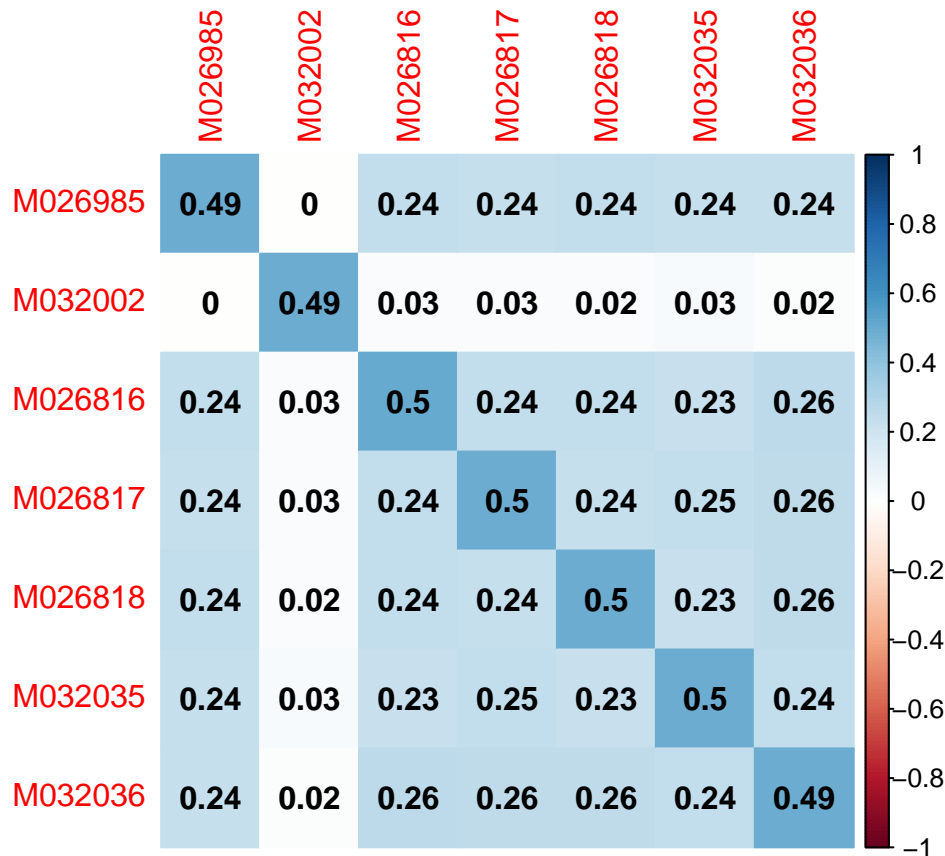
So the kids are indeed related (among themselves and with their mom). So either 1) he's just not the father, or 2) we mislabeled the father's tube) !

We know that this guy (M032002) is supposed to have other kids (from different clutches) so let's see if 1) some individuals are related to him, 2) his genomic kinship matches the pedigree kinship from other kids he fathered:

```r
#who is related to this guy?
row.names(kas)[which(kas[high$col[7],] > 0.16)]
```

```
##  [1] "M022875" "M032002" "M022870" "M022871" "M022872" "M022873" "M022876"
##  [8] "M022899" "M026834" "M026836" "M026837" "M032031" "M032378" "M032379"
```

```r
#bunch of kids
ped[ped$animal == "M022875",] #kid
```

```
##       animal    sire      dam sex
## 5986 M022875 M032002 M026985   2
```

```r
ped[ped$animal == "M022870",] #kid
```

```
##       animal    sire      dam sex
## 5981 M022870 M032002 M026985   1
```

```r
ped[ped$animal == "M022871",] #kid
```

```
##       animal    sire      dam sex
## 5982 M022871 M032002 M026985   2
```

```r
ped[ped$animal == "M022872",] #kid
```

```
##        animal       sire       dam sex
## 5983 M022872 M032002 M026985    1
```

```r
ped[ped$animal == "M022873",] #kid
```

```
##        animal       sire       dam sex
## 5984 M022873 M032002 M026985    2
```

```r
ped[ped$animal == "M022876",] #kid
```

```
##        animal       sire       dam sex
## 5987 M022876 M032002 M026985    1
```

```r
ped[ped$animal == "M022899",] #kid
```

```
##        animal       sire       dam sex
## 6010 M022899 M032002 M026985    2
```

```r
ped[ped$animal == "M026834",] #We don't know but father ID is missing
```

```
##        animal sire       dam sex
## 6625 M026834 <NA> M022998    2
```

```r
ped[ped$animal == "M026836",] #We don't know but father ID is missing
```

```
##        animal sire       dam sex
## 6627 M026836 <NA> M022998    1
```

```r
ped[ped$animal == "M026837",] #We don't know but father ID is missing
```

```
##        animal sire       dam sex
## 6628 M026837 <NA> M022998    2
```

```r
ped[ped$animal == "M032031",] #We don't know but father ID is missing
```

```
##        animal sire       dam sex
## 7702 M032031 <NA> M022998    1
```

```r
ped[ped$animal == "M032378",] #kid
```

```
##        animal       sire       dam sex
## 7159 M032378 M032002 M026985    1
```

```r
ped[ped$animal == "M032379",] #kid
```

```
##        animal       sire       dam sex
## 7160 M032379 M032002 M026985    2
```

This is interesting, for a bunch of the related individuals, he is labelled as the father, but for four of them, we did not know who was the father (when we recorded the pedigree). For the kids we knew he was the father, does the genomic kinship matches the expected pedigree kinship (0.25)?

```r
corrplot(kas[rownames(kas) %in% c("M022875","M022870","M022871","M022872",
                                  "M022873","M022876","M032378","M032379"),
             colnames(kas) %in% c("M022875","M022870","M022871","M022872",
                                  "M022873","M022876","M032378","M032379")],
         method = "color", addCoef.col = "black")
```
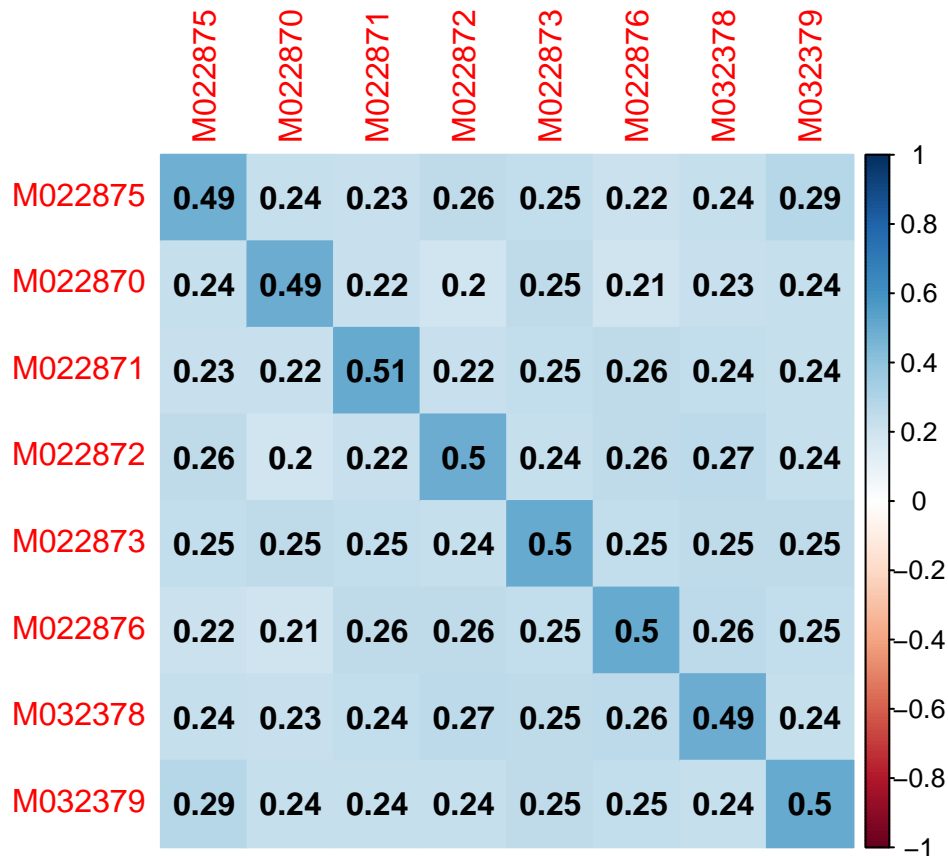
Yes it matches! Since this guy is the father of other kids and the pedigree and genomic kinship matches, we can say that it is not a labeling problem, he's just not the father of the clutch we were studying! We can also add him as a father from the rest of the kids which were related to him but for which the father was unknown!

Then for the clutch we were initially studying, who's the father? We'll start by looking for related individuals with one of the kids of the clutch:

```r
#who is related to the kids ?
row.names(kas)[which(kas[high$row[7],] > 0.16)]
```

```
##  [1] "M026985" "M022873" "M026816" "M026817" "M026818" "M032035" "M032036"
##  [8] "M032289" "M032463" "M032464" "M032465" "M032466" "M032495" "M032496"
## [15] "M032497" "M032498" "M032500"
```

```r
#bunch of kids
ped[ped$animal == "M026985",] # That's the mom
```

```
##       animal     sire     dam sex
## 7380 M026985 M026478 M027889   2
```

```r
ped[ped$animal == "M022873",] # sibling same "wrong" father
```

```
##       animal     sire     dam sex
## 5984 M022873 M032002 M026985   2
```

```r
ped[ped$animal == "M032289",] # offspring
```

```
##       animal     sire     dam sex
## 7076 M032289 M026816 M017312   2
```

```r
ped[ped$animal == "M032463",] # offspring
```

```
##       animal    sire      dam sex
## 7797 M032463 M026816 M032222   1
```

```r
ped[ped$animal == "M032464",] # offspring
```

```
##       animal    sire      dam sex
## 7798 M032464 M026816 M032222   1
```

```r
ped[ped$animal == "M032465",] # offspring
```

```
##       animal    sire      dam sex
## 7799 M032465 M026816 M032222   1
```

```r
ped[ped$animal == "M032466",] # offspring
```

```
##       animal    sire      dam sex
## 7800 M032466 M026816 M032222   1
```

```r
ped[ped$animal == "M032495",] # offspring
```

```
##       animal    sire      dam sex
## 7827 M032495 M026816 M017312   1
```

```r
ped[ped$animal == "M032496",] # offspring
```

```
##       animal    sire      dam sex
## 7828 M032496 M026816 M017312   2
```

```r
ped[ped$animal == "M032497",] # offspring
```

```
##       animal    sire      dam sex
## 7829 M032497 M026816 M017312   1
```

```r
ped[ped$animal == "M032498",] # offspring
```

```
##       animal    sire      dam sex
## 7830 M032498 M026816 M017312   1
```

```r
ped[ped$animal == "M032500",] # offspring
```

```
##       animal    sire      dam sex
## 7207 M032500 M026816 M017312   2
```

The only individuals which came out as related to the kid were one sibling and his offspring. This sibling (M022873) has the same parents but did not come out as weird when we identified samples with large discrepancies between pedigree and genomic kinship. So this is weird, let's check what his kinship with the "dad" (M032002) is:

```r
#genomic kinship with the "dad"?
kas[rownames(kas) == "M022873", colnames(kas) == "M032002"]
```

```
## [1] 0.2424666
```

The genomic kinship is around 0.25 indicating that M032002 probably is the dad of this individual. We'll check whether him (M022873) and the other kids (M026816, M026817, M026818, M032035, M032036) are from the same clutch. if there are not, them we'll just assume that the father was wrongly identified for one clutch but was correct for the other. If they are from the same clutch, it means that we failed to document some cross-fostering events:

```r
birds$BornClutchId[birds$RingId == "M022873"]
```

```
## [1] 1411
```

```r
birds$BornClutchId[birds$RingId == "M026816"]
```
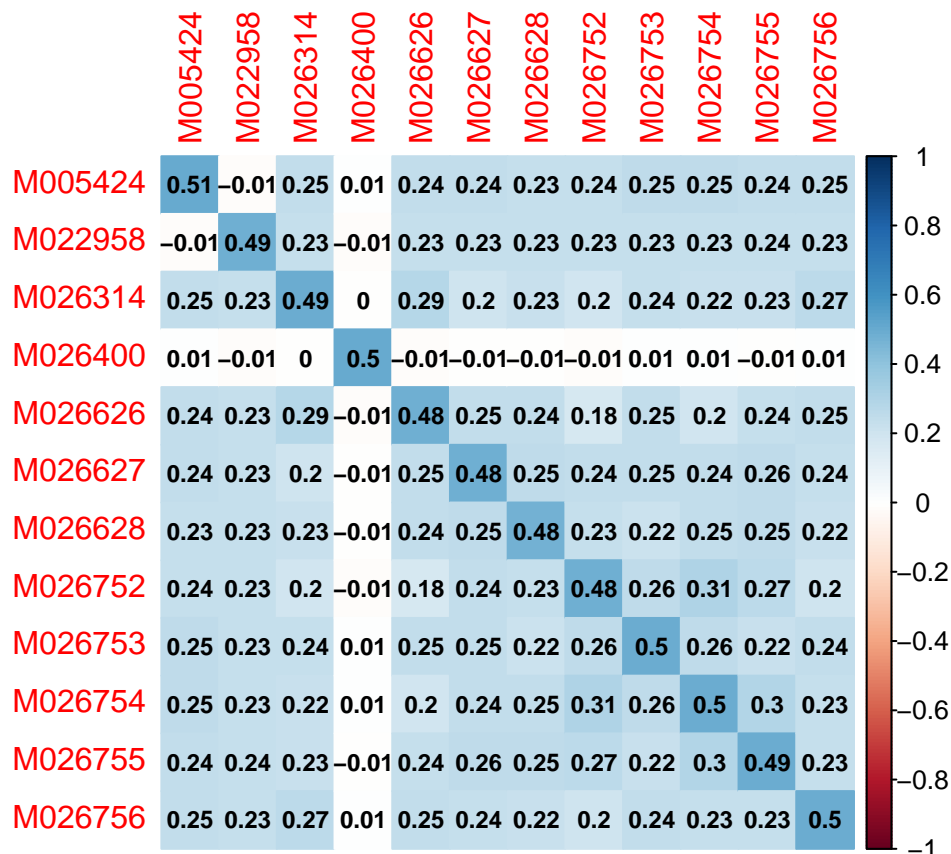
```
## [1] 1389
```

They're not from the same clutch! We'll keep all individuals, we'll just correct the pedigree: remove M032002 as a father of clutch 1389 and add him as the father of the other clutch! We could not identify the true father of clutch 1389, we'll just leave it as NA.K

group 5

The next rows are one individual M026400 and many individuals: M005424: his supposed dad; M022958: his supposed mom; and many supposed siblings : M026626, M026627, M026628, M026752, M026753, M026754, M026755, M026756 and M026314. We'll start by visualizing the genomic kinship matrix of this family:

```r
#family
listofFAM = c(high$RingId1[15], high$RingId2[15], high$RingId2[23],
              high$RingId1[25:32], high$RingId2[24])

#corplot
corrplot(as.matrix(kas[row.names(kas) %in% listofFAM,
                       row.names(kas) %in% listofFAM ]),
         method = 'color',addCoef.col = 'black', number.cex = .75)
```

	M005424	M022958	M026314	M026400	M026626	M026627	M026628	M026752	M026753	M026754	M026755	M026756
M005424	0.51	−0.01	0.25	0.01	0.24	0.24	0.23	0.24	0.25	0.25	0.24	0.25
M022958	−0.01	0.49	0.23	−0.01	0.23	0.23	0.23	0.23	0.23	0.23	0.24	0.23
M026314	0.25	0.23	0.49	0	0.29	0.2	0.23	0.2	0.24	0.22	0.23	0.27
M026400	0.01	−0.01	0	0.5	−0.01	−0.01	−0.01	−0.01	0.01	0.01	−0.01	0.01
M026626	0.24	0.23	0.29	−0.01	0.48	0.25	0.24	0.18	0.25	0.2	0.24	0.25
M026627	0.24	0.23	0.2	−0.01	0.25	0.48	0.25	0.24	0.25	0.24	0.26	0.24
M026628	0.23	0.23	0.23	−0.01	0.24	0.25	0.48	0.23	0.22	0.25	0.25	0.22
M026752	0.24	0.23	0.2	−0.01	0.18	0.24	0.23	0.48	0.26	0.31	0.27	0.2
M026753	0.25	0.23	0.24	0.01	0.25	0.25	0.22	0.26	0.5	0.26	0.22	0.24
M026754	0.25	0.23	0.22	0.01	0.2	0.24	0.25	0.31	0.26	0.5	0.3	0.23
M026755	0.24	0.24	0.23	−0.01	0.24	0.26	0.25	0.27	0.22	0.3	0.49	0.23
M026756	0.25	0.23	0.27	0.01	0.25	0.24	0.22	0.2	0.24	0.23	0.23	0.5

This guy (M026400) is indeed unrelated to all his supposed family. But can we find individuals he is related to (and should not be according for pedigree):

```r
#Who is he related to?
row.names(kas)[which(kas[high$row[15],] > 0.16)]
```

```
## [1] "M011685" "M011817" "M026400"
```
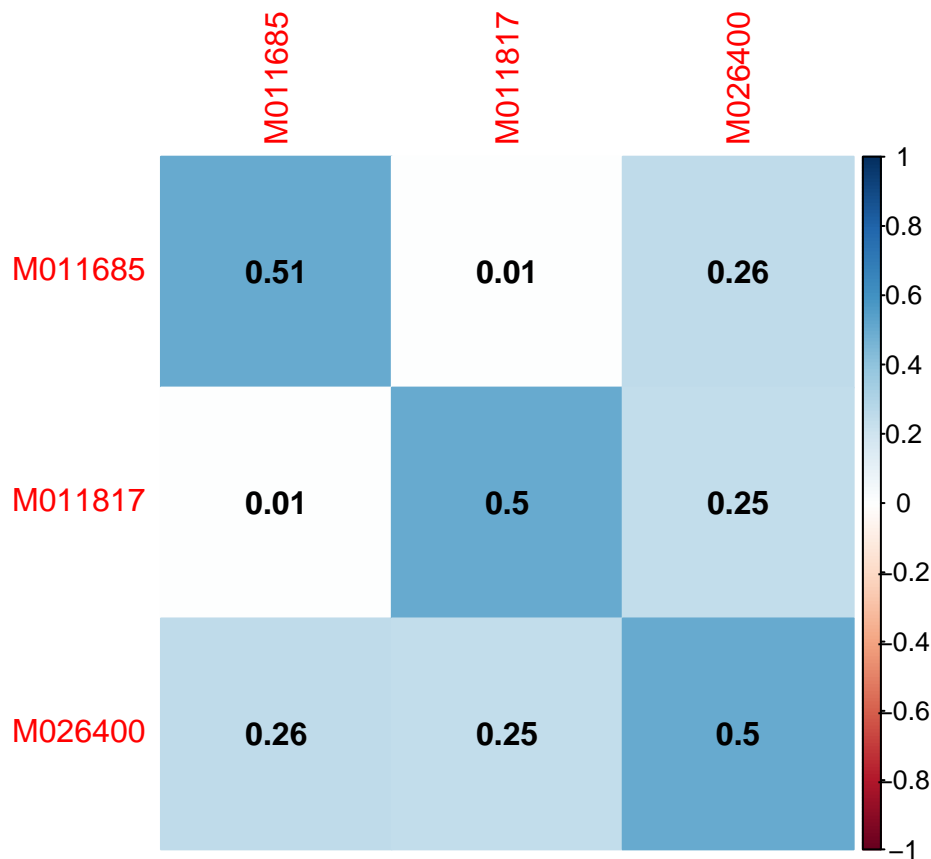
```r
#What are the relatedness with these guys

#pedigree
BetaPEDsuborder[rownames(BetaPEDsuborder) %in% c("M011685","M011817"),
                colnames(BetaPEDsuborder) %in% c("M026400")]
```

```
## M011685 M011817
##       0       0
```

```r
#genomic
corrplot(kas[rownames(kas) %in% c("M011685","M011817","M026400"),
             colnames(kas) %in%  c("M011685","M011817","M026400")],
         method = 'color',addCoef.col = 'black')
```



```r
#Do these individuals have other kids ?
ped$animal[(ped$sire == "M011685") & (ped$dam == "M011817") &
             (!is.na(ped$sire)) & (!is.na(ped$dam))]
```

```
## [1] "M022398" "M022400" "M026002" "M026004" "M026006" "M026008"
```

```r
#Did we sequence them ?
"M022398" %in% rownames(kas)
```

```
## [1] FALSE
```

```
"M022400" %in% rownames(kas)
```

```
## [1] FALSE
```

```
"M026002" %in% rownames(kas)
```

```
## [1] FALSE
```

```
"M026004" %in% rownames(kas)
```

```
## [1] FALSE
```

```
"M026006" %in% rownames(kas)
```

```
## [1] FALSE
```

```
"M026008" %in% rownames(kas)
```

```
## [1] FALSE
```

We found two individuals which might be his parents, especially since they have other kids with similar names (could be a typo when we wrote the names). Unfortunately we did not sequence any of their kids so we'll just have to remove individual M026400 from our dataset.

group 6

The next group is composed of M031719 and three individuals: his supposed dad: M032487 and two of his supposed siblings: M037505 and M037507. We'll start by looking at their genomic kinship:

```
#family link
listofFAM = c(high$RingId1[12], high$RingId2[12], high$RingId1[43], high$RingId1[44])

#corplot
corrplot(as.matrix(kas[row.names(kas) %in% listofFAM,
                       row.names(kas) %in% listofFAM ]),
         method = 'color',addCoef.col = 'black')
```
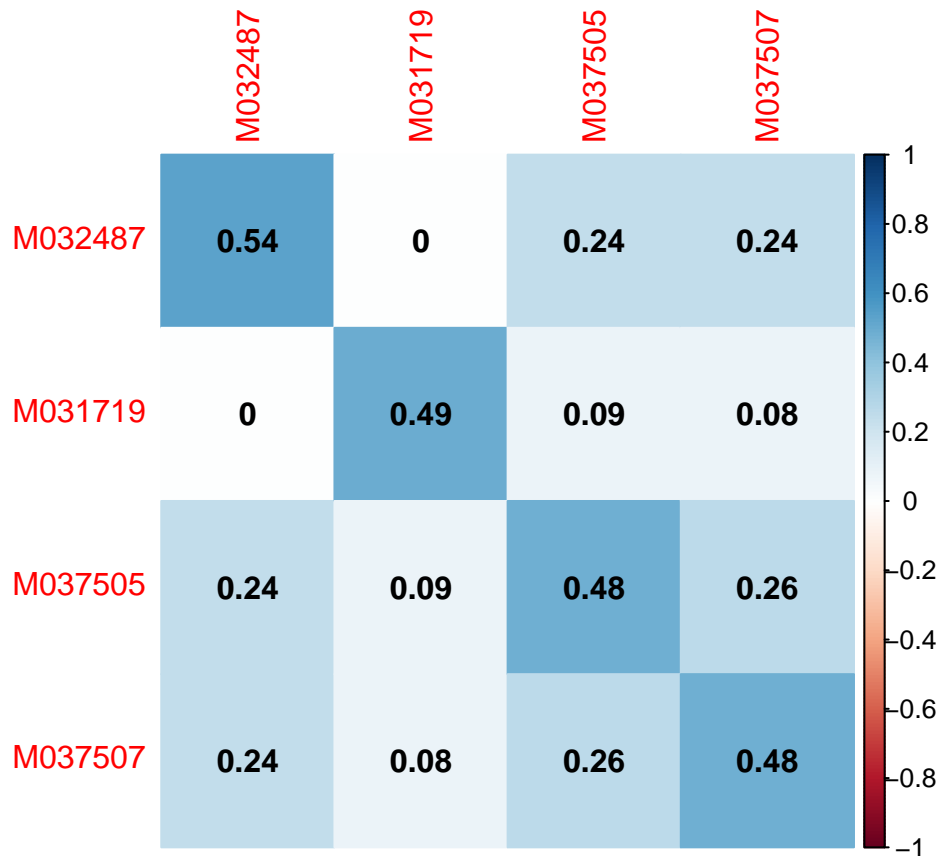
Individual M031719 is not related to his supposed father but his somewhat related to his supposed siblings: genomic kinship close to 0.1, which is larger than 0 (expected between unrelated) but lower than 0.25 (expected between siblings). The other siblings on he other hand seem to be the offspring of the dad (genomic kinship close to 0.25). What about the mom ?
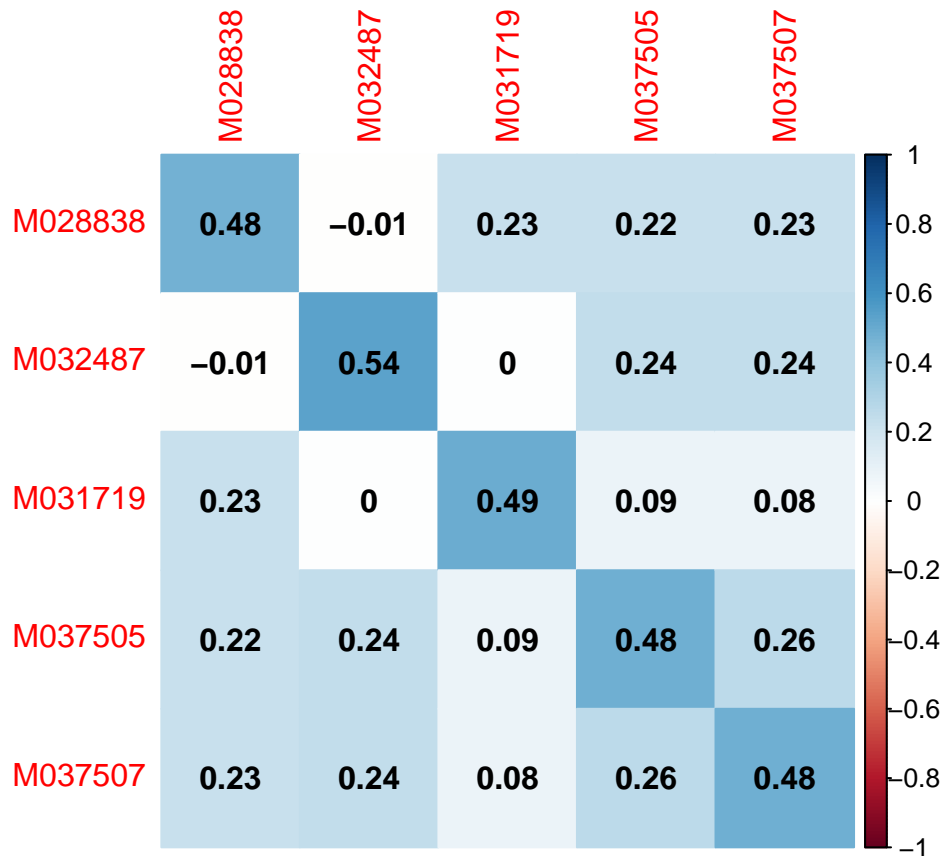
```
#What about his mom ?
ped[ped$animal == "M031719",]
```

```
##        animal     sire      dam sex
## 6812 M031719 M032487 M028838   2
```

```
listofFAM = c(high$RingId1[12], high$RingId2[12], high$RingId1[43],
              high$RingId1[44], "M028838")
```

```
#correlation plot
corrplot(as.matrix(kas[row.names(kas) %in% listofFAM,
                   row.names(kas) %in% listofFAM ]),
         method = 'color',addCoef.col = 'black')
```

There is no problem with the mom kinship: she has a genomic kinship close to 0.25 with the three of the kids! It seems that this kid (M031719) is the kid of the mom but not the dad. He's then half-sibling with the two others kids. We'll first check that he does not come from the same clutch (it would be weird if he did).

```
#same clutch ?
birds$BornClutchId[birds$RingId == "M031719"]
```

```
## [1] 1530
birds$BornClutchId[birds$RingId == "M037505"]
```

```
## [1] 1616
birds$BornClutchId[birds$RingId == "M037507"]
```

```
## [1] 1616
```

he does not come from the same clutch (good)! Now we'll then see if we can identify his true dad ?

```
#If we look for new links for this individuals
row.names(kas)[which(kas[high$row[12],] > 0.16)]
```

```
## [1] "M028838" "M031719" "M031855"
```

Other than the mom and himself, one extra individual pops up, it is probably the dad. We will keep all samples and fix the pedigree.

group 7

The next group is composed of one individual M043755 and five individuals: M033855, M004209, M043755, M043755 and M043755. These individuals should have a kinship of 0.25 (according to pedigree) but their

genomic kinship are lower.

```
listofFAM = c(high$RingId1[13], high$RingId2[13], high$RingId1[68],
              high$RingId1[69], high$RingId1[70], high$RingId2[50])

#corplot
corrplot(as.matrix(kas[row.names(kas) %in% listofFAM,
                   row.names(kas) %in% listofFAM ]),
         method = 'color',addCoef.col = 'black')
```
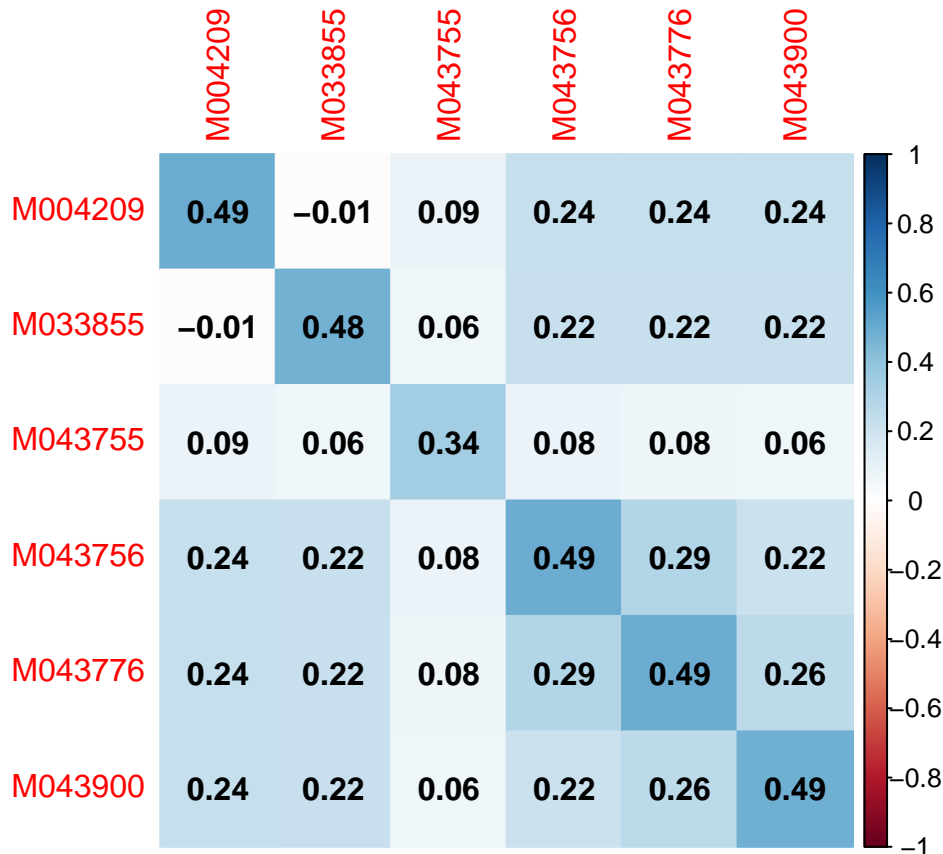

From this correlation plot, we can see that this guy has a weird self kinship (0.342) which could explain the lower kinship with the rest of his family... Anyway it's weird and might indicate that we cannot trust the genomic data of this individual so we'll just remove him.

group 8

The group is composed of two individuals: a dad: M005093 and it's supposed kid: M011301. The pedigree kinship is 0.25 but the genetic kinship are close to 0 which indicates that some link is wrong. let's first check the genomic kinship.

```
#individuals in this family
listofFAM = c(high$RingId2[14], high$RingId1[14])

#correlation plot
corrplot(as.matrix(kas[row.names(kas) %in% listofFAM,
                   row.names(kas) %in% listofFAM ]),
         method = 'color',addCoef.col = 'black')
```
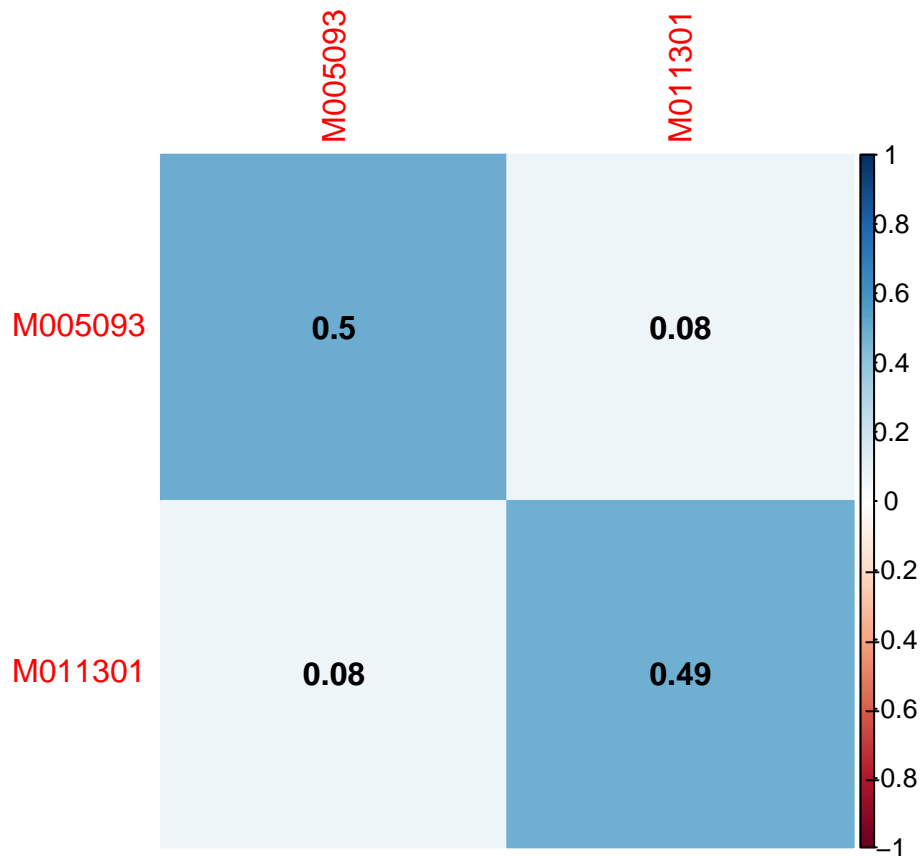
The dad (M005093) is not related to this particular kid in this clutch. We'll first see if this clutch contained any other kids and if we sequenced them:

```
#other kids ?
ped[(ped$sire == high$RingId2[14]) & (!is.na(ped$sire)),]
```

```
##         animal     sire    dam sex
## 4860 M005680 M005093 M000376   1
## 4861 M005681 M005093 M000376   2
## 4862 M005682 M005093 M000376   1
## 4863 M005683 M005093 M000376   1
## 5115 M011182 M005093 M000376   1
## 5126 M011194 M005093 M000376   2
## 5192 M011301 M005093 M000376   1
## 5193 M011302 M005093 M000376   2
## 5194 M011303 M005093 M000376   1
## 5278 M011713 M005093 M000376   2
## 5280 M011715 M005093 M000376   1
## 5609 M011995 M005093 M000376   1
```

```
#Yes, did we sequenced the kids ?
"M005680" %in% colnames(kas)
```

```
## [1] FALSE
```

```
"M005681" %in% colnames(kas)
```

```
## [1] FALSE
```

```r
"M005682" %in% colnames(kas)
```

```
## [1] FALSE
```

```r
"M005683" %in% colnames(kas)
```

```
## [1] FALSE
```

```r
"M011182" %in% colnames(kas)
```

```
## [1] FALSE
```

```r
"M011194" %in% colnames(kas)
```

```
## [1] FALSE
```

```r
"M011302" %in% colnames(kas)
```

```
## [1] FALSE
```

```r
"M011303" %in% colnames(kas)
```

```
## [1] FALSE
```

```r
"M011713" %in% colnames(kas)
```

```
## [1] FALSE
```

```r
"M011715" %in% colnames(kas)
```

```
## [1] FALSE
```

```r
"M011995" %in% colnames(kas)
```

```
## [1] TRUE
```

Fortunately, we did sequence one of them: M011995. We'll now check his genomic kinship with both the dad (M005093) and the kid (M011301). We will also investigate the mom: did we sequenced her and what is her genomic kinship with the kid?
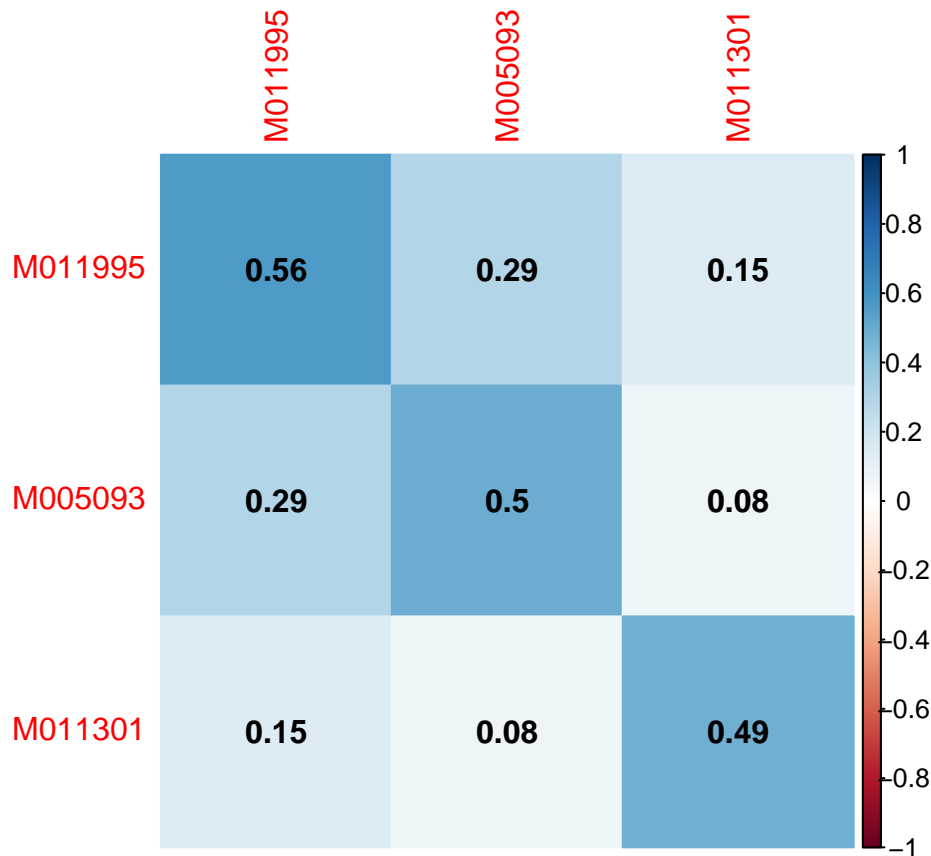
```r
#Did we sequenced the mom ?
"M000376" %in% colnames(kas)
```

```
## [1] FALSE
```

```r
#We did not sequence her

#What about the other sibling ?
#correlation plot
corrplot(as.matrix(kas[row.names(kas) %in% c("M011995",listofFAM),
                        row.names(kas) %in% c("M011995",listofFAM)]),
         method = 'color',addCoef.col = 'black')
```

We did not sequence the mom and the other kid we sequenced (M011995) has a genomic kinship with the dad (M005093) close to 0.25 suggesting that their relation is correct. So we don't have any other choice but to remove the first kid (M011301) from our data set. He (M011301) seems to be a half-sibling with his supposed sibling (M011995): the genomic kinship is around 0.15 but we cannot know who he really is. Since the dad kinship with the second kid (M011995) was correct, we'll keep him in the data set.

group 9

The next group is composed of five individuals: a dad: M022672 and it's supposed four kids: M031576, M031577, M031578, M031579. The pedigree kinship is 0.25 but the genetic kinship are close to 0 which indicates that some link is wrong. let's first check the genomic kinship.

```r
#individuals in this family
listofFAM = c(high$RingId2[16], high$RingId1[16:19])

#corplot
corrplot(as.matrix(kas[row.names(kas) %in% listofFAM,
                       row.names(kas) %in% listofFAM ]),
         method = 'color',addCoef.col = 'black')
```

The dad (M022672) is not related to any of the kids in this clutch. So either he is not the dad, or we did not sequence the correct individual. To check for this, we'll see if this individual (M022672) has any other kids or siblings and if we sequenced them:

```
#other kids ?
ped[(ped$sire == high$RingId2[16]) & (!is.na(ped$sire)),]
```

```
##        animal     sire      dam sex
## 7467 M031576 M022672 M031529   1
## 7468 M031577 M022672 M031529   1
## 7469 M031578 M022672 M031529   2
## 7470 M031579 M022672 M031529   2
```

```
#no
```

```
#Who are his parents ?
ped[ped$animal == high$RingId2[16],]
```

```
##        animal     sire     dam sex
## 5913 M022672 M001670 898630   1
```

```
#Does he have siblings ?
ped[(ped$sire == "M001670") & (ped$dam == "898630") &
    (!is.na(ped$sire)) & (!is.na(ped$dam)),]
```

```
##        animal     sire     dam sex
## 5672 M022028 M001670 898630   2
## 5673 M022029 M001670 898630   1
## 5674 M022030 M001670 898630   2
```

```
## 5675 M022031 M001670 898630    1
## 5676 M022032 M001670 898630    1
## 5738 M022095 M001670 898630    2
## 5739 M022096 M001670 898630    1
## 5740 M022097 M001670 898630    2
## 5741 M022098 M001670 898630    2
## 5743 M022100 M001670 898630    1
## 5804 M022174 M001670 898630    1
## 5805 M022175 M001670 898630    1
## 5806 M022176 M001670 898630    1
## 5807 M022177 M001670 898630    2
## 5838 M022359 M001670 898630    2
## 5909 M022668 M001670 898630    1
## 5910 M022669 M001670 898630    1
## 5911 M022670 M001670 898630    2
## 5912 M022671 M001670 898630    1
## 5913 M022672 M001670 898630    1
```

```r
#Did we sequence them ??
"M022028" %in% rownames(kas)
```

```
## [1] FALSE
```

```r
"M022029" %in% rownames(kas)
```

```
## [1] FALSE
```

```r
"M022030" %in% rownames(kas)
```

```
## [1] FALSE
```

```r
"M022031" %in% rownames(kas)
```

```
## [1] FALSE
```

```r
"M022032" %in% rownames(kas)
```

```
## [1] FALSE
```

```r
"M022095" %in% rownames(kas)
```

```
## [1] FALSE
```

```r
"M022096" %in% rownames(kas) #YES
```

```
## [1] TRUE
```

```r
"M022097" %in% rownames(kas)
```

```
## [1] FALSE
```

```r
"M022098" %in% rownames(kas) #YES
```
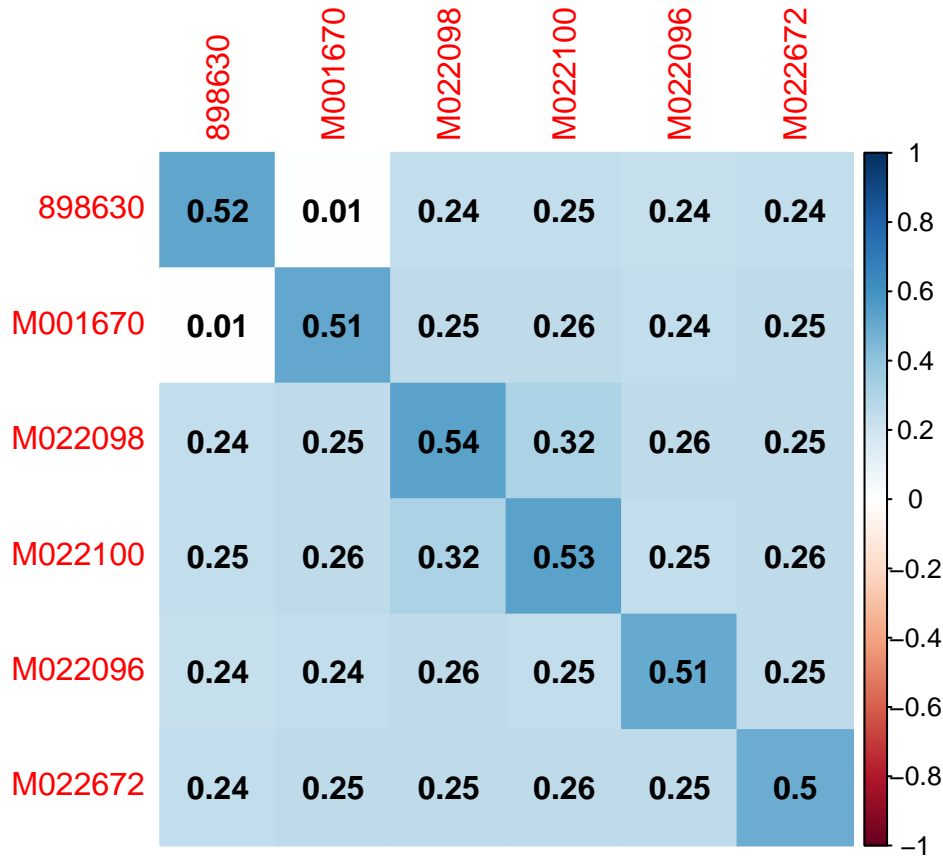
```
## [1] TRUE
```

```r
"M022100" %in% rownames(kas) #YES
```

```
## [1] TRUE
```

We sequenced three of his siblings, now we just need to check it's genomic kinship with both parents and siblings and see it it matches the expected values:

```
#Check relatedness
corrplot(as.matrix(kas[row.names(kas) %in% c("M022672", "M001670", "898630",
                                              "M022096", "M022098", "M022100"),
                        row.names(kas) %in% c("M022672", "M001670", "898630",
                                              "M022096", "M022098", "M022100")]),
         method = 'color',addCoef.col = 'black')
```

	898630	M001670	M022098	M022100	M022096	M022672
898630	0.52	0.01	0.24	0.25	0.24	0.24
M001670	0.01	0.51	0.25	0.26	0.24	0.25
M022098	0.24	0.25	0.54	0.32	0.26	0.25
M022100	0.25	0.26	0.32	0.53	0.25	0.26
M022096	0.24	0.24	0.26	0.25	0.51	0.25
M022672	0.24	0.25	0.25	0.26	0.25	0.5

The genomic kinship looks good (it matches the expected kinship around 0.25). So it is not a individual problem, he's just not the father ! We'll keep this guy but try to correct the pedigree by looking for the real father:

```
listofFAM = c(high$RingId1[16:19])
```

```
#check for related individuals for the first kid
row.names(kas)[which(kas[high$row[16],] > 0.16)]
```

```
##  [1] "M022673" "M032022" "M032023" "M032024" "M031529" "M031576" "M031577"
##  [8] "M031578" "M031579" "M031998" "M031999" "M032000" "M037885" "M038001"
```

```
#correlation plot
corrplot(kas[which(kas[high$row[16],] > 0.16),
             which(kas[high$row[16],] > 0.16)],
         method = 'color',addCoef.col = 'black', number.cex = .7)
```

	M022673	M032022	M032023	M032024	M031529	M031576	M031577	M031578	M031579	M031998	M031999	M032000	M037885	M038001
M022673	0.54	0.39	0.37	0.38	0.01	0.26	0.26	0.25	0.26	0.26	0.26	0.26	0.01	0.26
M032022	0.39	0.66	0.38	0.38	0.01	0.18	0.15	0.17	0.21	0.17	0.22	0.19	0.01	0.23
M032023	0.37	0.38	0.63	0.34	0.01	0.2	0.2	0.21	0.16	0.19	0.18	0.2	0.01	0.21
M032024	0.38	0.38	0.34	0.63	0	0.17	0.16	0.16	0.2	0.17	0.17	0.21	−0.01	0.19
M031529	0.01	0.01	0.01	0	0.5	0.25	0.25	0.24	0.24	0.24	0.25	0.25	0.25	0.25
M031576	0.26	0.18	0.2	0.17	0.25	0.5	0.27	0.27	0.25	0.22	0.25	0.27	0.17	0.21
M031577	0.26	0.15	0.2	0.16	0.25	0.27	0.5	0.27	0.25	0.24	0.25	0.26	0.09	0.25
M031578	0.25	0.17	0.21	0.16	0.24	0.27	0.27	0.49	0.26	0.28	0.22	0.23	0.11	0.26
M031579	0.26	0.21	0.16	0.2	0.24	0.25	0.25	0.26	0.5	0.22	0.28	0.24	0.13	0.24
M031998	0.26	0.17	0.19	0.17	0.24	0.22	0.24	0.28	0.22	0.49	0.22	0.21	0.11	0.27
M031999	0.26	0.22	0.18	0.17	0.25	0.25	0.25	0.22	0.28	0.22	0.5	0.27	0.15	0.28
M032000	0.26	0.19	0.2	0.21	0.25	0.27	0.26	0.23	0.24	0.21	0.27	0.5	0.12	0.24
M037885	0.01	0.01	0.01	−0.01	0.25	0.17	0.09	0.11	0.13	0.11	0.15	0.12	0.5	0.12
M038001	0.26	0.23	0.21	0.19	0.25	0.21	0.25	0.26	0.24	0.27	0.28	0.24	0.12	0.5

Ok so this group of individual is mess but we have (in the upper left triangle of the matrix) one inbred family: M022673 is the dad of three inbred kids: M032022, M032023 and M032024. It looks like he is also the dad of the initial clutch we were studying including kids: M031576, M031577, M031578 and M031579 as well as the dad of another clutch including kids: M031998, M031999, M032000. The "wrong" dad in the initial clutch we were studying had a RingId of M022672, one number away of the "true" dad we identified (M022673). We we believe that this is a typo problem and that we mis-typed the dad ID when registering the clutch into the database.

So we kept all the individuals and the pedigree with new father ID: M022673.

group 10

The next group contains four individual: one mom: M038109 and her three kids M022941, M022942, M022943. As before, the pedigree kinship is 0.25 but the genetic kinship is close to zero indicating that the link is wrong. We'll start by looking at the genomic kinship matrix.

```
listofFAM = c(high$RingId1[20],high$RingId2[20:22])

#corplot
corrplot(as.matrix(kas[row.names(kas) %in% listofFAM,
                   row.names(kas) %in% listofFAM ]),
     method = 'color',addCoef.col = 'black')
```

Just like before, the female is not related to any of the three offspring, but the offspring are related among themselves, as siblings (close to 0.25). We are not sure if the individuals we sequenced is not who we think she is or if she's just not the mom. So we'll look for siblings of this female:

```r
#Who are his parents ?
ped[ped$animal == high$RingId1[20],]
```

```
##       animal sire  dam sex
## 8988 M038109 <NA> <NA>   2
```
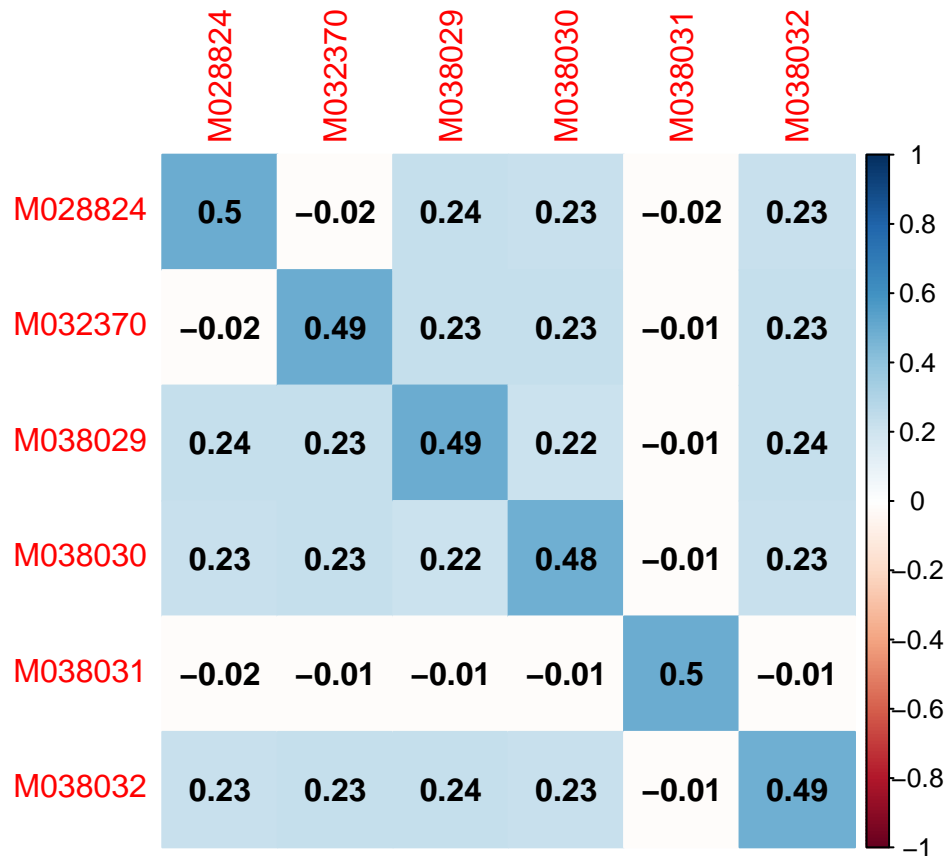
We don't know the parents so we cannot look for siblings. However, in the database, there is something about this clutch mentioning that the ring ID of the female was wrongly copied . . . What it means is that the DNA we sequenced comes from the correct individual, she is just not the mother of this clutch. So we fixed the pedigree with correct mom ID: M038110.

group 11

The next group is composed of individuals M038031 (kid) and it's mom: M028824; dad: M032370 and three siblings: M038029, M038030 and M038032. The pedigree kinship is around 0.25 but the genetic kinship are close to 0 so once again, the link is wrong! As usual, we'll strat by checking the genomic kinship:

```r
listofFAM = c(high$RingId1[c(33,56)],high$RingId2[c(33,49,54,55)])

#corplot
corrplot(as.matrix(kas[row.names(kas) %in% listofFAM,
                  row.names(kas) %in% listofFAM ]),
        method = 'color',addCoef.col = 'black')
```

	M028824	M032370	M038029	M038030	M038031	M038032
M028824	0.5	–0.02	0.24	0.23	–0.02	0.23
M032370	–0.02	0.49	0.23	0.23	–0.01	0.23
M038029	0.24	0.23	0.49	0.22	–0.01	0.24
M038030	0.23	0.23	0.22	0.48	–0.01	0.23
M038031	–0.02	–0.01	–0.01	–0.01	0.5	–0.01
M038032	0.23	0.23	0.24	0.23	–0.01	0.49

So he's definitively not in the family ! But does he have new links with other individuals ?

```r
#check for related individuals
row.names(kas)[which(kas[high$row[33],] > 0.16)]
```

```
## [1] "M038031" "M040503" "M040505"
```

```r
#who are they
ped[ped$animal == "M040503",]
```

```
##        animal    sire     dam sex
## 8178 M040503 M028912 M038112   2
```

```r
ped[ped$animal == "M040505",]
```

```
##        animal    sire     dam sex
## 8180 M040505 M028912 M038112   2
```

```r
#correlation plot
corrplot(kas[rownames(kas) %in% c("M038031","M040503","M040505"),
             colnames(kas) %in% c("M038031","M040503","M040505")],
         method = 'color',addCoef.col = 'black')
```
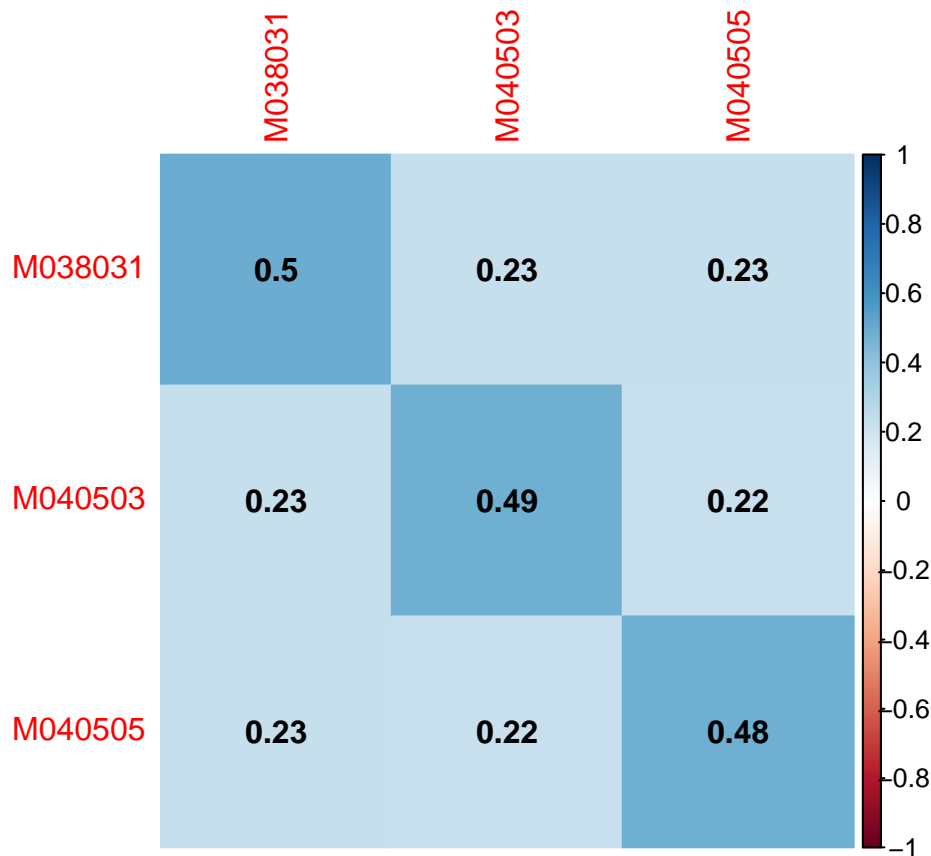
We identified two individuals which are siblings and they have a genomic kinship around 0.25 with M038031. In fact, these two guys also have a "kinship problem" with their supposed mom (M038112): the pedigree kinship is around 0.25 but the genomic kinship is close to 0.

```r
#/!\ we print the DIFFERENCE in kinship between pedigree and genomic here
diff_kin[rownames(diff_kin) %in% c("M040503","M040505","M038112"),
         colnames(diff_kin) %in% c("M040503","M040505","M038112")]
```

```
##              M038112    M040503 M040505
## M038112           NA         NA      NA
## M040503    0.2568685         NA      NA
## M040505    0.2581563 0.02546257      NA
```
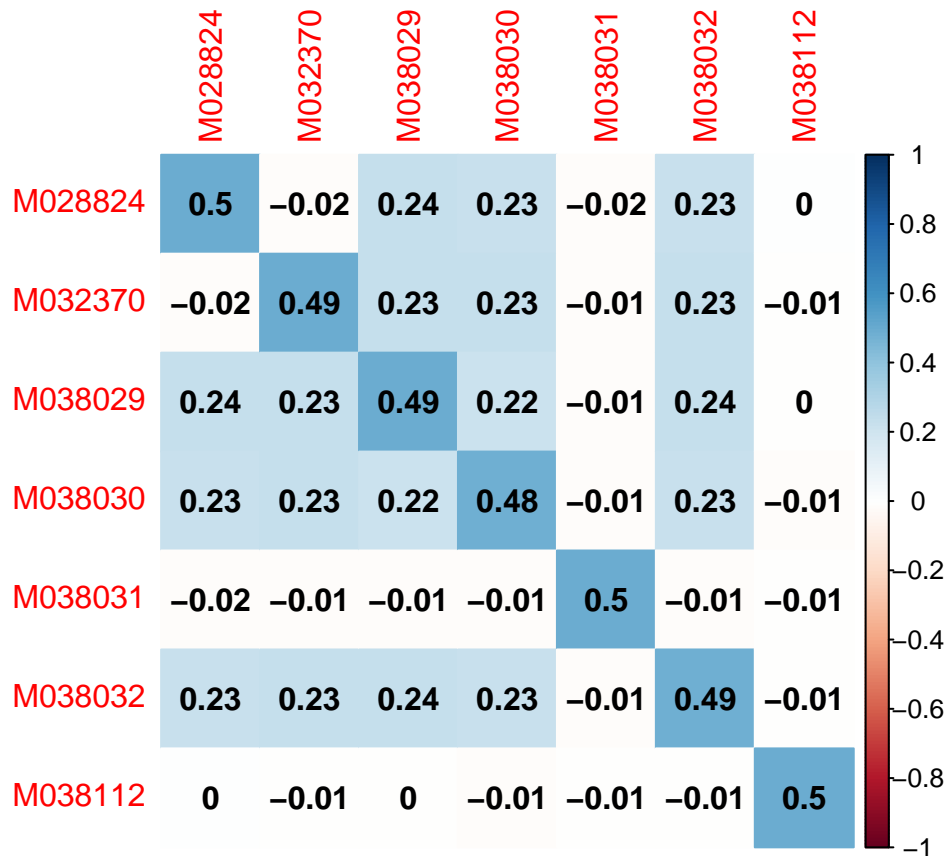
```r
#correlation plot
corrplot(kas[rownames(kas) %in% c("M040503","M040505","M038112"),
             colnames(kas) %in% c("M040503","M040505","M038112")],
         method = 'color',addCoef.col = 'black')
```
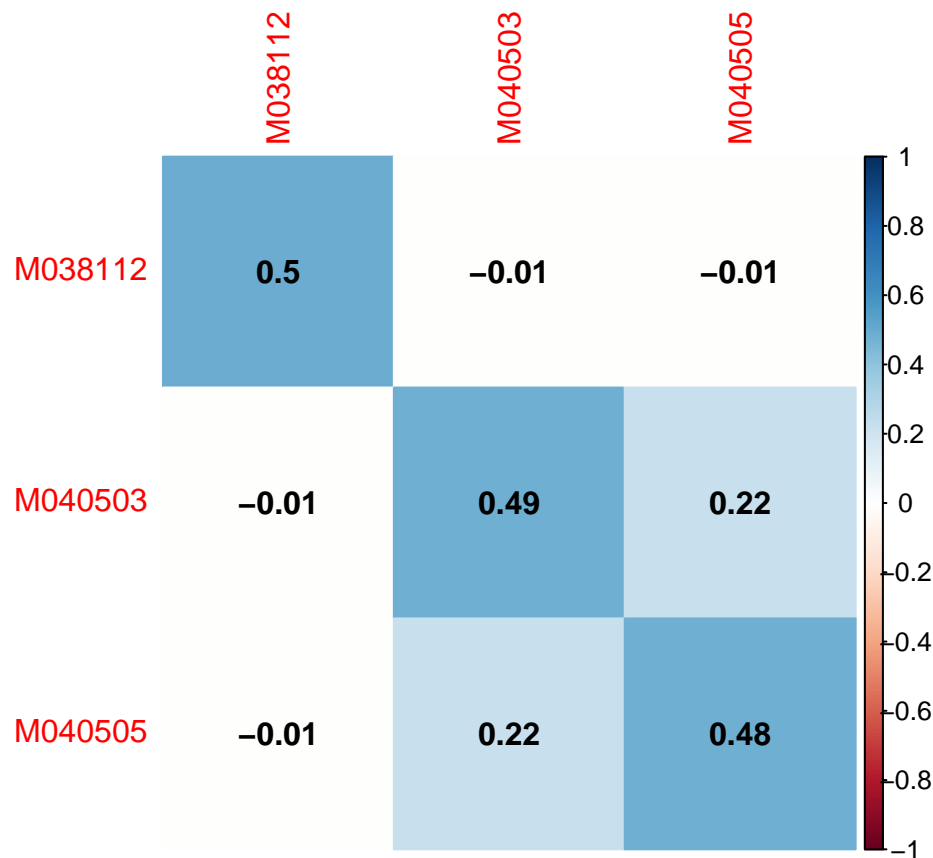
So we want to know whether there has been a sample swap between M038031 and M038112. We'll check if M038112 is related to the first family we were checking:

```r
#correlation plot
corrplot(as.matrix(kas[row.names(kas) %in% c("M038112",listofFAM),
                       row.names(kas) %in% c("M038112",listofFAM)]),
         method = 'color',addCoef.col = 'black')
```

	M028824	M032370	M038029	M038030	M038031	M038032	M038112
M028824	0.5	–0.02	0.24	0.23	–0.02	0.23	0
M032370	–0.02	0.49	0.23	0.23	–0.01	0.23	–0.01
M038029	0.24	0.23	0.49	0.22	–0.01	0.24	0
M038030	0.23	0.23	0.22	0.48	–0.01	0.23	–0.01
M038031	–0.02	–0.01	–0.01	–0.01	0.5	–0.01	–0.01
M038032	0.23	0.23	0.24	0.23	–0.01	0.49	–0.01
M038112	0	–0.01	0	–0.01	–0.01	–0.01	0.5

```
#She is not related to the first family we were investigating
# so it's NOT a tube swap ...
```

M038112 is not related to the first family either so it is not a tube/sample swap . . . We'll try to look for individuals with whom M038112 is related and investigate who they are:

```
#Who is she related to ?
row.names(kas)[which(kas[high$col[59],] > 0.16)]
```

```
## [1] "M031557" "M031558" "M038033" "M038034" "M038112"
```

```
#4 individuals, who are they ?
ped[ped$animal == "M031557",]
```

```
##        animal sire  dam sex
## 7920 M031557 <NA> <NA>   2
```

```
ped[ped$animal == "M031558",]
```

```
##        animal sire  dam sex
## 7921 M031558 <NA> <NA>   1
```

```
ped[ped$animal == "M038033",]
```

```
##        animal    sire     dam sex
## 8692 M038033 M031558 M031557   1
```
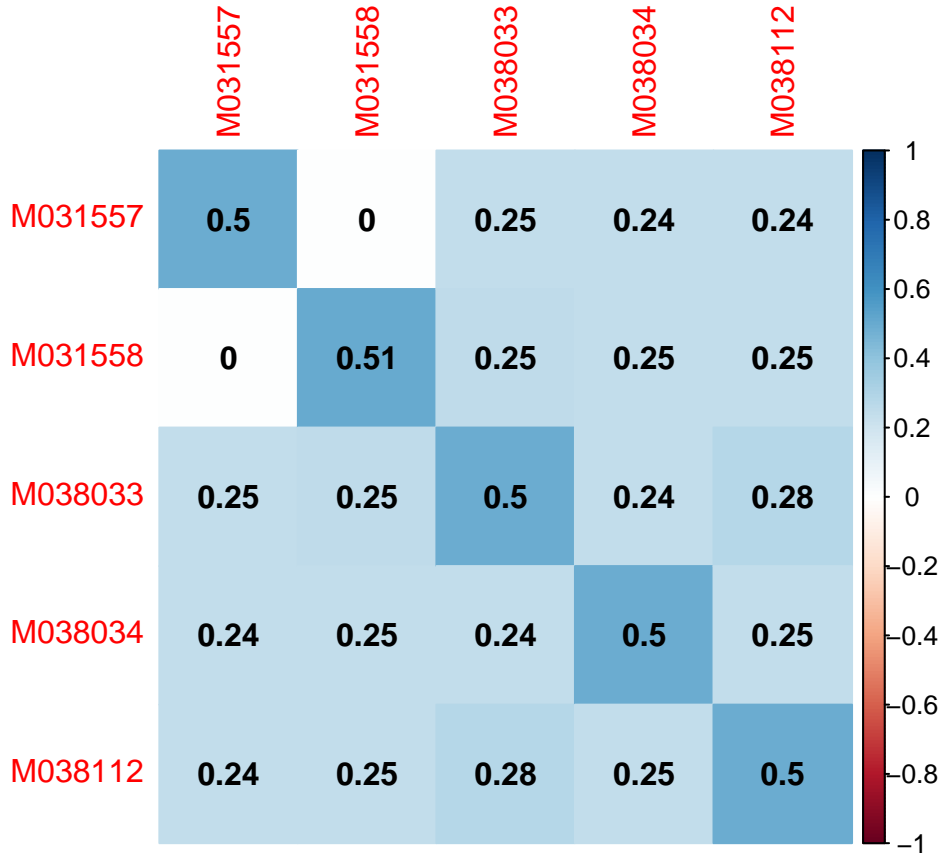
```
ped[ped$animal == "M038034",]
```

```
##        animal    sire     dam sex
## 8693 M038034 M031558 M031557   2
```

```
corrplot(kas[rownames(kas) %in% c("M038112","M031557", "M031558",
                                   "M038033", "M038034"),
           colnames(kas) %in% c("M038112","M031557", "M031558",
                                 "M038033", "M038034")],
         method = 'color',addCoef.col = 'black')
```


M038112 is further related to two parents and two of their offspring. We will now look for additional members of this family and see whether we sequenced them:

```
#Any one else in the family ?
ped[(ped$sire == "M031558") & (ped$dam == "M031557") &
    (!is.na(ped$sire)) & (!is.na(ped$dam)),]
```

```
##        animal    sire     dam sex
## 7246 M037517 M031558 M031557   2
## 7579 M031704 M031558 M031557   1
## 8692 M038033 M031558 M031557   1
## 8693 M038034 M031558 M031557   2
## 8694 M038035 M031558 M031557   1
## 8695 M038036 M031558 M031557   1
## 8696 M038037 M031558 M031557   1
```

```
#five other siblings, did we sequence the ?
"M037517" %in% rownames(kas)
```

```
## [1] FALSE
```

```
"M031704" %in% rownames(kas)
```

```
## [1] FALSE
```
```
"M038035" %in% rownames(kas)
```

```
## [1] TRUE
```
```
"M038036" %in% rownames(kas)
```
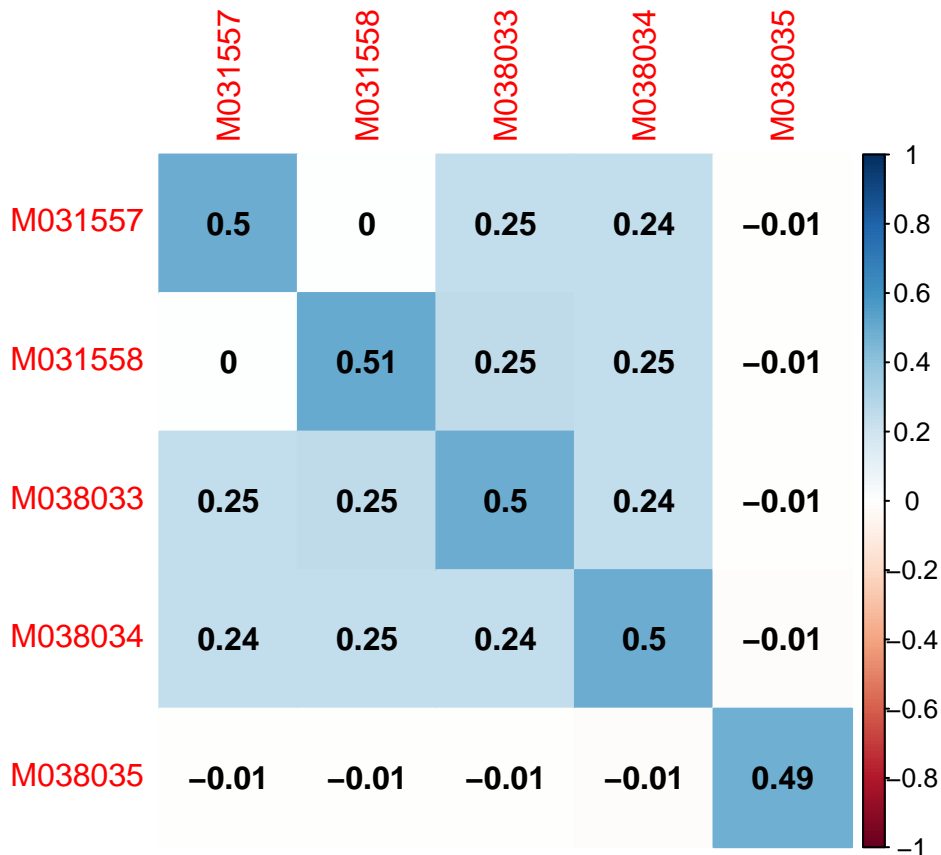
```
## [1] FALSE
```
```
"M038037" %in% rownames(kas)
```
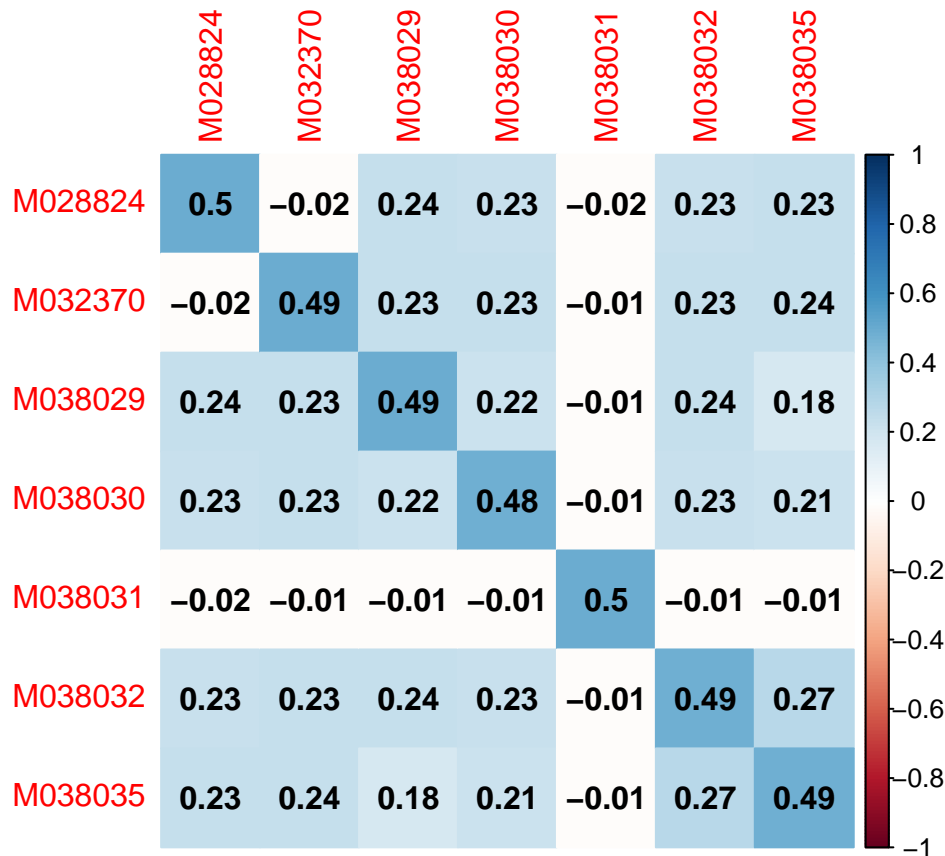
```
## [1] FALSE
```

We sequenced one extra offspring from this family: M038035. How does he correlate with the rest of his supposed family?

```
#correlation plot
corrplot(as.matrix(kas[row.names(kas) %in% c("M031557","M031558","M038033",
                                             "M038034","M038035"),
                      row.names(kas) %in% c("M031557","M031558","M038033",
                                             "M038034","M038035")]),
         method = 'color', addCoef.col = 'black')
```


This guy is not related to the family it should be related to! But what about the two previous families?

```
#correlation plot first family (supposedly M038031 family):
corrplot(as.matrix(kas[row.names(kas) %in% c("M038035",listofFAM),
                      row.names(kas) %in% c("M038035",listofFAM)]),
         method = 'color',addCoef.col = 'black')
```

	M028824	M032370	M038029	M038030	M038031	M038032	M038035
M028824	**0.5**	**−0.02**	**0.24**	**0.23**	**−0.02**	**0.23**	**0.23**
M032370	**−0.02**	**0.49**	**0.23**	**0.23**	**−0.01**	**0.23**	**0.24**
M038029	**0.24**	**0.23**	**0.49**	**0.22**	**−0.01**	**0.24**	**0.18**
M038030	**0.23**	**0.23**	**0.22**	**0.48**	**−0.01**	**0.23**	**0.21**
M038031	**−0.02**	**−0.01**	**−0.01**	**−0.01**	**0.5**	**−0.01**	**−0.01**
M038032	**0.23**	**0.23**	**0.24**	**0.23**	**−0.01**	**0.49**	**0.27**
M038035	**0.23**	**0.24**	**0.18**	**0.21**	**−0.01**	**0.27**	**0.49**

He's correlating with the first family (supposedly M038031 family). So we beleive that it is a three samples swap: M038035 is M038031; M038112 is M038035 and M038031 is M038112! If we look at the "corrected" family correlation plots:
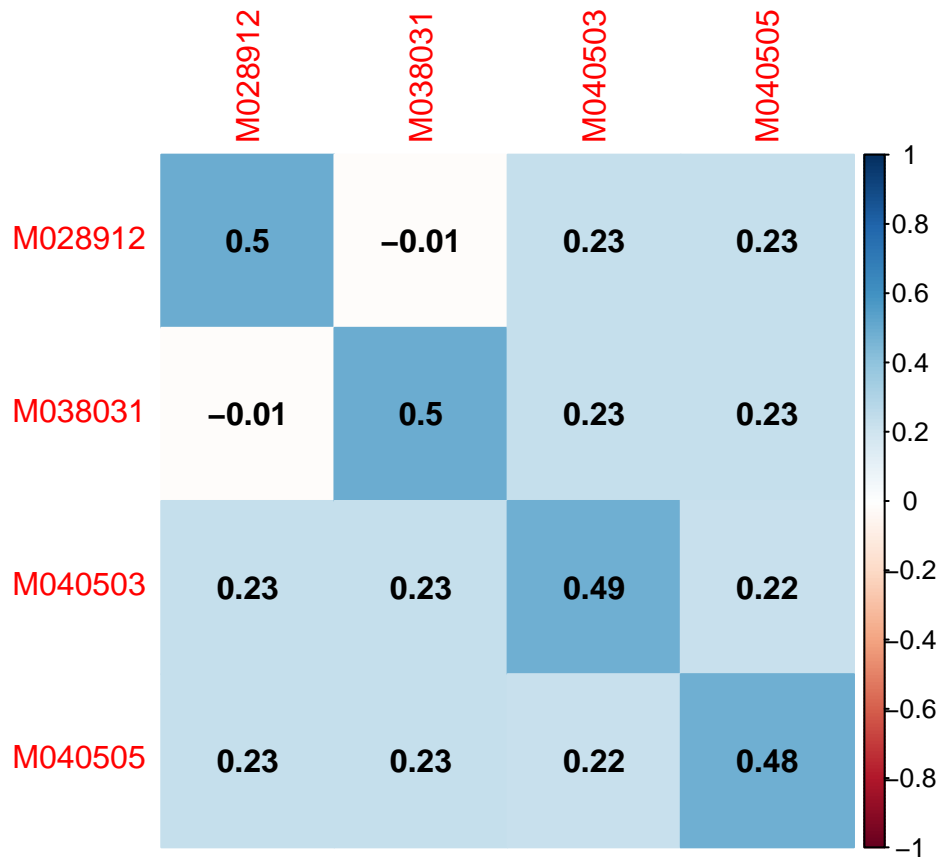
```r
#First family
corrplot(as.matrix(kas[row.names(kas) %in% c("M038035","M038032","M028824",
                                             "M032370","M038029","M038030"),
                   row.names(kas) %in% c("M038035","M038032","M028824",
                                         "M032370","M038029","M038030")]),
         method = 'color',addCoef.col = 'black')
```
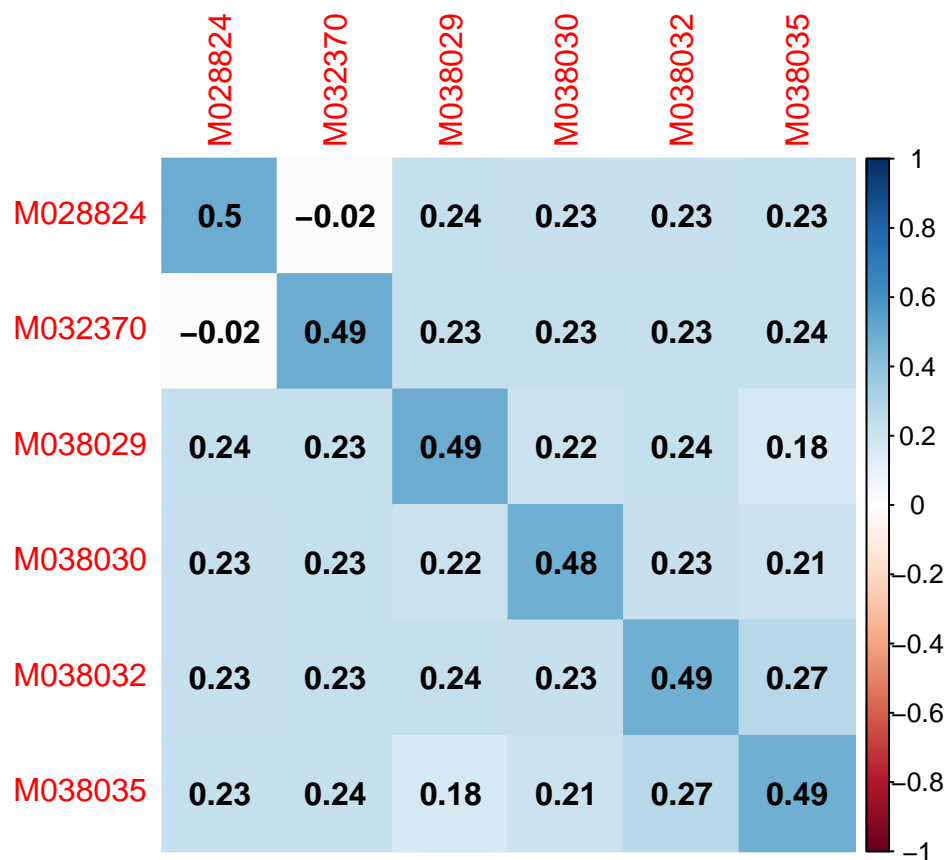
```
#good

#Second family
corrplot(as.matrix(kas[row.names(kas) %in% c("M040503","M040505","M038031","M028912"),
                      row.names(kas) %in% c("M040503","M040505","M038031","M028912")]),
        method = 'color',addCoef.col = 'black')
```

```
#good

#Third family
corrplot(as.matrix(kas[row.names(kas) %in% c("M038035","M038032","M028824",
                                    "M032370","M038029","M038030"),
                        row.names(kas) %in% c("M038035","M038032","M028824",
                                    "M032370","M038029","M038030")]),
         method = 'color',addCoef.col = 'black')
```
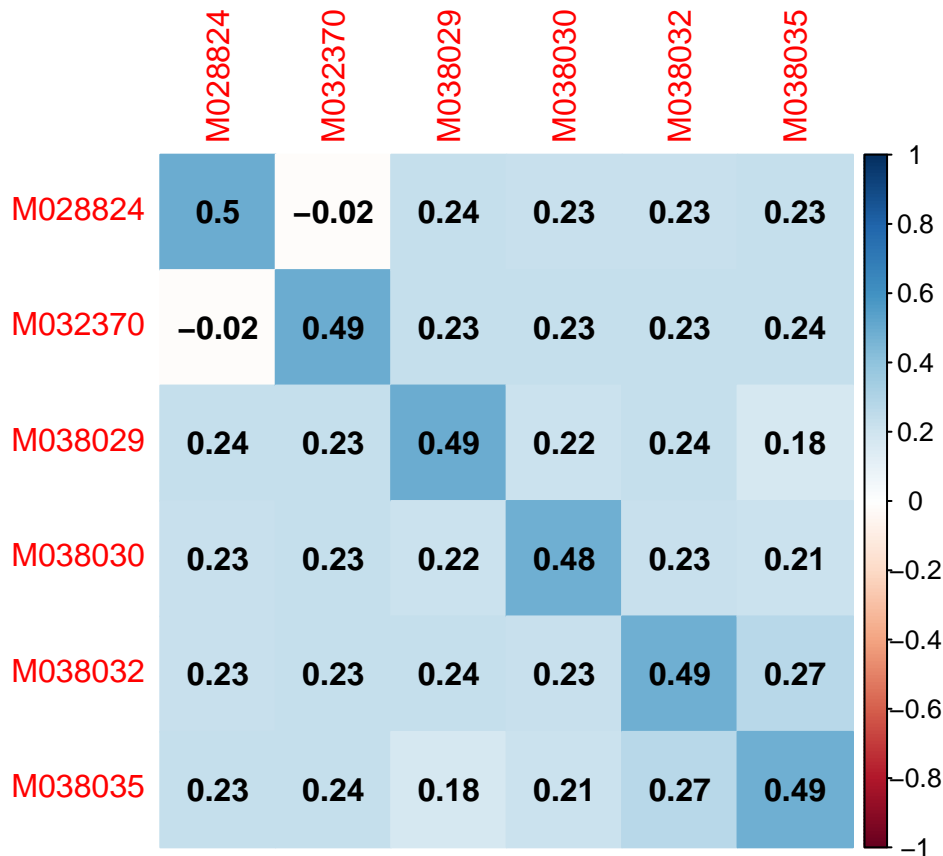
```
#good
```

So we conclude that this is a three samples swap: M038035 actually is M038031; M038112 actually is M038035 and M038031 actually is M038112. So we'll correct the individuals names in the VCF.

group 12

The next group is composed of three individuals: a female: M031570 and her supposed dad: M032313 and sibling: M031938. As usual, let's first check their genomic kinship:

```
listofFAM = c(high$RingId1[c(36,37)],high$RingId2[36])

#corplot
corrplot(as.matrix(kas[row.names(kas) %in% listofFAM,
                       row.names(kas) %in% listofFAM ]),
         method = 'color',addCoef.col = 'black')
```
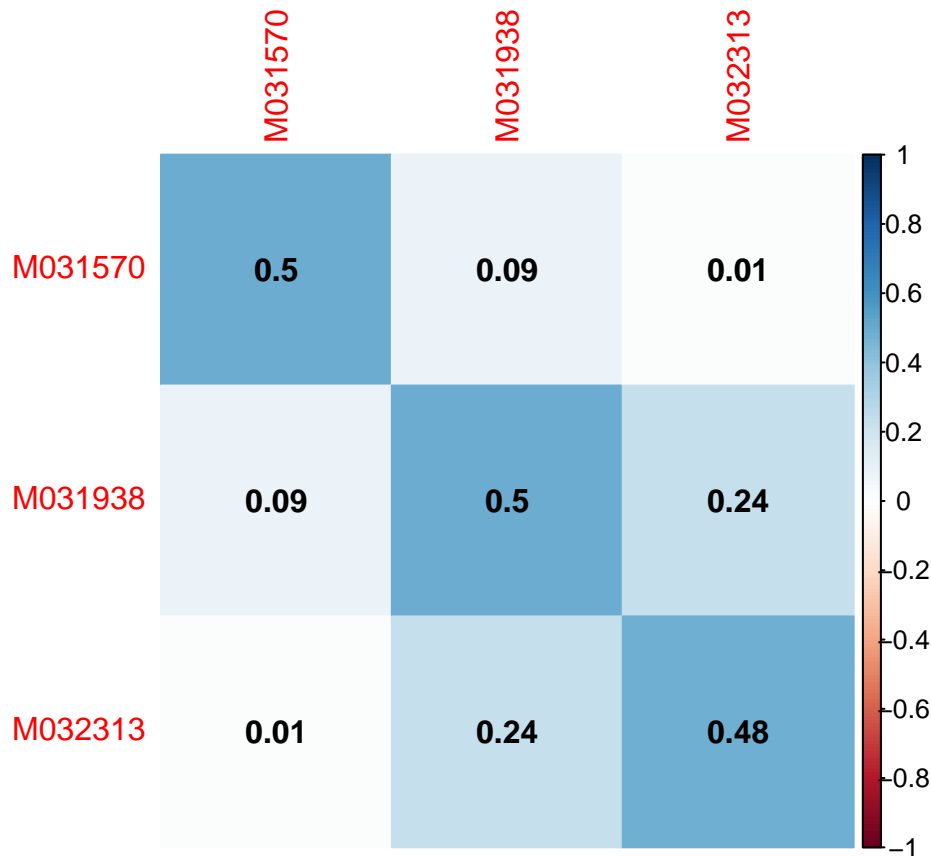
M031570 is lowly related to her sibling (genomic kinship close to 0.1) and not related to her dad. Let's see if she is related to anyone else we sequenced:

```r
#Who is she related to ?
row.names(kas)[which(kas[high$col[36],] > 0.16)]
```

```
## [1] "M031570" "M032426"
```

```r
#parents ?
ped[ped$animal == high$RingId2[36],]
```

```
##       animal    sire      dam sex
## 7462 M031570 M032313 M032426   2
```

She is related to another individual which is actually her mom (M032426). First she is not from the same clutch as her supposed sibling (M031938) so it is possible that they don't have the same father. Well now add the mom and check relatedness:

```r
listofFAM = c(high$RingId1[c(36,37)],high$RingId2[36], "M032426")

#corplot
corrplot(as.matrix(kas[row.names(kas) %in% listofFAM,
                  row.names(kas) %in% listofFAM ]),
       method = 'color',addCoef.col = 'black')
```
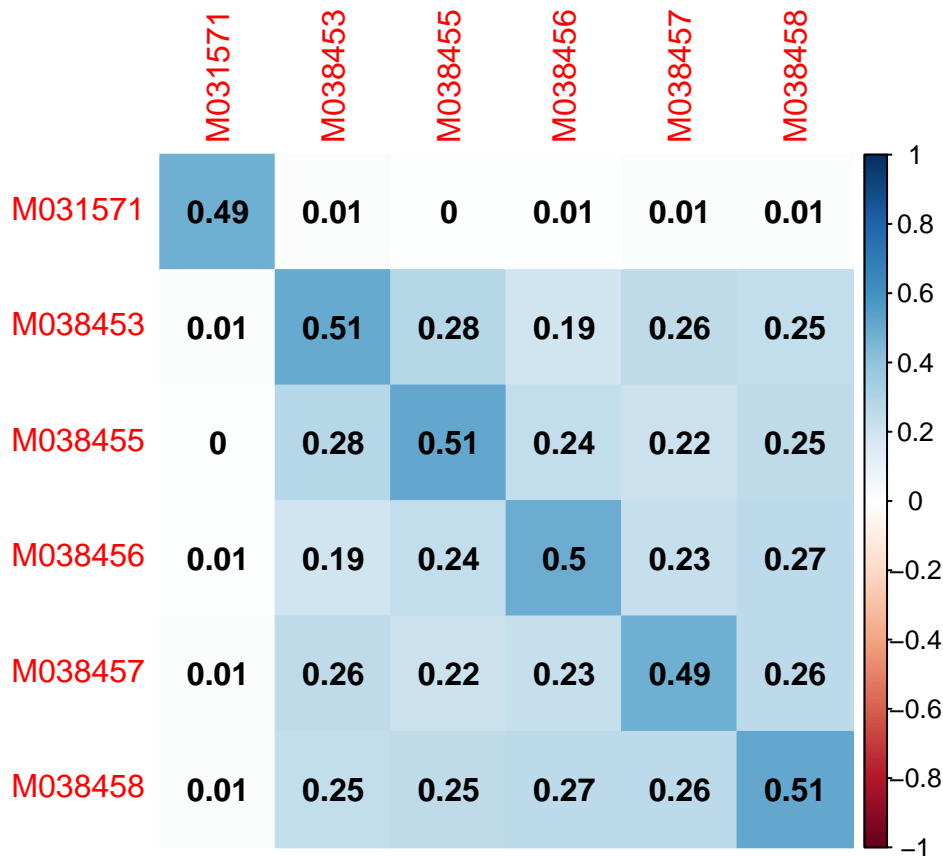
58
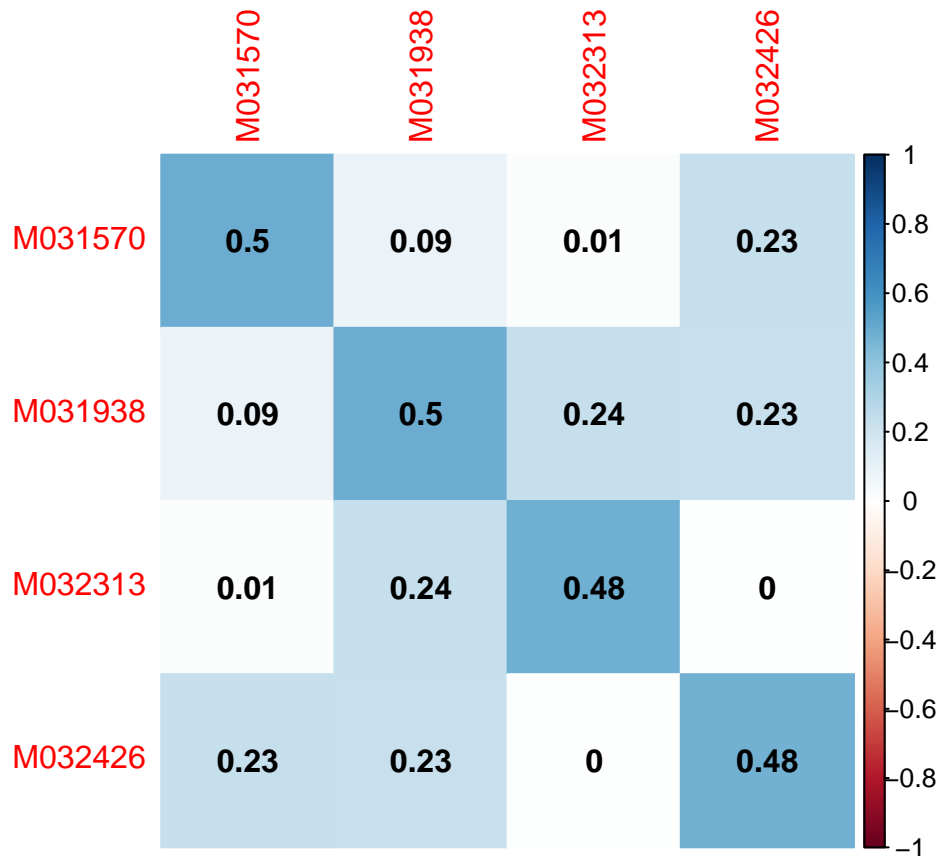
The mom is related to both of her offspring. We'll conclude that we sequenced the correct sample but that her dad was wrongly assigned and correct the pedigree.

group 13

The next group is composed of six individuals: a female: M031571 and her supposed offspring: M038453, M038455, M038456, M038457, M038458. Let's look at the genomic kinship:

```
listofFAM = c(high$RingId2[38],high$RingId1[38:42])

#corplot
corrplot(as.matrix(kas[row.names(kas) %in% listofFAM,
                    row.names(kas) %in% listofFAM ]),
        method = 'color',addCoef.col = 'black')
```

	M031571	M038453	M038455	M038456	M038457	M038458
M031571	0.49	0.01	0	0.01	0.01	0.01
M038453	0.01	0.51	0.28	0.19	0.26	0.25
M038455	0	0.28	0.51	0.24	0.22	0.25
M038456	0.01	0.19	0.24	0.5	0.23	0.27
M038457	0.01	0.26	0.22	0.23	0.49	0.26
M038458	0.01	0.25	0.25	0.27	0.26	0.51

All the siblings are correctly related so we can be sure about them. Now, we need to check if the RingID corresponds to the one we sequenced. We'll check whether she (M031571) has other links which could confirm her identity:

```r
#Who is she related to ?
row.names(kas)[which(kas[high$col[38],] > 0.16)]
```

```
## [1] "M031564" "M031571" "M031811" "M031812" "M031813"
```

```r
#who are they =
ped[ped$animal == "M031564",] #unknown
```

```
##        animal sire  dam sex
## 7922 M031564 <NA> <NA>   2
```

```r
ped[ped$animal == "M031811",] #kids
```

```
##        animal    sire     dam sex
## 6892 M031811 M032150 M031571   2
```

```r
ped[ped$animal == "M031812",] #kids
```

```
##        animal    sire     dam sex
## 6893 M031812 M032150 M031571   2
```

```r
ped[ped$animal == "M031813",] #kids
```

```
##        animal    sire     dam sex
## 6894 M031813 M032150 M031571   2
```

We identified three of her kids from another clutch! Let's check her kinship with them and if it's around 0.25

then we can conclude that she truly is who we think she is:

```r
listofFAM = c(high$RingId2[38], "M031811", "M031812", "M031813")

#corplot
corrplot(as.matrix(kas[row.names(kas) %in% listofFAM,
                       row.names(kas) %in% listofFAM ]),
         method = 'color',addCoef.col = 'black')
```
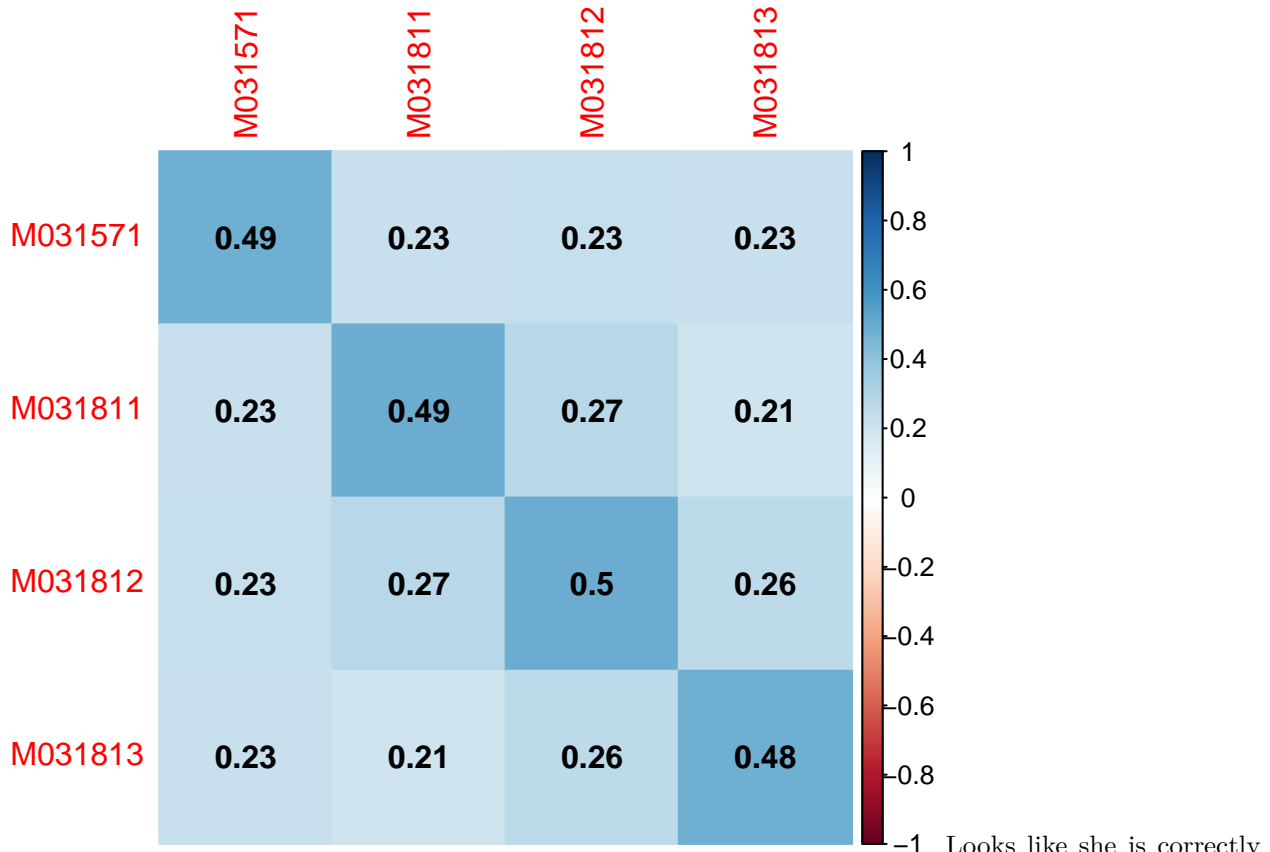

Looks like she is correctly related to her other kids! Now who is the mother of the first clutch we were investigating? We'll try to find related individuals in the data:

```r
#Who is she related to ?
row.names(kas)[which(kas[high$row[38],] > 0.16)]
```

```
## [1] "M031150" "M038453" "M038455" "M038456" "M038457" "M038458" "M042765"
```

```r
ped[ped$animal == "M031150",]
```

```
##       animal sire  dam sex
## 7414 M031150 <NA> <NA>   2
```
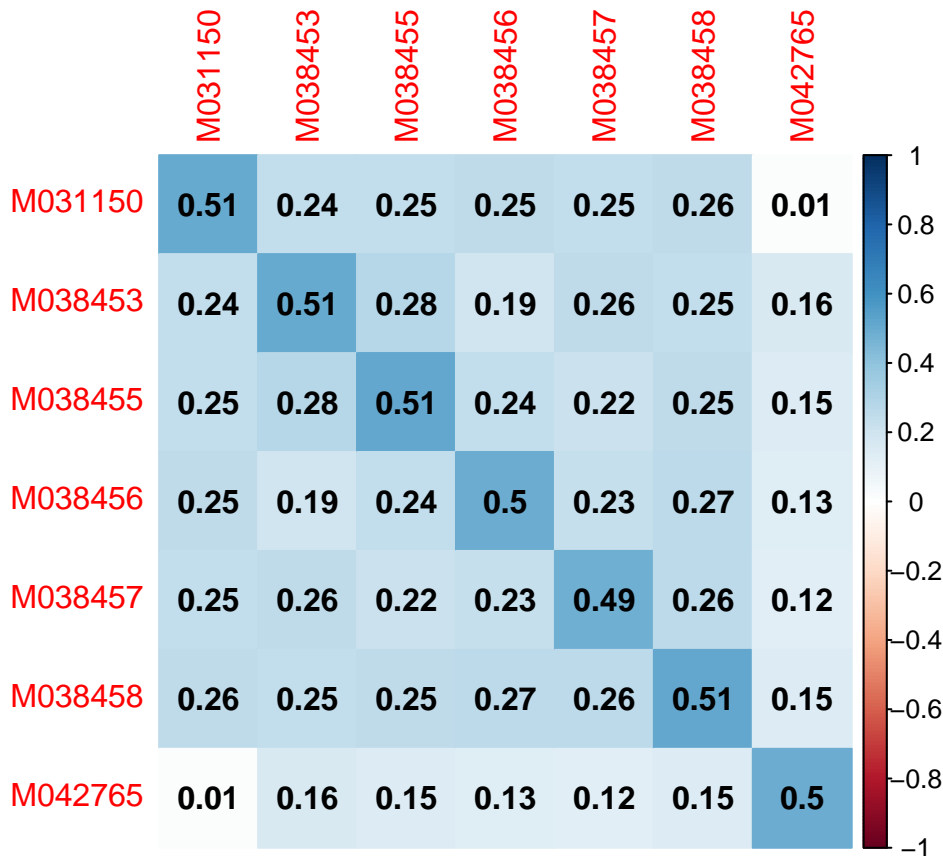
```r
ped[ped$animal == "M042765",]
```

```
##       animal sire  dam sex
## 8274 M042765 <NA> <NA>   1
```

We find the siblings plus two individuals: we'll see the genomic kinship matrix:

```r
listofFAM = c(high$RingId1[38:42], "M031150", "M042765")
```

```
#correlation plot
corrplot(as.matrix(kas[row.names(kas) %in% listofFAM,
                       row.names(kas) %in% listofFAM ]),
         method = 'color',addCoef.col = 'black')
```

	M031150	M038453	M038455	M038456	M038457	M038458	M042765
M031150	0.51	0.24	0.25	0.25	0.25	0.26	0.01
M038453	0.24	0.51	0.28	0.19	0.26	0.25	0.16
M038455	0.25	0.28	0.51	0.24	0.22	0.25	0.15
M038456	0.25	0.19	0.24	0.5	0.23	0.27	0.13
M038457	0.25	0.26	0.22	0.23	0.49	0.26	0.12
M038458	0.26	0.25	0.25	0.27	0.26	0.51	0.15
M042765	0.01	0.16	0.15	0.13	0.12	0.15	0.5

M031150 might be the true mom; M042765 is related but not enough to be the a parent so we won't conclude anything about this individual. We can check whether we sequenced the dad of this clutch:

```
ped[ped$animal == high$RingId1[38],]
```

```
##         animal sire     dam sex
## 8133 M038453 <NA> M031571   1
```

We don't even know who he is so there is really nothing more we can do. We'll just keep all the samples but correct the pedigree!

group 14

The next group is composed of two individuals: a female: M031798 and her supposed kid: M038047. Let's look at the genomic kinship:

```
listofFAM = c(high$RingId2[45], high$RingId1[45])

#corplot
corrplot(as.matrix(kas[row.names(kas) %in% listofFAM,
                       row.names(kas) %in% listofFAM ]),
         method = 'color',addCoef.col = 'black')
```
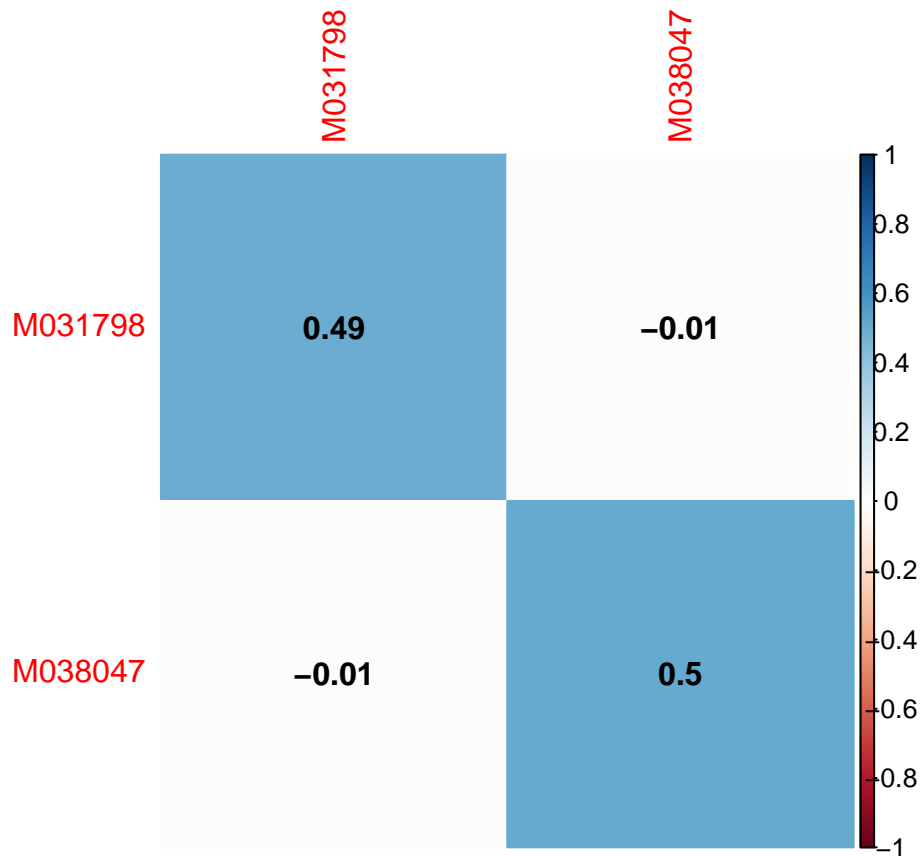
They're not related at all. Let's check if the mom truly is who we think she is. We'll see if the mom has other offspring, if we sequenced them and what their relatedness with her look like:

```
#pedigree
ped[(ped$dam == high$RingId2[45]) & (!is.na(ped$dam)),]
```

```
##        animal    sire     dam sex
## 6872 M031785 M034837 M031798   2
## 6873 M031786 M034837 M031798   2
## 6874 M031787 M034837 M031798   2
## 6875 M031788 M034837 M031798   1
## 6876 M031789 M034837 M031798   2
## 6877 M031790 M034837 M031798   2
## 7415 M031281 M034837 M031798  NA
## 8705 M038047    <NA> M031798   2
## 8706 M038048    <NA> M031798   2
## 8707 M038049    <NA> M031798   1
```

```
#Who are they ?
"M031785" %in% rownames(kas) #yes
```

```
## [1] TRUE
```

```
"M031786" %in% rownames(kas) #yes
```

```
## [1] TRUE
```

```
"M031787" %in% rownames(kas) #yes
```

```
## [1] TRUE
```

```
"M031788" %in% rownames(kas) #yes
```

```
## [1] TRUE
```
```
"M031789" %in% rownames(kas) #yes
```

```
## [1] TRUE
```
```
"M031790" %in% rownames(kas) #yes
```

```
## [1] TRUE
```
```
"M031281" %in% rownames(kas) #no
```

```
## [1] FALSE
```
```
"M038048" %in% rownames(kas) #no
```

```
## [1] FALSE
```
```
"M038049" %in% rownames(kas) #no
```

```
## [1] FALSE
```
```
#correlation plot for the one we sequenced
corrplot(kas[rownames(kas) %in% c("M031785","M031786","M031787","M031788",
                                   "M031789","M031790", high$RingId2[45]),
              colnames(kas) %in% c("M031785","M031786","M031787","M031788",
                                   "M031789","M031790", high$RingId2[45])],
         method = 'color',addCoef.col = 'black')
```


The mom had other offspring we sequenced and which has a relatedness with her around 0.25 so we can be confident that she is
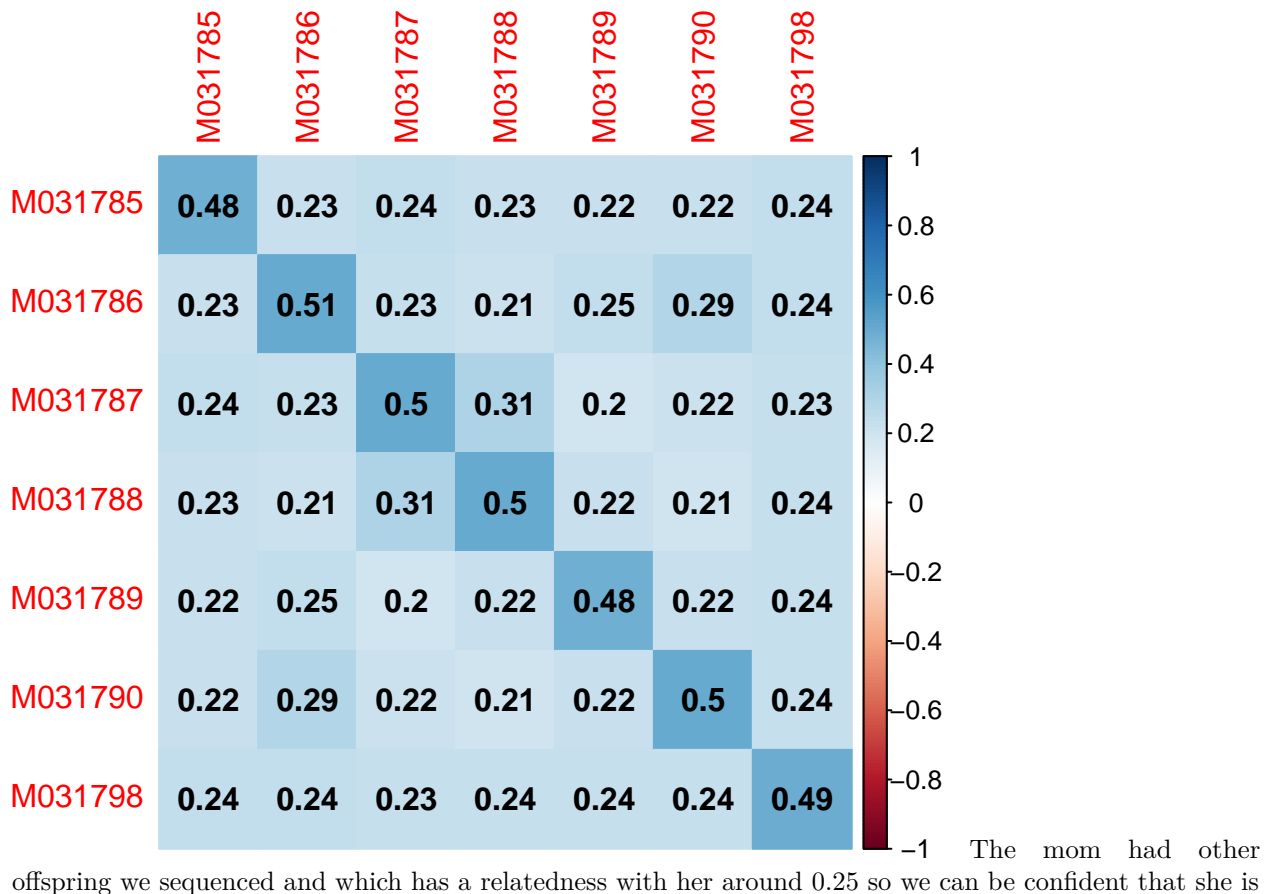
who we think she ism she's just not the mother of M038047.

Let's see if the kid (M038047) has siblings or is related to anyone else which we could use to either confirm his identity of find his parents:

```
#pedigree
ped[ped$animal == high$RingId1[45],]
```

```
##        animal sire     dam sex
## 8705 M038047 <NA> M031798   2
```

```
#new links ?
row.names(kas)[which(kas[high$row[45],] > 0.16)]
```

```
## [1] "M021087" "M031774" "M037858" "M038047"
```

```
#Who are they ?
ped[ped$animal == "M021087",] #unknown
```

```
##        animal sire  dam sex
## 5642 M021087 <NA> <NA>   2
```

```
ped[ped$animal == "M031774",] #unknown
```

```
##        animal sire  dam sex
## 7836 M031774 <NA> <NA>   1
```

```
ped[ped$animal == "M037858",] #unknown
```

```
##        animal sire  dam sex
## 8959 M037858 <NA> <NA>   1
```

```
corrplot(kas[rownames(kas) %in% c("M021087","M031774","M037858","M038047"),
           colnames(kas) %in% c("M021087","M031774","M037858","M038047")],
       method = 'color',addCoef.col = 'black')
```
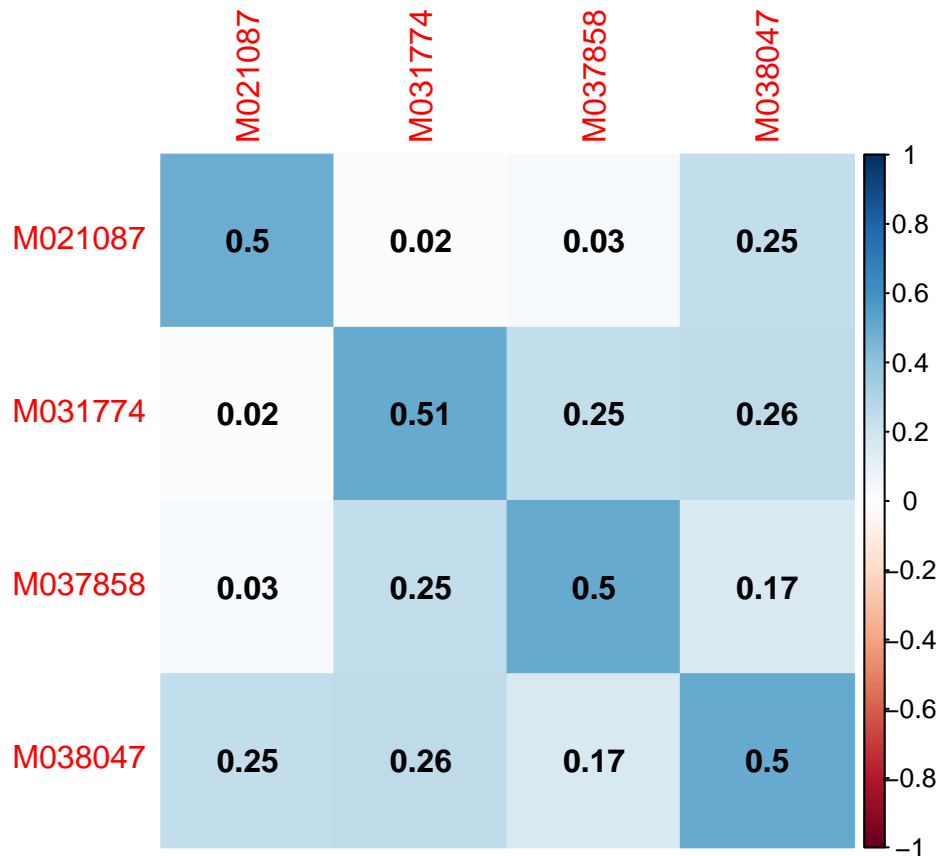
The kid has a bunch of unexpected links. Two unrelated individuals (among themselves) which could be his parents and one individuals with a genomic kinship of 0.17 which is only related to one of the potential parents and could be a half-sibling but we cannot really conclude anything so we'll remove the kid (M038047) from the data set.

group 15

The next group is composed of three individuals: a supposed dad: M032249 and two of his supposed offspring: M037801, M037802. Lets see what the genomic kinship matrix looks like:

```
listofFAM = c(high$RingId2[46], high$RingId1[46:47])

#corplot
corrplot(as.matrix(kas[row.names(kas) %in% listofFAM,
                   row.names(kas) %in% listofFAM ]),
     method = 'color',addCoef.col = 'black')
```
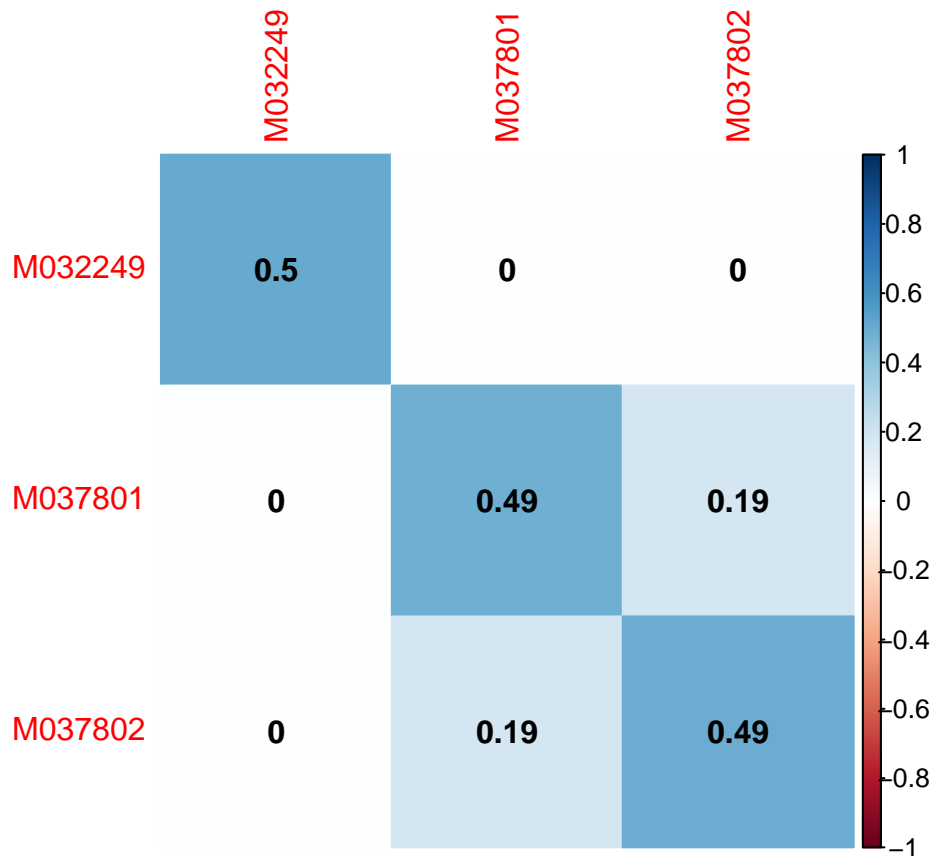
The kids are related with a genomic kinship of 0.19 but not related with the supposed dad (genomic kinship = 0). Let's first see if the dad has had any other offspring which we sequenced and which we could use to confirm his identity:

```
#individuals with the same dad:
ped[(ped$sire == high$RingId2[46]) & (!is.na(ped$sire)),]
```

```
##        animal    sire      dam sex
## 7062 M032275 M032249 M032248  NA
## 7074 M032287 M032249 M022098   1
## 7145 M032363 M032249 M032248   2
## 7146 M032364 M032249 M032248   2
## 7147 M032365 M032249 M032248   1
## 7148 M032366 M032249 M032248   2
## 7786 M032452 M032249 M022098  NA
## 7787 M032453 M032249 M022098   2
## 7788 M032454 M032249 M022098   1
## 7789 M032455 M032249 M022098   2
## 7790 M032456 M032249 M022098  NA
## 7791 M032457 M032249 M022098   1
## 7792 M032458 M032249 M022098   2
## 8510 M037801 M032249 M032248   1
## 8511 M037802 M032249 M032248   2
```

```
#did we sequence them?
"M032275" %in% rownames(kas)
```

```
## [1] FALSE
```

```r
"M032287" %in% rownames(kas)
```

```
## [1] FALSE
```
```r
"M032363" %in% rownames(kas)
```

```
## [1] FALSE
```
```r
"M032364" %in% rownames(kas)
```

```
## [1] TRUE
```
```r
"M032365" %in% rownames(kas)
```

```
## [1] FALSE
```
```r
"M032366" %in% rownames(kas)
```

```
## [1] FALSE
```
```r
"M032452" %in% rownames(kas)
```

```
## [1] FALSE
```
```r
"M032453" %in% rownames(kas)
```

```
## [1] FALSE
```
```r
"M032454" %in% rownames(kas)
```

```
## [1] TRUE
```
```r
"M032455" %in% rownames(kas)
```

```
## [1] FALSE
```

We sequenced two of them: M032454 and M032364. Let's see what their kinship with their dad is:

```r
listofFAM = c(high$RingId2[46], "M032454", "M032364")

#corplot
corrplot(as.matrix(kas[row.names(kas) %in% listofFAM,
                       row.names(kas) %in% listofFAM ]),
         method = 'color', addCoef.col = 'black')
```
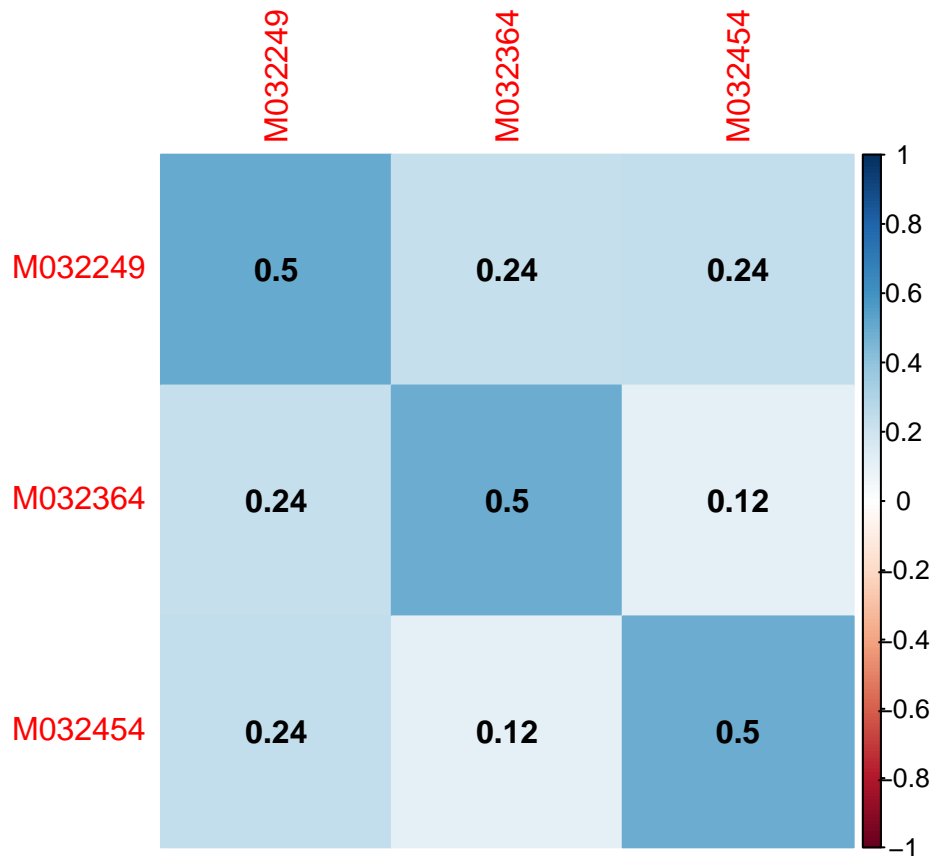
They both have a genomic kinship close to 0.25 with their supposed dad and of 0.12 between themselves which is expected as we know that they are half-siblings. So we can confirm that M032249 truly is himself. Now what about the kids ? We'll first try to see if we sequenced their supposed mom. If we did, we can check their kinship with her:

```
#individuals with the same dad:
ped[(ped$animal == high$RingId1[46]),]
```
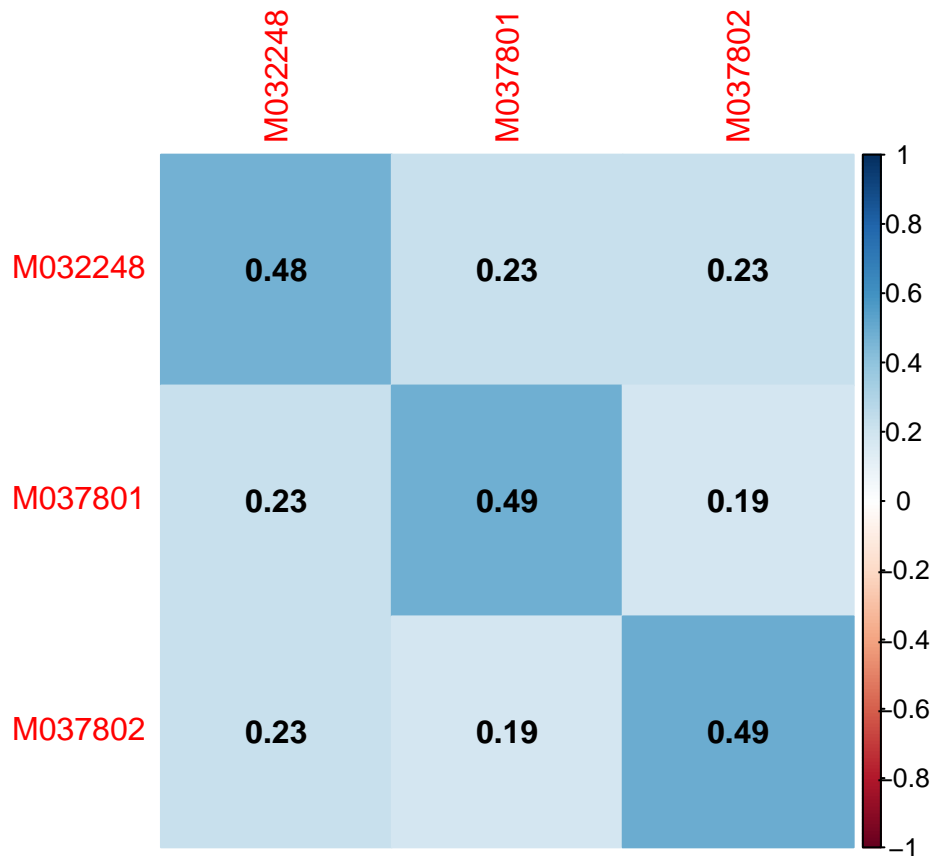
```
##        animal    sire     dam sex
## 8510 M037801 M032249 M032248   1
```

```
#did we sequence the mom?
"M032248" %in% rownames(kas)
```

```
## [1] TRUE
```

```
listofFAM = c(high$RingId1[46:47], "M032248")

#corplot
corrplot(as.matrix(kas[row.names(kas) %in% listofFAM,
                   row.names(kas) %in% listofFAM ]),
       method = 'color', addCoef.col = 'black')
```

They both have a kinship around 0.25 with their mom which is good. Now let's see if we could identify their true father by looking for new links:

```r
#new links ?
row.names(kas)[which(kas[high$row[46],] > 0.16)]
```

```
## [1] "M031665" "M031666" "M031667" "M031668" "M032248" "M032449" "M037801"
## [8] "M037802"
```

```r
#who are they ?
ped[ped$animal == "M031665",] #same mom
```

```
##       animal    sire    dam sex
## 7542 M031665 M032449 M032248   2
```

```r
ped[ped$animal == "M031666",] #same mom
```

```
##       animal    sire    dam sex
## 7543 M031666 M032449 M032248   1
```

```r
ped[ped$animal == "M031667",] #same mom
```

```
##       animal    sire    dam sex
## 7544 M031667 M032449 M032248   1
```

```r
ped[ped$animal == "M031668",] #same mom
```

```
##       animal    sire    dam sex
## 7545 M031668 M032449 M032248   1
```

```r
ped[ped$animal == "M032248",] #it's the mom
```

```
##         animal sire  dam sex
## 7870 M032248 <NA> <NA>   2
```

```r
ped[ped$animal == "M032449",] #unknown
```

```
##         animal    sire     dam sex
## 7783 M032449 M004206 M026849   1
```
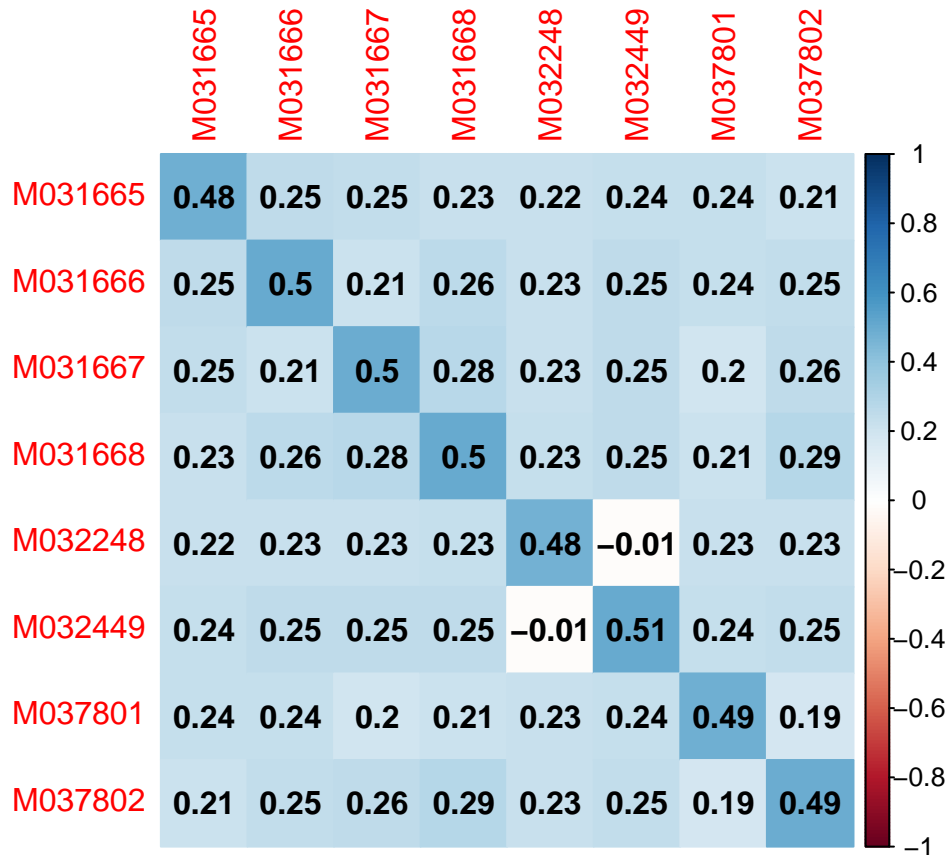
We identified six new individuals: the mom (M032248), four kids (which share the same mom): M031665, M031666, M031667, and M031668 and one extra individual which actually is the father of the four kids we also identified: M032449! What is interesting is that this individual has a RingId VERY close to the supposed father... So this is probably a typo when we entered the clutch father ID... Let's visualize the "corrected kinship matrix":

```r
#individuals
listofFAM = c(high$RingId1[46:47], "M031665", "M031666", "M031667", "M031668", "M032248", "M032449")

#corplot
corrplot(as.matrix(kas[row.names(kas) %in% listofFAM,
                       row.names(kas) %in% listofFAM ]),
         method = 'color', addCoef.col = 'black')
```

|	M031665	M031666	M031667	M031668	M032248	M032449	M037801	M037802
M031665	0.48	0.25	0.25	0.23	0.22	0.24	0.24	0.21
M031666	0.25	0.5	0.21	0.26	0.23	0.25	0.24	0.25
M031667	0.25	0.21	0.5	0.28	0.23	0.25	0.2	0.26
M031668	0.23	0.26	0.28	0.5	0.23	0.25	0.21	0.29
M032248	0.22	0.23	0.23	0.23	0.48	−0.01	0.23	0.23
M032449	0.24	0.25	0.25	0.25	−0.01	0.51	0.24	0.25
M037801	0.24	0.24	0.2	0.21	0.23	0.24	0.49	0.19
M037802	0.21	0.25	0.26	0.29	0.23	0.25	0.19	0.49

The new family looks good. So we'll keep all the samples and correct the father ID!

group 16

The next group is composed of two individuals: a supposed dad: M032348 and one of his supposed offspring: M037684. Let's see what the genomic kinship matrix looks like:

```
listofFAM = c(high$RingId2[48], high$RingId1[48])

#corplot
corrplot(as.matrix(kas[row.names(kas) %in% listofFAM,
                       row.names(kas) %in% listofFAM ]),
         method = 'color',addCoef.col = 'black')
```
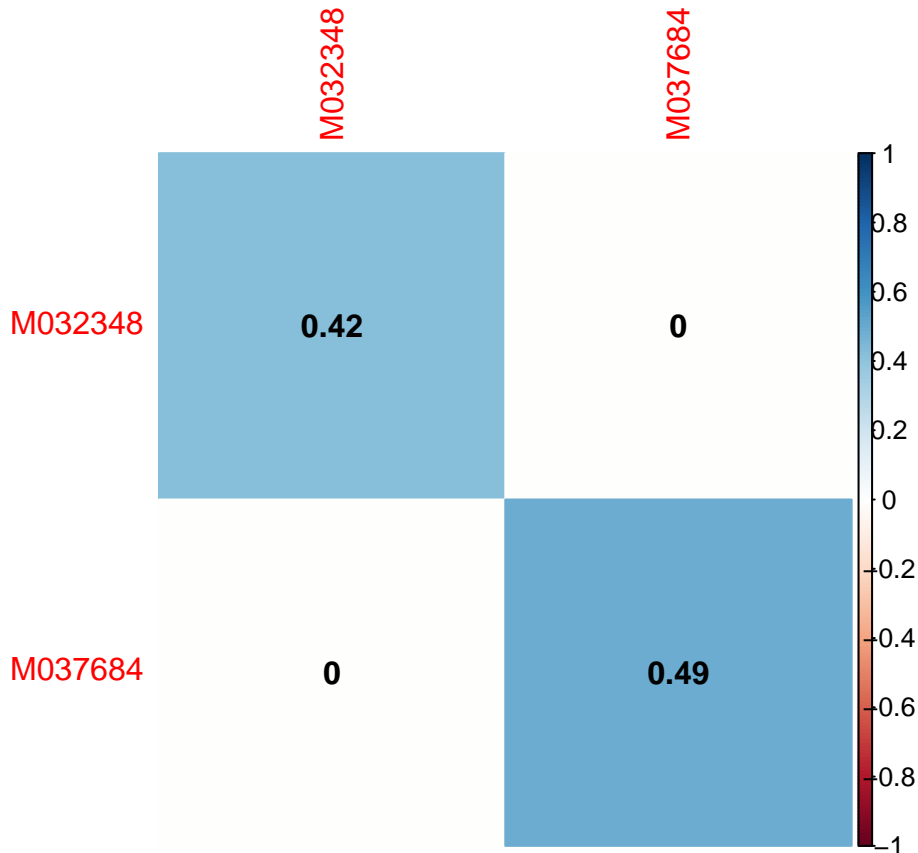

First, we can see that the self-kinship of the supposed dad (M032348) is around 0.42 which is fairly low... Let's start with him: does he have other offspring? Did we sequenced them? What is their kinship with him?

```
#other kids?
ped[(ped$sire == high$RingId2[48]) & (!is.na(ped$sire)),]
```

```
##         animal    sire     dam sex
## 6923 M031845 M032348 M032129    1
## 6924 M031846 M032348 M032129    2
## 6925 M031847 M032348 M032129    1
## 6926 M031848 M032348 M032129    2
## 6927 M031849 M032348 M032129    1
## 7103 M032318 M032348 M032231    2
## 7104 M032319 M032348 M032231    2
## 7105 M032320 M032348 M032231    2
## 7106 M032321 M032348 M032231    1
## 7107 M032322 M032348 M032231    1
```

```
## 7593 M031909 M032348 M032129   2
## 8398 M037683 M032348 M032129   1
## 8399 M037684 M032348 M032129   1
## 8400 M037685 M032348 M032129   1
## 8401 M037686 M032348 M032129   2
## 8402 M037687 M032348 M032129   1
## 8697 M038038 M032348 M032129   1
## 8698 M038039 M032348 M032129   2
## 8699 M038040 M032348 M032129   2
## 8700 M038041 M032348 M032129   2
## 8701 M038042 M032348 M032129   1
```

```r
#did we sequence them?
"M031845" %in% rownames(kas)
```

```
## [1] TRUE
```

```r
"M031846" %in% rownames(kas)
```

```
## [1] FALSE
```

```r
"M031847" %in% rownames(kas)
```

```
## [1] TRUE
```

```r
"M031848" %in% rownames(kas)
```

```
## [1] TRUE
```

```r
"M031849" %in% rownames(kas)
```

```
## [1] TRUE
```

```r
"M032318" %in% rownames(kas)
```

```
## [1] TRUE
```

```r
"M032319" %in% rownames(kas)
```

```
## [1] TRUE
```

```r
"M032320" %in% rownames(kas)
```

```
## [1] TRUE
```

```r
"M032321" %in% rownames(kas)
```

```
## [1] TRUE
```

```r
"M032322" %in% rownames(kas)
```

```
## [1] TRUE
```

```r
"M031909" %in% rownames(kas)
```

```
## [1] TRUE
```

```r
"M037683" %in% rownames(kas) #same clutch as M037684
```

```
## [1] TRUE
```

```r
"M037685" %in% rownames(kas) #same clutch as M037684
```

```
## [1] TRUE
```

```
"M037686" %in% rownames(kas) #same clutch as M037684
```

```
## [1] TRUE
```

```
"M037687" %in% rownames(kas) #same clutch as M037684
```

```
## [1] TRUE
```

```
"M038038" %in% rownames(kas)
```

```
## [1] FALSE
```

```
"M038039" %in% rownames(kas)
```

```
## [1] FALSE
```

```
"M038040" %in% rownames(kas)
```

```
## [1] TRUE
```

```
"M038041" %in% rownames(kas)
```

```
## [1] TRUE
```

```
"M038042" %in% rownames(kas)
```

```
## [1] TRUE
```

That guy had a LOT of offspring most of which we sequenced. Some of these offspring actually belong to the same clutch as his supposed kid! Let's see what the genomic kinship matrix looks like for those which we sequenced:

```
listofFAM = c(high$RingId2[48], high$RingId1[48], "M031845", "M031847",
              "M031847", "M031848", "M031849", "M032318", "M032319",
              "M032320", "M032321", "M032322", "M031909", "M037683",
              "M037685", "M037686", "M037687", "M038040", "M038041",
              "M038042")

#corplot
corrplot(as.matrix(kas[row.names(kas) %in% listofFAM,
                       row.names(kas) %in% listofFAM ]),
         method = 'color',addCoef.col = 'black', number.cex = .5)
```
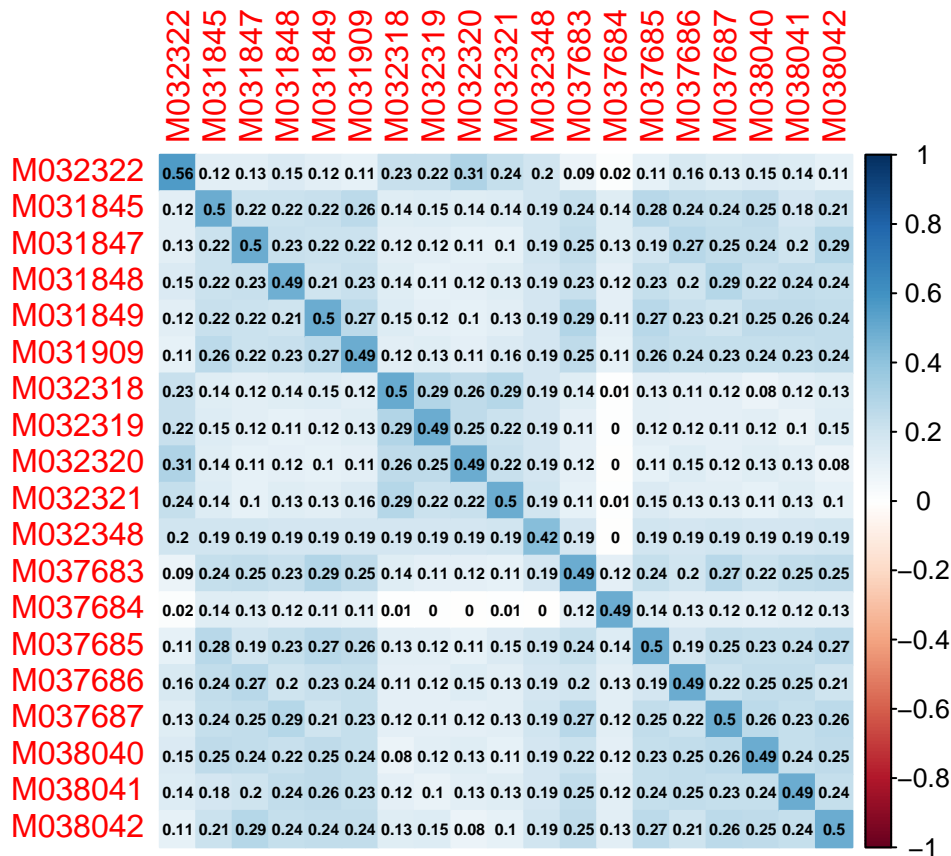
The kinship matrix is big but we can clearly see 1) that the dad is related to everyone except the offspring we were investigating (M037684) so we can be pretty confident that he is who we think he is; 2) the offspring we were investigating (M037684) is completely unrelated with some of the other offspring but it looks like he is related as half-siblings should be (genomic kinship around 0.12) with two other clutches.

Let's investigate the supposed offspring now. First let's see who his mom is and whether we sequenced her:

```
#other kids?
ped[(ped$animal == high$RingId1[48]),]
```

```
##          animal    sire      dam sex
## 8399 M037684 M032348 M032129   1
```

```
#did we sequence the mom?
"M032129" %in% rownames(kas)
```

```
## [1] FALSE
```

We did not sequence the mom so we can't confirm anything about this individual. Let's see who he is related to:

```
#new links?
row.names(kas)[which(kas[high$row[48],] > 0.16)]
```

```
## [1] "M031652" "M037684" "M038130" "M038132" "M038133" "M038134"
```

```
#who are they?
ped[ped$animal == "M031652",] #unknown
```

```
##          animal sire   dam sex
## 7936 M031652 <NA> <NA>   1
```

```
ped[ped$animal == "M038130",] #same mom
```

```
##        animal     sire      dam sex
## 8769 M038130 M031652 M032129   2
```

```
ped[ped$animal == "M038132",] #same mom
```

```
##        animal     sire      dam sex
## 8771 M038132 M031652 M032129   1
```
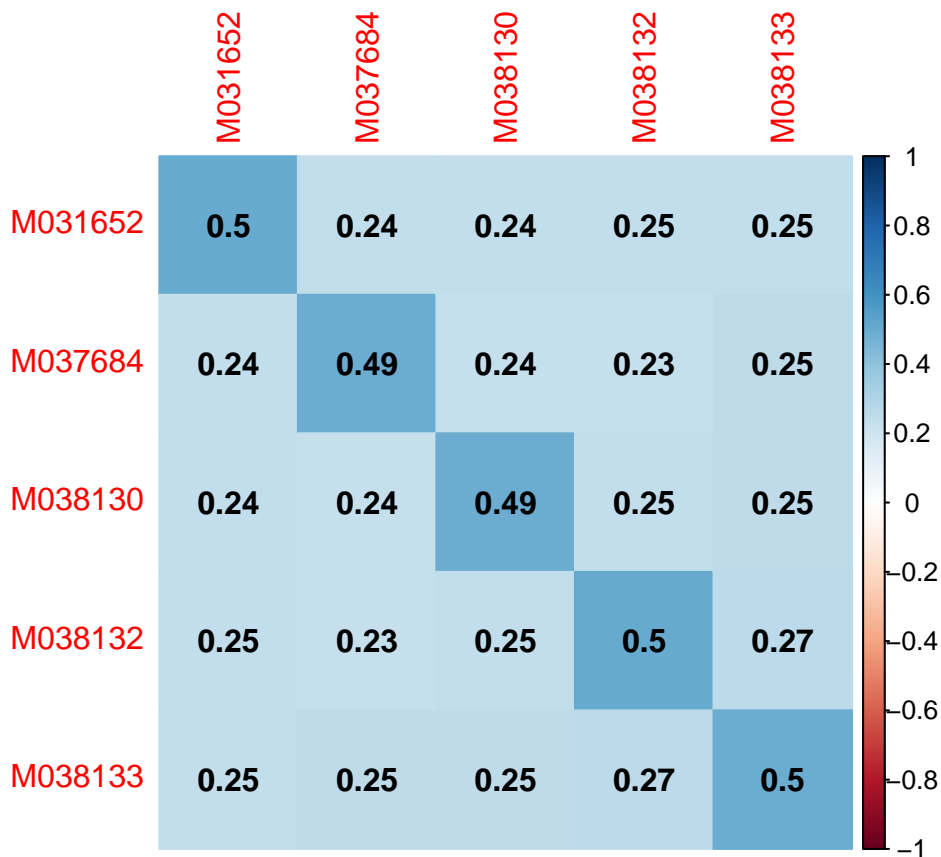
```
ped[ped$animal == "M038133",] #same mom
```

```
##        animal     sire      dam sex
## 8772 M038133 M031652 M032129   1
```

```
ped[ped$animal == "M038134",] #same mom
```

```
##        animal     sire      dam sex
## 8773 M038134 M031652 M032129   1
```

```
corrplot(kas[rownames(kas) %in% c("M031652","M038130","M038132",
                                  "M038133","M038133","M037684"),
         colnames(kas) %in% c("M031652","M038130","M038132",
                              "M038133","M038133","M037684")],
     method = 'color',addCoef.col = 'black')
```


This kid is related to five new individuals: four offspring with the same mom (which we did not sequenced) and their dad (M031652). It looks like this individual (M037684) comes from another clutch or extra pair copulation but we cannot know. Since we cannot really know who this individual is, we'll remove him from

the data set.

group 17

The next group is composed of four individuals: a supposed dad: M036777 and three of his supposed offspring: M043707, M043708, M043709. Let's see what the genomic kinship matrix looks like:

```
listofFAM = c(high$RingId2[51], high$RingId1[51:53])

#corplot
corrplot(as.matrix(kas[row.names(kas) %in% listofFAM,
                       row.names(kas) %in% listofFAM ]),
         method = 'color',addCoef.col = 'black')
```
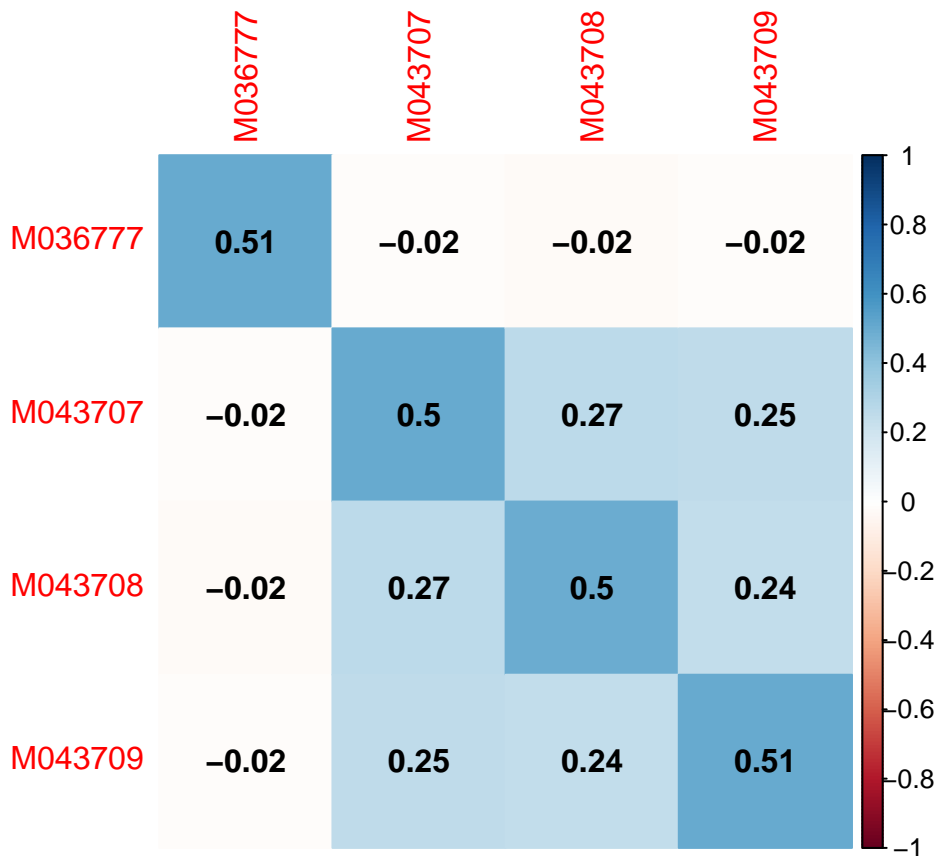

All offspring have a sibling genomic kinship (close to 0.25), only the supposed dad (M036777) is unrelated to all of them. As usual we need to know whether we did not sequence the individual we think we did or if he's just not the father of the clutch! So we'll start by looking for other offspring he could have:

```
#other offspring
ped[(ped$sire == high$RingId2[51]) & (!is.na(ped$sire)),]

##          animal    sire     dam sex
## 8328 M043705 M036777 M037679   2
## 8329 M043706 M036777 M037679   2
## 8330 M043707 M036777 M037679   2
## 8331 M043708 M036777 M037679   2
## 8332 M043709 M036777 M037679   2
```

```r
#did we sequence them?
"M043705" %in% rownames(kas)
```

```
## [1] TRUE
```

```r
"M043706" %in% rownames(kas)
```

```
## [1] TRUE
```

```r
#clutch ID?
birds$BornClutchId[birds$RingId == "M043705"]
```

```
## [1] 1952
```

```r
birds$BornClutchId[birds$RingId == "M043706"]
```

```
## [1] 1952
```

```r
birds$BornClutchId[birds$RingId == "M043707"]
```

```
## [1] 1952
```

```r
birds$BornClutchId[birds$RingId == "M043708"]
```

```
## [1] 1952
```

```r
birds$BornClutchId[birds$RingId == "M043709"]
```

```
## [1] 1952
```

We did sequence two of his other offspring but weirdly they come from the same clutch as the three offspring we are investigating (M043707, M043708 and M043709). Let's see if we sequenced the mom first and then look at the entire family (with the five offspring, the dad and the mom) kinship matrix:
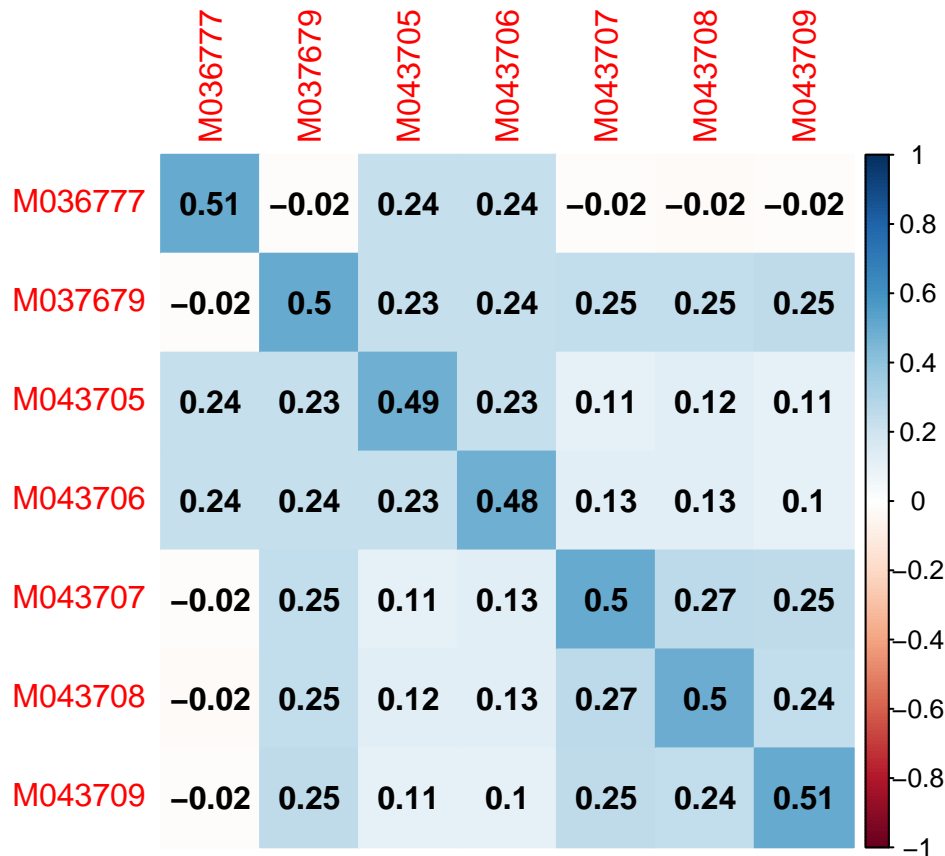
```r
#sequenced the mom
"M037679" %in% rownames(kas)
```

```
## [1] TRUE
```

```r
#family list
listofFAM = c(high$RingId2[51], high$RingId1[51:53], "M043705", "M043706", "M037679")

#corplot
corrplot(as.matrix(kas[row.names(kas) %in% listofFAM,
                   row.names(kas) %in% listofFAM ]),
       method = 'color',addCoef.col = 'black')
```

Okay so the mom is the mom of all the offspring but the dad is only the dad of two of them. The three offspring we are investigating (M043707, M043708, M043709) and the two new offspring (M043705 and M043706) are half-sibling. What it looks like is extra-pair copulation from the female which resulted in half of the clutch fathered by one individual and the rest by another. The question now is can we identify the true father of the three offspring?

```r
#new links?
row.names(kas)[which(kas[high$row[51],] > 0.16)]
```

```
## [1] "M032332" "M037679" "M043707" "M043708" "M043709"
```
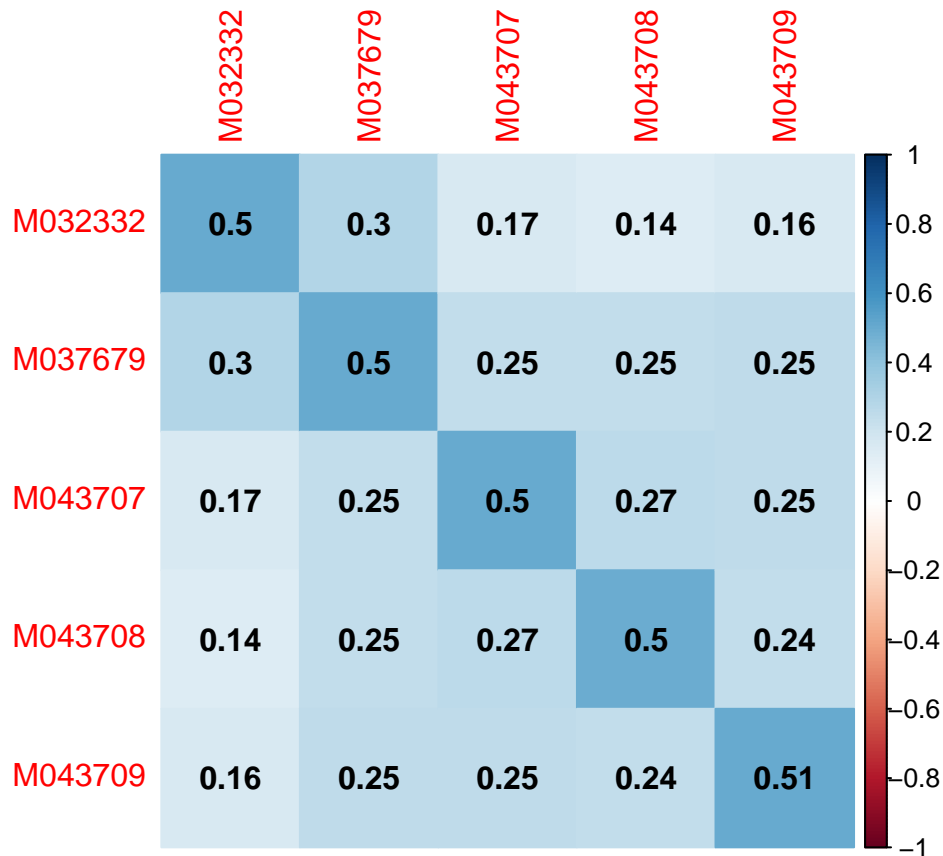
```r
#who are they?
ped[ped$animal == "M032332",] #unknown
```

```
##        animal    sire     dam sex
## 7117 M032332 M026075 M032230   2
```

```r
ped[ped$animal == "M037679",] #the mom
```

```
##        animal sire  dam sex
## 8394 M037679 <NA> <NA>   2
```

```r
#correlation plot
corrplot(kas[rownames(kas) %in% c("M032332","M037679","M043707",
                                  "M043708", "M043709"),
           colnames(kas) %in% c("M032332","M037679","M043707",
                                  "M043708", "M043709")],
         method = 'color',addCoef.col = 'black')
```

That guy is more related to the mom that to the kids so he's not the father but probably of relative of the mom. However, since we don't know the mom parents we can't be sure who he is. So what we'll do in the end is to keep all the samples and correct the pedigree.

group 18

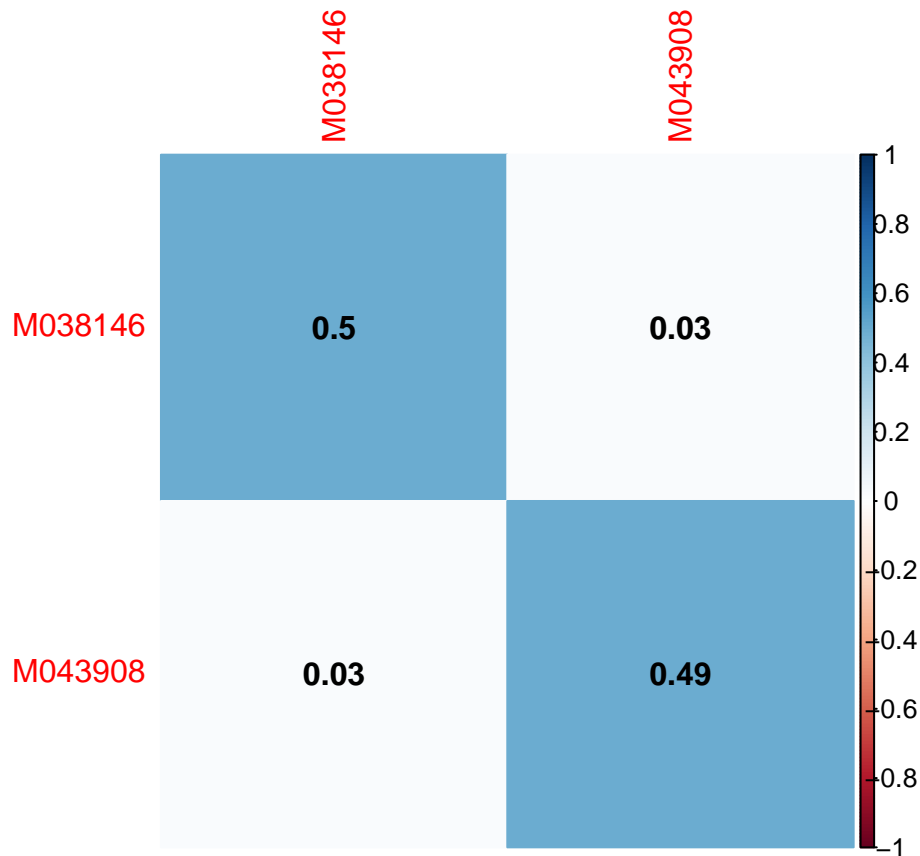The next group is composed of two individuals: a supposed dad: M038146 and one of his supposed offspring: M043908. Let's see what the genomic kinship matrix looks like:

```
listofFAM = c(high$RingId2[61], high$RingId1[61])

#corplot
corrplot(as.matrix(kas[row.names(kas) %in% listofFAM,
                       row.names(kas) %in% listofFAM ]),
         method = 'color',addCoef.col = 'black')
```

The supposed dad and kid are not related. As usual we'll first tart by investigating the dad: any other offspring or relatives?

```
#any other kids
ped[(ped$sire == high$RingId2[61]) & (!is.na(ped$sire)),]
```

```
##          animal    sire     dam sex
## 9106   M043910 M038146 M028824   2
## 9107   M043911 M038146 M028824   1
## 9108   M043912 M038146 M028824   2
## 9109   M043908 M038146 M028824   2
## 9110   M043909 M038146 M028824   2
## 9493   M040891 M038146 M040764   1
## 9494   M040892 M038146 M040764   2
## 9495   M040893 M038146 M040764   1
## 9496   M040705 M038146 M040764   2
## 9497   M040706 M038146 M040764   2
## 10701 M049381 M038146 M048285   2
## 10702 M049382 M038146 M048285   1
## 10703 M049390 M038146 M048285   2
```

```
#did we sequence them?
"M043910" %in% rownames(kas)
```

```
## [1] TRUE
```

```
"M043911" %in% rownames(kas)
```

```
## [1] TRUE
```

```r
"M043912" %in% rownames(kas)
```

```
## [1] TRUE
```

```r
"M043909" %in% rownames(kas)
```

```
## [1] TRUE
```

```r
"M040891" %in% rownames(kas)
```

```
## [1] TRUE
```

```r
"M040892" %in% rownames(kas)
```

```
## [1] TRUE
```

```r
"M040893" %in% rownames(kas)
```

```
## [1] TRUE
```

```r
"M040705" %in% rownames(kas)
```

```
## [1] TRUE
```

```r
"M040706" %in% rownames(kas)
```

```
## [1] TRUE
```

He has many other offspring which we did sequence. What are their kinship?

```r
listofFAM = c(high$RingId2[61], "M043910", "M043911", "M043912",
              "M043909", "M040891", "M040892", "M040893",
              "M040705", "M040706")

#corplot
corrplot(as.matrix(kas[row.names(kas) %in% listofFAM,
                       row.names(kas) %in% listofFAM ]),
         method = 'color',addCoef.col = 'black')
```

	M038146	M040705	M040706	M040891	M040892	M040893	M043909	M043910	M043911	M043912
M038146	0.5	0.24	0.25	0.25	0.24	0.25	0.24	0.25	0.25	0.24
M040705	0.24	0.49	0.24	0.23	0.2	0.23	0.13	0.11	0.12	0.14
M040706	0.25	0.24	0.5	0.23	0.27	0.25	0.12	0.14	0.12	0.12
M040891	0.25	0.23	0.23	0.5	0.25	0.2	0.14	0.15	0.14	0.15
M040892	0.24	0.2	0.27	0.25	0.5	0.21	0.13	0.14	0.15	0.14
M040893	0.25	0.23	0.25	0.2	0.21	0.5	0.14	0.14	0.15	0.16
M043909	0.24	0.13	0.12	0.14	0.13	0.14	0.49	0.22	0.26	0.25
M043910	0.25	0.11	0.14	0.15	0.14	0.14	0.22	0.5	0.25	0.21
M043911	0.25	0.12	0.12	0.14	0.15	0.15	0.26	0.25	0.5	0.24
M043912	0.24	0.14	0.12	0.15	0.14	0.16	0.25	0.21	0.24	0.49

This individual (M038146) is correctly related (around 0.25) with all his supposed offspring so we can be pretty confident about the fact that he is who we think he is. We'll keep him, he's just not the father of this kid. This kid might actually be the problem let's investigate whether he has some siblings / related individuals.

```r
#his parents
ped[ped$animal == high$RingId1[61],]
```

```
##         animal     sire     dam sex
## 9109 M043908 M038146 M028824   2
```

```r
#any guys from the same clutch?
ped[(ped$sire == high$RingId2[61]) & (!is.na(ped$sire)) &
      (ped$dam == "M028824") & (!is.na(ped$dam)),]
```

```
##         animal     sire     dam sex
## 9106 M043910 M038146 M028824   2
## 9107 M043911 M038146 M028824   1
## 9108 M043912 M038146 M028824   2
## 9109 M043908 M038146 M028824   2
## 9110 M043909 M038146 M028824   2
```

```r
#did we sequence them?
"M043910" %in% rownames(kas)
```

```
## [1] TRUE
```

```r
"M043911" %in% rownames(kas)
```
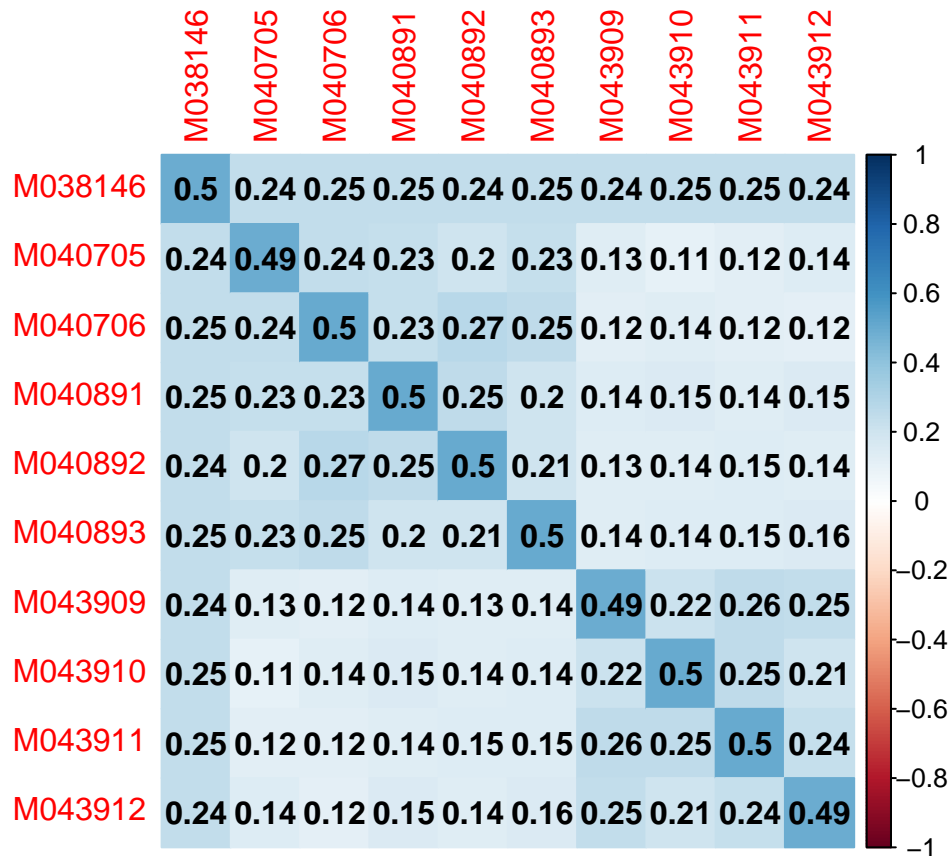
```
## [1] TRUE
```

```
"M043912" %in% rownames(kas)
```
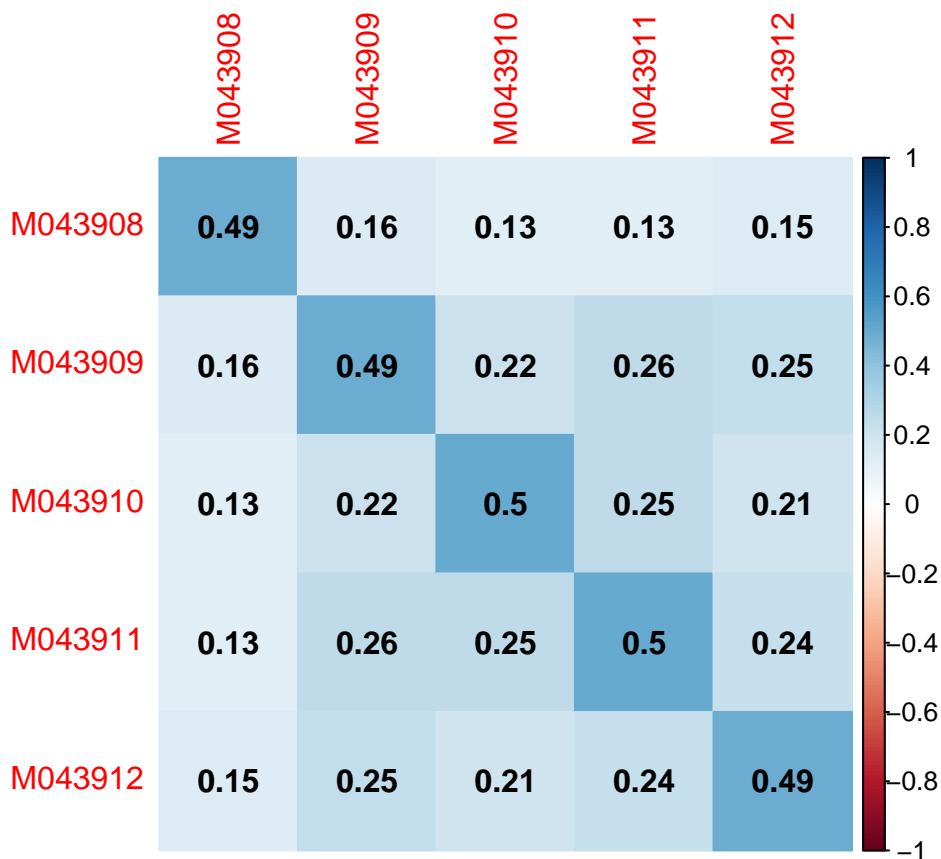
[1] TRUE

```
"M043909" %in% rownames(kas)
```

[1] TRUE

We identified four offspring which came from the same clutch and which we already identified before as kids of the dad. What are they genomic kinship?

```
listofFAM = c(high$RingId1[61], "M043909", "M043910", "M043911", "M043912")

#corplot
corrplot(as.matrix(kas[row.names(kas) %in% listofFAM,
                        row.names(kas) %in% listofFAM ]),
         method = 'color',addCoef.col = 'black')
```

	M043908	M043909	M043910	M043911	M043912
M043908	0.49	0.16	0.13	0.13	0.15
M043909	0.16	0.49	0.22	0.26	0.25
M043910	0.13	0.22	0.5	0.25	0.21
M043911	0.13	0.26	0.25	0.5	0.24
M043912	0.15	0.25	0.21	0.24	0.49

It looks like our focal kid (M043908) is half-sibling with the rest of the clutch. Did we sequenced the mom and what is her kinship with this clutch?

```
"M028824" %in% rownames(kas)
```
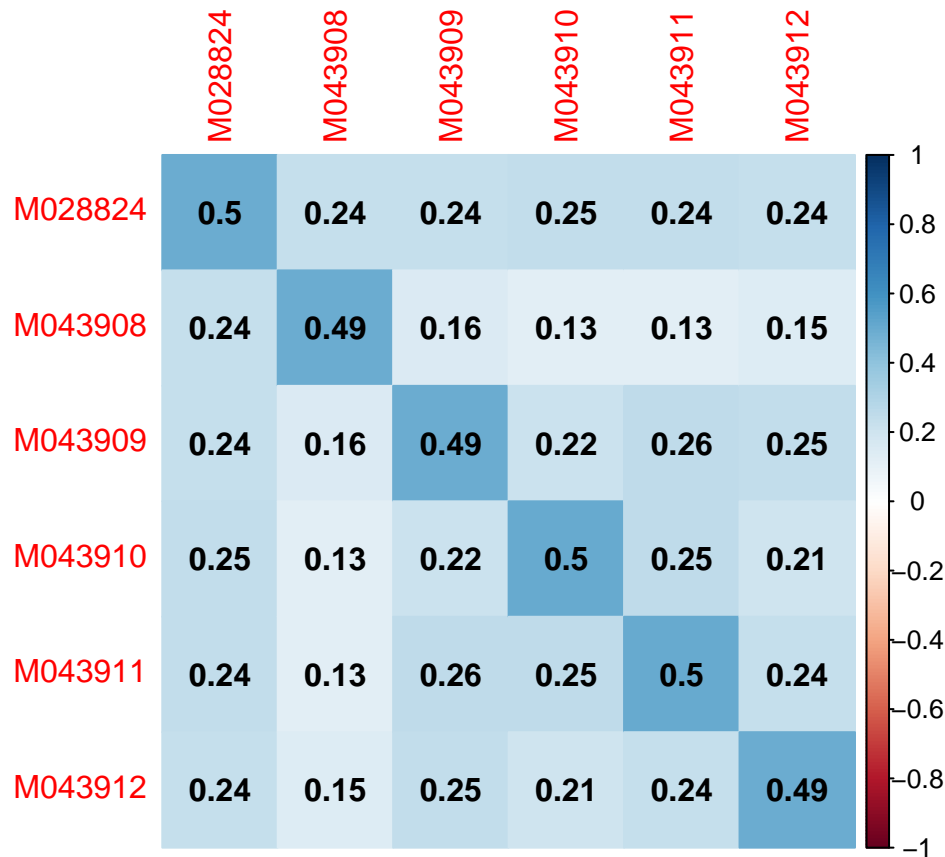
[1] TRUE

```
listofFAM = c(high$RingId1[61], "M043909", "M043910", "M043911", "M043912", "M028824")

#correlation plot
corrplot(as.matrix(kas[row.names(kas) %in% listofFAM,
                        row.names(kas) %in% listofFAM ]),
```

```
method = 'color',addCoef.col = 'black')
```

	M028824	M043908	M043909	M043910	M043911	M043912
M028824	0.5	0.24	0.24	0.25	0.24	0.24
M043908	0.24	0.49	0.16	0.13	0.13	0.15
M043909	0.24	0.16	0.49	0.22	0.26	0.25
M043910	0.25	0.13	0.22	0.5	0.25	0.21
M043911	0.24	0.13	0.26	0.25	0.5	0.24
M043912	0.24	0.15	0.25	0.21	0.24	0.49

She is the mother of all the clutch. So what do we do with this? Can we identify related individuals to this offspring (M043908)? What are there genomic kinship?

```
#new links?
row.names(kas)[which(kas[high$row[61],] > 0.16)]
```

```
## [1] "M028824" "M043672" "M043673" "M043674" "M043908"
```

```
#who are they?
ped[ped$animal == "M028824",]
```

```
##        animal sire  dam sex
## 7406 M028824 <NA> <NA>   2
```

```
ped[ped$animal == "M043672",]
```

```
##        animal    sire     dam sex
## 9257 M043672 M031833 M028824   2
```

```
ped[ped$animal == "M043673",]
```

```
##        animal    sire     dam sex
## 9258 M043673 M031833 M028824   2
```
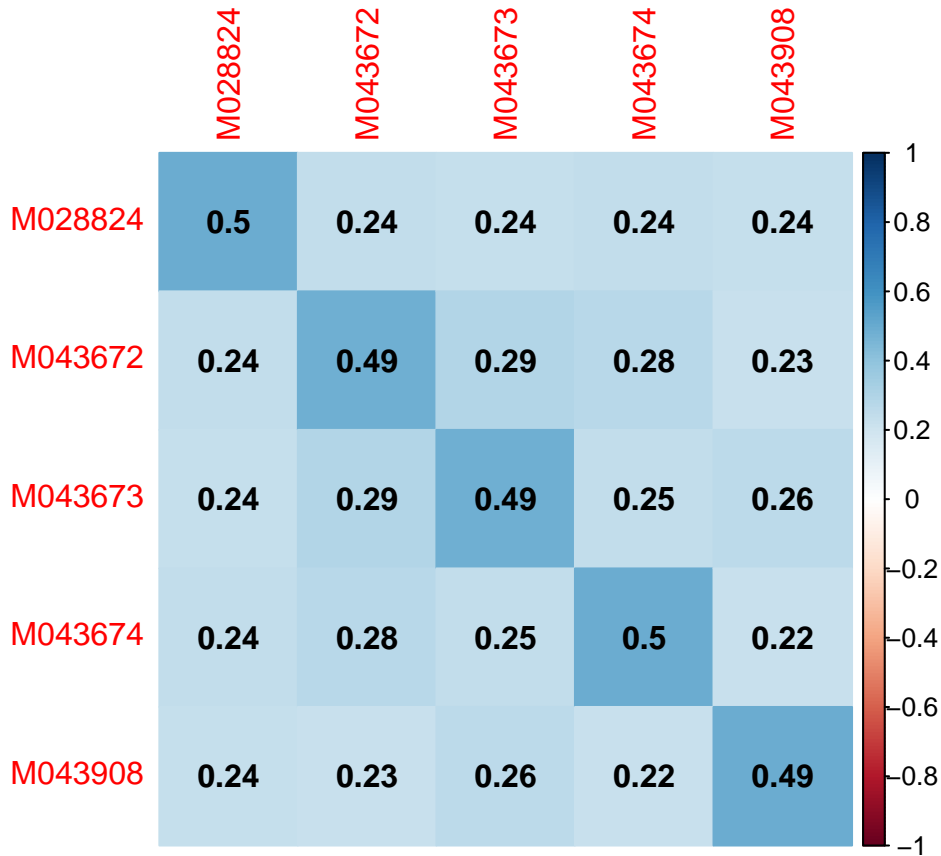
```
ped[ped$animal == "M043674",]
```

```
##        animal    sire     dam sex
## 9259 M043674 M031833 M028824   1
```

```
#correlation plot
corrplot(as.matrix(kas[row.names(kas) %in% c("M028824", "M043672", "M043673",
                                              "M043674", "M043908"),
                        row.names(kas) %in% c("M028824", "M043672", "M043673",
                                              "M043674", "M043908")]),
         method = 'color',addCoef.col = 'black')
```
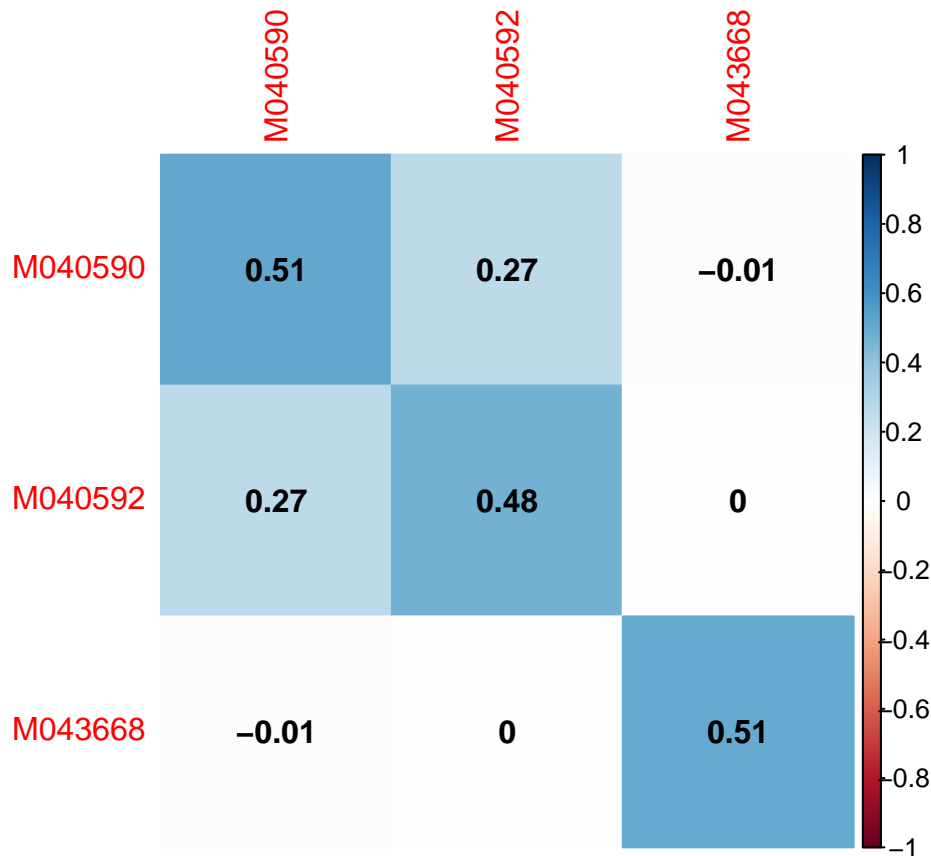
	M028824	M043672	M043673	M043674	M043908
M028824	0.5	0.24	0.24	0.24	0.24
M043672	0.24	0.49	0.29	0.28	0.23
M043673	0.24	0.29	0.49	0.25	0.26
M043674	0.24	0.28	0.25	0.5	0.22
M043908	0.24	0.23	0.26	0.22	0.49

This guy is really the "true" sibling of the others. So we cannot be sure who he is so we'll remove him (M043908) from the data set.

group 19

The next group is composed of two individuals: a supposed dad: M043668 and two of his supposed offspring: M040590 and M040592. Let's see what the genomic kinship matrix looks like:

```
listofFAM = c(high$RingId1[62], high$RingId2[62:63])

#correlation plot
corrplot(as.matrix(kas[row.names(kas) %in% listofFAM,
                       row.names(kas) %in% listofFAM ]),
         method = 'color',addCoef.col = 'black')
```

M040590 — M040592 — M043668 (column labels)

	M040590	M040592	M043668
M040590	0.51	0.27	−0.01
M040592	0.27	0.48	0
M043668	−0.01	0	0.51

Color scale: 1, 0.8, 0.6, 0.4, 0.2, 0, −0.2, −0.4, −0.6, −0.8, −1

The two kids are related (genomic kinship = 0.27) like siblings but not related to the dad (genomic kinship = 0). Let's start by looking at whether the dad has other offspring, whether we sequenced them and what is their kinship with him?

```
#other offspring ?
ped[(ped$sire == high$RingId1[62]) & (!is.na(ped$sire)),]
```

```
##          animal    sire     dam sex
## 9111   M040590 M043668 M028977   1
## 9112   M040591 M043668 M028977   2
## 9113   M040592 M043668 M028977   1
## 10217  M040990 M043668 M048274   2
## 10218  M040991 M043668 M048274   2
## 10219  M040992 M043668 M048274   2
## 10220  M040993 M043668 M048274   1
## 10221  M048328 M043668 M048274   2
## 11158  M049426 M043668 M040832   2
## 11159  M049427 M043668 M040832   2
## 11160  M049428 M043668 M040832   2
## 11161  M049429 M043668 M040832   1
## 11162  M049430 M043668 M040832   2
## 11163  M049431 M043668 M040832   2
```

```
#Did we sequence them ?
"M040590" %in% rownames(kas)
```

```
## [1] TRUE
```

```r
"M040591" %in% rownames(kas)
```

```
## [1] FALSE
```

```r
"M040592" %in% rownames(kas)
```

```
## [1] TRUE
```

```r
"M040990" %in% rownames(kas)
```

```
## [1] FALSE
```

```r
"M040991" %in% rownames(kas)
```

```
## [1] FALSE
```

```r
"M040992" %in% rownames(kas)
```

```
## [1] FALSE
```

```r
"M040993" %in% rownames(kas)
```

```
## [1] FALSE
```

```r
"M048328" %in% rownames(kas)
```

```
## [1] FALSE
```

```r
"M049426" %in% rownames(kas)
```

```
## [1] FALSE
```

```r
"M049427" %in% rownames(kas)
```

```
## [1] FALSE
```

```r
"M049428" %in% rownames(kas)
```

```
## [1] FALSE
```

```r
"M049429" %in% rownames(kas)
```

```
## [1] FALSE
```

```r
"M049430" %in% rownames(kas)
```

```
## [1] FALSE
```

```r
"M049431" %in% rownames(kas)
```

```
## [1] FALSE
```

```r
#correlation plot
corrplot(as.matrix(kas[row.names(kas) %in% c("M040590", "M040592",
                                             high$RingId1[62]),
                   row.names(kas) %in% c("M040590", "M040592",
                                             high$RingId1[62])]),
         method = 'color',addCoef.col = 'black')
```
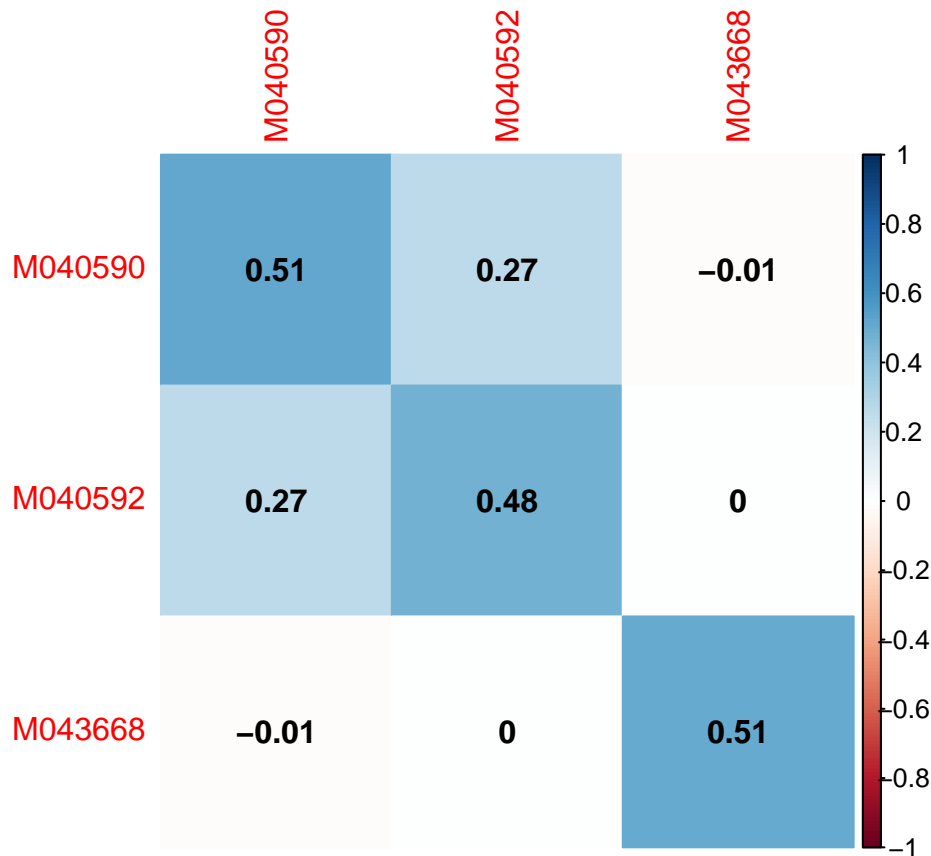
The supposed dad (M043668) is not related to any of his supposed offspring. So we can conclude that he is not who we think he is. Thus, we'll remove him from our data set.

Now about the two offspring (M040590 and M040592), are they who we think they are ? To answer this question, we'll use the mom: did we sequenced her and how is she related to them?
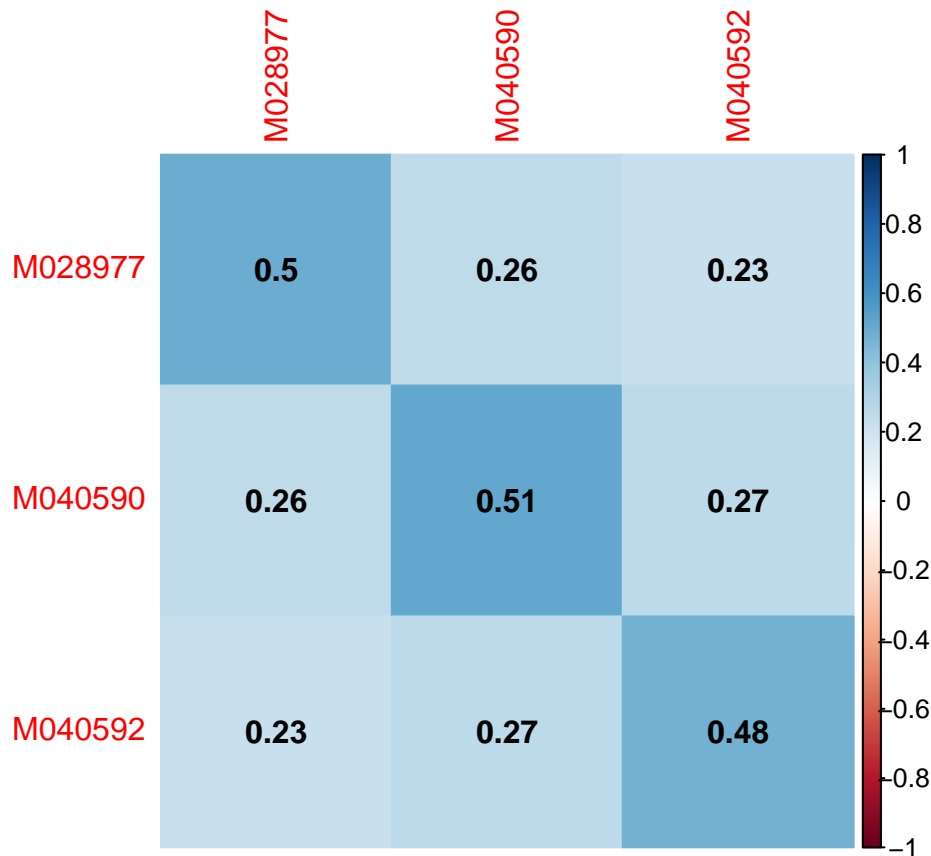
```
#other offspring ?
ped[ped$animal == high$RingId2[62],]
```

```
##         animal    sire     dam sex
## 9111 M040590 M043668 M028977   1
```

```
#did we sequence her?
"M028977" %in% rownames(kas)
```

```
## [1] TRUE
```

```
#correlation plot
corrplot(as.matrix(kas[row.names(kas) %in% c("M028977", high$RingId2[62],
                                      high$RingId2[63]),
                 row.names(kas) %in%  c("M028977", high$RingId2[62],
                                      high$RingId2[63])]),
      method = 'color',addCoef.col = 'black')
```

We did sequence her and she's correctly related to them (genomic kinship around 0.25) so we'll keep both kids.

group 20

The next group is a composed of four individuals: the supposed mom: M043948, the supposed dad: M043949 and two of their supposed offspring: M043647 and M043649. What does the kinship matrix look like?

```
listofFAM = c(high$RingId1[64], high$RingId1[65],
              high$RingId2[64], high$RingId2[66])


#corplot
corrplot(as.matrix(kas[row.names(kas) %in% listofFAM,
                       row.names(kas) %in% listofFAM ]),
         method = 'color',addCoef.col = 'black')
```
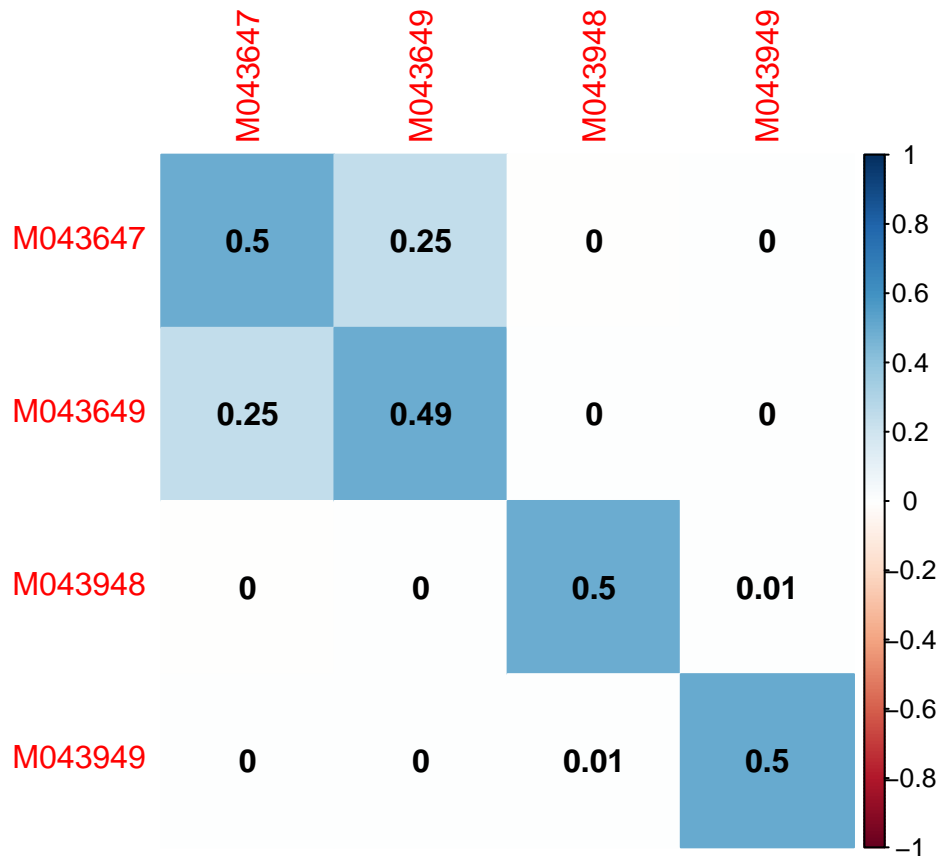
From the matrix the siblings are siblings but are not related to any of the parents... Are they related to anyone else? And if yes how?

```r
#new links?
row.names(kas)[which(kas[high$col[64],] > 0.16)]
```

```
## [1] "EA711589" "M043647"  "M043649"  "M043946"
```

```r
#who are they ?
ped[ped$animal == "EA711589",]
```

```
##          animal sire  dam sex
## 9570 EA711589 <NA> <NA>   2
```

```r
ped[ped$animal == "M043946",]
```

```
##          animal sire  dam sex
## 9645 M043946 <NA> <NA>   1
```

```r
#new list of individuals
listofFAM = c(high$RingId1[64], high$RingId1[65], "EA711589",
              high$RingId2[64], high$RingId2[66], "M043946")

#correlation plot
corrplot(as.matrix(kas[row.names(kas) %in% listofFAM,
                   row.names(kas) %in% listofFAM ]),
         method = 'color',addCoef.col = 'black')
```
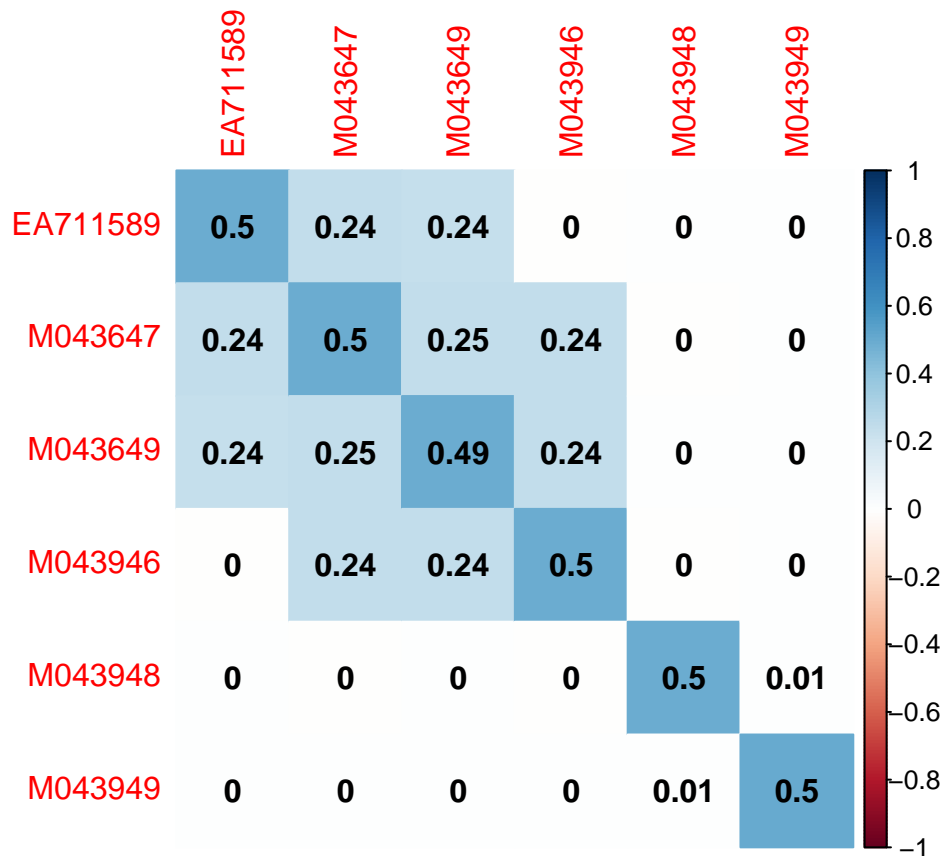
The siblings (M043647 and M043649) are related to two new owls which could be the parents. To further investigate this we'll see if these two owls have kids and if they really are the parents these others kids should be half-siblings with our focal kids (M043647 and M043649):

```r
#new links?
row.names(kas)[which(kas[rownames(kas) == "EA711589",] > 0.16)]
```

```
## [1] "EA711589" "M043647"  "M043649"
```

```r
row.names(kas)[which(kas[rownames(kas) == "M043946",] > 0.16)]
```

```
## [1] "M043647" "M043649" "M043946"
```

No other individuals are related to these individuals so we cannot confirm the kids identity. Consequently, we'll remove both (M043647 and M043649) from our data set.

Now what about the dad? Can we confirm who he is? To do this, we'll look for extra offspring he should have:

```r
#other offspring ?
ped[(ped$sire == high$RingId1[65]) & (!is.na(ped$sire)),]
```

```
##         animal    sire    dam sex
## 9335  M043647 M043949 M043948   1
## 9336  M043649 M043949 M043948   1
## 9337  M043648 M043949 M043948   1
## 9941  M047573 M043949 M043598   1
## 9942  M047574 M043949 M043598   2
## 9943  M047575 M043949 M043598   2
## 9944  M047576 M043949 M043598   2
```

```
## 10129 M047698 M043949 M047616    2
## 10130 M047699 M043949 M047616    2
## 10131 M047700 M043949 M047616    2
## 10132 M047701 M043949 M047616    2
## 10133 M047702 M043949 M047616    1
## 10776 M047900 M043949 M047694    2
## 10777 M047801 M043949 M047694    2
## 10778 M047802 M043949 M047694    1
## 10779 M047803 M043949 M047694    1
## 11165 M049818 M043949 M043309    1
## 11166 M049820 M043949 M043309    1
## 11167 M049819 M043949 M043309    2
## 11168 M049821 M043949 M043309    1
```

```r
#Did we sequence them ?
"M043648" %in% rownames(kas)
```

```
## [1] FALSE
```

```r
"M047573" %in% rownames(kas)
```

```
## [1] TRUE
```

```r
"M047574" %in% rownames(kas)
```

```
## [1] TRUE
```

```r
"M047575" %in% rownames(kas)
```

```
## [1] TRUE
```

```r
"M047576" %in% rownames(kas)
```

```
## [1] FALSE
```

```r
"M047698" %in% rownames(kas)
```

```
## [1] FALSE
```

```r
"M040993" %in% rownames(kas)
```

```
## [1] FALSE
```

```r
"M047699" %in% rownames(kas)
```

```
## [1] FALSE
```

```r
"M047700" %in% rownames(kas)
```

```
## [1] FALSE
```

```r
"M047701" %in% rownames(kas)
```

```
## [1] FALSE
```

```r
"M047702" %in% rownames(kas)
```

```
## [1] FALSE
```

```r
"M047900" %in% rownames(kas)
```

```
## [1] FALSE
```

```r
"M047801" %in% rownames(kas)
```

```
## [1] FALSE
```
```r
"M047802" %in% rownames(kas)
```
```
## [1] FALSE
```
```r
"M047803" %in% rownames(kas)
```
```
## [1] FALSE
```
```r
"M049818" %in% rownames(kas)
```
```
## [1] FALSE
```
```r
"M049819" %in% rownames(kas)
```
```
## [1] FALSE
```
```r
"M049820" %in% rownames(kas)
```
```
## [1] FALSE
```
```r
"M049821" %in% rownames(kas)
```
```
## [1] FALSE
```
```r
#correlation plot
corrplot(as.matrix(kas[row.names(kas) %in% c("M047573", "M047574",
                                             "M047575", high$RingId1[65]),
                row.names(kas) %in% c("M047573", "M047574",
                                      "M047575", high$RingId1[65])]),
      method = 'color',addCoef.col = 'black')
```
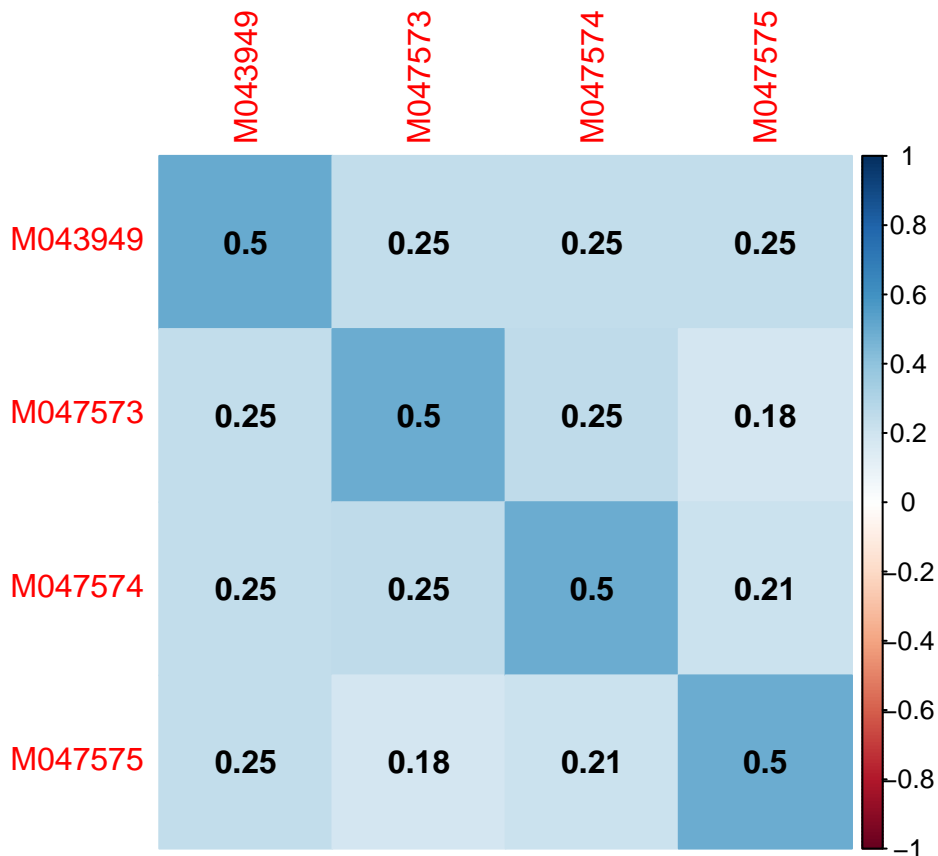
The dad is the individual we though he was: he has a genomic kinship of 0.25 with his other supposed offspring! He's just not the dad of the kids we were investigating (M043647 and M043649) ! So we will keep this guy in the data set.

What about the mom? Can we make sure she is who we think she is? To do this, we'll look for any other offsprings or siblings and check what kinship she has with them:

```
#other offspring ?
ped[(ped$dam == high$RingId1[64]) & (!is.na(ped$dam)),]
```

```
##        animal      sire       dam sex
## 9335 M043647 M043949 M043948   1
## 9336 M043649 M043949 M043948   1
## 9337 M043648 M043949 M043948   1
```

```
#Did we sequence them ?
"M043648" %in% rownames(kas) #no
```

```
## [1] FALSE
```

```
#any siblings ?
ped[ped$animal == high$RingId1[64],]
```

```
##        animal sire   dam sex
## 9565 M043948 <NA> <NA>   2
```

```
#any related individuals?
row.names(kas)[which(kas[high$row[64],] > 0.16)]
```

```
## [1] "M043948"
```

She does not have any other kid or siblings or related individuals... We cannot confirm anything about her so we will remove her from the data set.

group 21

The next group is composed of two individuals: a supposed dad: M043950 and his supposed offspring: M047563. Let's see what the genomic kinship matrix looks like:

```
listofFAM = c(high$RingId2[71], high$RingId1[71])
```

```
#correlation plot
corrplot(as.matrix(kas[row.names(kas) %in% listofFAM,
                       row.names(kas) %in% listofFAM ]),
        method = 'color',addCoef.col = 'black')
```
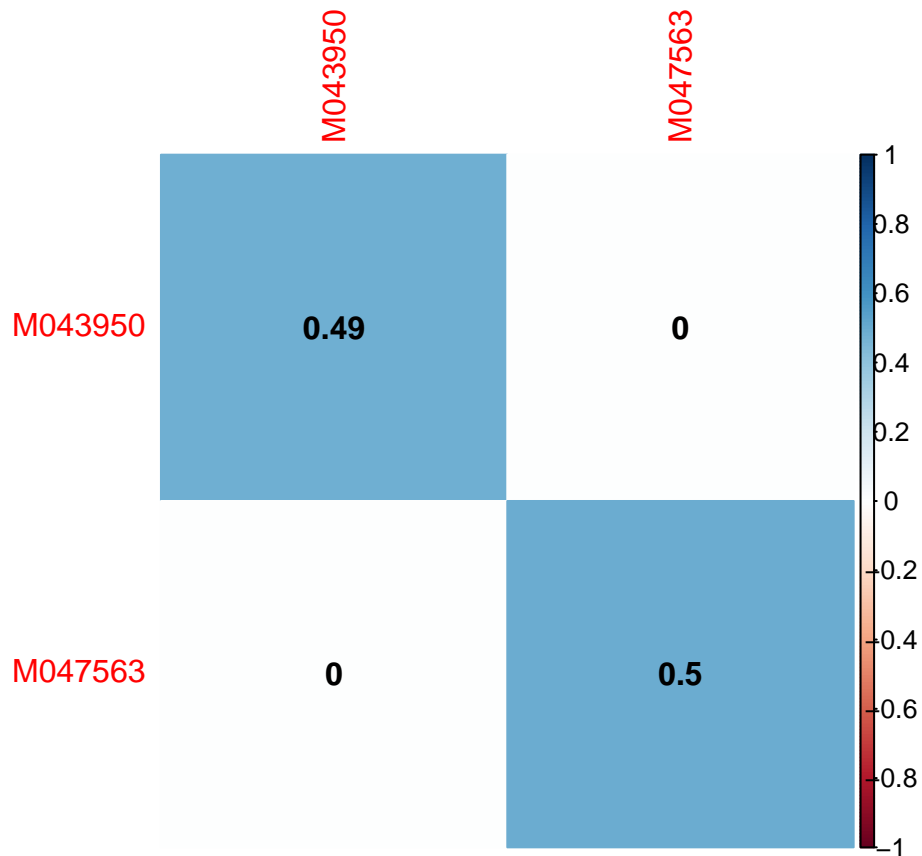
They are genomically unrelated. Now we'll need to confirm wether the dad is really the individual we think he is by looking for other of his offspring:

```
#other offspring ?
ped[(ped$sire == high$RingId2[71]) & (!is.na(ped$sire)),]
```

```
##          animal    sire      dam sex
## 9365 M043012 M043950 M043628   1
## 9366 M043013 M043950 M043628   2
## 9367 M043014 M043950 M043628   1
## 9368 M043015 M043950 M043628   1
## 9465 M043046 M043950 M043628   2
## 9716 M047598 M043950 M043628   2
## 9987 M047561 M043950 M043628   1
## 9988 M047563 M043950 M043628   1
## 9989 M047562 M043950 M043628   1
```

```
#Did we sequence them ?
"M043012" %in% rownames(kas)
```

```
## [1] TRUE
```

```
"M043013" %in% rownames(kas)
```

```
## [1] TRUE
```

```
"M043014" %in% rownames(kas)
```

```
## [1] TRUE
```

```
"M043015" %in% rownames(kas)
```

```
## [1] TRUE
```

```
"M043046" %in% rownames(kas)
```

```
## [1] TRUE
```

```
"M047598" %in% rownames(kas)
```

```
## [1] TRUE
```
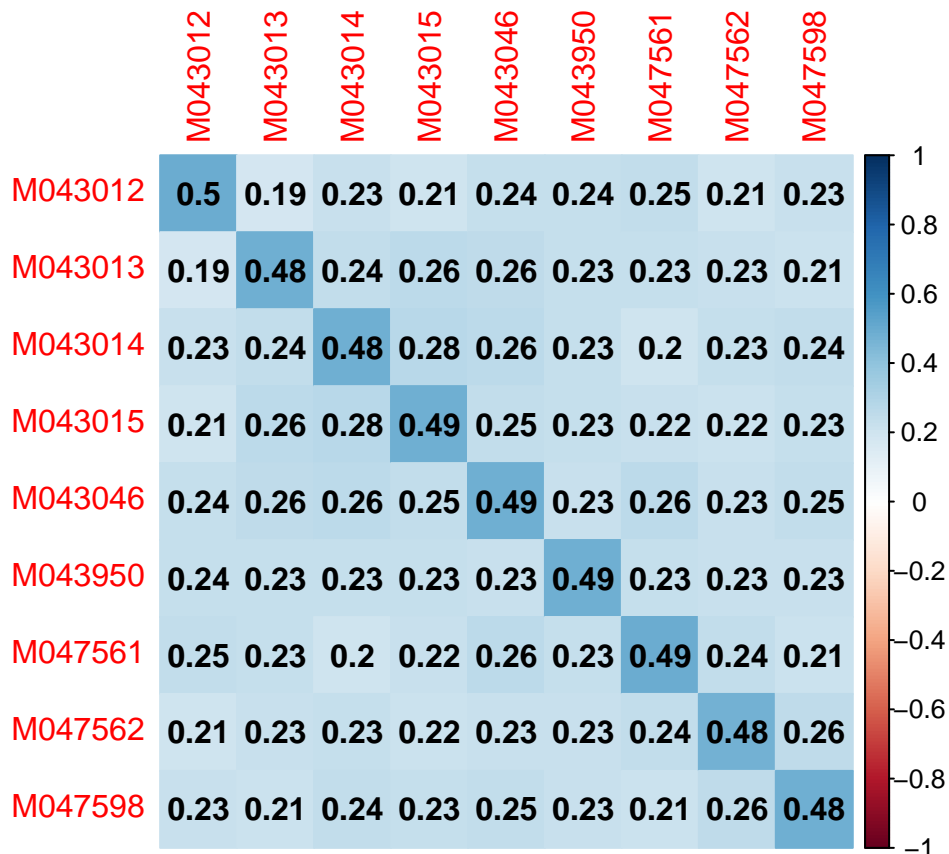
```
"M047561" %in% rownames(kas)
```

```
## [1] TRUE
```

```
"M047562" %in% rownames(kas)
```

```
## [1] TRUE
```

```
#correlation plot
corrplot(as.matrix(kas[row.names(kas) %in% c(high$RingId2[71], "M043012",
                                             "M043013", "M043014", "M043015",
                                             "M043046", "M047598", "M047561",
                                             "M047562"),
                  row.names(kas) %in% c(high$RingId2[71], "M043012",
                                             "M043013", "M043014", "M043015",
                                             "M043046", "M047598", "M047561",
                                             "M047562")]),
         method = 'color', addCoef.col = 'black')
```

The dad is who we think he is, since he really is the dad of the rest of his supposed offspring. Now what about the kid?

What about the kid M047563? Does he have any siblings?

```r
ped[ped$animal == high$RingId1[71],]
```

```
##         animal    sire      dam sex
## 9988 M047563 M043950 M043628   1
```

```r
#any siblings with the same parents?
ped[(ped$sire == high$RingId2[71]) & (!is.na(ped$sire)) &
      (ped$dam == "M043628") & (!is.na(ped$dam)),]
```

```
##         animal    sire      dam sex
## 9365 M043012 M043950 M043628   1
## 9366 M043013 M043950 M043628   2
## 9367 M043014 M043950 M043628   1
## 9368 M043015 M043950 M043628   1
## 9465 M043046 M043950 M043628   2
## 9716 M047598 M043950 M043628   2
## 9987 M047561 M043950 M043628   1
## 9988 M047563 M043950 M043628   1
## 9989 M047562 M043950 M043628   1
```

```r
#Did we sequence them ?
"M043012" %in% rownames(kas)
```

```
## [1] TRUE
```

```r
"M043013" %in% rownames(kas)
```

```
## [1] TRUE
```

```r
"M043014" %in% rownames(kas)
```

```
## [1] TRUE
```

```r
"M043015" %in% rownames(kas)
```

```
## [1] TRUE
```

```r
"M043046" %in% rownames(kas)
```

```
## [1] TRUE
```

```r
"M047598" %in% rownames(kas)
```

```
## [1] TRUE
```

```r
"M047561" %in% rownames(kas)
```

```
## [1] TRUE
```

```r
"M047562" %in% rownames(kas)
```

```
## [1] TRUE
```

```r
#correlation plot
corrplot(as.matrix(kas[row.names(kas) %in% c(high$RingId1[71], "M043012",
                                             "M043013", "M043014", "M043015",
                                             "M043046", "M047598", "M047561",
                                             "M047562"),
```

```r
                        row.names(kas) %in% c(high$RingId1[71], "M043012",
                                              "M043013", "M043014", "M043015",
                                              "M043046", "M047598", "M047561",
                                              "M047562")]),
        method = 'color', addCoef.col = 'black')
```
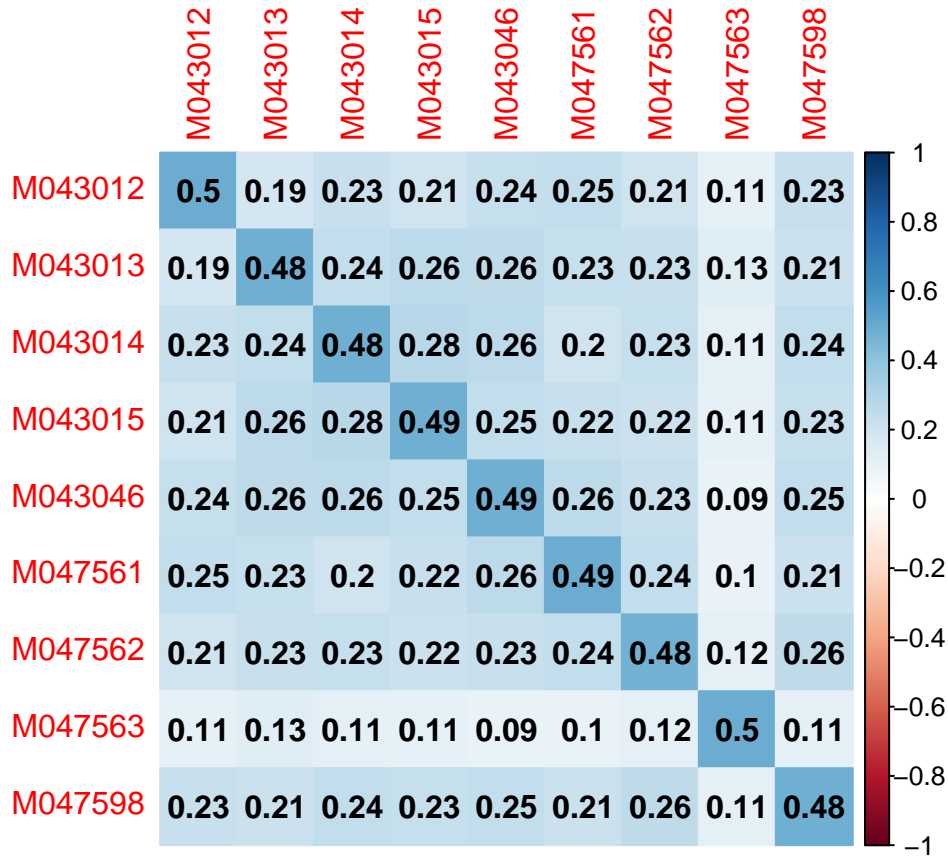
	M043012	M043013	M043014	M043015	M043046	M047561	M047562	M047563	M047598
M043012	0.5	0.19	0.23	0.21	0.24	0.25	0.21	0.11	0.23
M043013	0.19	0.48	0.24	0.26	0.26	0.23	0.23	0.13	0.21
M043014	0.23	0.24	0.48	0.28	0.26	0.2	0.23	0.11	0.24
M043015	0.21	0.26	0.28	0.49	0.25	0.22	0.22	0.11	0.23
M043046	0.24	0.26	0.26	0.25	0.49	0.26	0.23	0.09	0.25
M047561	0.25	0.23	0.2	0.22	0.26	0.49	0.24	0.1	0.21
M047562	0.21	0.23	0.23	0.22	0.23	0.24	0.48	0.12	0.26
M047563	0.11	0.13	0.11	0.11	0.09	0.1	0.12	0.5	0.11
M047598	0.23	0.21	0.24	0.23	0.25	0.21	0.26	0.11	0.48

Seems like he is the half-sibling of all the other kids (which are the kids we identified as the dad's other kids btw). What about his mom? Did we sequence her and is he related to her? (Probably since he's half sibling with the rest of the clutch but let's check).

```r
"M043628" %in% rownames(kas)
```
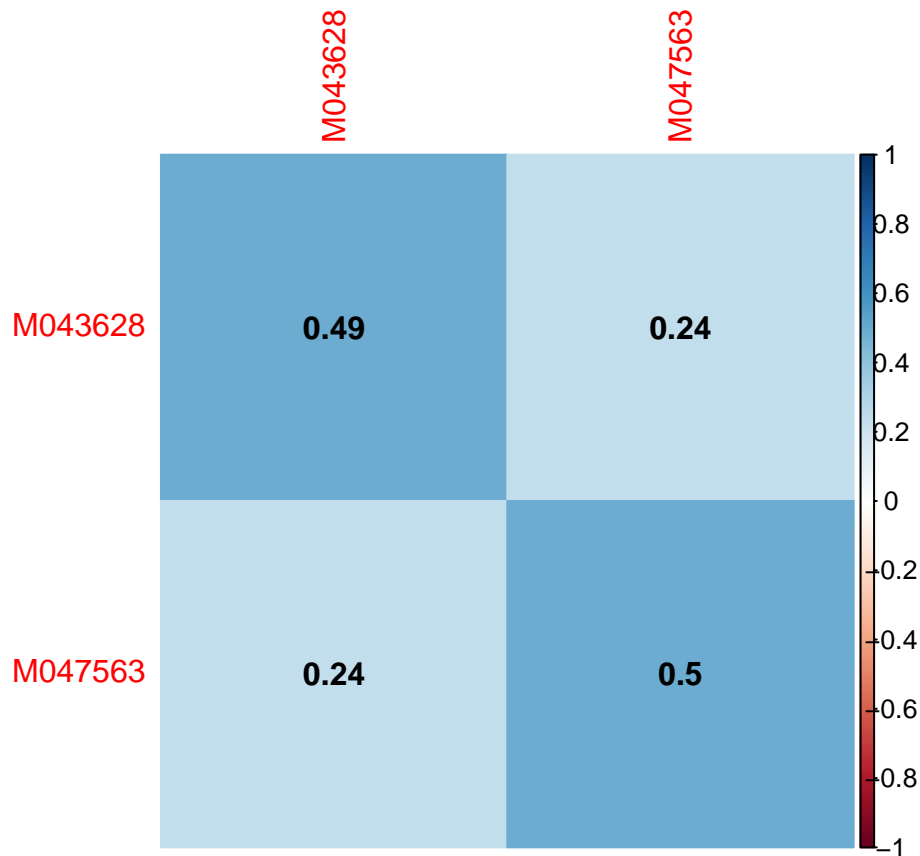
```
## [1] TRUE
```

```r
#correlation plot
corrplot(as.matrix(kas[row.names(kas) %in% c(high$RingId1[71], "M043628"),
                   row.names(kas) %in% c(high$RingId1[71], "M043628")]),
        method = 'color', addCoef.col = 'black')
```

He's related to the mom! Can we try to identify who the dad might be by looking for new links?

```r
#new links?
row.names(kas)[which(kas[high$row[71],] > 0.16)]
```

```
## [1] "M043628" "M047563"
```

```r
#genomic kinship
kas[rownames(kas) == high$RingId1[71],
    colnames(kas) == "M043628"]
```

```
## [1] 0.2403869
```

```r
#who are they?
ped[ped$animal == "M043628",]
```

```
##        animal sire  dam sex
## 9569 M043628 <NA> <NA>   2
```

He's related to someone new. This guy could be the true father.

EXTRA PAIR COPULATION AGAIN ????

CHECK with ALEX !!!!

Round 2

After visual checking, we noticed that it would be great to also check pairs for which we have more than 0.1 difference between (corrected) pedigree and new (after individuals trimming) AS genomic kinship. So we'll read the new genomic matrix, check difference with the corrected pedigree and start round 2 of correction:

```r
#read new AS after correction
grmAS = readRDS("~/PhD/Barn_owl/ID_in3Kowls/data/All3085_AUTOSAUMES_RP502SNPs.RDS")
kasCORRECTED = grm2kinship(grmAS)


## PEDIGREE

pedCORRECTED = ped

## group 1

#the father is not the father
pedCORRECTED$sire[pedCORRECTED$animal %in% c(high$RingId1[1], high$RingId1[2], high$RingId1[3])] = "M026

## group 2

#both parents are not the parents (probably cross-fostering errors)
pedCORRECTED$sire[pedCORRECTED$animal == high$RingId2[4]] = NA
pedCORRECTED$dam[pedCORRECTED$animal == high$RingId2[4]] = NA

## group 4

#the dad is not the dad
pedCORRECTED$sire[pedCORRECTED$animal == high$RingId1[7]] = NA
pedCORRECTED$sire[pedCORRECTED$animal == high$RingId1[8]] = NA
pedCORRECTED$sire[pedCORRECTED$animal == high$RingId1[9]] = NA
pedCORRECTED$sire[pedCORRECTED$animal == high$RingId1[10]] = NA
pedCORRECTED$sire[pedCORRECTED$animal == high$RingId1[11]] = NA

#but he's the dad from some other clutches
pedCORRECTED$sire[pedCORRECTED$animal == "M026834"] = "M032002"
pedCORRECTED$sire[pedCORRECTED$animal == "M026836"] = "M032002"
pedCORRECTED$sire[pedCORRECTED$animal == "M026837"] = "M032002"
pedCORRECTED$sire[pedCORRECTED$animal == "M032031"] = "M032002"

## group 6

#the dad is not the dad
#the siblings are actually half-siblings
pedCORRECTED$sire[pedCORRECTED$animal == high$RingId1[12]] = NA

## group 9

#the dad is not the dad
pedCORRECTED$sire[pedCORRECTED$animal %in% high$RingId1[16:19]] = "M022673"

## group 10

#DATABASE comment: the mom is not the mom
pedCORRECTED$dam[pedCORRECTED$animal %in% high$RingId2[20:22]] = "M038110"

## group 12

#the dad is not the dad
```

```r
pedCORRECTED$sire[pedCORRECTED$animal == high$RingId2[36]] = NA

## group 13

#the mom is not the mom
pedCORRECTED$dam[pedCORRECTED$animal %in% high$RingId1[38:42]] = NA

## group 15

#the dad is not the dad
pedCORRECTED$sire[pedCORRECTED$animal %in% high$RingId1[46:47]] = "M032449"

## group 17

#extra pair copulation, we found the true father:
pedCORRECTED$sire[pedCORRECTED$animal %in% high$RingId1[51:53]] = NA

### FORMAT PEDIGREE

colnames(pedCORRECTED) = c("id","dadid","momid", "sex")
#sex NA to 3
pedCORRECTED$sex[is.na(pedCORRECTED$sex)] = 3
#if only one parent know -> both NA
pedCORRECTED$dadid[is.na(pedCORRECTED$momid)] = NA
pedCORRECTED$momid[is.na(pedCORRECTED$dadid)] = NA

#pass to pedigree object
pedCORRECTED.2 = kinship2::pedigree(id = pedCORRECTED$id, dadid = pedCORRECTED$dadid, momid = pedCORRECT
#recalculate pedigree kinship
PedKinCORRECTED = kinship2::kinship(pedCORRECTED.2)

#Subset individuals in kasCORRECTED matrix
PedKinCORRECTEDsub = PedKinCORRECTED[rownames(PedKinCORRECTED) %in% rownames(kasCORRECTED),
                                     colnames(PedKinCORRECTED) %in% colnames(kasCORRECTED)]


#order so that individuals' order matches AS kinship matrix order
orderCOL = match(colnames(kasCORRECTED), colnames(PedKinCORRECTEDsub))
PedKinCORRECTEDsubOrder = PedKinCORRECTEDsub[orderCOL,orderCOL]

#sanity check
unique(colnames(kasCORRECTED) == colnames(PedKinCORRECTEDsubOrder))
```

```
## [1] TRUE
```

```r
unique(rownames(kasCORRECTED) == rownames(PedKinCORRECTEDsubOrder))
```

```
## [1] TRUE
```

Just like before we'll identify pairs for which the difference is above 0.1:

```r
#### Recalculate difference
diff_kinCORRECTEDmat = as.matrix(PedKinCORRECTEDsubOrder) - as.matrix(kasCORRECTED)
diff_kinCORRECTEDmat.2 = diff_kinCORRECTEDmat
diff_kinCORRECTEDmat.2[upper.tri(diff_kinCORRECTEDmat.2)] = NA
```

```r
diag(diff_kinCORRECTEDmat.2) = NA

#identify weird pairs
WeirdHighKin <- as.data.frame(which(diff_kinCORRECTEDmat.2 > 0.1, arr.ind=T))
rownames(WeirdHighKin) <- NULL
WeirdHighKin$dif<- NA
WeirdHighKin$p <- NA
WeirdHighKin$b <- NA

#Loop through difference to add indv RingId
for(i in 1:nrow(WeirdHighKin)){
  WeirdHighKin$dif[i] <- diff_kinCORRECTEDmat.2[WeirdHighKin$row[i],WeirdHighKin$col[i]]
  WeirdHighKin$p[i] <- kasCORRECTED[WeirdHighKin$row[i],WeirdHighKin$col[i]]
  WeirdHighKin$b[i] <- PedKinCORRECTEDsubOrder[WeirdHighKin$row[i],WeirdHighKin$col[i]]
}

#Set row.names and col.names
WeirdHighKin$RingId1 <- row.names(diff_kinCORRECTEDmat.2)[WeirdHighKin$row]
WeirdHighKin$RingId2 <- colnames(diff_kinCORRECTEDmat.2)[WeirdHighKin$col]

length(unique(c(WeirdHighKin$row,WeirdHighKin$col)))
```

```
## [1] 23
```

group 22

The group is composed of four individuals: 898166 and three of his half-siblings: 877125, 877175, 877177. Let's see what the kinship matrix looks like:

```r
listofFAM = c(WeirdHighKin$RingId1[1], WeirdHighKin$RingId2[1:3])

#correlation plot
corrplot(as.matrix(kasCORRECTED[row.names(kasCORRECTED) %in% listofFAM,
                  row.names(kasCORRECTED) %in% listofFAM ]),
         method = 'color',addCoef.col = 'black')
```
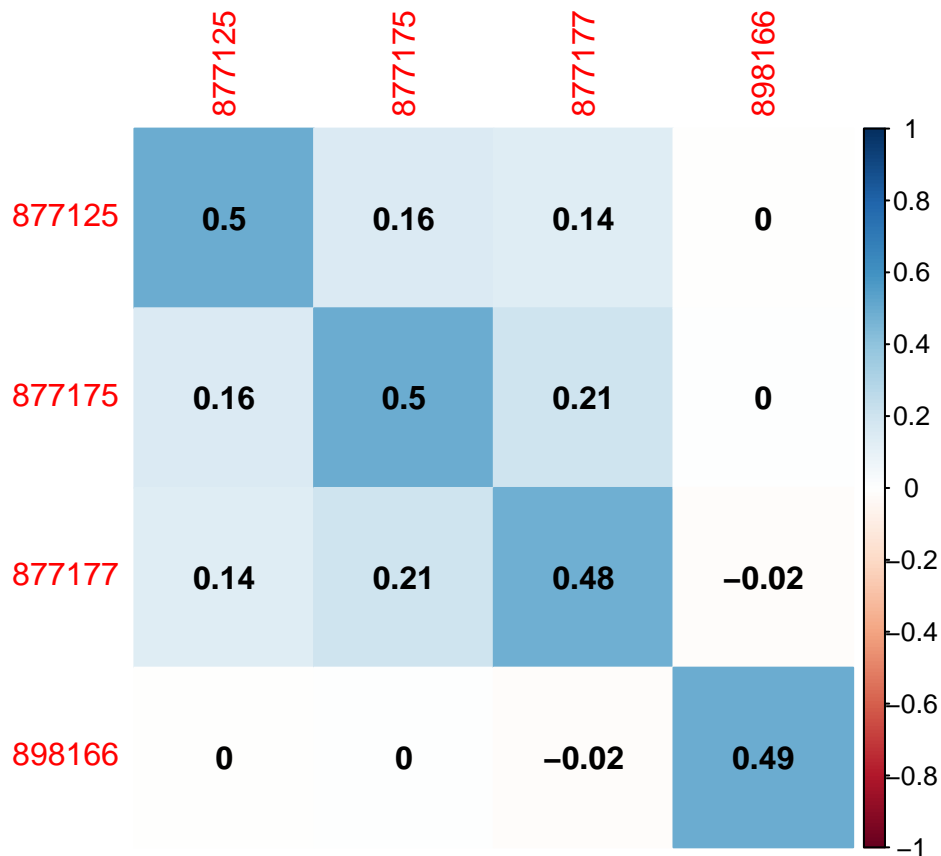
877125, 877175, 877177 look good, their kinship matches what we expect from siblings and hal-siblings, only 898166 is not related to any of them! What about their dad (or supposed dad for 898166): did we sequence him ?

```
pedCORRECTED[pedCORRECTED$id == "898166",]
```

```
##           id  dadid  momid sex
## 3865 898166 866268 889528    2
```

```
#did we sequence his dad?
"866268" %in% rownames(kasCORRECTED)
```

```
## [1] FALSE
```

Nop, let's look for his mom, siblings or kids to confirm his identity:

```
#did we sequenced his mom?
"889528" %in% rownames(kasCORRECTED)
```

```
## [1] FALSE
```

```
#no

#any siblings?
pedCORRECTED[(pedCORRECTED$momid == "889528") & (!is.na(pedCORRECTED$momid)) &
             (pedCORRECTED$dadid == "866268") & (!is.na(pedCORRECTED$dadid)),]
```

```
##           id  dadid  momid sex
## 3861 898162 866268 889528    1
## 3862 898163 866268 889528    2
## 3863 898164 866268 889528    2
```

```
## 3864 898165 866268 889528   1
## 3865 898166 866268 889528   2
## 3866 898167 866268 889528   1
## 4296 899066 866268 889528   1
## 4297 899067 866268 889528   2
## 4298 899068 866268 889528   2
## 4299 899069 866268 889528   1
## 4300 899070 866268 889528   2
## 4301 899071 866268 889528   2
```

```r
#did we sequence them?
"898162" %in% rownames(kasCORRECTED)
```

```
## [1] FALSE
```

```r
"898163" %in% rownames(kasCORRECTED)
```

```
## [1] FALSE
```

```r
"898164" %in% rownames(kasCORRECTED)
```

```
## [1] FALSE
```

```r
"898165" %in% rownames(kasCORRECTED)
```

```
## [1] FALSE
```

```r
"898167" %in% rownames(kasCORRECTED)
```

```
## [1] FALSE
```

```r
"899066" %in% rownames(kasCORRECTED)
```

```
## [1] FALSE
```

```r
"899067" %in% rownames(kasCORRECTED)
```

```
## [1] FALSE
```

```r
"899068" %in% rownames(kasCORRECTED)
```

```
## [1] FALSE
```

```r
"899069" %in% rownames(kasCORRECTED)
```

```
## [1] FALSE
```

```r
"899070" %in% rownames(kasCORRECTED)
```

```
## [1] FALSE
```

```r
"899071" %in% rownames(kasCORRECTED)
```

```
## [1] FALSE
```

```r
#We did not sequence any of them...

#Any kids ?
pedCORRECTED[(pedCORRECTED$momid == "898166") & (!is.na(pedCORRECTED$momid)),]
```

```
##           id  dadid  momid sex
## 4750 M005328 889045 898166   1
## 4751 M005329 889045 898166   1
## 4752 M005330 889045 898166   2
```

```
## 4753 M005331 889045 898166     2
## 4754 M005332 889045 898166     2
```

```r
#did we sequence them?
"M005328" %in% rownames(kasCORRECTED)
```

```
## [1] FALSE
```

```r
"M005329" %in% rownames(kasCORRECTED)
```

```
## [1] FALSE
```

```r
"M005330" %in% rownames(kasCORRECTED)
```

```
## [1] FALSE
```

```r
"M005331" %in% rownames(kasCORRECTED)
```

```
## [1] FALSE
```

```r
"M005332" %in% rownames(kasCORRECTED)
```

```
## [1] FALSE
```

We did not sequence any of his parents, siblings or kids which could be used to confirm his identity. So we'll have to remove that sample.
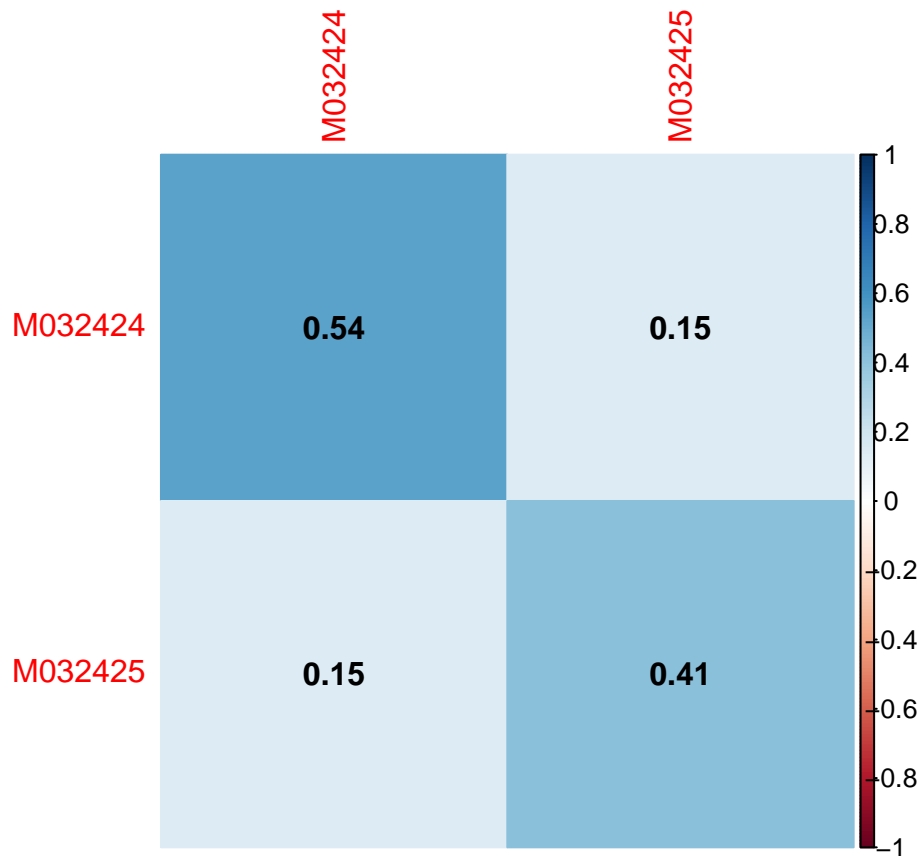
group 23

This group is composed of two supposed siblings: M032425 and M032424. Let's first check their genomic kinship:

```r
listofFAM = c(WeirdHighKin$RingId1[4], WeirdHighKin$RingId2[4])

#correlation plot
corrplot(as.matrix(kasCORRECTED[row.names(kasCORRECTED) %in% listofFAM,
                    row.names(kasCORRECTED) %in% listofFAM ]),
        method = 'color',addCoef.col = 'black')
```

From this kinship matrix, it looks like these guys are actually half-siblings. Let's see if we sequenced other siblings from this clutch, or their supposed parents:

```
pedCORRECTED[pedCORRECTED$id == WeirdHighKin$RingId1[4],]
```

```
##           id    dadid   momid sex
## 7202 M032425 M032203 M032228   2
```

```
pedCORRECTED[pedCORRECTED$id == WeirdHighKin$RingId2[4],]
```

```
##           id    dadid   momid sex
## 7201 M032424 M032203 M032228   2
```

```
#did we sequence the parents?
"M032203" %in% rownames(kasCORRECTED)
```

```
## [1] TRUE
```

```
"M032228" %in% rownames(kasCORRECTED)
```

```
## [1] TRUE
```

```
#any siblings?
pedCORRECTED[(pedCORRECTED$momid == "M032228") & (!is.na(pedCORRECTED$momid)) &
              (pedCORRECTED$dadid == "M032203") & (!is.na(pedCORRECTED$dadid)),]
```

```
##           id    dadid   momid sex
## 7070 M032283 M032203 M032228   1
## 7087 M032301 M032203 M032228   2
## 7088 M032302 M032203 M032228   2
## 7089 M032303 M032203 M032228   2
```

```
## 7090 M032304 M032203 M032228   2
## 7091 M032305 M032203 M032228   2
## 7199 M032422 M032203 M032228   2
## 7200 M032423 M032203 M032228   2
## 7201 M032424 M032203 M032228   2
## 7202 M032425 M032203 M032228   2
## 7203 M032426 M032203 M032228   2
## 7204 M032427 M032203 M032228   1
## 7205 M032428 M032203 M032228   2
```

```
#did we sequence them?
"M032283" %in% rownames(kasCORRECTED)
```

```
## [1] FALSE
```

```
"M032301" %in% rownames(kasCORRECTED)
```

```
## [1] FALSE
```

```
"M032302" %in% rownames(kasCORRECTED)
```

```
## [1] TRUE
```

```
"M032303" %in% rownames(kasCORRECTED)
```

```
## [1] TRUE
```

```
"M032304" %in% rownames(kasCORRECTED)
```

```
## [1] TRUE
```

```
"M032305" %in% rownames(kasCORRECTED)
```

```
## [1] TRUE
```

```
"M032422" %in% rownames(kasCORRECTED)
```

```
## [1] FALSE
```

```
"M032423" %in% rownames(kasCORRECTED)
```
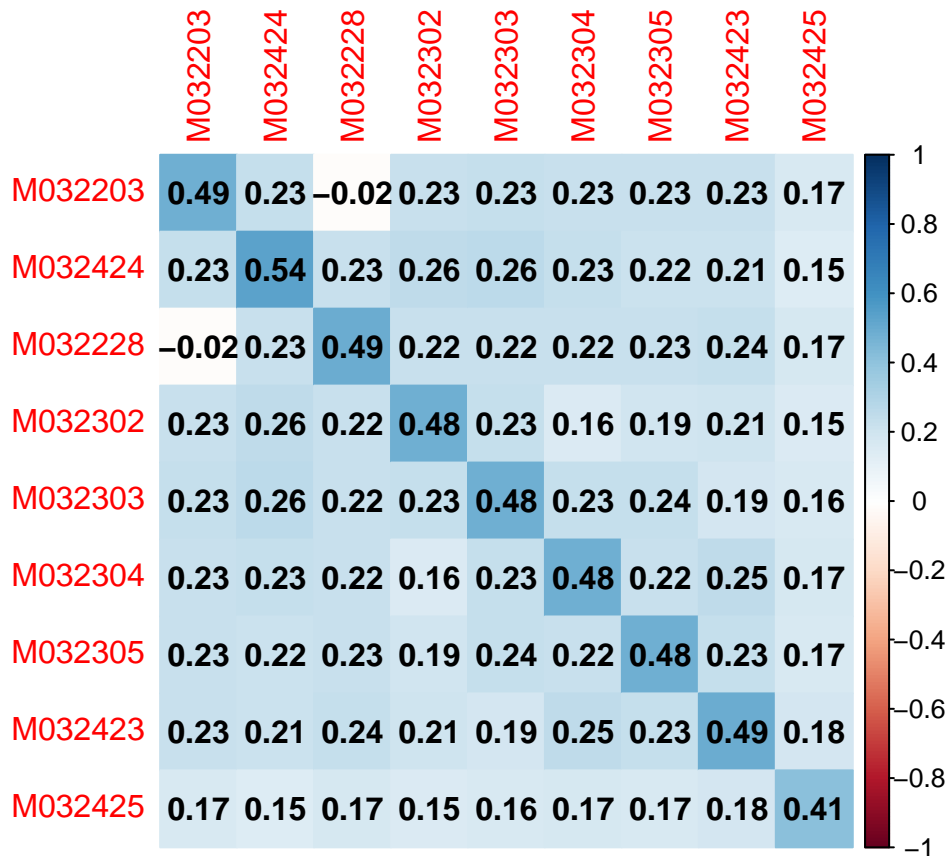
```
## [1] TRUE
```

Let's see the entire kinship matrix:

```
listofFAM = c("M032203", "M032228", "M032302", "M032303", "M032304",
              "M032305", "M032422", "M032423",
              WeirdHighKin$RingId1[4], WeirdHighKin$RingId2[4])

#correlation plot
corrplot(as.matrix(kasCORRECTED[row.names(kasCORRECTED) %in% listofFAM,
                    row.names(kasCORRECTED) %in% listofFAM ]),
        method = 'color',addCoef.col = 'black')
```

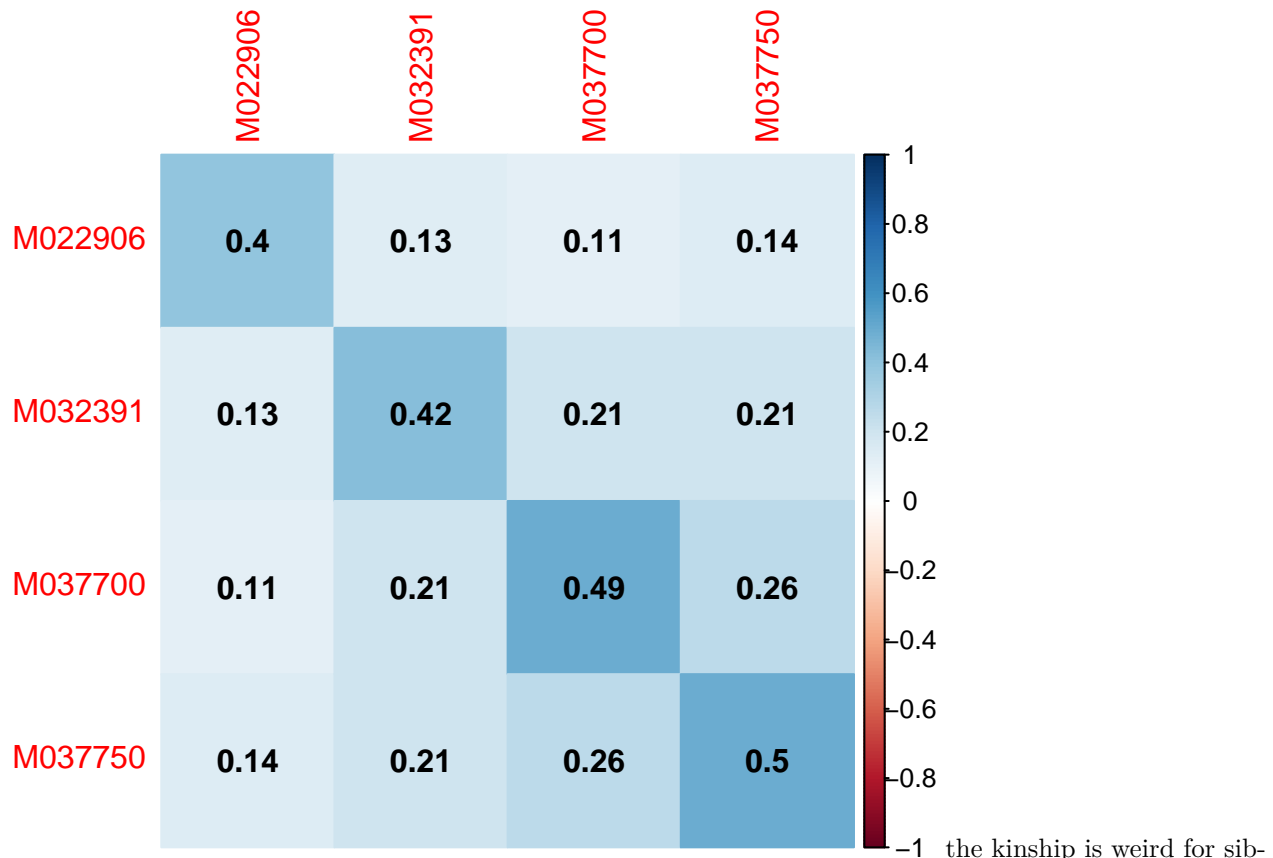|	M032203	M032424	M032228	M032302	M032303	M032304	M032305	M032423	M032425
M032203	**0.49**	**0.23**	**−0.02**	**0.23**	**0.23**	**0.23**	**0.23**	**0.23**	**0.17**
M032424	**0.23**	**0.54**	**0.23**	**0.26**	**0.26**	**0.23**	**0.22**	**0.21**	**0.15**
M032228	**−0.02**	**0.23**	**0.49**	**0.22**	**0.22**	**0.22**	**0.23**	**0.24**	**0.17**
M032302	**0.23**	**0.26**	**0.22**	**0.48**	**0.23**	**0.16**	**0.19**	**0.21**	**0.15**
M032303	**0.23**	**0.26**	**0.22**	**0.23**	**0.48**	**0.23**	**0.24**	**0.19**	**0.16**
M032304	**0.23**	**0.23**	**0.22**	**0.16**	**0.23**	**0.48**	**0.22**	**0.25**	**0.17**
M032305	**0.23**	**0.22**	**0.23**	**0.19**	**0.24**	**0.22**	**0.48**	**0.23**	**0.17**
M032423	**0.23**	**0.21**	**0.24**	**0.21**	**0.19**	**0.25**	**0.23**	**0.49**	**0.18**
M032425	**0.17**	**0.15**	**0.17**	**0.15**	**0.16**	**0.17**	**0.17**	**0.18**	**0.41**

Nothing seems wrong... Could it be because M032425 has low self kinship (0.41) which reduces its kinship with the rest of his family? It's coverage is 1.4145X.

pair 24

This group is composed of four individuals all siblings: M022906 and M032391, M037700, M037750. Let's first see the kinship matrix:

```
listofFAM = c(WeirdHighKin$RingId2[5], WeirdHighKin$RingId1[5:7])

#correlation plot
corrplot(as.matrix(kasCORRECTED[row.names(kasCORRECTED) %in% listofFAM,
                  row.names(kasCORRECTED) %in% listofFAM ]),
        method = 'color',addCoef.col = 'black')
```

the kinship is weird for siblings. Let's see if we sequenced other siblings from this clutch, or their supposed parents:

```
pedCORRECTED[pedCORRECTED$id == WeirdHighKin$RingId2[5],]
```

```
##          id    dadid    momid sex
## 6017 M022906 M028883 M022875   2
```

```
pedCORRECTED[pedCORRECTED$id %in% WeirdHighKin$RingId1[5:7],]
```

```
##          id    dadid    momid sex
## 7172 M032391 M028883 M022875   1
## 8415 M037700 M028883 M022875   2
## 8461 M037750 M028883 M022875   1
```

```
#did we sequence the parents?
"M028883" %in% rownames(kasCORRECTED) #no
```

```
## [1] FALSE
```

```
"M022875" %in% rownames(kasCORRECTED) #yes
```

```
## [1] TRUE
```

```
#any siblings?
pedCORRECTED[(pedCORRECTED$momid == "M022875") & (!is.na(pedCORRECTED$momid)) &
             (pedCORRECTED$dadid == "M028883") & (!is.na(pedCORRECTED$dadid)),]
```

```
##          id    dadid    momid sex
## 6012 M022901 M028883 M022875   2
## 6017 M022906 M028883 M022875   2
## 7068 M032281 M028883 M022875   2
```

```
## 7069 M032282 M028883 M022875   1
## 7172 M032391 M028883 M022875   1
## 7173 M032392 M028883 M022875   2
## 7174 M032393 M028883 M022875   2
## 8414 M037699 M028883 M022875   1
## 8415 M037700 M028883 M022875   2
## 8461 M037750 M028883 M022875   1
## 8860 M038223 M028883 M022875   2
## 8861 M038224 M028883 M022875   1
```

```r
#did we sequence them?
"M022901" %in% rownames(kasCORRECTED)
```

```
## [1] TRUE
```

```r
"M022906" %in% rownames(kasCORRECTED)
```

```
## [1] TRUE
```

```r
"M032281" %in% rownames(kasCORRECTED)
```

```
## [1] FALSE
```

```r
"M032282" %in% rownames(kasCORRECTED)
```

```
## [1] FALSE
```

```r
"M032392" %in% rownames(kasCORRECTED)
```

```
## [1] FALSE
```

```r
"M032393" %in% rownames(kasCORRECTED)
```

```
## [1] TRUE
```

```r
"M037699" %in% rownames(kasCORRECTED)
```

```
## [1] TRUE
```

```r
"M038223" %in% rownames(kasCORRECTED)
```

```
## [1] TRUE
```

```r
"M038224" %in% rownames(kasCORRECTED)
```
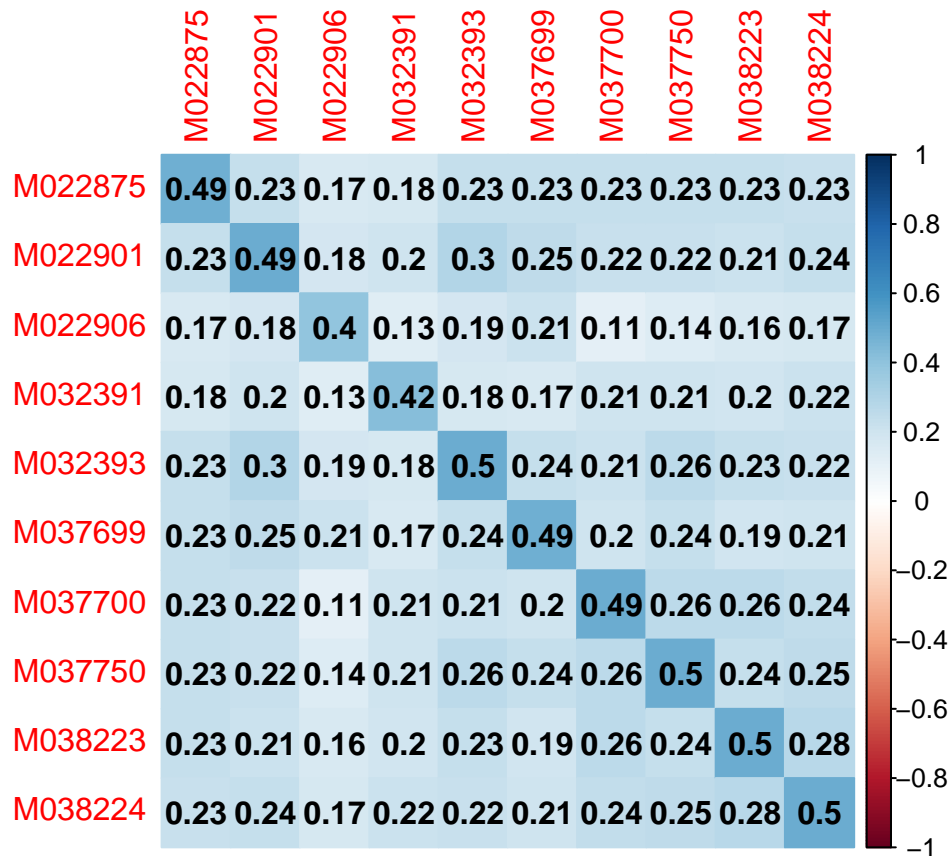
```
## [1] TRUE
```

Let's see the entire kinship matrix:

```r
listofFAM = c("M028883", "M022875", "M022901", "M022906", "M032393",
              "M037699", "M038223", "M038224",
              WeirdHighKin$RingId2[5], WeirdHighKin$RingId1[5:7])

#correlation plot
corrplot(as.matrix(kasCORRECTED[row.names(kasCORRECTED) %in% listofFAM,
                  row.names(kasCORRECTED) %in% listofFAM ]),
         method = 'color',addCoef.col = 'black')
```

Same as before, could it be because M022906 has low self kinship (0.40) which reduces its kinship with the rest of his family? It's coverage was 2.0392X.

pair 25

This group is composed of a male: M026444 and his mom: M026258. Let's first see the kinship matrix:

```
pedCORRECTED[pedCORRECTED$id == WeirdHighKin$RingId1[8], ]
```

```
##           id    dadid   momid sex
## 6326 M026444 M012595 M026258   1
```

```
pedCORRECTED[pedCORRECTED$id == WeirdHighKin$RingId2[8], ]
```
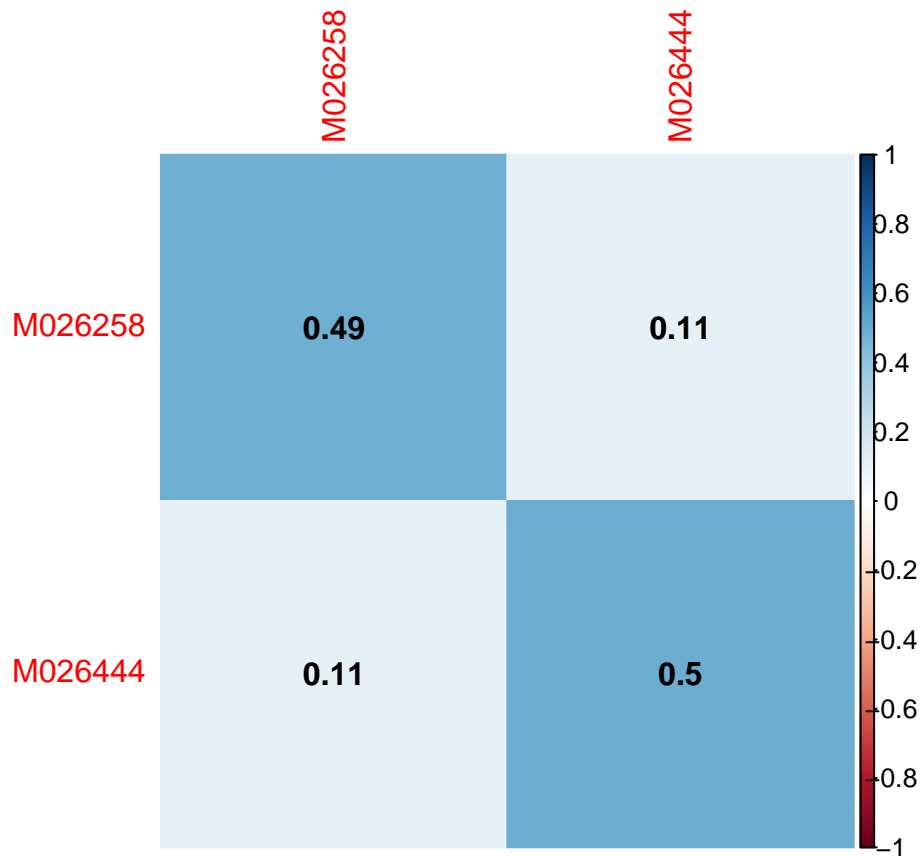
```
##           id    dadid   momid sex
## 6157 M026258 M022393 M022674   2
```

```
listofFAM = c(WeirdHighKin$RingId1[8], WeirdHighKin$RingId2[8])

#correlation plot
corrplot(as.matrix(kasCORRECTED[row.names(kasCORRECTED) %in% listofFAM,
                  row.names(kasCORRECTED) %in% listofFAM ]),
        method = 'color',addCoef.col = 'black')
```

They should be mom and son but their kinship is lower than that...

See what you can do here but I am not sure maybe check the dad ??

```
pedCORRECTED[pedCORRECTED$id == WeirdHighKin$RingId1[8], ]
```

```
##          id   dadid   momid sex
## 6326 M026444 M012595 M026258   1
```

```
pedCORRECTED[pedCORRECTED$id == WeirdHighKin$RingId2[8], ]
```

```
##          id   dadid   momid sex
## 6157 M026258 M022393 M022674   2
```

```
#did we sequence the parents?
"M012595" %in% rownames(kasCORRECTED)
```

```
## [1] TRUE
```

```
"M026258" %in% rownames(kasCORRECTED)
```

```
## [1] TRUE
```
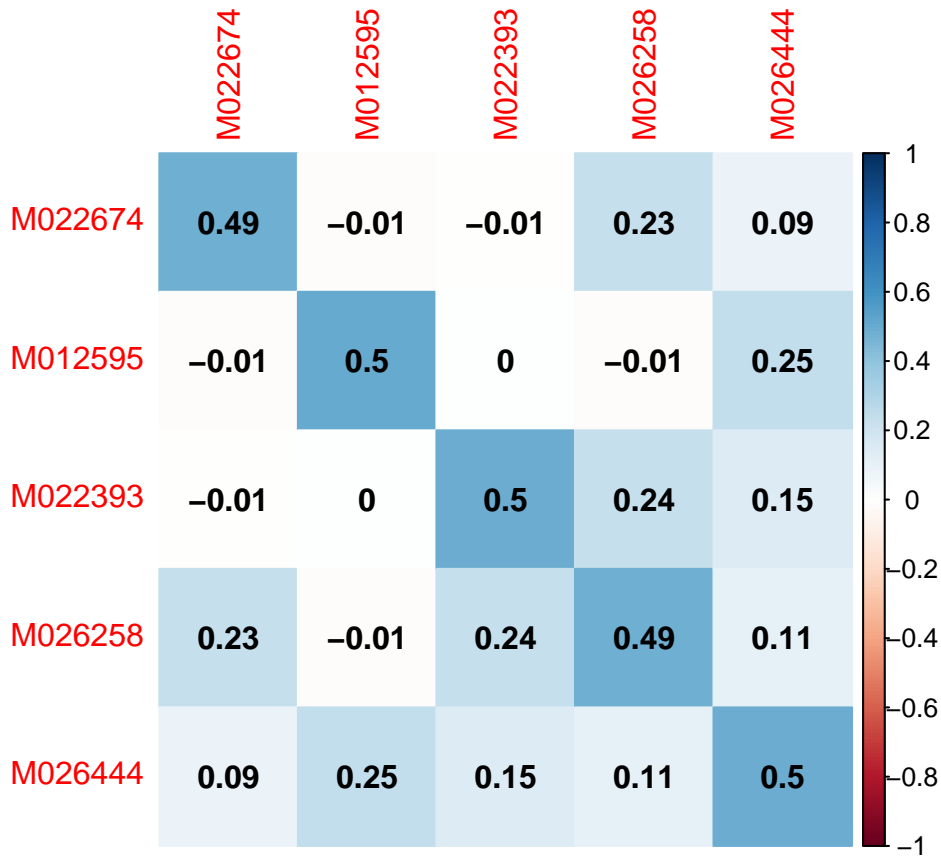
```
"M022393" %in% rownames(kasCORRECTED)
```

```
## [1] TRUE
```

```
"M022674" %in% rownames(kasCORRECTED)
```

```
## [1] TRUE
```

```
listofFAM = c(WeirdHighKin$RingId1[8], WeirdHighKin$RingId2[8],
              "M012595", "M022393", "M022674")

#correlation plot
corrplot(as.matrix(kasCORRECTED[row.names(kasCORRECTED) %in% listofFAM,
                   row.names(kasCORRECTED) %in% listofFAM ]),
         method = 'color',addCoef.col = 'black')
```


Okay so the the mom of