APIS: An Auto-Adaptive Parentage Inference Software tolerant to missing parents

Ronan GRIOT

2019-02-11

Description

This package include all the functions to assign with APIS (Griot & al., 2019). Parentage assignment is widely used for farmed and natural populations. As most of the likelihood software are based on simulation, the estimation of the simulation parameters is a key point for assignment reliability. Among those parameters, the proportion of missing parent is one of the most important. To avoid estimation of missing parents, we developed APIS (Auto-Adaptive Parentage Inference Software), based on observed average Mendelian transmission probabilities.

Install and load the package

```
# devtools::install_github("rgriot/APIS")
require(APIS)
#> Loading required package: APIS
#> Loading required package: doParallel
#> Warning: package 'doParallel' was built under R version 3.4.4
#> Loading required package: foreach
#> Warning: package 'foreach' was built under R version 3.4.3
#> Loading required package: iterators
#> Warning: package 'iterators' was built under R version 3.4.4
#> Loading required package: parallel
```

Formate your data

APIS requires matrices of characters. Each matrix has individuals as rows, markers as columns. The individual labels are set as rownames. Each cell is the genotype of one marker, coded "All1/All2". For example "A/A", "A/B", "B/B" for bi-allelic markers and "NA/NA" for missing value.

```
data("genotype_APIS")

head(off[,1:10])
#>       V2  V3  V4  V5  V6  V7    V8  V9 V10 V11
#> 100  A/B B/B A/A A/A B/B B/B NA/NA A/A A/A A/B
#> 1000 A/B B/B A/B A/A B/B B/B   B/B A/A A/A A/B
#> 1001 A/B A/B A/B A/A B/B B/B   B/B A/A A/A A/A
#> 1002 B/B B/B A/A A/A B/B B/B   B/B A/A A/B A/A
#> 1003 B/B A/B A/B A/A B/B B/B   B/B A/A A/B A/B
#> 1004 A/B B/B A/A A/A B/B B/B   B/B A/A B/B A/B
rownames(off[1:6,])
#> [1] "100"  "1000" "1001" "1002" "1003" "1004"
```

Prepare the inputs

APIS main function requires 4 inputs :

- off.genotype = matrix of offspring genotypes
- sire.genotype = matrix of sires genotypes
- dam.genotype = matrix of dams genotypes
- error = accepted assignment error rate (What is the maximum error rate I accept in my assignment results ?)

```
error <- 0.05 #I accept 5% of errors in the results
```

Running the assignment

The main function to perform parentage assignment with APis the "APIS" function

```
# result <- APIS(off.genotype = off.genotype,
#                sire.genotype = sire.genotype,
#                dam.genotype = dam.genotype,
#                error = error)
```

Analyse the results

APIS gives you 3 different outputs :

- pedigree
- log containing Mendelian transmission probabilities for the first 3 parent pairs
- graphs

Following graphs were drawn from an example dataset. The Mendelian transmission probabilities values are higher than a real-world example. However, the distribution shapes are realistic.

When all parents are present, the distribution of P1 has one mode and the distribution of P2 also has one mode

When all parents are missing, the distribution of P1 has two modes :

- The one with the highest values representing offspring with both parents present
- The one with the lowest values, under the P2 distribution, representing the offspring with missing pareny

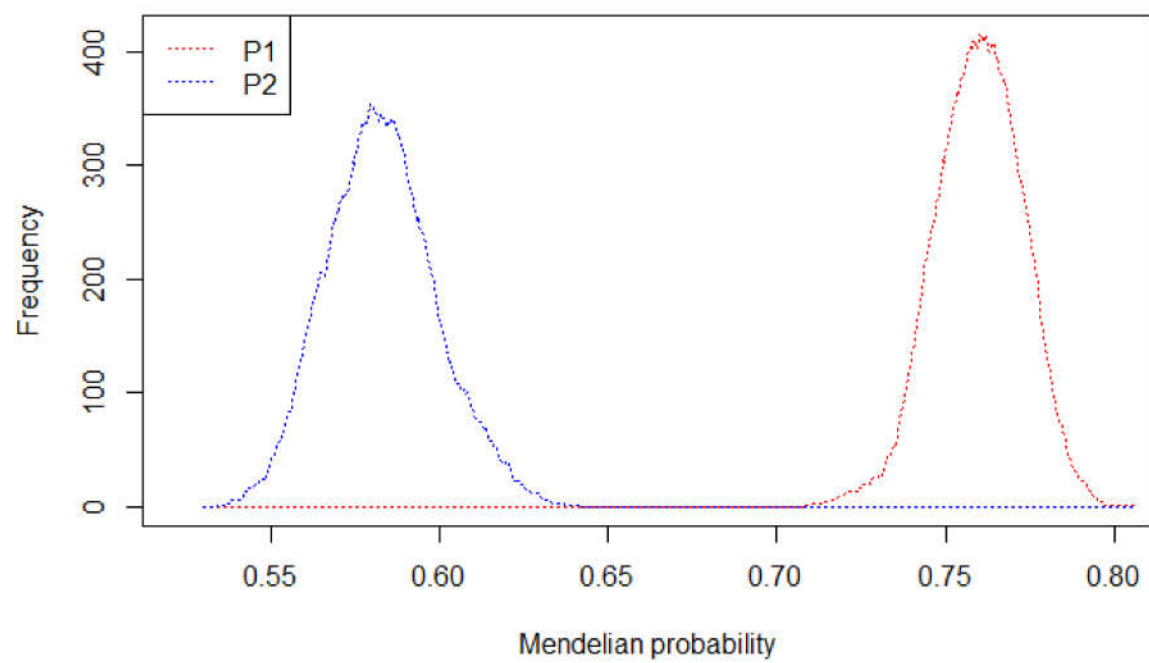and the distribution of P2 also has one mode.

Annexe

Figure 1: Average Mendelian transmission probability distributions calculated on 1000 offspring from 10 sires and 10 dams, genotyped on 400 markers
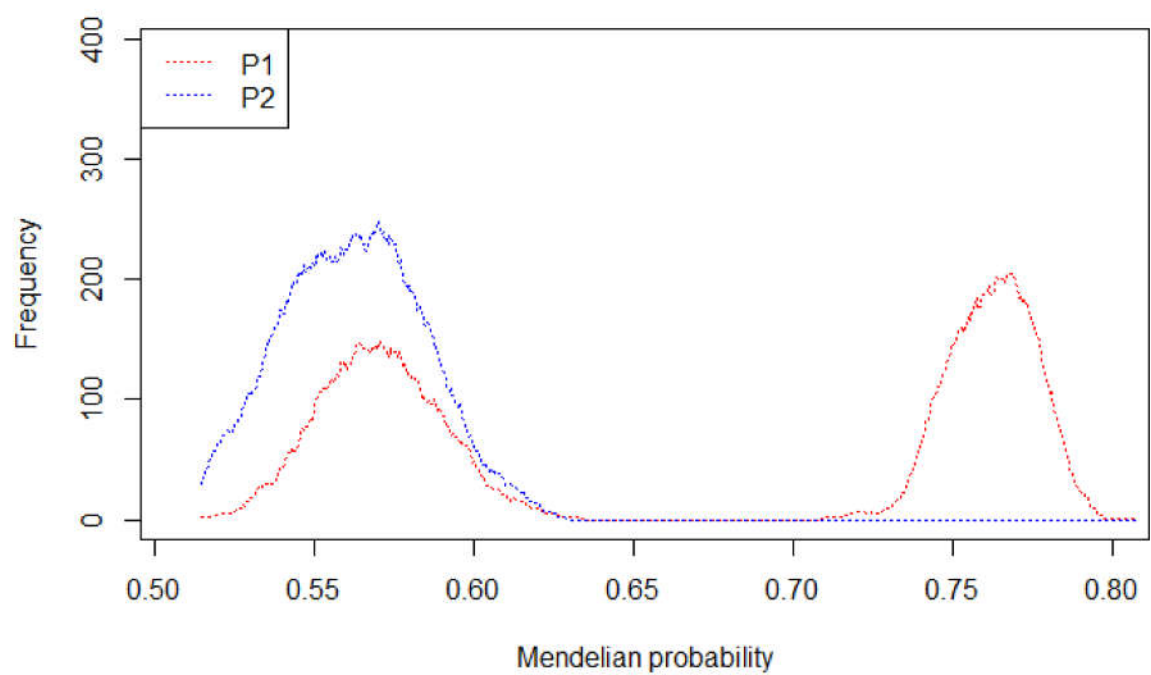
Figure 2: Average Mendelian transmission probability distributions calculated on 1000 offspring from 5 sires (+ 5 sires missing) and 10 dams, genotyped on 400 markers

```
print(sessionInfo(), locale=FALSE)
#> R version 3.4.1 (2017-06-30)
#> Platform: x86_64-w64-mingw32/x64 (64-bit)
#> Running under: Windows 7 x64 (build 7601) Service Pack 1
#>
#> Matrix products: default
#>
#> attached base packages:
#> [1] parallel   stats     graphics  grDevices utils     datasets  methods
#> [8] base
#>
#> other attached packages:
#> [1] APIS_0.1.0       doParallel_1.0.14 iterators_1.0.10  foreach_1.4.4
#>
#> loaded via a namespace (and not attached):
#>  [1] Rcpp_0.12.17      codetools_0.2-15 digest_0.6.18     magrittr_1.5
#>  [5] evaluate_0.10.1   stringi_1.1.7     rmarkdown_1.11    tools_3.4.1
#>  [9] stringr_1.3.1     yaml_2.1.19       compiler_3.4.1    htmltools_0.3.6
#> [13] knitr_1.20
```