



HAL
open science

Génotypage de triploïdes à partir de puces à ADN et Assignment de parenté

Julien Roche, Mathieu Besson, Pierre Patrice, Pierrick Haffray, François Allal,
Marc Vandeputte, Florence Phocas

► To cite this version:

Julien Roche, Mathieu Besson, Pierre Patrice, Pierrick Haffray, François Allal, et al.. Génotypage de triploïdes à partir de puces à ADN et Assignment de parenté. Journées Techniques Interfilières SYSAAF, Oct 2023, Rennes, France. hal-04251343

HAL Id: hal-04251343

<https://hal.inrae.fr/hal-04251343>

Submitted on 20 Oct 2023

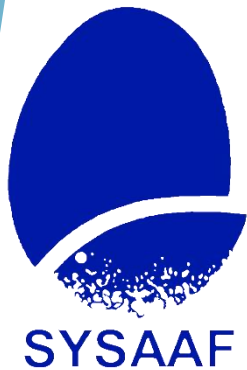
HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Génotypage de triploïdes à partir de puces à ADN et Assignation de parenté

Projet FEAMP HypoTemp

Julien Roche, Mathieu Besson, Pierre Patrice, Pierrick Haffray, François Allal, Marc Vandeputte, Florence Phocas



INRAE



GENTYANE
PLATE-FORME de
GÉNOTYPAGE et SÉQUENÇAGE en AUVERGNE

En 2020 : peu d'études publiées sur l'analyse génétique des polyploïdes

- Une étude américaine chez la truite arc-en-ciel montre **l'efficacité d'une sélection sur individus diploïdes pour améliorer les performances de triploïdes** ; la prise en compte de phénotypes de triploïdes améliorerait toutefois l'efficacité de la sélection (Leeds et al., 2019).
- **Corrélation génétique $2n/3n$ de 0.79 ± 0.05** du poids chez le saumon Atlantique (Kjøglum et al., 2019).
- Packages R *FitTetra* et *FitPoly* (Zych et al., 2019 - Wageningen)
- Méthode utilisant le package R *mclust* pour réaliser le clustering (Grashei et al., 2020)

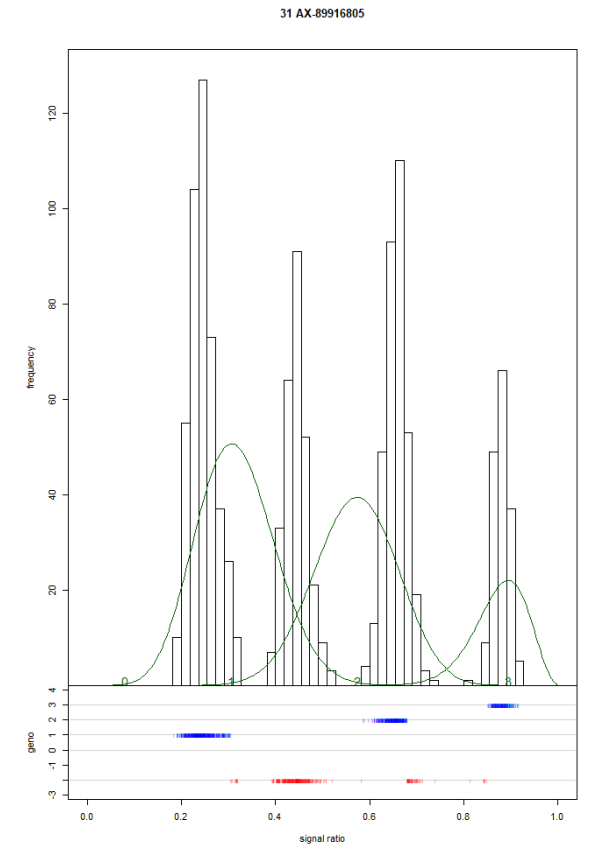
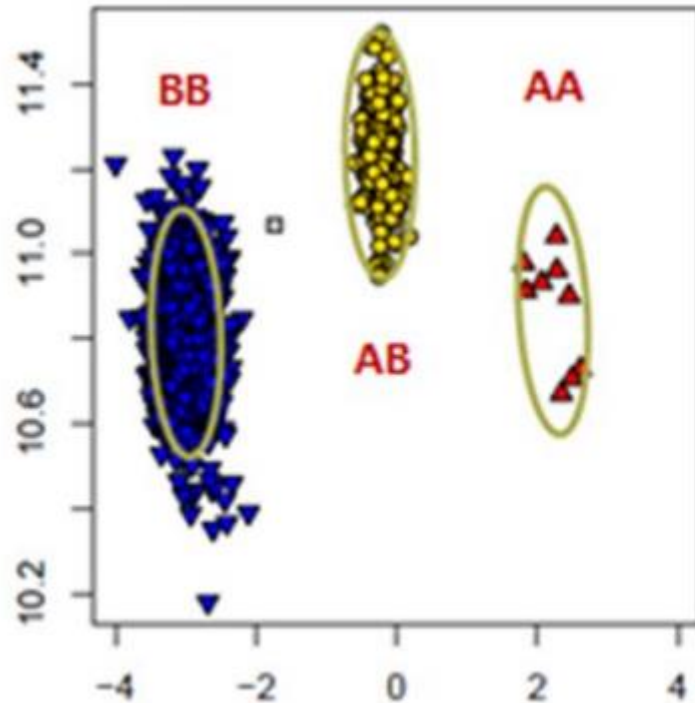


Image : fitPoly

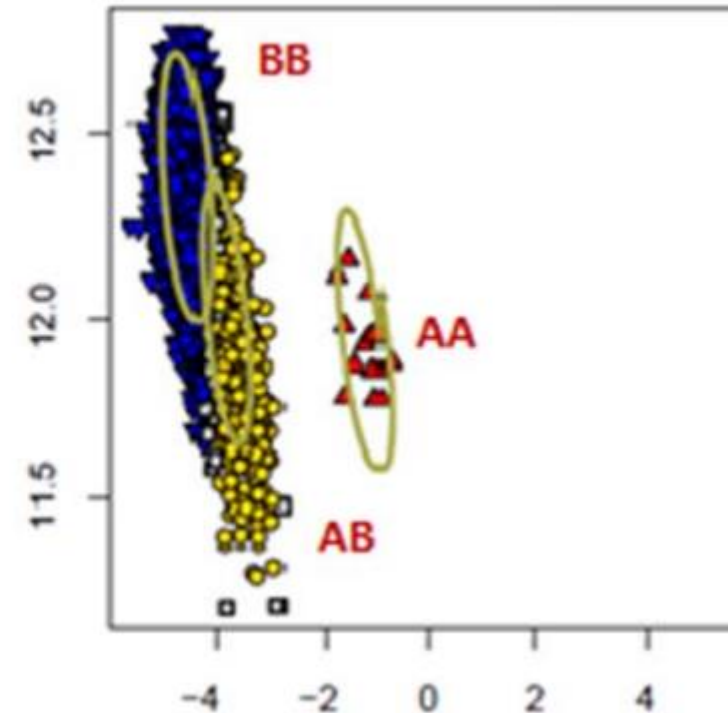
Le génotypage de diploïde : Axiom Analysis Suite (AXAS)

Bon génotypage d'un SNP diploïde



VS

Mauvais génotypage d'un SNP diploïde



Axiom Genotyping Solution Data Analysis Guide

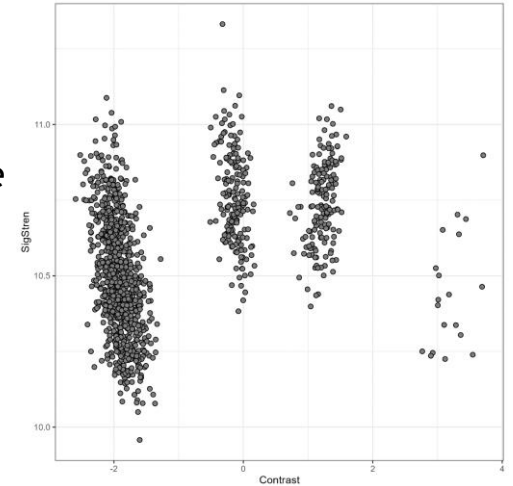
Diploïde : 2 chromosomes homologues → 2 allèles → 3 génotypes possibles (AA/AB/BB)

Triploïde : 3 chromosomes homologues → 3 allèles → 4 génotypes possibles (AAA/AAB/ABB/BBB)

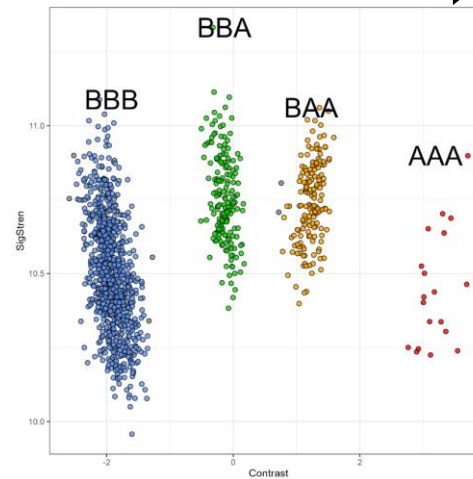
Sommaire

| | | | | | | |
|---------------|------------|------------|------------|------------|------------|------------|
| AX-89925273-A | 3647.50757 | 2166.61182 | 2300.07373 | 3123.87549 | 1658.13232 | 2462.44922 |
| AX-89925273-B | 969.42352 | 1539.82690 | 1679.93164 | 1006.64655 | 1508.46436 | 715.02588 |
| AX-89925274-A | 1100.43506 | 1442.76929 | 1214.14417 | 1659.19446 | 1076.88989 | 912.47174 |
| AX-89925274-B | 1836.24963 | 1346.85559 | 1834.99133 | 1252.59473 | 1808.80640 | 1681.55103 |
| AX-89925275-A | 1249.49304 | 1302.56702 | 1450.95996 | 1230.28479 | 1833.50183 | 912.04279 |
| AX-89925275-B | 1077.59521 | 1138.49060 | 572.14417 | 1069.59631 | 567.56757 | 1197.07886 |
| AX-89925276-A | 4307.21484 | 2228.09521 | 3891.22559 | 3371.29858 | 3470.70557 | 3344.26953 |
| AX-89925276-B | 660.15778 | 1416.87207 | 556.33911 | 515.86658 | 387.55569 | 608.75018 |
| AX-89925277-A | 4256.81885 | 4204.97705 | 4147.66699 | 2948.95337 | 2985.42041 | 3426.54419 |
| AX-89925277-B | 663.88446 | 645.76587 | 711.01196 | 675.90399 | 737.83990 | 762.83575 |
| AX-89925278-A | 401.44696 | 391.36746 | 809.20093 | 723.34503 | 420.85757 | 833.73065 |

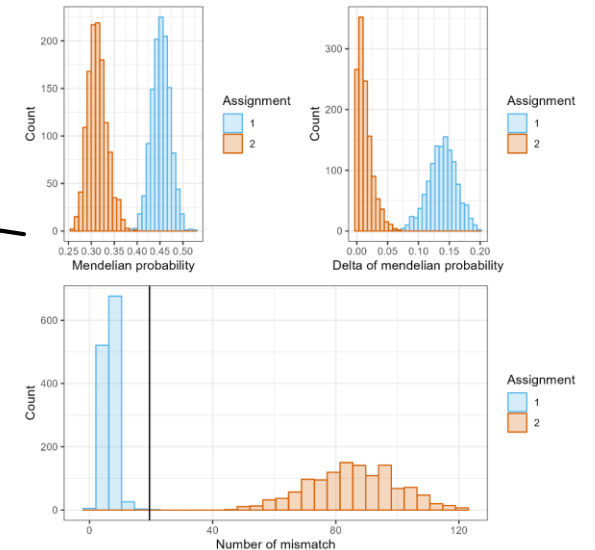
Traitement



Clustering & génotypage



Assignment de parenté : APIS



Pedigree

| offspring | sire | dam |
|-----------|------------|------------|
| EN0267701 | DEMO282569 | DEMO232961 |
| EN0267702 | DEMO313039 | DEMO232882 |
| EN0267703 | DEMO287496 | DEMO232768 |
| EN0267704 | DEMO287125 | DEMO287454 |
| EN0267705 | DEMO287170 | DEMO287354 |
| EN0267706 | DEMO282569 | DEMO232500 |
| EN0267707 | DEMO287253 | DEMO232419 |
| EN0267708 | DEMO287210 | DEMO232687 |
| EN0267709 | DEMO287210 | DEMO232687 |
| EN0267710 | DEMO287210 | DEMO232687 |

Sélection

- ▶ I. Génotypage de triploïdes
 - ▶ 1. Jeux de données
 - ▶ 2. Clustering
 - ▶ 3. Génotypage

- ▶ II. Assignment de parenté (triploid offspring)
 - ▶ 1. Matrices d'exclusion/de vraisemblance
 - ▶ 2. APIS 3n

I. Génotypage de triploïdes

1. Jeux de données

- ▶ Truites diploïdes & triploïdes de Bretagne-Truite
- ▶ Génotype de parents diploïdes (190 mères et 98 pères) - 57kSNP - AXAS
- ▶ Données brut d'AXAS pour 1232 descendants triploïdes - 38kSNP (57k réduit à 38k) :
 - ▶ Pour chaque marqueur et chaque individu : 2 valeurs S_A et S_B



Nom d'un individu → `probeset id` Demo267701 E21.CEL

| | |
|---------------|------------|
| AX-86908953-A | 805.31281 |
| AX-86908953-B | 1357.04797 |
| AX-86930057-A | 714.62933 |
| AX-86930057-B | 977.40704 |
| AX-86951728-A | 299.47256 |
| AX-86951728-B | 958.70886 |
| AX-86996661-A | 3134.04126 |
| AX-86996661-B | 906.70013 |
| AX-87024365-A | 1179.99280 |
| AX-87024365-B | 5087.75342 |
| AX-87082249-A | 2279.57764 |
| AX-87082249-B | 1900.69458 |
| AX-87133063-A | 1863.04016 |
| AX-87133063-B | 2303.40796 |

Exemples de différents couples de valeurs pour un individu

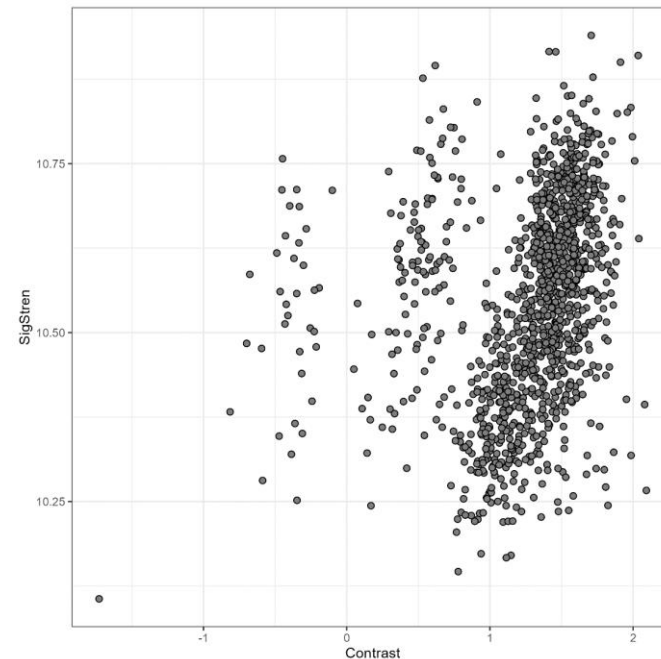
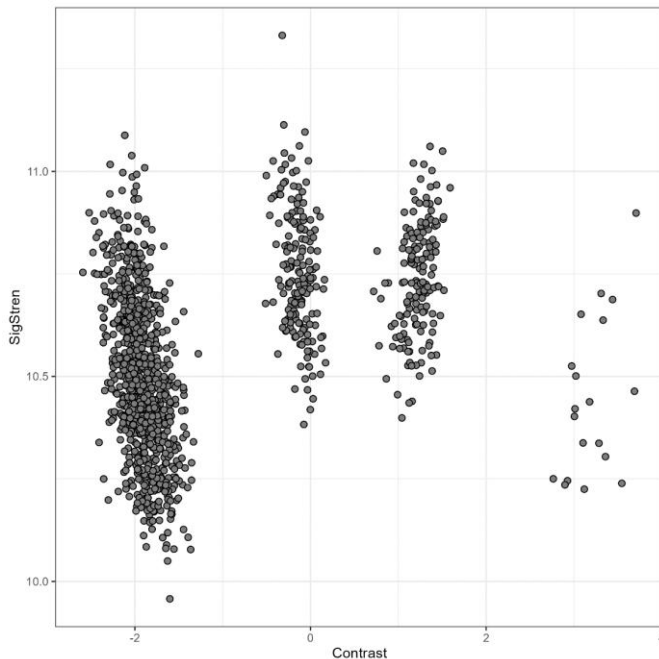
I. Génotypage de triploïdes

1. Jeux de données

- ▶ Objectif : Bonne discrimination des individus suivant leur génotype réel
- ▶ Création de variable (cf. logiciel de traitement des signaux de fluorescence AXAS - Axiom Analysis Suite) :

- ❖ $Contrast = \log_2\left(\frac{S_A}{S_B}\right)$

- ❖ $Signal\ Strength = \frac{\log_2(S_A) + \log_2(S_B)}{2}$



I. Génompage de triploïdes

2. Clustering

► Objectif :

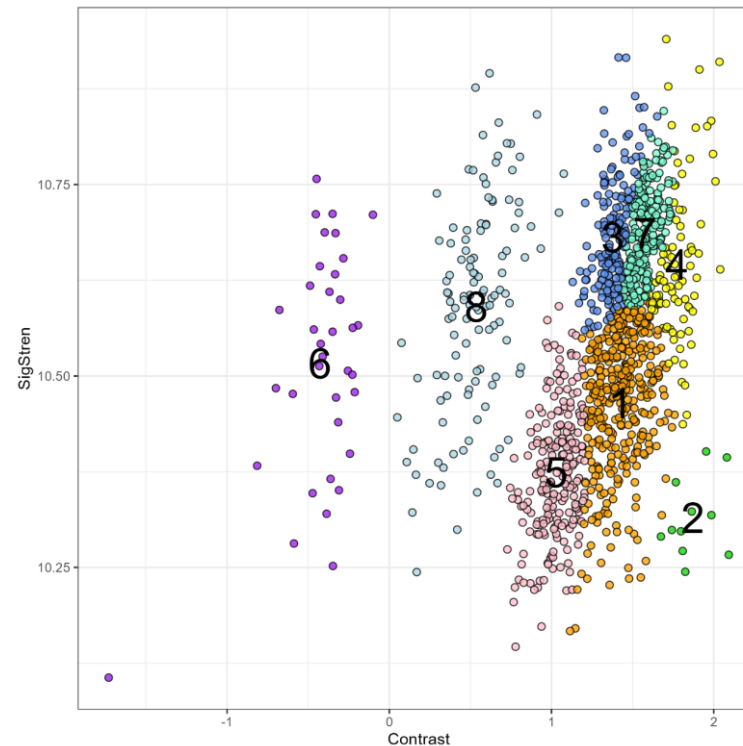
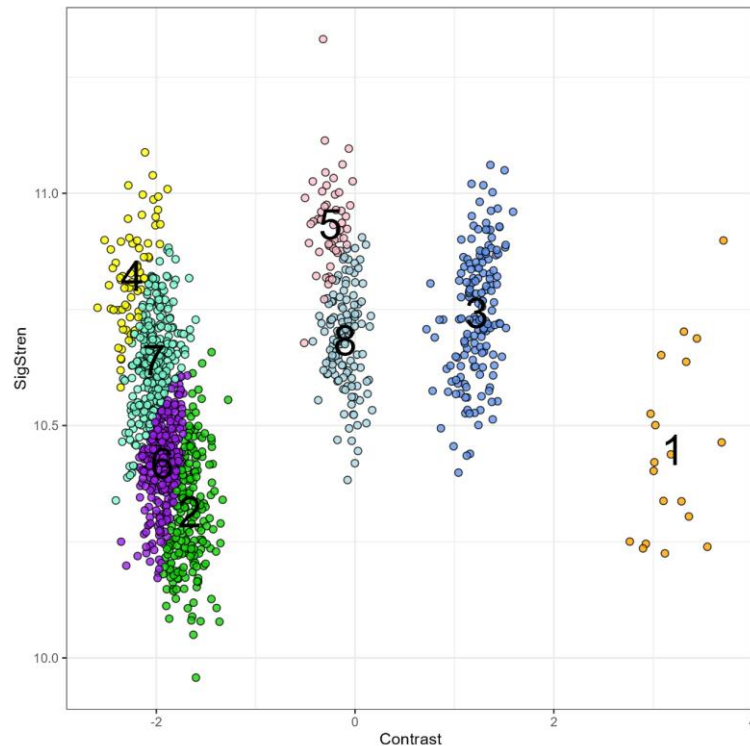
- ❖ Les génotypes avec peu d'individus (homozygotes de l'allèle mineur) doivent être identifiés

I. Génotypage de triploïdes

2. Clustering

► Objectif :

- ❖ Les génotypes avec peu d'individus doivent être identifiés
 - ✓ Demander à l'algorithme un nombre de clusters très supérieurs au nombre de génotypes différents attendus
- ❖ Les clusters doivent ne contenir qu'un seul génotype réel et chaque cluster doit être l'unique représentant d'un génotype donné

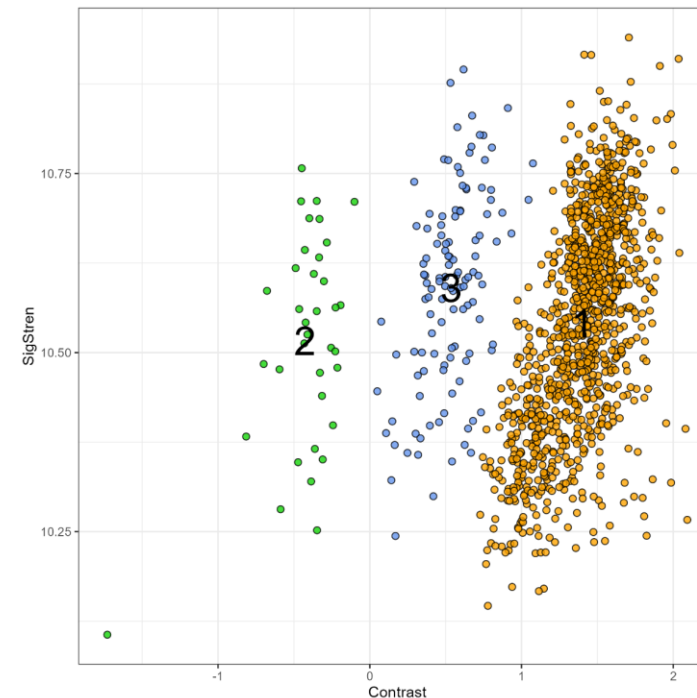
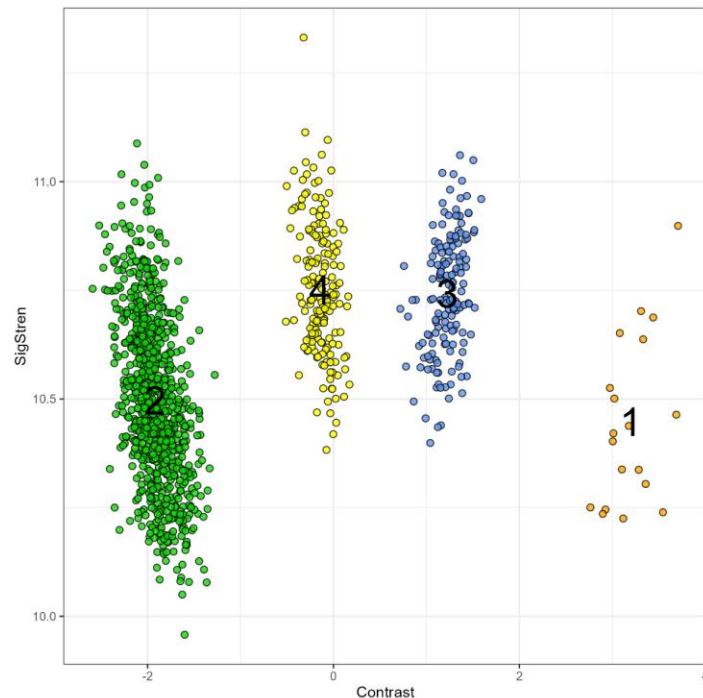


I. Génotypage de triploïdes

2. Clustering

► Objectif :

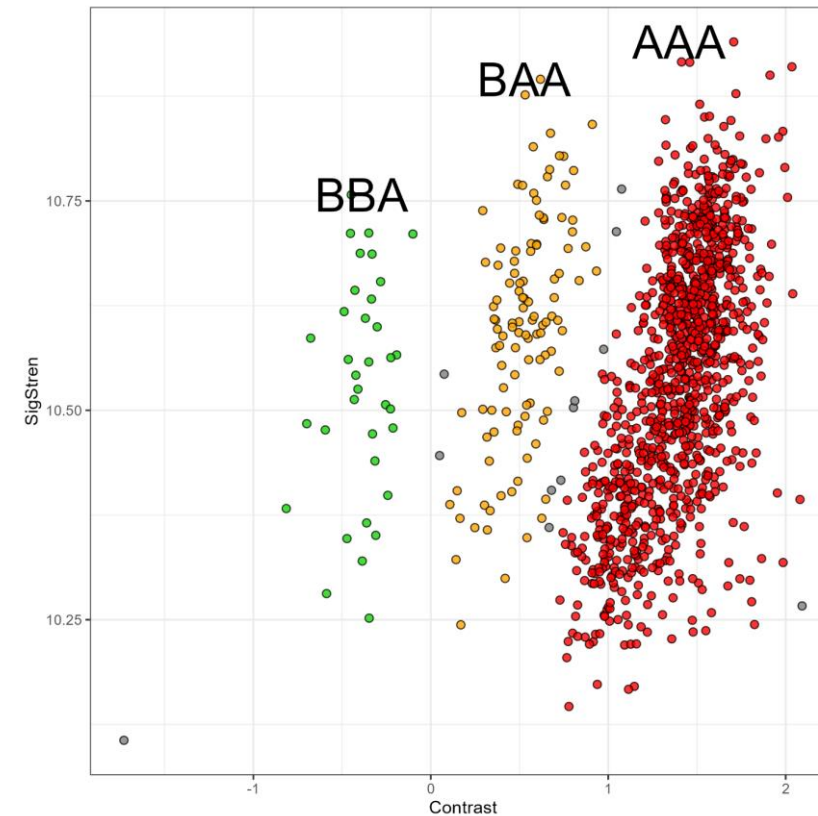
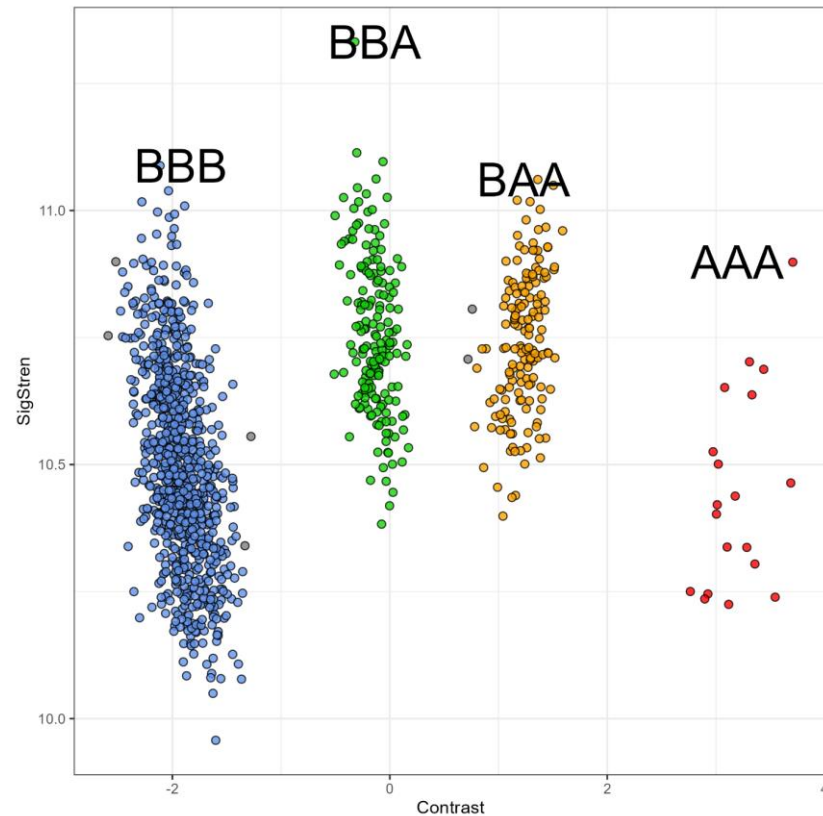
- ❖ Les génotypes avec peu d'individus doivent être identifiés
 - ✓ Demander un grand nombre de clusters à l'algorithme
- ❖ Chaque cluster doit être l'unique représentant d'un génotype
 - ✓ Regrouper les clusters suivant leurs proximités pour n'en laisser qu'au maximum 4 et continuer suivant règle de proximité : $Distance_{Clus1 \& Clus2} > 0.28 * (1 + \text{mean}(\text{Contrast}_{Clus1}, \text{Contrast}_{Clus2}))$



I. Génotypage de triploïdes

3. Génotypage

- ▶ Objectif : assigner à chaque cluster son génotype réel
 - ❖ Soit en regardant la position des clusters les uns par rapport aux autres
 - ❖ Soit en comparant les positions des clusters par rapport aux positions de clusters de référence

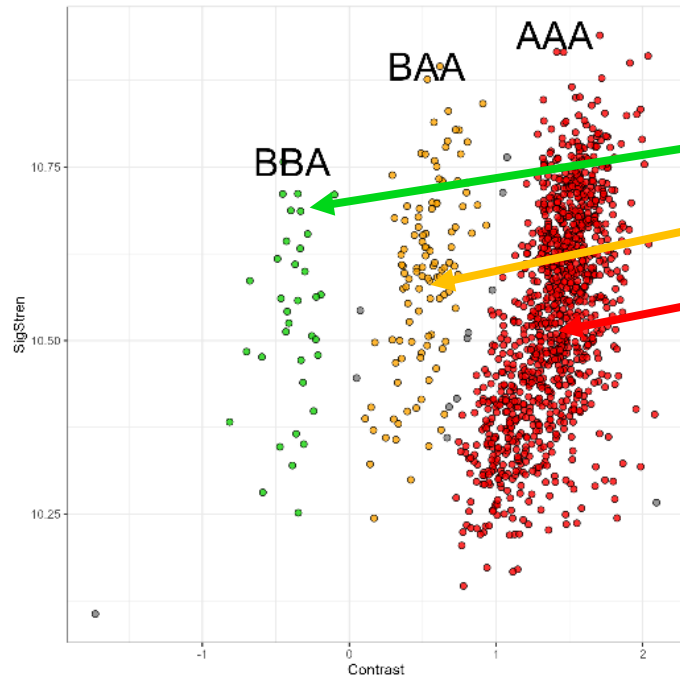


I. Génotypage de triploïdes

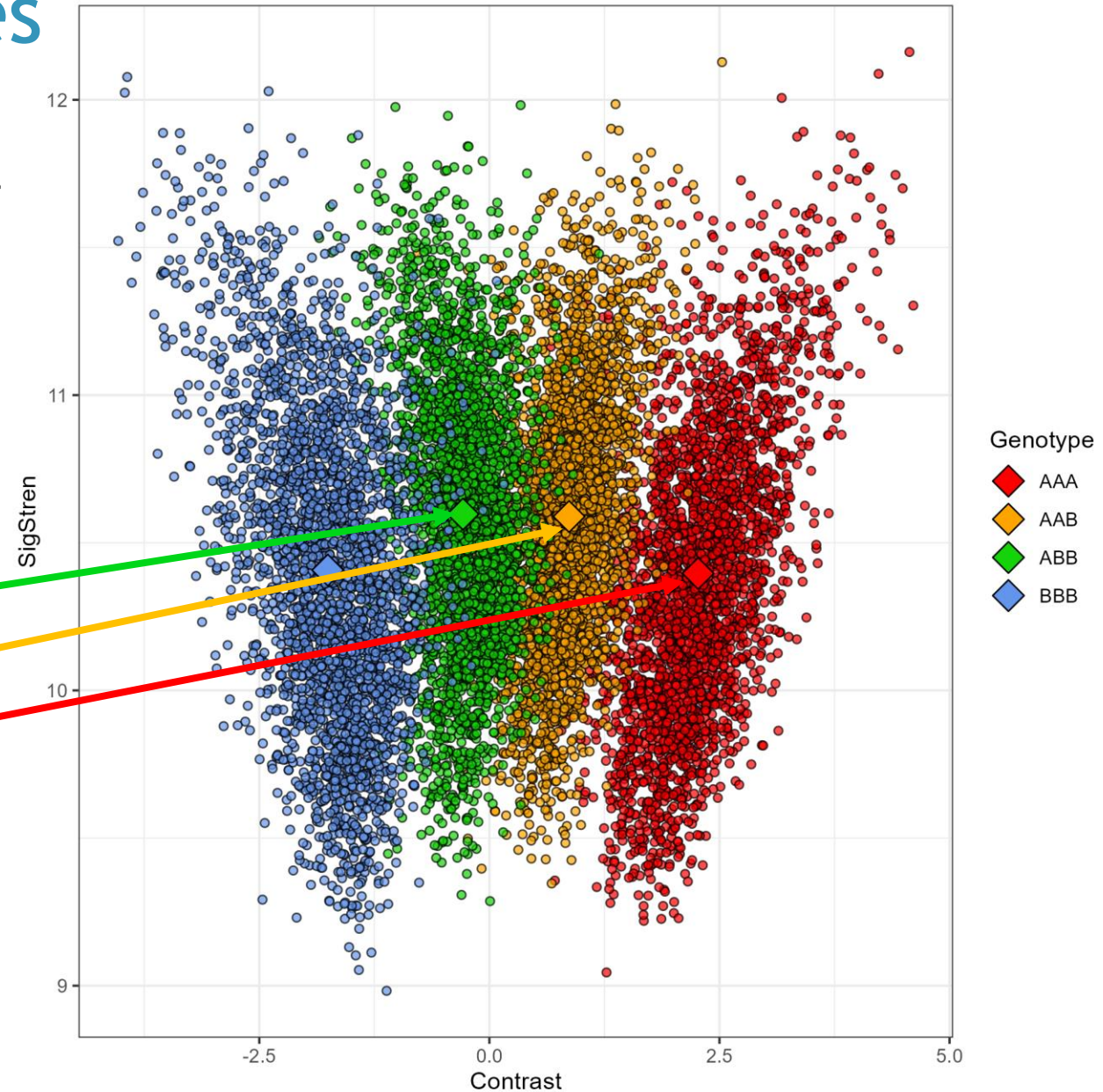
3. Génotypage

- ▶ Objectif : assigner à chaque cluster son génotype réel
 - ❖ Soit en regardant la position des clusters les uns par rapport aux autres
 - ❖ Soit en comparant les positions des clusters par rapport aux positions de clusters de référence

Signaux observés pour un SNP donné

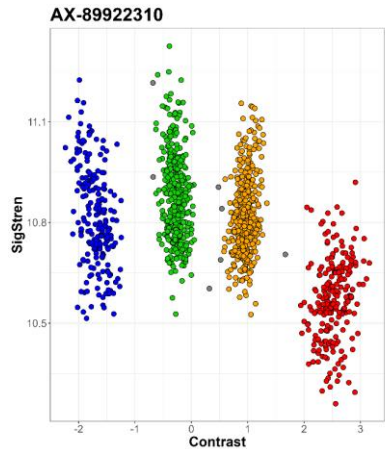


Positions des clusters de référence

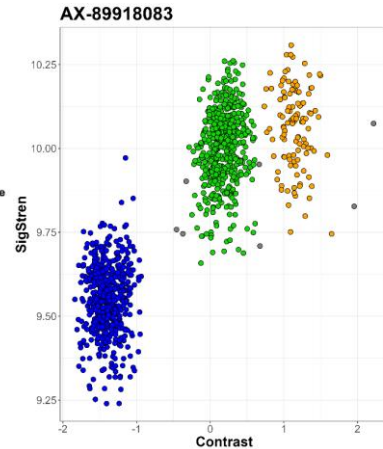


I. Génotypage de triploïdes

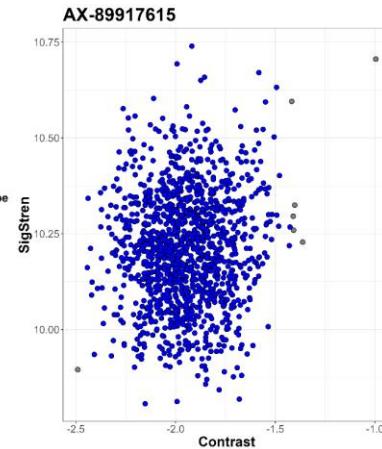
Conclusion



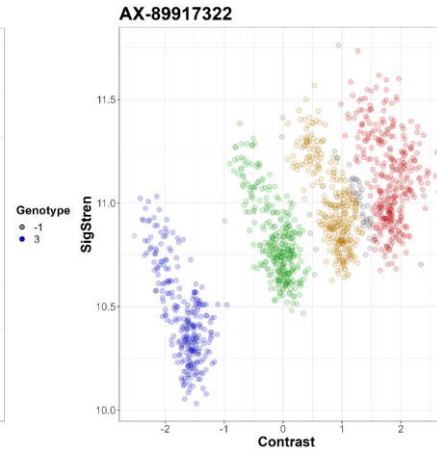
21715 PolyHighResolution



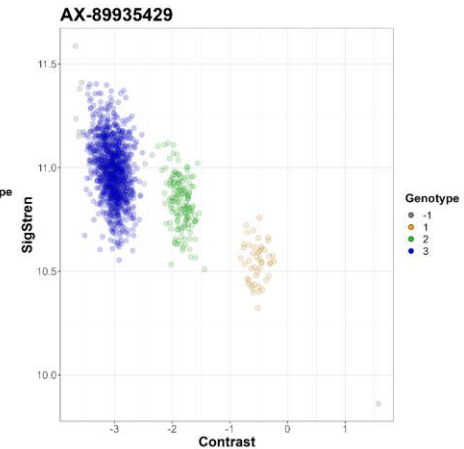
6233 NoMinorHomozygote



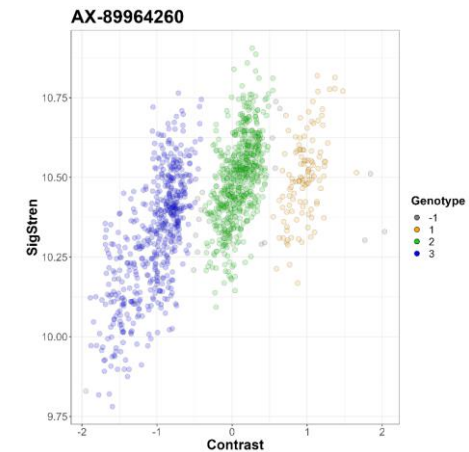
4377 MonoHighRes



4734 CallRateBelowThreshold



421 OffTargetVariant



553 Other

La méthode permet :

- ❖ D'identifier les génotypes des individus
- ❖ D'identifier les marqueurs à ne pas utiliser
- ❖ ET...

Package [GenoTriplo](#) sur le CRAN & [Papier Roche et al.](#) (en préparation)

II. Assignation de parenté

1. Matrices d'exclusion/de vraisemblance

Descendants Diploïdes - AA

| Sire \ Dam | AA | AB | BB | Missing |
|------------|----|----|----|---------|
| AA | 0 | 0 | 1 | 0 |
| AB | 0 | 0 | 1 | 0 |
| BB | 1 | 1 | 2 | 1 |
| Missing | 0 | 0 | 1 | 0 |

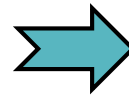


Descendants Triploïdes - AAA

| Sire \ Dam | AA | AB | BB | Missing |
|------------|----|----|----|---------|
| AA | 0 | 0 | 1 | 0 |
| AB | 0 | 0 | 1 | 0 |
| BB | 1 | 1 | 2 | 1 |
| Missing | 0 | 0 | 1 | 0 |

Descendants Diploïdes - AB

| Sire \ Dam | AA | AB | BB | AC | BC | CC | Missing |
|------------|----|----|----|----|----|----|---------|
| AA | 1 | 0 | 0 | 1 | 0 | 1 | 0 |
| AB | 0 | 0 | 0 | 0 | 0 | 1 | 0 |
| BB | 0 | 0 | 1 | 0 | 1 | 1 | 0 |
| AC | 1 | 0 | 0 | 1 | 0 | 1 | 0 |
| BC | 0 | 0 | 1 | 0 | 1 | 1 | 0 |
| CC | 1 | 1 | 1 | 1 | 1 | 2 | 1 |
| Missing | 0 | 0 | 0 | 0 | 0 | 1 | 0 |



Descendants Triploïdes - ABB

| Sire \ Dam | AA | AB | BB | AC | BC | CC | Missing |
|------------|----|----|----|----|----|----|---------|
| AA | 1 | 0 | 0 | 1 | 0 | 1 | 0 |
| AB | 1 | 0 | 0 | 1 | 0 | 1 | 0 |
| BB | 1 | 0 | 1 | 1 | 1 | 1 | 0 |
| AC | 1 | 0 | 0 | 1 | 0 | 1 | 0 |
| BC | 1 | 0 | 1 | 1 | 1 | 1 | 0 |
| CC | 2 | 1 | 1 | 2 | 1 | 2 | 1 |
| Missing | 1 | 0 | 0 | 1 | 0 | 1 | 0 |

II. Assignment de parenté

1. Matrices d'exclusion/de vraisemblance

Descendants Diploïdes - AA

| Sire \ Dam | AA | AB | BB | Missing |
|------------|-------|-----------|----|-----------|
| AA | 1 | 0,5 | e | f_A |
| AB | 0.5 | 0.25 | e | $0.5*f_A$ |
| BB | e | e | e | e |
| Missing | f_A | $0.5*f_A$ | e | f_A^2 |

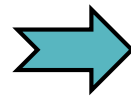


Descendants Triploïdes - AAA

| Sire \ Dam | AA | AB | BB | Missing |
|------------|-------|-------------|----|---------------------------------|
| AA | 1 | $0.5*r$ | e | $f_A^2+2f_A(1-f_A)*r*0.5$ |
| AB | 0.5 | $0.5*0.5*r$ | e | $(f_A^2+2f_A(1-f_A)*r*0.5)*0.5$ |
| BB | e | e | e | e |
| Missing | f_A | $0.5*r*f_A$ | e | $(1-r)f_A^3+r*f_A^2$ |

Descendants Diploïdes - AB

| Sire \ Dam | AA | AB | BB | AC | BC | CC | Missing |
|------------|-------|-----------|-------|-------|-----------|----|-----------------|
| AA | e | 0.5 | 1 | e | 0.5 | e | f_A |
| AB | 0.5 | 0.5 | 0.5 | 0.25 | 0.25 | e | $0.5*(f_A+f_B)$ |
| BB | 1 | 0.5 | e | 0.5 | e | e | f_A |
| AC | e | 0.25 | 0.5 | e | 0.25 | e | $0.5*f_B$ |
| BC | 0.5 | 0.25 | e | 0.25 | e | e | $0.5*f_A$ |
| CC | e | e | e | e | e | e | e |
| Missing | f_B | f_A+f_B | f_A | f_B | $0.5*f_B$ | e | $2*f_A*f_B$ |







Descendants Triploïdes - ABB

| Sire \ Dam | AA | AB | BB | AC | BC | CC | Missing |
|------------|----|-----------------------|-------|----|-------------|----|---|
| AA | e | $0.5*r$ | 1 | e | $0.5*r$ | e | $r*f_A f_B+f_B^2+r*f_B f_C$ |
| AB | e | $0.5-0.25*r$ | 0,5 | e | $0.25*r$ | e | $(1-0.5*r)*f_A f_B+0.5*f_B^2+0.5*r*f_B f_C$ |
| BB | e | $1-r$ | e | e | e | e | $2f_A f_B*(1-r)$ |
| AC | e | $0.25*r$ | 0,5 | e | $0.25*r$ | e | $0.5*r*f_A f_B+0.5*r*f_B f_C+0.5*f_B^2$ |
| BC | e | $0.5*(1-r)$ | e | e | e | e | $f_A f_B*(1-r)$ |
| CC | e | e | e | e | e | e | e |
| Missing | e | $0.5*r*f_A+(1-r)*f_B$ | f_A | e | $0.5*r*f_A$ | e | $f_A*(r*f_A f_B+f_B^2+r*f_B f_C)+f_B^2 f_A^2*(1-r)$ |

II. Assignation de parenté

2. APIS 3n

Parentage assignment with APIS    

By J.Roche

Choose your step

- Formatting
- APIS**
- Verification

Male and female parents are in the same dataset ?

Choose the ploidy level of the offspring ?
2

Dataset of OFFSPRING ?
Browse... No file selected

Dataset of PARENTS ?
Browse... No file selected

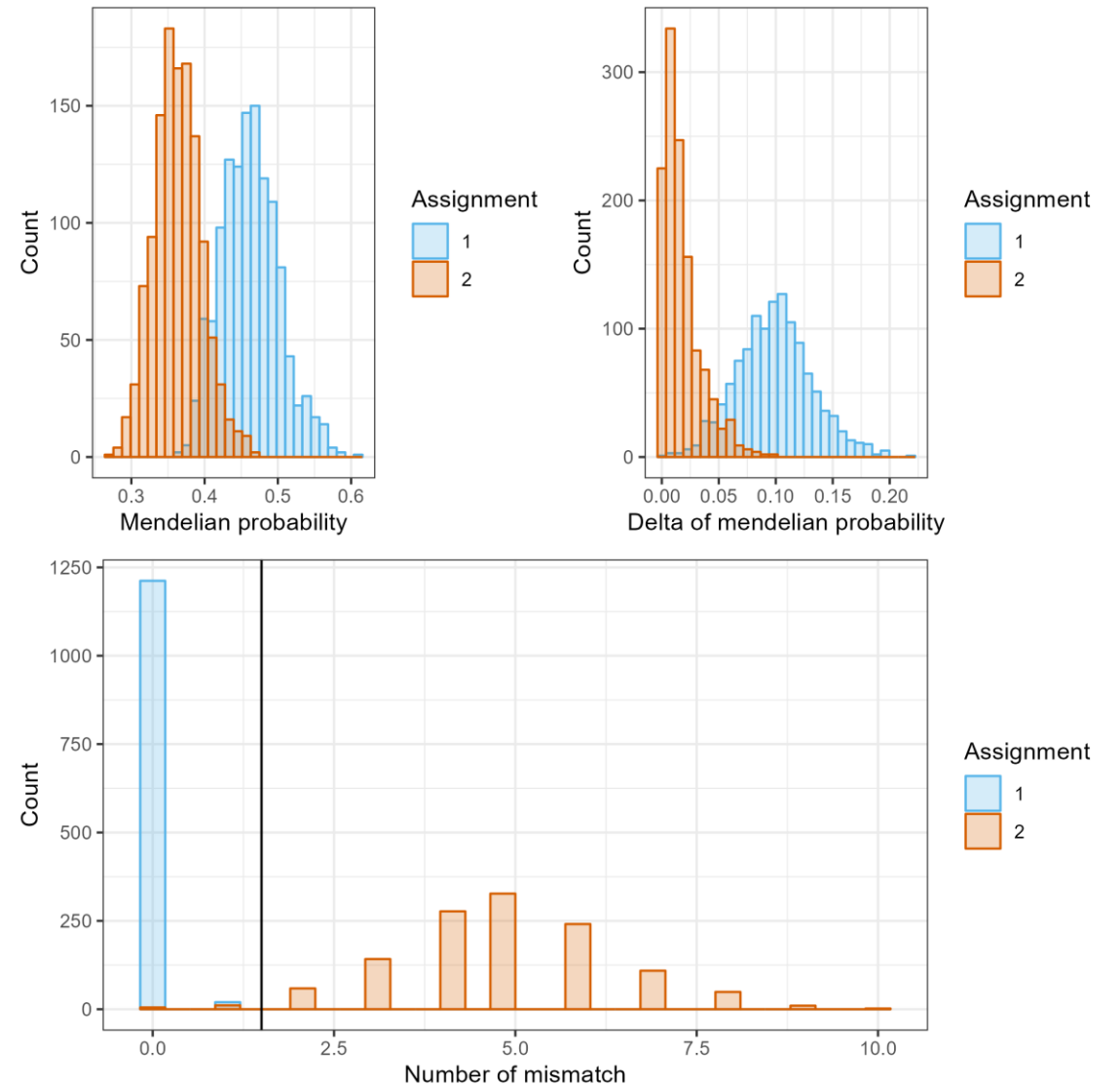
File (.txt) with the sex of each parents ?
Browse... No file selected

File (.txt) with the markers to use (optional) ?
Browse... No file selected

Warnings : Make sure your app is launched on the desired directory. If not, quit the app and use setwd() !

Input .Rdata or .txt format example
You can find 'log_APIS.txt' in './log' directory for more info.

Interface Shiny disponible



~96 SNP nécessaires pour 100% assignation correct

II. Assignation de parenté

Conclusion

APIS: Auto-Adaptive Parentage Inference Software Tolerant to Missing Parents

Parentage assignment package. Parentage assignment is performed based on observed average Mendelian transmission probability distributions or Exclusion. The main functions of this package are the function `APIS_2n()`, `APIS_3n` and `launch_APIShiny()`, which perform parentage assignment.

Version: 2.0.2
Depends: R ($\geq 3.5.0$), [shinyBS](#)
Imports: [cowplot](#), [data.table](#), [doParallel](#), [dplyr](#), [DT](#), [foreach](#), [ggplot2](#), [gridExtra](#), [htmltools](#), [methods](#), [plotly](#), [rlang](#), [shiny](#), [shinythemes](#)
Published: 2023-09-06
Author: Julien Roche [aut, cre], Ronan Griot [aut], Mathieu Besson [aut], François Allal [aut], Marc Vandeputte [aut], Jonathan D'Ambrosio [aut], Romain Morvezin [aut], Florence Phocas [aut], Sophie Brard-Fudulea [aut], Pierrick Haffray [aut]
Maintainer: Julien Roche <jjm.roche at gmail.com>
License: [GPL-2](#) | [GPL-3](#) [expanded from: GPL]
NeedsCompilation: yes
CRAN checks: [APIS results](#)

1 APIS: an updated parentage assignment software managing triploids 2 induced from diploid parents

3 Roche Julien, Griot Ronan, Allal François, Besson Mathieu, Patrice Pierre, Phocas Florence,
4 Vandeputte Marc

5 Introduction

6

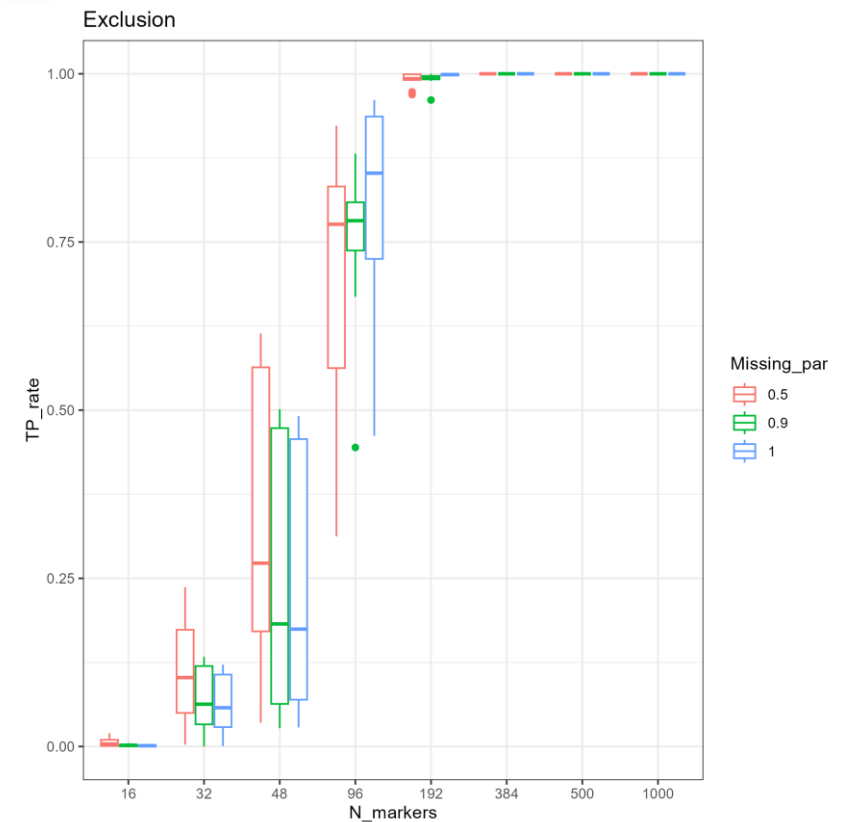
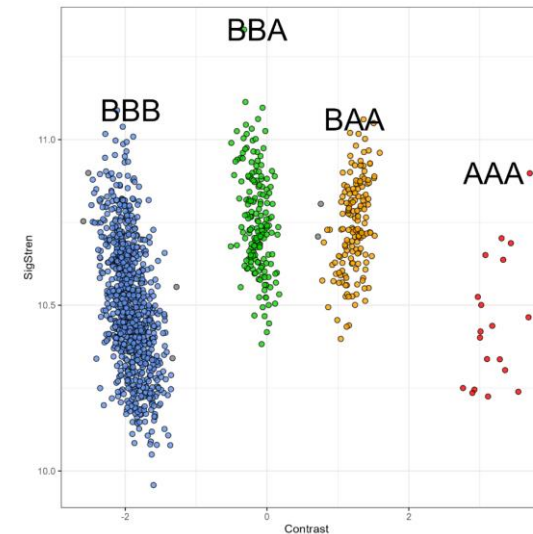
7 Triploids, which are individuals bearing three sets of chromosomes instead of two, are commonly used
8 in plant, fish and shellfish breeding (Pfeffer et al., 2009; Wang et al., 2016). The benefits of using
9 triploids are diverse and can imply higher growth due to larger cell size at least in plants (SUGIYAMA,
10 2005), although this effect does not seem to be general (Trukaya, 2013). One of the main interests of
11 triploids is that they are often sterile, which has several implications, (1) a higher growth, mainly after
12 the age at maturity, due to a lack of investment in gonads and reproduction (Quillet et al., 1988), (2) a
13 better organoleptic quality due to the absence of mature gonads (oyster), seeds (citrus) or to the
14 sparing of lipid and pigments, not transferred from the edible muscle to the gonads (trout), (3) a much
15 lower risk of genetic introgression of farmed genotypes into wild populations (Berley, 2016).

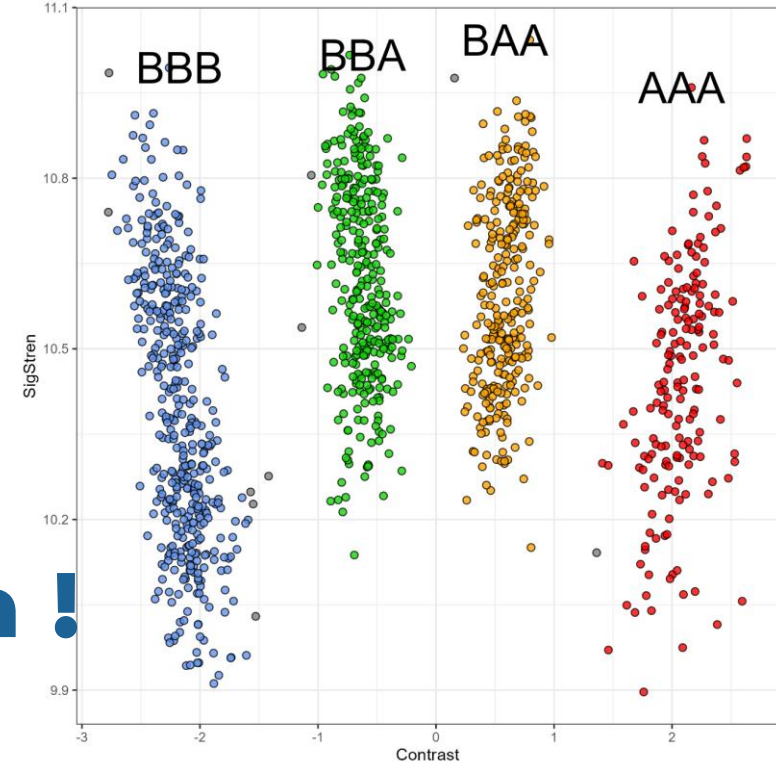
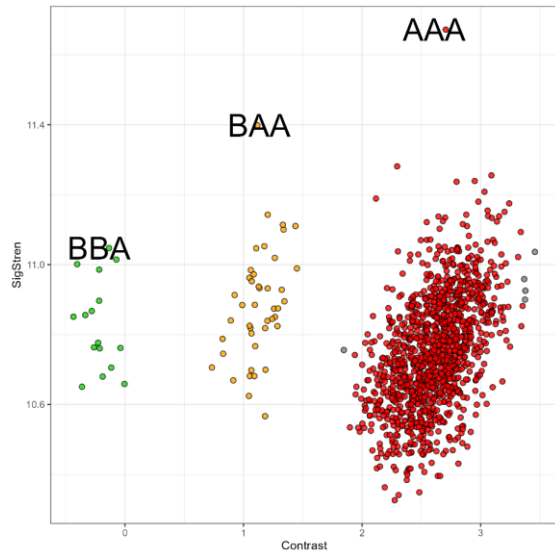
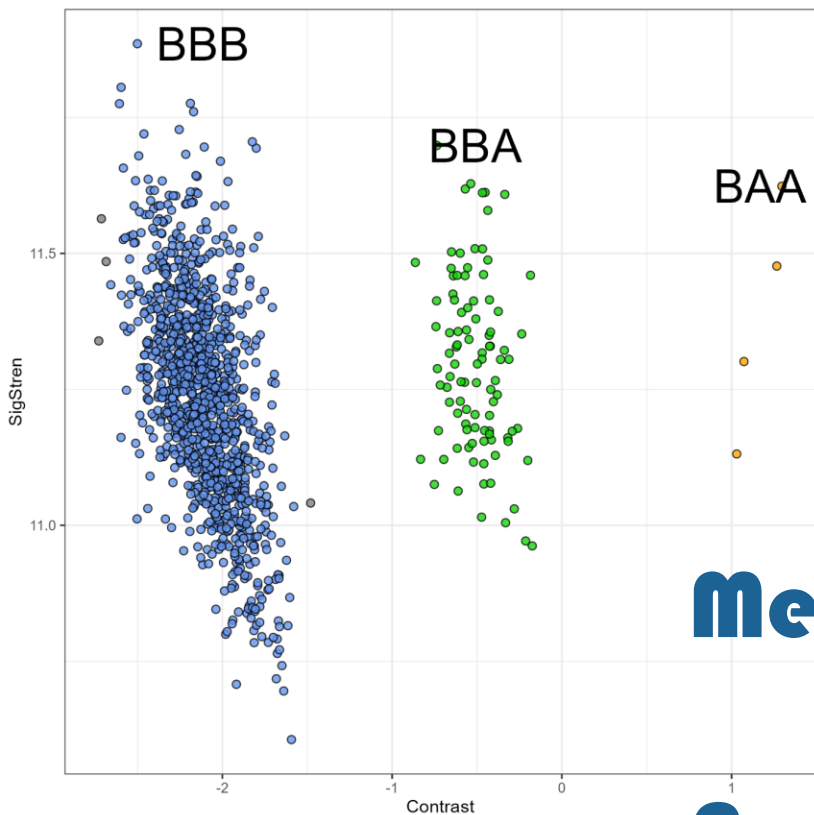
16 As triploids have benefits for farming, as any farmed genotype, they require genetic improvement,
17 which implies in many cases the establishment of pedigrees. When families and/or genotypes are
18 reared in a common garden to avoid confounding environmental effects, it is necessary to use genetic
19 markers to assign the tested genotype(s) to their parents (see Vandeputte and Haffray, 2014 for the
20 case of cultured fish).

Package APIS à jour et sur le CRAN & Papier Roche et al. (en préparation)

Conclusion et perspective

- ▶ Génotypage des triploïdes opérationnel
- ▶ Assignation de parenté descendants triploïdes efficace
- ▶ Possibilité de faire de la sélection génétique sur triploïde !
- ▶ CDD 4 mois INRAE-GABI-GenAqua :
 - Valorisation scientifique
 - Estimation des valeurs génétiques chez les descendants triploïdes





Merci pour votre attention !

Avez-vous des questions ?

