Laboratoire « Adaptation et Adaptabilité des Animaux et des Systèmes d'élevage »

COMBINING VITASSIGN AND COLONY: AN EFFICIENT PRACTICAL PROCEDURE FOR PARENTAL ASSIGNMENT WITH MISSING PARENTS

Doan Q.K.[1], Vandeputte M.[1, 2], Chatain B.[1], Vergnet A.[1], Allal F.[1§]

[1] *Ifremer, UMR 9190 MARBEC, chemin de Maguelone, 34250 Palavas-les-Flots, France*

[2] *INRA, UMR1313 Génétique Animale et Biologie Intégrative, 78350 Jouy-en-Josas, France*

International Symposium on Genetics in Aquaculture XII
SANTIAGO DE COMPOSTELA
2015

INTRODUCTION

The contribution of parentage assignment in selective breeding of aquaculture species is undeniable [1, 3, 7]. Parentage assignment is based on two computation methods, exclusion-based methods and likelihood-based methods [3]. Exclusion is very simple and makes no hypotheses other than Mendelian segregation of alleles, but is very sensitive to genotyping errors while likelihood methods use a different approach with probabilities. Likelihood methods generally give higher assignment rates than exclusion with low power marker sets but sometimes give inconsistent results. However, breeding programs often face practical management problems and it is not uncommon that some broodstock genotypes miss because of premature death, traceability problems or sample quality problems [5, 6, 7]. This may lead to unexpectedly low parentage assignment [5, 6] and decrease markedly the potential of genetic improvement. In this study, we explored the potential of combining two softwares, VITASSIGN (exclusion) and COLONY (likelihood) for obtaining parentage assignment in the case of a few missing parental genotypes in a full factorial mating design.

MATERIALS AND METHODS

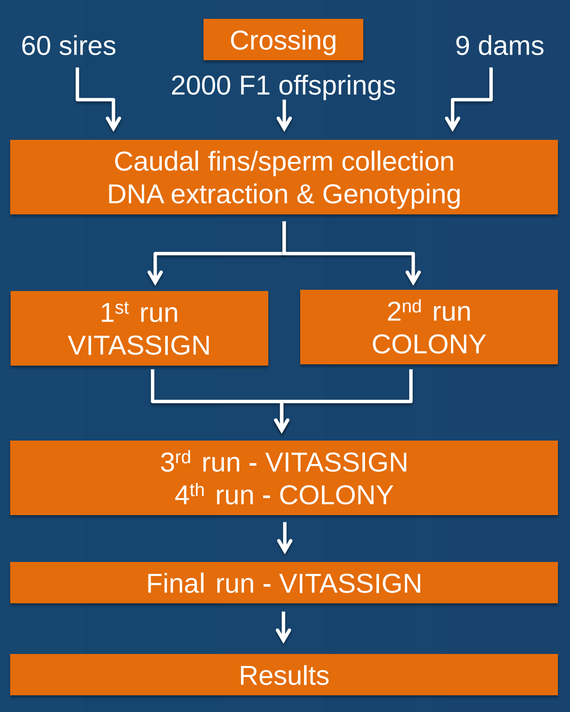

Figure 1: Study's procedure

The global process implemented in this study is depicted in figure 1.

❖ **Materials:** 60 wild sea bass sires were crossed with 9 wild sea bass dams in a full factorial mating scheme and 2000 offsprings were reared in a single batch. The caudal fins or sperm of parents were collected directly during the artificial mating while the caudal fins of the 2000 offsprings were collected at five months post-harvest. All were sent to LABOGENA (Jouy-en-Josas, France) for DNA extraction and genotyping of 12 microsatellite markers.

❖ **VITASSIGN:** running as described by Vandeputte et al. 2006 [5], allowing for up to two allelic mismatches between parents and offsprings.

❖ **COLONY:** running as described by Jones and Wang, 2010 [2, 9] without data for known or excluded parentalship and indication of the putative number of parents (60 sires , 9 dams).

❖ **Reconstructing missing genotypes and correcting genotyping errors:** lying on the genotypes inferred by COLONY, 2 potential dams showing posterior probabilities equal to 1, were chosen as alternative genotypes of the 2 missing dams. In addition, candidate sires and dams with missing loci or genotyping errors were corrected using genotypes inferred by COLONY with posterior probabilities equal to 1 (12 parent genotypes were corrected or completed for a total of 48 corrected alleles).

RESULTS AND DISCUSSION

All samples (60 sires, 9 dams and 2000 offsprings) were genotyped for 12 microsatellite loci. However, because of a low sample quality, 2 dams, 2 sires and 9 offspring could not be genotyped [3, 7]. Therefore, only 7 dams, 58 sires and 1991 offsprings were used for first pedigree assignment trials using VITASSIGN, an exclusion-based parentage assignment software [5]. However, because of genotyping errors and missing genotypes, only 40.6% of the offsprings were assigned to a single parent pair with perfect match (55.6% allowing up to 2 mismatches) (figure 2).

In order to identify the missing genotypes and genotyping errors, the same data set was processed with COLONY, a maximum likelihood parentage software [2, 9]. If highly probable pedigree was obtained for only 52.6% of the offsprings (figure 2), this run allowed identifying 252 additional potential dam genotypes. Two genotypes among those, displaying likelihood probabilities equal to 1, were suspected to correspond to the 2 missing dams. Due to the multiplication rule of the probabilities, the other 250 genotypes with mean posterior probabilities under 0.95 (real probability is lower than $0.95^{12}=0.54$) were discarded.

The next pedigree assignment, including the 2 dam genotypes inferred by COLONY, resulted in 78.0% of perfect match with VITASSIGN (92.4% allowing up to 2 mismatches) and in 77.1% of assignment with COLONY (figure 2). Later genotyping of alternative samples of the two missing dams, confirmed that the genotypes inferred by COLONY were exact. Nevertheless, because of missing loci and genotyping errors of some sires and dam, the proportion of parental assignment with perfect match remains lower than expected by VITASSIGN simulations [5]. These candidate sires and dams were corrected based on the genotypes inferred by COLONY (1 dam and 11 sire genotypes were corrected or completed, for a total of 29 corrected alleles). Finally, using VITASSIGN, 96.4% of the offsprings were uniquely assigned with assignment power >0.99 (86.1% with perfect match and 96.4% with up to 2 mismatches allowed). Only 3.4% of the offsprings could not be assigned (figure 2).

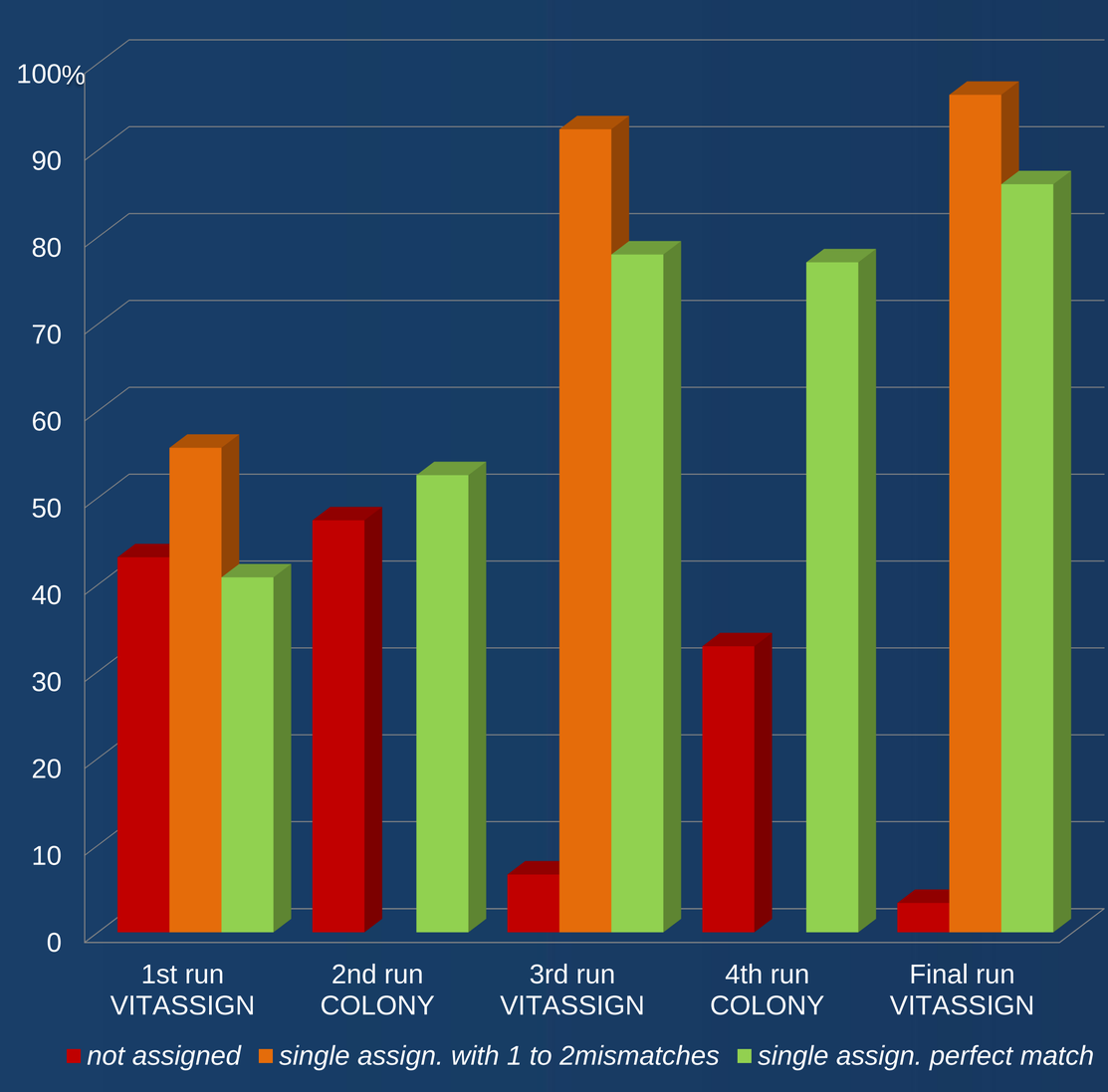

Figure 2: *Parentage assignment*

CONCLUSION

We demonstrated the power of combining VITASSIGN and COLONY for significantly improving pedigree assignments when parent genotypes are missing. In this study, the proportion of parentage assignment was increased from 40.6% to 96.4%. This improvement was allowed by combining the successful reconstructions of missing genotypes and genotyping-errors corrections using likelihood posterior probabilities calculated by COLONY and the exclusion-based assignment power of VITASSIGN.

REFERENCES

1. Garcia de Leon F. et al. (1996). *Aquaculture* **159**: 303-316
2. Jones O. and Wang J. (2010). *Mol. Ecol. Re.* **10**: 551-555
3. Vandeputte M. and Haffray P. (2014). *Front. Genet.* **5**:432
4. Vandeputte M. et al. (2011). *Aquaculture* **314**: 60-66
5. Vandeputte M. et al. (2006). *Mol. Ecol. Res.* **6**: 265–267
6. Villanueva B. et al. (2002). *Anim. Genet.* **33**, 33-41
7. Wang J. (2004). *Genetics* **166**: 1963-1979
8. Wang J. (2012). *Genetics* **191**: 183-194
9. Wang J. (2013). *Mol. Ecol. Res.* **13**: 734-739

ACKNOWLEDGEMENTS

This work was carried out in the frame of the FUI project RE-SIST funded by BPI and Languedoc-Roussillon and the PhD grant of M. DOAN Quoc Khanh supported by Vietnamese government. We wish to specially thank all partners for their active participation at all stages of the project.