Laboratoire « Adaptation et Adaptabilité des Systèmes d'élevage »

COMBINING VITASSIGN AND COLONY FOR PEDIGREE RECONSTRUCTION IN A CASE OF FACTORIAL MATING WITH MISSING PARENTAL GENOTYPES

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

[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*

INTRODUCTION

Marker-based pedigree reconstruction has become essential in aquaculture breeding as well as in experimental genetic studies [1, 2, 3]. For this purpose, several parentage assignment softwares have been developed and allow accurate pedigree recovery in most cases. However, parentage recovery with missing parental genotypes due to premature death, traceability or sample quality problems remains tricky [3, 4, 5]. In this study, we explored the potential of combining VITASSIGN [4], an exclusion-based parentage assignment software, and COLONY [6, 7], a maximum likelihood parentage software, for pedigree reconstruction of fish obtained from a mating where missing parental genotypes occurred.

MATERIALS AND METHODS

In this study, 60 wild sea bass sires were crossed with 9 wild sea bass dams in a full factorial mating scheme and 2000 offspring 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 offspring were collected at five months post-hatching. All were sent to LABOGENA (Jouy-en-Josas, France) for DNA extraction and genotyping of 12 microsatellite markers.

Two computation methods were used for parentage assignment. **VITASSIGN**, an exclusion-based method, making no hypotheses other than Mendelian segregation of alleles, but very sensitive to genotyping errors, was used as described by Vandeputte et al. 2006 [4], allowing for up to two allelic mismatches between parents and offspring. **COLONY**, a maximum likelihood-based method, was launched as described by Jones and Wang, 2010 [7] with priors on the putative number of parents (60 sires , 9 dams) but none concerning known or excluded parentalship.

Genotypes recovery and corrections of genotyping errors were done using the outputs of COLONY. The genotypes of potential alternative parents displaying average posterior probabilities equal to 1 were identified. Candidate sires and dams with missing loci or genotyping errors were corrected using genotypes inferred by COLONY when alternative alleles displayed posterior probabilities equal to 1.

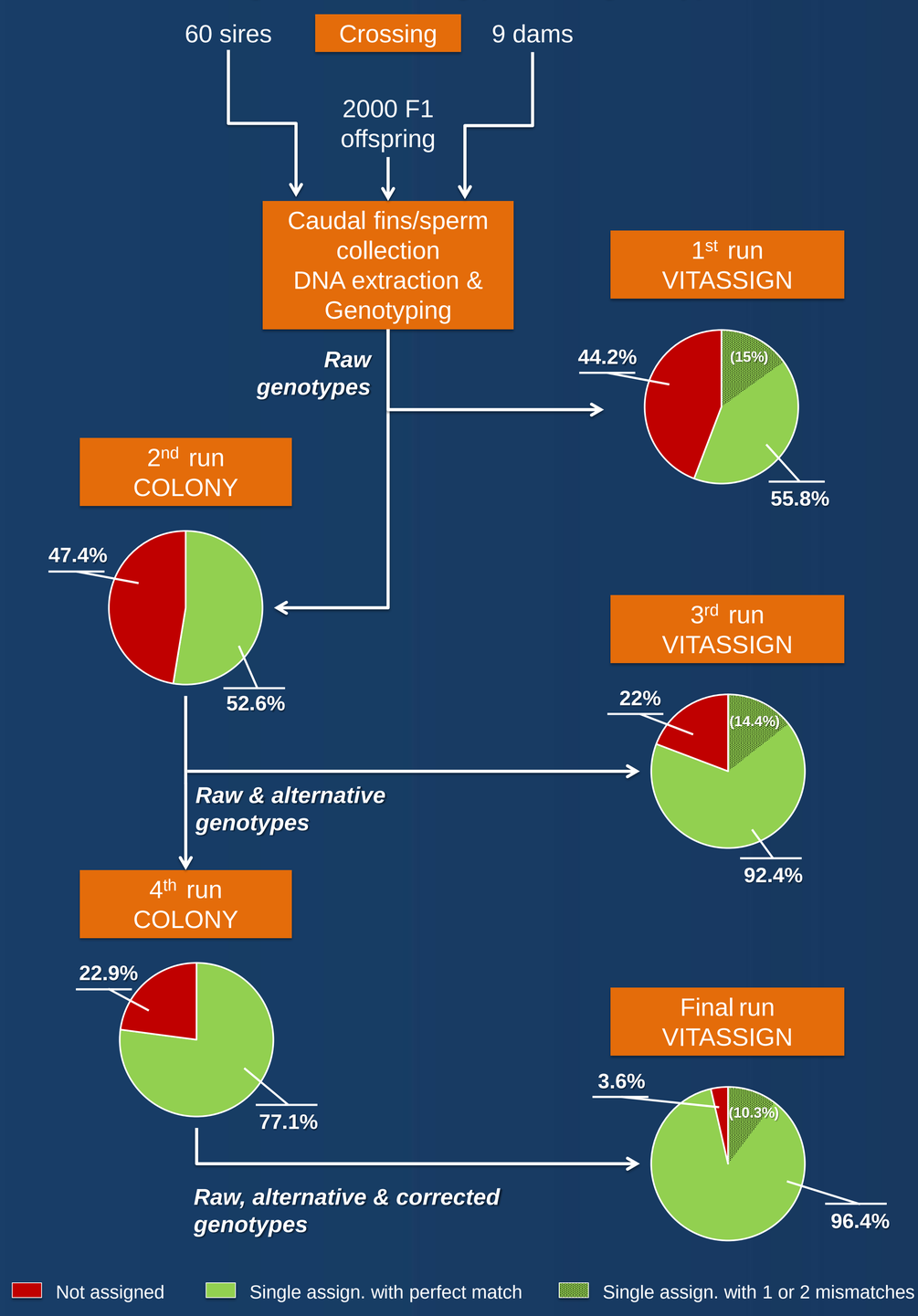

Figure 1: *Parentage assignment procedure and results*

RESULTS AND DISCUSSION

All samples (60 sires, 9 dams and 2000 offspring) were genotyped for 12 microsatellite loci. However, because of a low sample quality, 2 dams, 2 sires and 9 offspring could not be genotyped [2, 3]. Therefore, only 7 dams, 58 sires and 1991 offspring were used for first pedigree assignment trials using VITASSIGN [4]. Because of genotyping errors and missing genotypes, only 40.8% of the offspring were assigned to a single parent pair with perfect match (55.8% allowing up to 2 mismatches) (figure 1).

In order to identify the missing genotypes and genotyping errors, the same data set was processed with COLONY, a maximum likelihood parentage software [6, 7]. If highly probable pedigree was obtained for only 52.6% of the offspring (figure 1), this run allowed identifying 252 additional potential dam genotypes. Two genotypes among those, displaying posterior probabilities equal to 1, were suspected to correspond to the 2 missing dams.

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 1). 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 (99.5% of unique assignment) [4]. 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 48 corrected alleles). Finally, using VITASSIGN, 96.4% of the offspring were uniquely assigned (86.1% with perfect match and 96.4% with up to 2 mismatches allowed). Only 3.4% of the offspring could not be assigned (figure 1).

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 40.8% 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. (1998). *Aquaculture* **159**: 303-316
2. Vandeputte M. and Haffray P. (2014). *Front. Genet.* **5**:432
3. Wang J. (2004). *Genetics* **166**: 1963-1979
4. Vandeputte M. et al. (2006). *Mol. Ecol. Res.* **6:** 265–267
5. Villanueva B. et al. (2002). *Anim. Genet.* **33**, 33-41
6. Jones O. and Wang J. (2010). *Mol. Ecol. Re.* **10**: 551-555
7. 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.