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

Doan O.K.1, Vanderputte M.1, 2, Chatain B.1, Vergnet A.1, Allal F.1, 6
1 IFREMER, UMR 9190 MARBEC, chemin de Mauguiole, 34250 Palavas-les-Flots, France
2 INRA, UMR1313 Génétique Animale et Biologie Intégrative, 78350 Jouy-en-Josas, France

INTRODUCTION

The contribution of parentage assignment in effective breeding of aquaculture species is undeniable [1-7]. Parentage assignment is based on two computation methods: exclusion-based methods and likelihood-based methods [5]. Exclusion is very simple and makes no hypothesis other than Mendelian segregation of alleles but it 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 breeders lack genotypes miss because of premature death, traceability problems or sample quality problems [5-7]. This may lead to unexpectedly low parentage assignment [5-6] and decreases 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

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

Materials: 60 wild zebrafish lines were crossed with 1 wild zebrafish dam in a full factorial mating scheme and 2000 offspring were reared in a single batch. The causal father or sperm of parents were collected directly during the artificial mating while the causal fish of the 2000 offspring 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 was used to assign the parents to the offspring, their genotypes were available.

COLONY was used to assign the parents to the offspring, their genotypes were available.

Reconstructing missing genotypes and correcting genotyping errors: 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 2 dams and 2 dams with missing loci or genotyping errors were corrected using genotypes inferred by COLONY with posterior probabilities equal to 1 (2 parent genotypes were corrected or completed for a total of 46 corrected alleles).

RESULTS AND DISCUSSION

All samples (60 zebrafish dams and 2000 offspring) were genotyped for 12 microsatellite loci. However, because of a low sample quality, 2 dams, 2 sire and 1 offspring could not be genotyped [5-7]. Therefore only 1 dams, 25 sires and 198 offspring 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 10.0% of the offspring were assigned to a single parent pair with perfect match (59.9% 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 assignment software [2-6] if higher probable pedigree was obtained for only 1/3 error of the offspring (figure 2) the 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 had mean posterior probabilities under 0.1% (real probability is lower than 0.1%) were discarded.

The next pedigree assignment including the 2 dam genotypes inferred by COLONY, resulted in 10.0% of perfect match with VITASSIGN (12.9% 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 zones and dams the proportion of parental assignment with perfect match remains lower than expected by VITASSIGN simulations [5]. These candidate zones and dams were corrected based on the genotypes inferred by COLONY (1 dam and 1 sire genotypes were corrected or completed for a total of 24 corrected alleles) finally using VITASSIGN 44.6% of the offspring were uniquely assigned with assignment power 90.1% (46.6% with perfect match and 44.1% with up to 2 mismatches allowed). Only 54% of the offspring could not be assigned (figure 2).

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 parental assignment was increased from 10.0% to 44.6%. This improvement was allowed by combining the successful reconstitutions of missing genotypes and genotyping errors corrections in likelihood probabilities calculated by COLONY and the exclusion-based assignment power of VITASSIGN.

ACKNOWLEDGEMENTS

This work was carried out in the frame of the FIR project REELT funded by BPI and Longitude. Research and the PhD grant of M. DONAQ Que Khanh supported by Vietnam government. We wish to specially thank all workers for their active participation at all stages of the project.

REFERENCES