

Identification of SNPs and candidate genes associated with nutritional composition traits by GWAS and transcriptome analysis in blunt snout bream (*Megalobrama amblycephala*)



Ming Zhao^a, Yan-Jing Wu^a, Rui-Hong Lu^a, Qian-Qian Guan^a, Shi-Long Wang^a, Ao Gao^a, Wu-Ying Chu^b, Xin Zhu^b, Ze-Xia Gao^{a,c,d}, Lu-Sha Liu^{a,c*}

^a College of Fisheries, Key Lab of Freshwater Animal Breeding, Ministry of Agriculture/Key Lab of Agricultural Animal Genetics, Breeding and Reproduction of Ministry of Education/ Huazhong Agricultural University, Wuhan 430070, China

^b College of Biological and Chemical Engineering, Changsha University, Changsha 410022, China

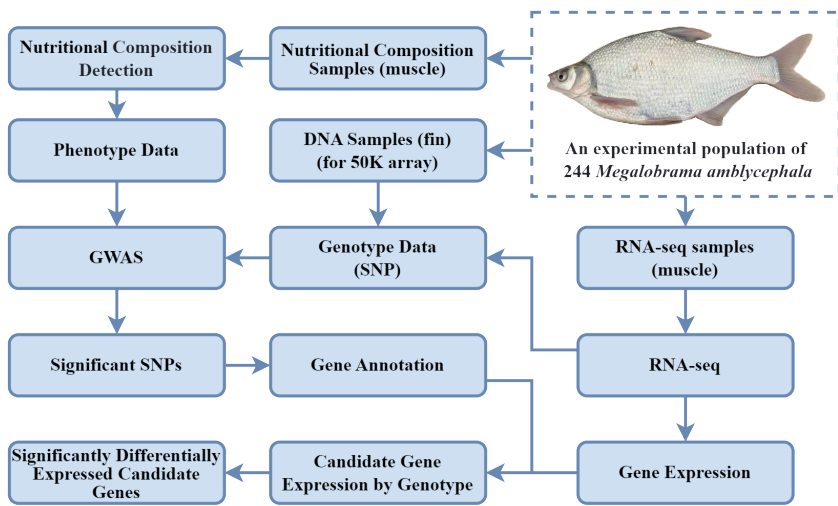
^c Hubei Hongshan Laboratory, Wuhan 430070, China

^d Engineering Technology Research Center for Fish Breeding and Culture in Hubei Province, Wuhan 430070, China

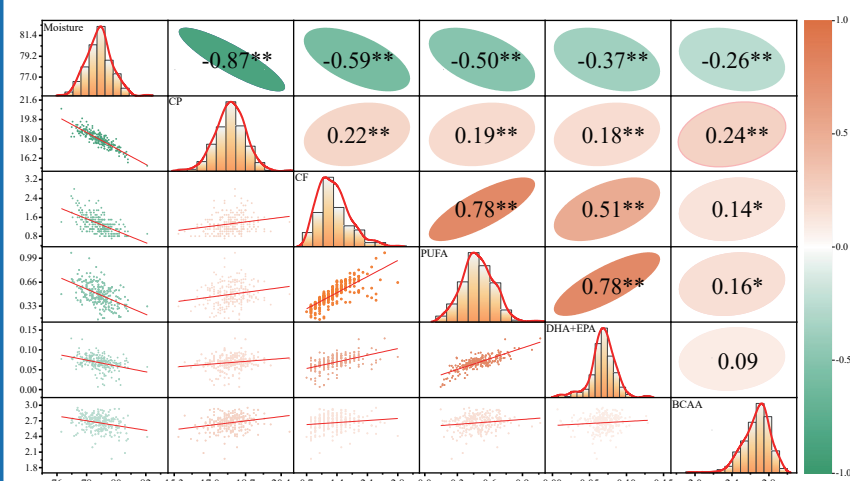
INTRODUCTION

Fish is a vital source of high-quality protein and fatty acids, with its nutritional composition determining nutritional value, and economic worth. *Megalobrama amblycephala* (blunt snout bream), an important freshwater aquaculture species in China, plays a key role in the freshwater aquaculture. However, the genetic mechanisms regulating its muscle nutritional traits remain unclear, and targeted breeding for quality improvement is slow. This study used 244 blunt snout bream, integrating 50K SNP array genotyping and dorsal muscle transcriptomics. It conducted GWAS for 6 nutritional traits, analyzed differential gene expression, and aimed to identify candidate genes, providing molecular markers for high-quality breeding.

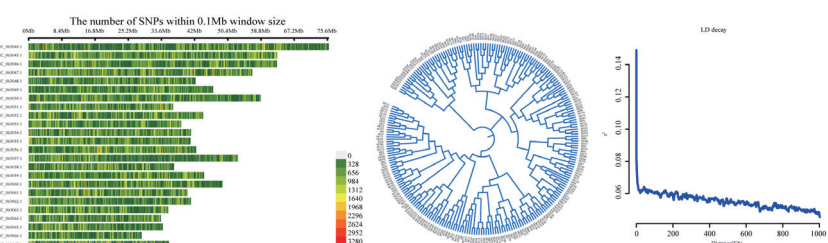
MATERIALS AND METHODS



PHENOTYPE AND GENOTYPE

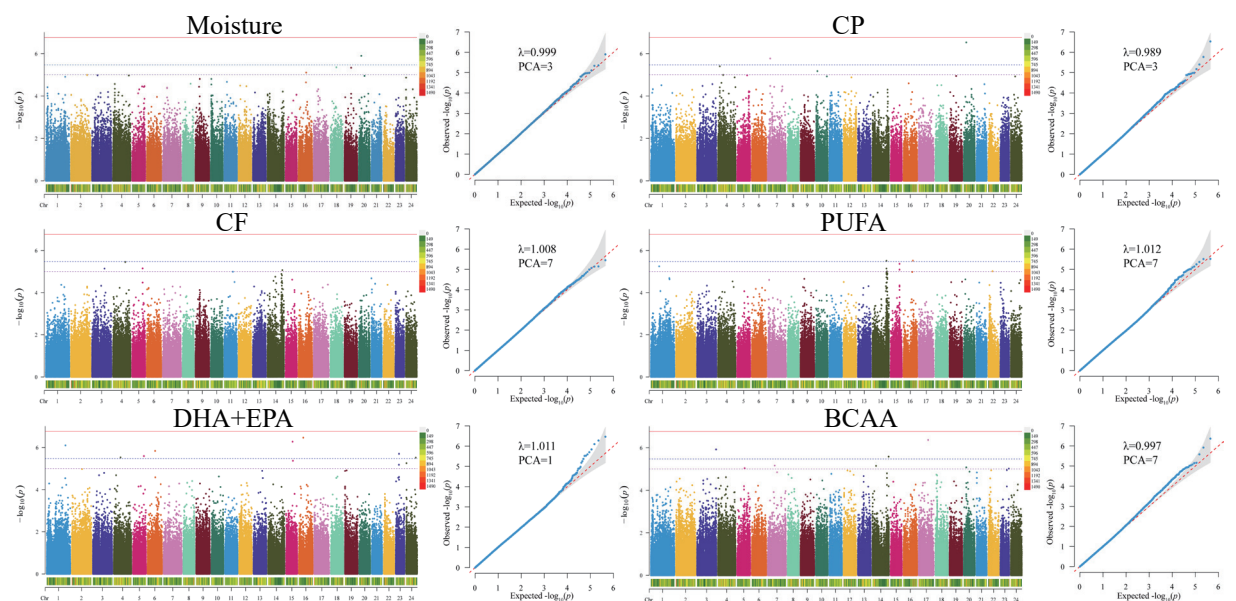


Nutritional composition detection and statistical analysis were performed on the dorsal muscles of 244 *M. amblycephala*. Statistical tests showed that six traits: Moisture, crude protein (CP), crude fat (CF), polyunsaturated fatty acids (PUFA), DHA+EPA, and branched-chain amino acids (BCAA)-all had a near-normal distribution.



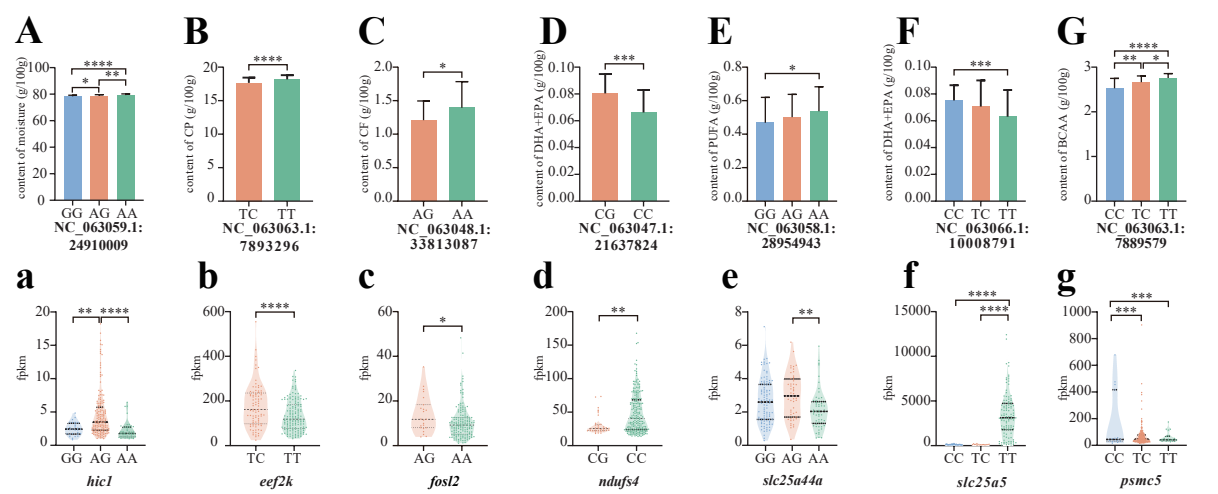
SNP genotyping was conducted using 50K array and transcriptome sequencing data and a total of 5,239,984 SNPs were obtained from a population of 244 *M.amblycephala* individuals, with uniform distribution across chromosomes.

GWAS



Employing the Farm CPU model, GWAS were conducted with 3, 3, 7, 7, 1, and 7 PCs included, respectively. This identified 4, 4, 4, 8, 11, and 7 significant loci, and subsequent annotation of their 675kb upstream/downstream regions yielded 848 candidate genes.

CANDIDATE GENES



Differential analysis of candidate gene expression was performed based on significant SNP genotypes and a total of 133 candidate genes with significant differences were screened out. Integrating gene function, expression levels, and their associations with phenotypes, potential candidate genes regulating each nutritional trait in *M. amblycephala* were inferred: *hic1* (Moisture), *eef2k* (CP), *fosl2* (CF), *ndufs4* and *slc25a5* (DHA+EPA), *slc25a44a* (PUFA), and *psmc5* (BCAA).

CONCLUSION

This study explored the genetic basis of muscle nutritional traits in *M. amblycephala* and addressed gaps in targeted nutritional quality selection. Using a 244-individual population, GWAS was conducted on 6 traits (moisture, CP, CF, PUFA, DHA+EPA, BCAA), identifying 4, 4, 4, 8, 11, 8 significant SNPs respectively. Transcriptomic integration and genotype-specific expression analysis revealed 133 highly differentially expressed genes. Core candidates were inferred: *hic1* (Moisture), *eef2k* (CP), *fosl2* (CF), *ndufs4*/*slc25a5* (DHA+EPA), *slc25a44a* (PUFA), *psmc5* (BCAA). These findings enhance understanding of *M. amblycephala* nutritional trait genetics and provide markers and genes for meat quality breeding. Future work may validate genes and expand populations.



华中农业大学
HUAZHONG AGRICULTURAL UNIVERSITY



水产学院
College of Fisheries



湖北洪山实验室
HUBEI HONGSHAN LABORATORY