Exploration of Fatty Acid-Related Genes in Large Yellow Croaker

Shisheng Wang [a], Zhiyong Wang [a,*] , Ming Fang [a,*]

[a] Key Laboratory of Healthy Mariculture for the East China Sea, Ministry of Agriculture and Rural Affairs, Jimei University, Xiamen, China

Introduction

The large yellow croaker (Larimichthys crocea) is greatly loved due to its rich nutrition and delicious taste，and it is the largest scale of marine fish cultivation in China. The synthesis and conversion abilities of fatty acids in the large yellow croaker are important traits that have significant economic value. These abilities not only affect the nutrition and flavor profile of the fish itself but also the enhanced fatty acid synthesis capability can reduce feed costs. Furthermore, the exploration of fatty acid-related genes provides insights into the fundamental energy metabolism and growth development of fish.
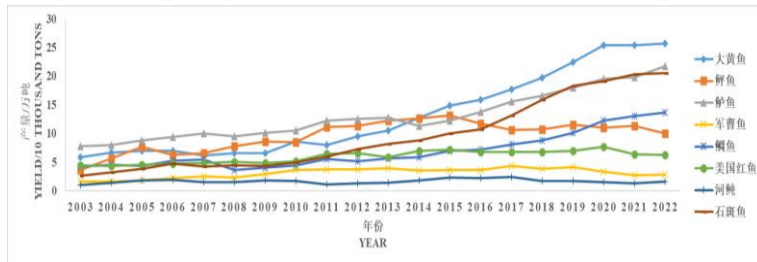

Fig. 1 Changes in the yield of main mariculture fish in China from 2003 to 2022

A genome-wide association study (GWAS) was conducted on the traits of 37 fatty acid contents in the dorsal muscle tissues of 672 large yellow croakers. The reference genome was assembled by our laboratory. Co-localization was performed using the COLOC method with multi-tissue cis-eQTL data from subpopulations of the same group to screen for fatty acid relevant genes.

Methods

Genome-Wide Association Study (GWAS)

In this study, a genome-wide association analysis (GWAS) was conducted using the GEMMA software through a mixed linear model.

Cis expression Quantitative Trait Locus mapping (Cis-eQTL)

Using normalized expression data, an analysis was conducted through TensorQTL to identify which genotypic variations are associated with changes in gene expression levels. The scope of the analysis encompassed a range of 1 Mb upstream and downstream of the gene transcription start sites.

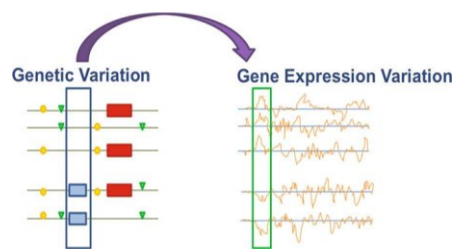Fig. 2 Genome-Wide Association Study (GWAS) work flow.

Fig. 3 Basic Concept of QTL Mapping

Co-localization analysis

COLOC employs an approximate Bayesian factor to estimate the posterior probability that a variation is causal. It is one of the most widely used co-localization methods currently in use.
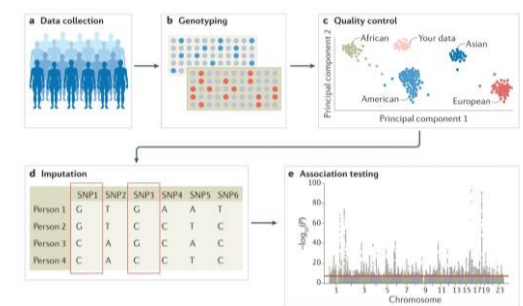
Results

The Bonferroni multiple testing correction was applied to determine significant thresholds in GWAS, with the significant threshold set at a P-value of 1 divided by the number of SNPs, and the highly significant threshold set at a P-value of 0.05 divided by the number of SNPs. In the GWAS results for the content of various fatty acids in crude fat, significant signals were found for traits such as C16:0 (Palmitic Acid), C17:0 (Margaric Acid), C20:4n6 (Arachidonic Acid), among others. Partial results are as follows:
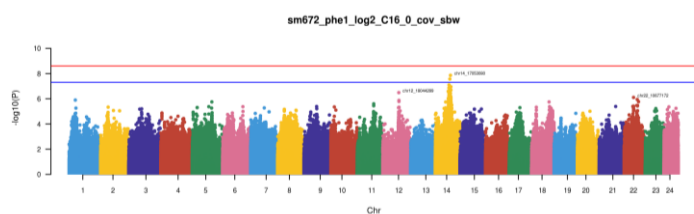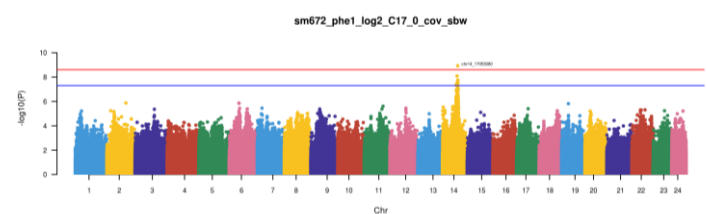

Fig. 5 Manhattan plot of C17:0(Margaric Acid).

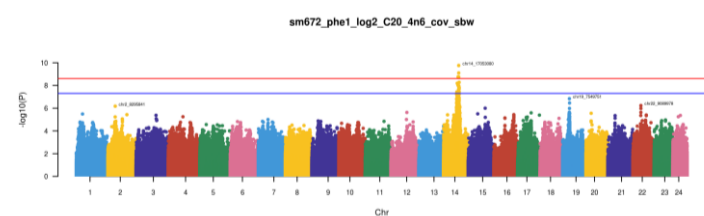Fig. 4 Manhattan plot of C16:0(Palmitic Acid).

Fig. 6 Manhattan plot of C20:4n6(Arachidonic Acid).

Results

Partial Results from Colocalization:

GWAS data and cis-eQTL data from brain tissues were analyzed using the COLOC colocalization method. It was found that the gene with the identifier LYC00146 has a posterior probability (H4) of a common causal variant at 0.75, indicating that this gene may influence the content of caprylic acid (C8:0) in the crude fat of large yellow croaker muscle tissue.
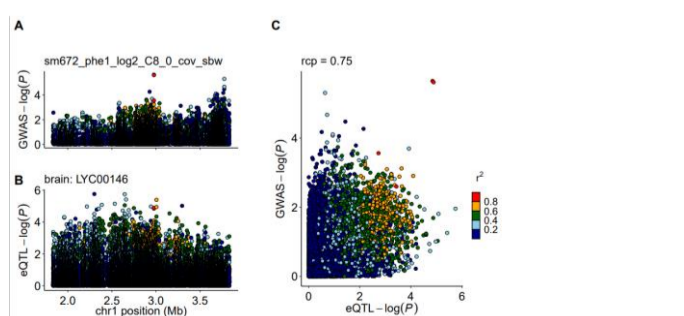
Similarly, the gene with the identifier LYC15757 has an H4 posterior probability of 0.74, suggesting that this gene may affect the content of erucic acid (C22:1n9) in the crude fat of large yellow croaker muscle tissue.
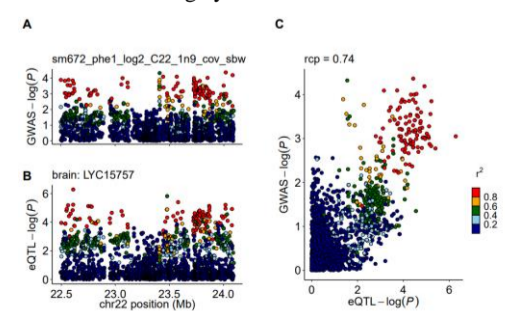

Fig. 7 C8:0 colocalization plot of LYC00146 in brain.

Fig. 8 C22:1n9 colocalization plot of LYC15157 in brain.

Bibliography

1. A Taylor-Weiner, F Aguet, NJ Haradhvala, et al. Scaling computational genomics to millions of individuals with GPUs[J]. Genome Biology, 2019.
2. Zhou X, Stephens M. Genome-wide efficient mixed-model analysis for association studies[J]. Nature Genetics, 2012, 44(7): 821–824.
3. Giambartolomei C, Vukcevic D, Schadt E E, et al. Bayesian Test for Colocalisation between Pairs of Genetic Association Studies Using Summary Statistics[J]. PLOS Genetics, 2014, 10(5): e1004383.