# PANGENOMIC ANALYSIS PROVIDES INSIGHT INTO THE PRECISE IDENTIFICATION OF GERMPLASM IN COMMON CARP

Jian Xu*, Qinglei Xu, Shangqi Li, Li Feng, Linyan Zhou

Fisheries Engineering Institute, Chinese Academy of Fishery Sciences, Beijing, China

*Correspondence: xuj@cafs.ac.cn

## THE DEMAND

Currently trait dissection and breeding analysis in aquaculture species mainly rely on SNPs generated through aligning sequencing data to a single reference genome. However, it is far from enough to cover all genetic variations of germplasm (especially large structural variations). Pangenomic analysis played a key role in recent years in germplasm identification and breeding of plants and livestock animals, which used graphic genomes to accurately screen trait associated genes and variations.

## Methods

This study has completed the sequencing and assembly of the third-generation genomes of blood, muscle, and liver samples from 7 carp germplasms. The third-generation sequencing was performed using the Pacbio platform. Illumina small fragment library was constructed using the Illumina platform. The genome of carp "Longke 12" was assembled using T2T (telomeric to telomeric) strategy. Combined with genome data of 3 other germplasms, a graph pangenome was constructed covering 10 germplasms using the vg giraffe software. Structural variations were also detected using vg software, and PAVs (presence or absence variations) were paid more attention and further analyzed due to the importance in domestication and trait determination.

## THE RESULTS

This study constructed 7 chromosome-level genomes of common carp, then annotations and variation identification were carried out. Combined with the existing genomes of Songpu mirror carp, Heilongjiang carp, and Hebao red carp, a high-precision graphic pan-genome covering 10 species of common carp was constructed.

Table 1. Genome assembly of 7 germplasms of common carp

| Germplasm | Genome size (Mbp) | Contig Number | Contig N50 length (Mbp) |
|---|---|---|---|
| Mirror carp | 1584.90 | 284 | 28.56 |
| Huanghe carp | 1552.37 | 228 | 27.83 |
| Longke 12 carp(YZ) | 1605.69 | 430 | 23.95 |
| Furui carp | 1565.84 | 228 | 29.33 |
| Jian carp | 1581.97 | 257 | 29.21 |
| Longke 11 carp(KB) | 1560.95 | 220 | 29.65 |
| Yuanjiang carp | 1547.67 | 222 | 29.82 |

Totally 40,852 genes were annotated, including core genome of 17,745 genes, dispensable genome of 21,714 genes, and germplasm specific 1,393 genes. Moreover, we identified 7098 PAVs (2738 deletions, 4360 insertions) in KB carp, which were associated with 3286 genes, while 37985 PAVs (16520 deletions, 21465 insertions) were identified in YZ carp, which were associated with 8117 genes. Functional annotation of these genes revealed that the PAV associated genes of KB carp were enriched in immune related pathways such as Th1 and Th2 cell differentiation, JAK-STAT signaling pathway, etc., while the PAV associated genes of YZ carp were enriched in various lipid metabolism related pathways such as steroid biosynthesis and inositol phosphate metabolism.

In summary, this study conducted comparative genomic analysis, annotation of large segment variations, whole genome association analysis, and gene functional identification, identifying functional variations and gene elements related to traits such as disease resistance, intramuscular fat, providing a solid foundation for subsequent breeding of common carp.

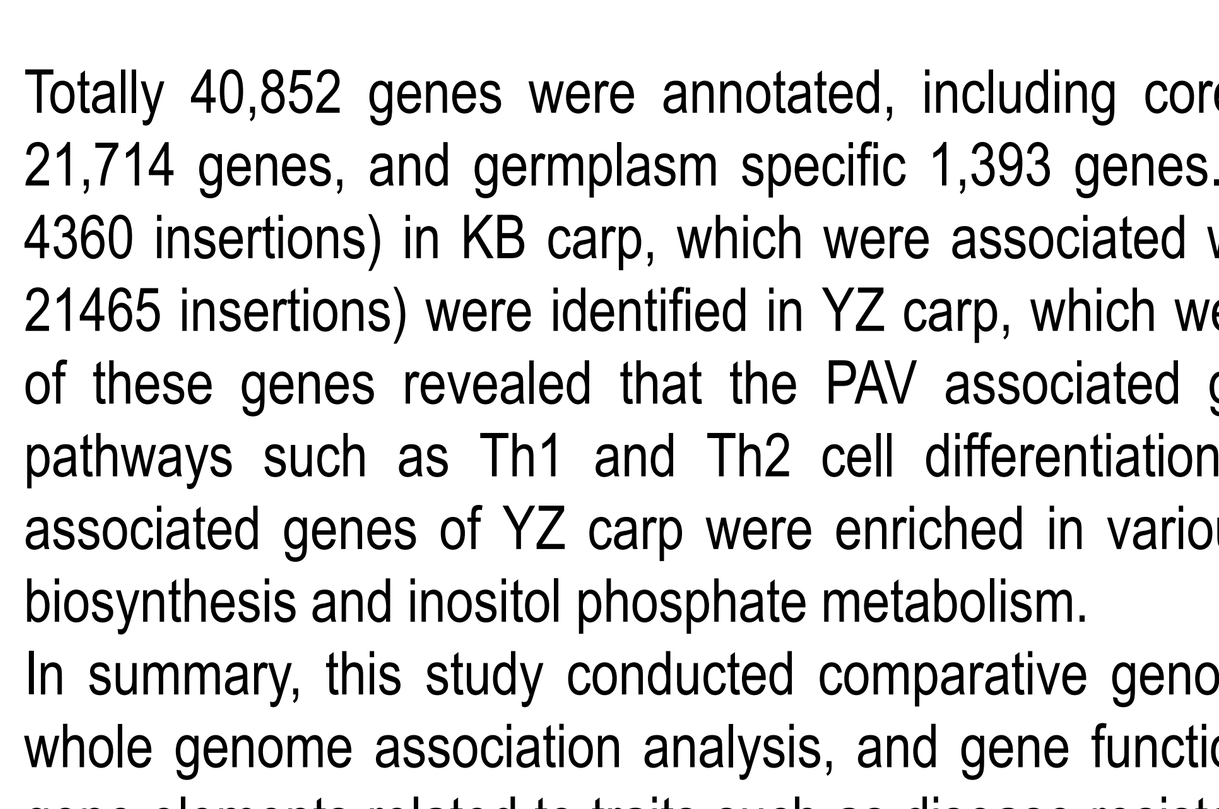


Figure 1. Pangenome covering 10 germplasms of common carp

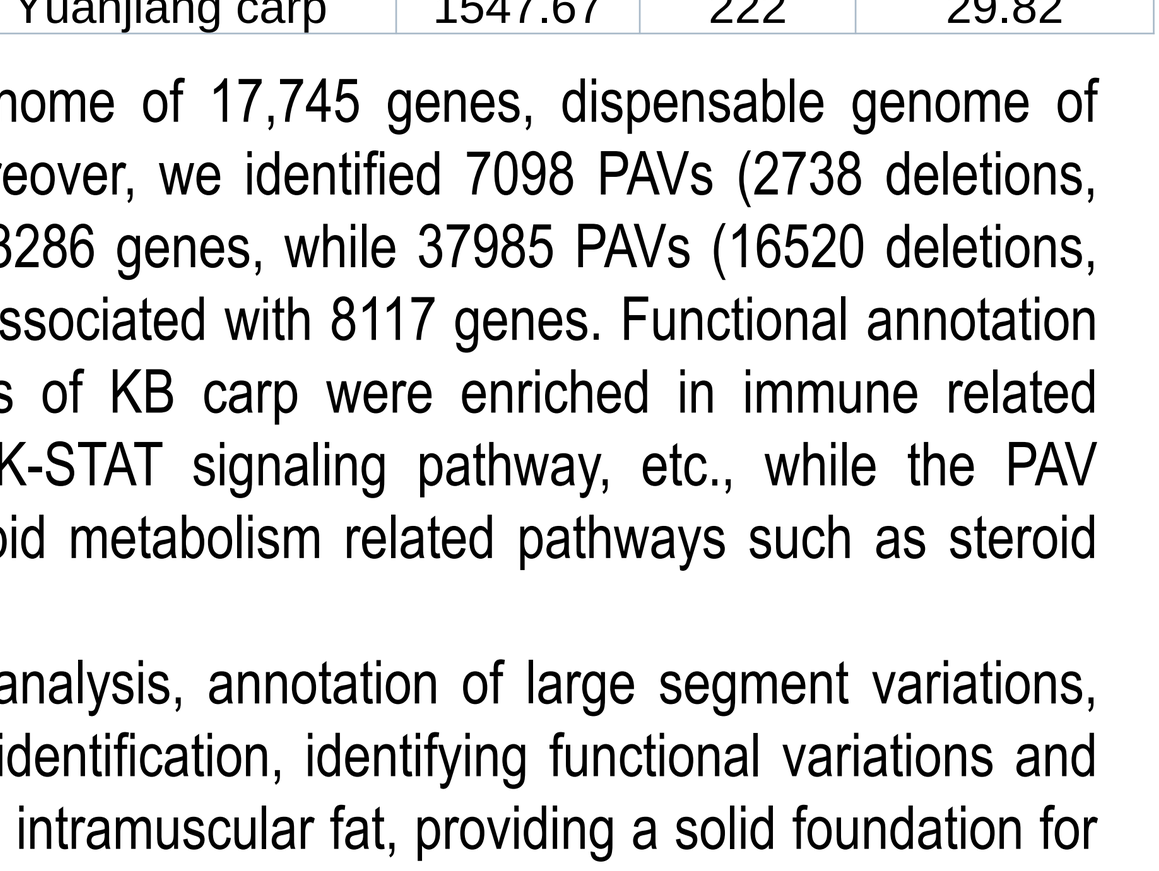


Figure 2. Functional annotation of specific genes and PAVs

## Contribution

中国水产科学研究院 渔业工程研究所 FISHERIES ENGINEERING INSTITUTE

中国水产科学研究院 黑龙江水产研究所 HEILONGJIANG RIVER FISHERIES RESEARCH INSTITUTE

中国水产科学研究院 淡水渔业研究中心 FRESHWATER FISHERIES RESEARCH CENTER,CAFS