Fishery Technology 62 (2025) : 357 - 368

Identification of Immune Related Genes and Associated Simple Sequence Repeats from Kidney Transcriptome of *Cyprinus carpio,* Exposed to Two Different Temperature Conditions

Nishita Chauhan[1], Vindhya Mohindra[1]*, Labrechai Mog Chowdhury[1], Alisha Paul[1], Rajesh Kumar[1], Ajay Kumar Singh[1], Mohd. Imran[2], and J. K. Jena[3]

[1] *ICAR-National Bureau of Fish Genetic Resources, Canal Ring Road, P.O. Dilkusha, Lucknow- 226 002, Uttar Pradesh, India.*

[2] *Institute of Bio-sciences and Technology, Shri Ramswaroop Memorial University, Barabanki- 225003, Uttar Pradesh, India.*

[3] *Indian Council of Agricultural Research (ICAR), Krishi Anusandhan Bhawan - II, New Delhi 110 012, India*

Abstract

Common carp is a widely farmed and economically important freshwater fish species that thrives in diverse aquatic habitats. In the present study, two temperature conditions were used to maximise and capture the transcript repertoire. High-throughput transcriptome sequencing data were used to identify transcriptome and immune related genes, as well as associated simple sequence repeats (SSRs). A total of 67.44 GB of high-quality paired-end reads were generated, which were then assembled and clustered into 133,419 transcripts with a GC content of 51.63%. SSRs were identified in 6,245 genes, with the highest number of dinucleotide repeats (33.42%), followed by tri-nucleotides repeats (13.24%) and tetra/penta/hexa/septa/nona/deca nucleotide repeats (2.43%). Further classification analysis identified 1,355 immune genes, of which the largest groups under innate immune response were co-stimulatory molecules (234), followed by cytokines (203) and RIG-I-like receptors (186). Adaptive immune genes included Immunoglobulins (104) and Interleukin (106), followed by negative regulators of immune system (60). Among the immune categories, 142 SSRs were identified in innate immune genes, 67 in adaptive immune genes and 6 in negative regulators. The generated data of transcriptome-derived SSRs in this study offer a wealth of marker resources related to immune mechanisms for molecular marker development and characterization of germplasm, which will aid in understanding genetic diversity and their application in molecular marker-assisted breeding of *C. carpio* for improving aquaculture sustainability.

Keywords: RNA-seq, transcriptomic analysis, temperature, immune response, SSR markers

Introduction

Fish in aquaculture ponds or tanks can experience sudden changes in their environment, such as temperature fluctuations, diseases, over-population and deteriorating food quality. These elements can trigger a stress reaction and disrupt the equilibrium between the physiology of fish and their environment (Canosa & Bertucci, 2023). This may lead to the depletion of energy reserves and hormone imbalances, which suppress their immune system and increase susceptibility to infectious diseases (Schreck & Tort, 2016). It has also been observed that the immune function is closely linked to the immune response of immune organs (Jiang et al., 2017). The ability to coordinate the innate immune system's response to stress and adapt to changes in the external environment depends on the head kidney, which is the primary lymphatic immune organ in fish (Shved et al., 2009; Bjørgen & Koppang, 2022).

Received 10 February 2025; Revised 14 April 2025; Accepted 16 April 2025

Handling editor: Dr. B. Madhusudana Rao

**Email:* *vindhyamohindra@gmail.com*

The common carp (*Cyprinus carpio* L.), is farmed in most countries and is a significant food fish (Bongers, Sukkel, Gort, Komen, & Richter, 1998). However, in recent years, intensive culture practices have impacted the production of this species due to severe infectious disease outbreaks (Li et al., 2015). Low temperatures suppress most of the fish host's immunity, favouring the growth of the Epizootic Ulcerative Syndrome (EUS) pathogen, an oomycete fungus (Kumar et al., 2020). Despite this, the common carp is recognised as a species naturally resistant to this disease (OIE, 2019) even at lower temperatures (Billard, 1995). Therefore, the generation of basic information on the transcriptome profile to understand the repertoire of expressed immune related genes at different temperatures along with their associated markers, has become necessary. Common carp genetic research has been focused on creating genetic markers for breeding and genetic assessment (Tóth et al., 2020; Sumana et al., 2024), transcriptome profiles (Ji et al., 2012; Kolder et al., 2016), creation of genetic and physical maps (Xu et al., 2011; Zhang, Luan, Cao, & Hu, 2021) and genome sequences (Xu et al., 2014; Wang et al., 2024). However, compared to its genetic investigations, there are relatively few immunogenic studies (Kongchum, Palti, Hallerman, Hulata, & David, 2010; Wang, Zhou, & Jiang, 2024) and these have not yet been thoroughly characterized through immunogenetic approaches with their associated molecular markers, which could aid in future marker-assisted selection programs (Eze, 2019).

Recent developments in next-generation sequencing (NGS) technologies have made it possible to investigate the transcriptome of a species of interest for aquaculture in a high-throughput fashion. High-throughput RNA sequencing (RNA-Seq) provides an effective method for the generation and analysis of whole transcriptome, aiding in identification and quantification of RNA transcripts (Moyerbrailean et al., 2015) and has been used in a variety of aquatic animals to investigate a range of traits, such as growth (Mohindra, Chowdhury, Chauhan, Maurya, & Jena, 2022) and immunological responses (Pietrzak, Mazurkiewicz, & Slawinska., 2020; Montero et al., 2022; Baloch et al., 2023; Wang et al., 2024). Unlike genomic SSRs, EST-SSRs (from genes) can help identify potential functional genes, which can increase the efficacy of marker-assisted selection (Zhao et al., 2019).

In order to obtain a thorough understanding and create a genome resource with a large number of effective genic-SSR molecular markers associated with the immune system, *de novo* transcriptome sequencing of kidney tissues was carried out using Illumina paired-end sequencing at two different temperatures, to maximise the gene expression and the number of transcripts obtained. The data generated in this study provide valuable genetic/genomic resources of the expressed genome, including those related to immune mechanisms, which may be useful for characterization of germplasm and molecular marker-assisted breeding in *C. carpio*.

Materials and Methods

Ten *C. carpio* specimens (16-19 cm, 91-133 g) from the farm ponds of ICAR- NBFGR, Lucknow, were acclimated at room temperature for a month and then exposed to 30°C and 20°C (five each) for another month. This exposure allowed the capture of maximum transcripts expressed at both temperatures. The fishes from both groups were euthanized with MS222 (Sigma Aldrich, USA) and kidney samples were collected, snap-frozen in liquid nitrogen and stored at -80°C in the laboratory.

The total RNA was extracted from the kidney tissues of ten specimens based on the TRI Reagent method (Rio, Ares, Hannon, & Nilsen, 2010) and treated with 10 units of DNase I (TaKaRa, Dalian, China) for 1 hour at 37°C to denature genomic DNA. The quality and quantity of the RNA were assessed using a BioAnalyzer 2100 (Agilent Technologies, Santa Clara, CA) and a NanoDrop 2000 spectrophotometer (ThermoFisher Scientific, CA), respectively. The RNA Integrity Number (RIN) of the purified RNA was determined using an Agilent Tapestation (Agilent Technologies, USA). RNA samples (RIN between 7.0 and 9.5) from five biological replicates of both groups were used to prepare libraries using an Illumina paired-end library sequencing kit, Illumina Hiseq2500 in 125PE mode (Illumina Inc., San Diego, CA, USA).

Transcriptome sequencing was performed by using the Illumina HiSeq2500 (Illumina Inc., USA). Quality assessment of the raw FASTQ reads from the samples was performed using FastQC v.0.11.9 (Andrews, 2017). The raw FASTQ reads were pre-processed using Fastp v.0.20.1 (Chen, Zhou, Chen, & Gu, 2018), followed by quality re-assessment using FastQC and summarization with MultiQC

(Ewels, Magnusson, Lundin, & Käller, 2016). The trimmed fastp reads were aligned against the silva database to filter the rRNA reads using bowtie2 v2.4.2 (Langmead & Salzberg, 2012). The reads that were not aligned to the silva database were subjected to *de novo* assembly using Trinity v.2.13.2 (Grabherr et al., 2011). The Trinity assembled transcript FASTA file was further clustered using CD-HIT-EST (Li & Godzik, 2006; Huang, Niu, Gao, Fu, & Li, 2010). For gene prediction, AUGUSTUS (Stanke & Morgenstern, 2005) was used. The completeness of assembly and filtered transcriptome was assessed using gVolante (Nishimura, Hara, & Kuraku, 2019) against the core Actinopterygii, Vertebrate, and Eukaryote genes databases, using the BUSCO orthology pipeline using gVolante..

The predicted genes from AUGUSTUS were further functionally annotated using Diamond BlastP (OmicsBox, 2019) against the NR database, with reference to Actinopterygii database and with an e-value ≤10^{-5}, similarity score ≥40%. Gene ontology (GO) terms were characterized and classified into biological process (BP), molecular function (MF) and cellular component (CC) using Omics box (OmicsBox, 2019) Further analysis of molecular pathways was performed using the KAAS (KEGG Automatic Annotation Server, Moriya Itoh, Okuda, Yoshizawa, & Kanehisa, 2007) database.

The Microsatellite Identification tool (MISA, version 1.0, http://pgrc.ipkatersleben.de/misa/) was used to identify SSRs with a minimum repeat number of six for dinucleotides and five for tri-, tetra-, penta-, and hexa-nucleotide microsatellites. For downstream analysis, mono-nucleotide SSRs were excluded to avoid non-genomic, artificial SSR loci resulting from the in-depth sequencing of the transcripts. Additionally, immune genes containing SSRs were blasted against the *C. carpio* genome (ASM1834038v1, NCBI RefSeq assembly: GCF_018340385.1) to identify polymorphism in repeats within these immune genes.

Results and Discussion

Transcriptome analysis is a powerful technique used to gain deeper insights into the fundamental mechanisms that govern a host's cell fate, development, and disease progression. Additionally, it is a quick and efficient method for surveying the genome and identifying large numbers of functional genes and molecular markers (Morozova, Hirst, & Marra, 2009). Simple sequence repeats (SSR) are tandem repeats of two to six nucleotides found throughout the genome. These are valuable tools for genomic research, as they are among the most adaptable genetic markers (Zhu et al., 2017). The genetic variety of common carp strains results in unique phenotypes in terms of growth rate, scale pattern, body colour, body form, and disease response and tolerance (Qiao et al., 2022). To identify microsatellite markers and create a reasonably large transcriptome library for future studies, we sequenced ten cDNA libraries, yielding a total of 528.6 million raw reads resulting in a total of 84.47 GB reads. After trimming and removal of contaminants, 477 million trimmed reads, with 67.44 GB high-quality paired-end reads were obtained and used for further bioinformatics analysis (Supplementary Table 1).

De novo transcriptome assembly from the cleaned data resulted in 15,36,525 assembled transcripts,

Table 1. Assembly and functional annotation of *Cyprinus carpio* kidney tissue transcriptome generated in the present study.

	Denovo Assembled transcripts		
1	Total transcripts		15,36,525
2	Transcripts after clustering with CD-HIT		12,75,165
3	Functional Transcripts through AUGUSTS gene prediction		1,33,419
5	Annotated Transcripts through NCBI BLAST		1,21,867
7	No. of Unique genes		39,039
8	No. of significant Gene Ontology (GO) Terms		
		Biological process (BP)	5474
		Cellular component (CC)	1157
		Molecular function (MF)	2979

with the longest transcript being 23,532 bp (Supplementary Table 2, Supplementary Fig. 1). Further clustering resulted in 12,75,165 transcripts, and AGUSTUS predicted 1,33,419 transcripts to have functional annotations, which confirmed 97.40% completeness of the transcriptome through BUSCO analysis (Supplementary Table 3, Supplementary Fig. 2). In previous studies on European common carp, gene prediction based on both RNA-Seq data and *de novo* gene prediction identified the presence of 50,527 genes and approximately similar number of 52,610 genes were predicted on Songpu (*C. carpio* Songpu) (Kolder et al., 2016). A prior study by Ji et al. (2012) characterized the common carp transcriptome and identified 21,16,226 raw sequencing reads with an average length of 331 bp. After removing ambiguous nucleotides, low-quality sequences (quality scores<20), contaminated microbial sequences, ribosomal RNA sequences, and common carp mitochondrial genome sequences, a total of 14,18,591 cleaned reads with an average length of 321 bp and a median length of 328 bp, were harvested using Roche 454 sequencing technology. NCBI accessions for SRA and transcriptome for the present study are given in the Data Availability section.

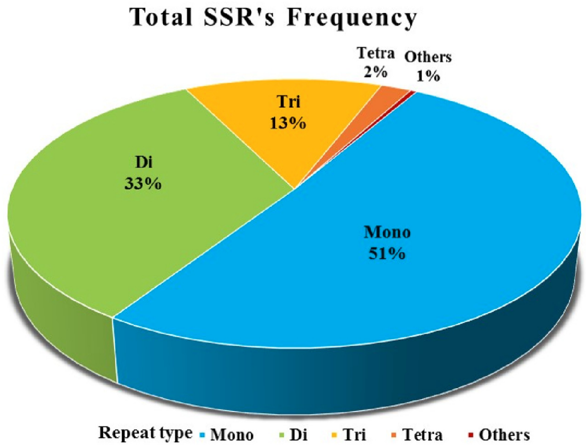

Fig. 1. Frequency of simple sequence repeats (SSR) distribution in unique genes identified from *Cyprinus carpio* kidney transcriptome

In present study, a total of 1,33,419 transcripts were characterized, resulting in the annotation of 80,378 transcripts, corresponding to 39,031 unique genes. Of the remaining,, 32,351 transcripts were hypothetical protein, 4,599 as uncharacterized proteins, and 3,630 low quality proteins, and 908 under unnamed protein product, based on Blastp against the Actinopterygii database with significant simi-

Table 2. Immune genes classified from functionally annotated transcriptome of *Cyprinus carpio* kidney tissue

Immune gene	Numbers	Polymorphic	%
Toll-like receptors (TLR)	23	2	8.70
NOD-like receptors (NLRs)	69	8	11.59
RIG-I-like receptors (RLRs)	186	10	5.38
Complement system	84	4	4.76
Costimulatory molecules	234	11	4.70
Chemokines	40	3	7.50
Lectin family members	115	6	5.22
Interferons	102	13	12.75
Cytokines	203	20	9.85
Total Innate genes	1056		
Immunoglobulins	104	0	0.00
Major Histocompatibility Complex (MHC)	29	0	0.00
Interleukin	106	0	0.00
Total Adaptive genes	239	0	0.00
Negative regulator of immune	60	0	0.00
Total genes	1355		

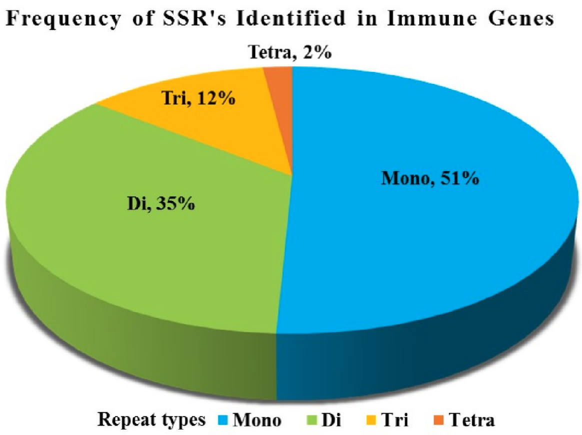

Fig. 2. Frequency of simple sequence repeats (SSR) distribution in unique immune genes identified in *Cyprinus carpio* kidney transcriptome

larities (E value 1×10^{-5}). Additionally, 39,039 unique genes were also annotated using InterProScan and of these 37,149 genes were annotated with GO mapping, and 36,488 genes with GO annotations, respectively. These unique genes yielded 18,701 homology hits in a BLAST search against *C. carpio,* followed by 2,186 and 2,055 genes under *Sinocyclocheilus grahami* and *S. rhinocerous*, respectively (Table 1, Supplementary Fig. 3).

Gene Ontology (GO) databases were used to identify the putative functions of the unigenes. The functions of genes and their products were assigned the GO classification, resulting in 36,488 genes annotated with GO terms, with 145 GO terms in BP category, with the highest number of 7,732 genes under regulation of transcription by signalling; 71

Table 3. Frequency of simple sequence repeats (SSRs) identified from immune related genes of *Cyprinus carpio* kidney tissue transcriptome

Repeats	No. of repeats							Total	Frequency
	5	6	7	8	9	10	>10		
AC/GT	-	30	14	11	14	8	61	138	22.81
AG/CT	-	17	3	5	4	1	16	46	7.60
AT/AT	-	10	2	3	3	2	8	28	4.63
Total di-nucleotide								212	35.04
ATC/ATG	5	2	5	1	2		2	17	2.81
AAC/GTT	9	2	3				0	14	2.31
AGC/CTG	9	3		1			0	13	2.15
AAT/ATT	5	4	1	1			0	11	1.82
AGG/CCT	6	2					0	8	1.32
AAG/CTT	2	2	1				0	5	0.83
ACC/GGT	1	1	1				0	3	0.50
ACG/CGT	1	1					0	2	0.33
CCG/CGG	1						0	1	0.17
Total tri nucleotide								74	12.23
AAAC/GTTT	2						0	2	0.33
AAAG/CTTT	1						0	1	0.17
AAAT/ATTT	1						0	1	0.17
ACAT/ATGT	1						0	1	0.17
ACTC/AGTG		2					0	2	0.33
AGAT/ATCT	1	1				1	2	5	0.83
Total tetra nucleotide								12	1.98
Total	45	77	30	22	23	84	324	605	
Frequency	7.43	12.72	4.95	3.63	3.80	13.88	53.55		

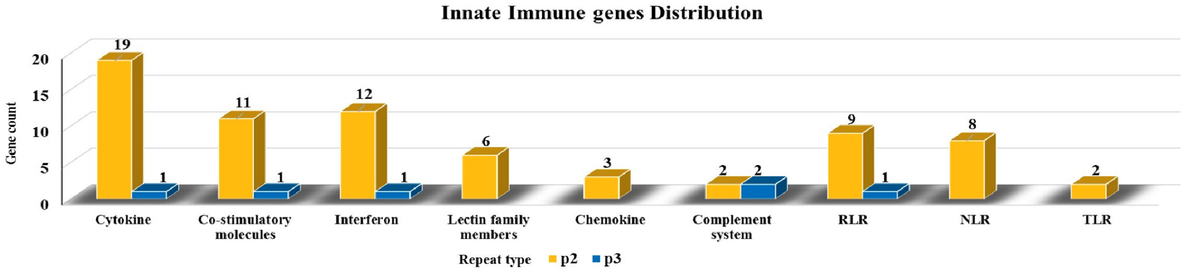

Fig. 3. Innate immune genes with polymorphic simple sequence repeats (SSRs) identified in *Cyprinus carpio* kidney tissue transcriptome

terms in CC category, with the highest number of 8,225 genes under nucleus; and 102 terms under MF category, with the highest number of 5,704 genes under transferase activity (Supplementary Table 4, Supplementary Fig. 4). According to GO distribution at level 2, 28,437 genes were found under GO:0110165: cellular anatomical structure (CC), followed by 24,284 genes under GO:0009987: cellular process (BP) and 12,629 genes under GO:0003824: catalytic activity (MF) (Supplementary Fig. 5). As reported in previous transcriptome analyses of the naked carp (*Gymnocypris przewalskii*) the results showed that the annotated unigenes were assigned to similar terms of the biological process category (Tong, Zhang, Zhang, & Zhao, 2015).

Functional annotation against the KEGG database identified a total of 9,436 genes with KO terms, of which 1,778 genes were involved in organismal systems, with 630 genes under the immune system; 1,474 in environmental information processing, with the highest number of 1,230 genes under signal transduction; 1,446 in cellular process, with the highest number of 641 genes under transport and catabolism; 1,165 genes under genetic information processing, with 378 genes under folding, sorting and degradation; and 1,136 in metabolism, with 242 genes under lipid metabolism (Supplementary Table 5). In the immune system category, 102 genes were associated with the chemokine signalling pathway [PATH: ko04062] and in the signal transduction category, 208 genes were associated with the PI3K-Akt signalling pathway [PATH: ko04151].

Polymorphic SSR markers are useful for various applications, such as genetic diversity studies, genetic linkage mapping, comparative genomics, and trait-association analysis research. In the analysis of putative SSRs in the present study, mononucleotide repeats were excluded due to the differentiation of polyadenylation sites and false-positive mononucleotide repeats caused by sequencing errors.

In the present study, a total of 19,154 genes were identified, with 2,083 SSRs in compound formation and 6,245 genes with the highest number of 6,403 (33.42%) di-nucleotide repeats; followed by 2,536 (13.24%) tri-nucleotide repeats and 467 (2.43%) tetra/penta/hexa/septa/nona/deca-nucleotide repeats (Supplementary Table 6, Fig. 1). A previous study by Wang, Liao, Cheng, Yu, and Tong (2007), in common carp identified novel EST-SSR markers in 10,088 ESTs, of which 555 (about 5.5%) contained SSRs within 400 singletons. Dinucleotide and trinucleotide repeats made up the majority of these common carp EST-SSRs. In the present study, the most abundant repeats were dinucleotide motifs with four types of nucleotide repeats, among which AC/GT and AG/CT repeats were the most abundant, attributing 21.42% (4102) and 7.92% (1517), respectively. Dinucleotides are the predominant type of microsatellite repeats in most aquaculture species, though trinucleotide repeats are more abundant in plants (Kantety La Rota, Matthews, & Sorrells, 2002; Chen, Zhou, Choi, Huang, & Gmitter, 2006). In this

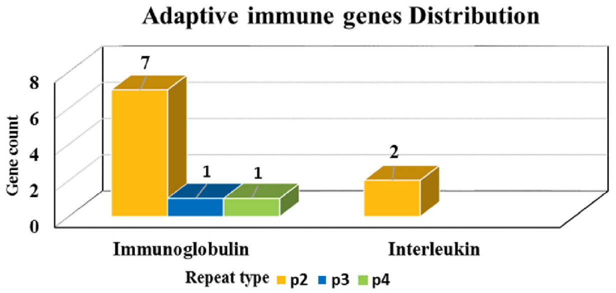

Fig. 4. Adaptive immune genes with polymorphic simple sequence repeats (SSRs) identified in *Cyprinus carpio* kidney tissue transcriptome

Table 4. Polymorphic SSRs identified in innate immune related genes in kidney tissue transcriptome of *Cyprinus carpio,* blasted against genomes

Innate Immune genes	Gene symbol	Gene description	SSR type	Position	SSR	Repeat count	Blast with genome	Repeat number in genome
Chemokine	*loc109100617*	Chemokine XC receptor 1-like	p2	3′UTR	CA	8	NC_056573.1	9
	ccl-c5a	Chemokine CCL-c5a-like protein	p2	3′UTR	TG	6	NC_056601.1	10
	loc109106102	Chemokine- like receptor 1	p2	3′UTR	TG	11	NC_056612.1	14
Complement system	*c1qtnf4*	Complement C1q tumour necrosis factor-related protein 4	p2	5′UTR	GA	6	NC_056578.1	13
	c5-2	Complement component C5-2	p2	5′UTR	TG	12	NC_056601.1	35
	a1cf	Predicted: APOBEC1 complementation factor-like	p3	5′UTR	TTG	5	NC_056583.1	7
	loc109076170	Complement system-related protein c1qr-1	p3	3′UTR	TTG	7	NC_056613.1	4
Co-stimulatory molecule	*loc109058010*	T-cell surface glycoprotein CD3 zeta chain-like	p2	5′UTR	AC	9	NC_056572.1	8
	loc109088766	Cdc42 effector protein 2-like	p2	3′UTR	TA	6	NC_056576.1	9
	loc109049851	T-cell surface glycoprotein CD4-like	p2	3′UTR	TG	12	NC_056587.1	8
	cdc42se2	CDC42 small effector protein 2	p2	3′UTR	GT	7	NC_056587.1	8
	loc132142960	T-cell surface glycoprotein CD3 zeta chain-like isoform X1	p2	3′UTR	AC	11	NC_056597.1	14
	loc132142960	T-cell surface glycoprotein CD3p2 zeta chain-like isoform X1	5′UTR	CA	12	NC_056597.1		14
	loc109058010	T-cell surface glycoprotein CD3 zeta chain-like	p2	5′UTR	TG	7	NC_056597.1	14
	loc122136444	CD9 antigen-like isoform X1	p2	3′UTR	CT	11	NC_056599.1	12
	loc109079512	T-cell surface glycoprotein CD3 epsilon chain-like	p2	3′UTR	TG	14	NC_056611.1	19
	cd276	CD276 antigen isoform X1	p2	5′UTR	CA	12	NW_024879200.1	6
	loc109054117	CD276 antigen isoform X3	p2	5′UTR	CA	12	NW_024879200.1	6
	loc109054500	CD209 antigen-like protein D	p3	5′UTR	TTA	6	NC_056612.1	8
Cytokine	*tnip2*	TNFAIP3-interacting protein 2	p2	3′UTR	GT	9	NC_056572.1	5
	loc113055050	Suppressor of cytokine signalling 7-like isoform X2	p2	3′UTR	AT	12	NC_056582.1	27
	socs7	PREDICTED: suppressor of cytokine signalling 7	p2	3′UTR	AT	12	NC_056582.1	27
	loc113055050	PREDICTED: suppressor of cytokine signalling 7-like isoform X2	p2	3′UTR	AT	12	NC_056582.1	27
	dock2	Dedicator of cytokinesis 2	p2	5′UTR	TG	10	NC_056585.1	17
	dock3	Dedicator of cytokinesis protein 3 isoform X1, partial	p2	5′UTR	AC	7	NC_056593.1	6
	dock4b	Dedicator of cytokinesis protein 4bp2 isoform X1, partial	3′UTR	CA	8	NC_056596.1		11
	dock4	Dedicator of cytokinesis protein 4	p2	3′UTR	CA	8	NC_056596.1	11
	loc109104913	Cytokine receptor-like factor 1 isoform X1	p2	3′UTR	TG	12	NC_056598.1	13
	loc109111375	Cytokine receptor-like factor 1 isoform X2	p2	5′UTR	TG	12	NC_056598.1	13
	si:ch211-202h22.7	Lipopolysaccharide-induced tumor necrosis factor-alpha factor homolog	p2	3′UTR	AC	18	NC_056599.1	5
	loc109110284	Dedicator of cytokinesis protein 7-like isoform X2	p2	3′UTR	TG	6	NC_056599.1	6
	loc109063179	TNF receptor-associated factor 2	p2	5′UTR	AT	6	NC_056601.1	5
	loc109095670	Dedicator of cytokinesis 9-like protein	p2	5′UTR	GT	12	NC_056605.1	15
	loc109098971	Tumor necrosis factor receptor super family member 9-like isoform X2	p2	5′UTR	GT	7	NC_056607.1	9
	loc109102414	TNFAIP3-interacting protein 1-like isoform X1	p2	3′UTR	AT	10	NC_056610.1	8
	loc109082862	Tumor necrosis factor receptor superfamily member 19-like	p2	3′UTR	GT	11	NC_056611.1	12
	tnfaip8l3	Tumor necrosis factor alpha-induced protein 8-like protein 3	p2	5′UTR	AG	16	NC_056614.1	11
	dock4	Dedicator of cytokinesis protein 4	p2	5′UTR	TG	6	NC_056621.1	7
	loc109097924	Tumor necrosis factor alpha-induced protein 8-like protein 1 isoform X2	p3	3′UTR	TTG	5	NC_056593.1	5

Interferon	*loc109065755*	Interferon regulatory factor 4-like	p2	3′UTR	TG	8	NC_056573.1	36
	irf5	Interferon regulatory factor 5	p2	5′UTR	AC	6	NC_056575.1	7
	loc109067118	Interferon regulatory factor 6-like isoform X1	p2	5′UTR	AC	6	NC_056575.1	7
	loc109082506	Interferon alpha/beta receptor 1a-like isoform X3	p2	3′UTR	CA	11	NC_056580.1	23
	loc109065507	Interferon alpha/beta receptor 1b-like	p2	5′UTR	AC	9	NC_056580.1	6
	prkra	Interferon-inducible double-stranded RNA-dependent protein kinase activator A homolog isoform X3	p2	5′UTR	GT	20	NC_056580.1	16
	loc116332750	Interferon-induced protein with tetratricopeptide repeats 5-like	p2	3′UTR	AC	11	NC_056583.1	9
	loc109085153	Interferon-induced GTP-binding protein Mx2-like	p2	5′UTR	GT	27	NC_056596.1	12
	prkra	Interferon-inducible double-stranded RNA-dependent protein kinase activator A homolog isoform X2	p2	5′UTR	GT	20	NC_056605.1	37
	loc109111911	Interferon gamma receptor 1-like	p2	3′UTR	TC	16	NC_056609.1	9
	irf-1	Interferon regulatory factor 1	p2	3′UTR	AT	9	NC_056617.1	8
	loc109070803	Interferon regulatory factor 7 isoform X1	p2	3′UTR	CA	11	NC_056621.1	7
	isg20l2	PREDICTED: interferon-stimulated 20 kda exonuclease-like 2	p3	3′UTR	AAT	6	NC_056587.1	5
Lectin family members	*loc109078724*	Fucolectin-like	p2	5′UTR	GT	14	NC_056576.1	9
	loc109072365	Sialic acid-binding Ig-like lectin 14	p2	5′UTR	AT	10	NC_056587.1	13
	loc109083161	Galectin-related protein-like	p2	5′UTR	AC	12	NC_056609.1	6
	preb	Prolactin regulatory element-binding protein	p2	3′UTR	AT	6	NC_056612.1	7
	loc109112338	Plectin isoform X13	p2	5′UTR	GT	8	NC_056612.1	6
	loc109078027	Galectin-8-like isoform X1	p2	3′UTR	AT	9	NC_056616.1	12
NLR	*loc122139768*	NACHT, LRR and PYD domains-containing protein 1a allele 5-like isoform X3	p2	5′UTR	GT	16	NC_056582.1	37
	loc122134236	NACHT, LRR and PYD domains-containing protein 12-like isoform X1	p2	5′UTR	TA	7	NC_056583.1	45
	loc109092179	PREDICTED: protein NLRC3-like isoform X3	p2	5′UTR	AC	17	NC_056595.1	15
	loc122137612	PREDICTED: NACHT, LRR and PYD domains-containing protein 6-like	p2	5′UTR	TG	12	NC_056600.1	8
	nlrx1	NLR family member X1 isoform X3	p2	5′UTR	GT	12	NC_056611.1	28
	loc109103209	NLR family member X1 isoform X1	p2	5′UTR	TA	6	NC_056611.1	11
	loc109053867	NACHT, LRR and PYD domains-containing protein 12 isoform X1	p2	3′UTR	TA	11	NC_056612.1	10
	loc109113594	PREDICTED: protein NLRC3-like isoform X1	p2	5′UTR	AC	17	NC_056620.1	10
RLR	*loc109067147*	Probable ATP-dependent RNA helicase DHX58	p2	5′UTR	GT	20	NC_056574.1	8
	dhx36	ATP-dependent DNA/RNA helicase DHX36	p2	5′UTR	TG	12	NC_056589.1	8
	loc109085958	Retinoic acid-induced protein 2-like isoform X1	p2	5′UTR	GA	11	NC_056594.1	12
	loc109085957	Retinoic acid-induced protein 2-like	p2	5′UTR	GA	11	NC_056594.1	13
	loc109076721	Putative ATP-dependent RNA helicase DDX23	p2	5′UTR	TA	8	NC_056594.1	29
	rarga	Retinoic acid receptor gamma-2, partial	p2	5′UTR	TC	10	NC_056594.1	18
	loc109101281	Retinoic acid receptor alpha-A-like isoform X2	p2	5′UTR	AC	11	NC_056608.1	15
	loc109057926	Retinoic acid receptor alpha-like isoform X3	p2	5′UTR	TG	12	NC_056615.1	22
	loc109110237	PREDICTED: retinoic acid receptor gamma-A isoform X3	p2	3′UTR	AC	6	NC_056619.1	28
	loc109045792	Probable ATP-dependent RNA helicase DDX31	p3	3′UTR	TAA	6	NC_056576.1	7
TLR	*loc109045012*	Toll-like receptor 8	p2	3′UTR	GT	11	NC_056605.1	12
	loc109067352	Toll-like receptor 13	p2	3′UTR	AT	12	NC_056612.1	14

study, TG dinucleotides were predominant, followed by AC repeats. Similarly, a previous study by Wang et al. (2007) also found dinucleotide repeats to be the most abundant in common carp microsatellite analysis based on EST sequences. Among the trinucleotide repeats, ATC/ATG and AGG/CCT were the most common, with 509 (2.66%) and 476 (2.49%) repeats, respectively (Supplementary Fig. 6, Supplementary Table 7).

Immune genes classification identified a total of 1,295 genes. Among these, in the innate immune response, 23 genes were associated with Toll-like receptor, 69 genes with the NOD-like receptor, 186 gene with the RIG-I-like receptor, 40 genes with Chemokines, 84 genes with the Complement system, 203 genes with Cytokines, 102 genes with Interferon, 115 genes with the Lectin family, and 234 genes with Co-stimulatory molecules (Table 2).

In the adaptive immune response, 104 genes were associated with Immunoglobulins, 29 genes with MHC class I and II, 106 genes with Interleukin and 60 genes with negative regulator of the immune system (Table 2).

A total of 605 SSRs were identified in immune genes, with 212 dinucleotide repeats. Among these, majority of repeats were AC/GT (138, 22.8%), followed by AG/CT (48, 7.6%) and 74 trinucleotide repeats, with majority of repeats being ATC/ATG (17, 2.81%), followed by AAC/GTT (14, 2.31%) (Table 3, Fig. 2, Supplementary Table 8, Supplementary Fig. 7). Of these SSRs identified in the genes of immune categories, 142 were found within innate immune genes, 67 within adaptive immune genes, and 6 within negative regulators of the immune system (Supplementary Fig. 8). In innate immune genes, 16 genes belonging to the RIG-I-like receptors (RLRs) category contained 20 di-nucleotide repeats,

Table 5. Polymorphic SSR identified in Adaptive immune related genes in kidney tissue transcriptome of *Cyprinus carpio* blasted against genomes

Adaptive Immune genes	Gene name	Sequence description	SSR type raw	Position	SSR raw	Repeat count transcriptome	Blast with genome	Repeat number genome
Immunoglobulin	*loc109070054*	Leucine-rich repeat and immunoglobulin-like domain-containing nogo receptor-interacting protein 3	p2	5′UTR	TC	8	NC_056573.1	9
	loc109098546	Immunoglobulin superfamily member 21-like isoform X2	p2	5′UTR	GA	12	NC_056582.1	10
			p2	5′UTR	GA	20	NC_056582.1	24
	loc109083627	Leucine-rich repeats and immunoglobulin-like domains protein 1						
	loc109100532	V-type immunoglobulin domain-containing suppressor of T-cell activation-like	p2	3′UTR	AC	7	NC_056584.1	9
	loc109079510	Immunoglobulin superfamily DCC subclass member 4-like	p2	5′UTR	TG	6	NC_056603.1	5
	loc109095485	Immunoglobulin-like domain-containing receptor 1	p2	3′UTR	TG	10	NC_056605.1	8
	loc109082378	Leucine-rich repeat and immunoglobulin-like domain-containing nogo receptor-interacting protein 1 isoform X2	p2	3′UTR	CA	7	NC_056621.1	6
	rbpja	Recombination signal binding protein for immunoglobulin kappa J region a isoform X1	p3	5′UTR	TAT	8	NC_056572.1	9
	loc109094002	PREDICTED: high affinity immunoglobulin epsilon receptor subunit gamma-like	p4	5′UTR	GTGA	6	NC_056573.1	5
Interleukin	*loc109094694*	Interleukin-17 receptor C-like	p2	5′UTR	AC	8	NC_056579.1	5
	loc109084617	Interleukin-1 receptor accessory protein-like 1-B	p2	3′UTR	TG	12	NC_056582.1	16
	il-6a	Interleukin-6a	p2	3′UTR	AC	10	NC_056590.1	16
	il7r	Interleukin-7 receptor subunit alpha	p2	3′UTR	TA	11	NC_056592.1	9

13 tri-nucleotide repeats, and 2 tetra-nucleotide repeats. The costimulatory molecule category had 20 SSRs, consisting of 21 di-nucleotide repeats and 13 tri-nucleotide repeats. In the cytokine category, 33 genes had 29 di-nucleotide and 5 tri-nucleotide repeats. Within adaptive immune genes, 16 genes under immunoglobulins contained 9 di-, 7 tri-, and 1 tetra-nucleotide repeats, while the genes under interferons contained 16 di-, 5 tri- and 1 tetra-nucleotide repeats (Supplementary Table 9).

For further analysis to detect polymorphisms, immune genes identified were blasted against the *C. carpio* genome (ASM1834038v1, NCBI RefSeq assembly: GCF_018340385.1). The results revealed that 78 polymorphic SSRs were found in genes within innate immune categories, of which 72 were di-nucleotide and 6 were tri-nucleotide repeats. Among these, 20 polymorphic (19 di- and 1 tri repeats) SSRs were identified within Cytokines (Table 3, Figure 3). In the adaptive immune category, polymorphic SSRs were found in 13 genes, with 9 (7 di-, 1 tri-, 1 tetra repeats) in immunoglobulin genes (Table 4, Fig. 4). These 78 polymorphic SSRs were located in the 5′UTR (42) and 3′UTR regions (36), and polymorphisms were observed only in the innate immune genes, with the highest frequency found in interferons (12.75%) (Table 1 and 4). These identified polymorphic SSR loci offer a wealth of marker resources for mapping genes of significant economic traits. They also have the potential for use in genetic resource conservation, the study of immune related gene markers and association studies in *C. carpio*.

In conclusion, a comprehensive genomic resource was generated in this study through the sequencing and *de novo* assembly of the *C. carpio* kidney transcriptome, which yielded 12,75,165 transcripts. A total of 39,031 unique genes were annotated against the Actinopterygii database. Among these, 1,355 immune genes (1295 innate and adaptive immune genes, 60 negative regulators of immune system) were identified, and SSRs were identified in 605 immune genes, dinucleotide repeats being the dominant. Polymorphism was observed in SSRs of 75 innate and 13 adaptive immune genes. Thus, the findings from the present study highlight the effectiveness of utilizing the expressed genome related to immune mechanisms to mine for SSRs, revealing a wealth of genomic information that can be harnessed for various applications. These applications include marker-assisted breeding, population and evolutionary studies, germplasm characterisation, and gene associations of *C. carpio*.

Availability of data and materials

All data generated during this study have been submitted to NCBI, at DDBJ/EMBL/GenBank under transcriptome accession number GKQF00000000. The version described in this paper is the first version, GKQF00000000.

Acknowledgments

The authors are grateful to the Director, ICAR-NBFGR, Lucknow, India for providing the facilities for this work. This work was carried out under ICAR-NBFGR component of ICAR-Consortium Research Platform (CRP) on Genomics, New Delhi, India, and financial support by ICAR-CRP Genomics is duly acknowledged. Dr. Rajeev K. Singh and Dr. B. Kushwaha are acknowledged for their support.

Ethics statement

All the protocols followed were approved by the Institute Animal Ethics Committee (IAEC), ICAR-NBFGR, Lucknow, India, vide G/IAEC/2022/6, Dated 25-05-2022. All methods were performed in accordance with the relevant guidelines and regulations.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.56093/ft.v62i3.164657.

References

Andrews, S. (2017). *FASTQC: a quality control tool for high throughput sequence data*, Babraham Bioinformatics. https://www.bioinformatics.babraham.ac.uk/projects/fastqc/.

Baloch, A. A., Steinhagen, D., Gela, D., Kocour, M., Piaèková, V., & Adamek, M. (2023). Immune responses in carp strains with different susceptibility to carp edema virus disease. *PeerJ*, *11*, Article e15614. https://doi.org/10.7717/peerj.15614.

Billard, R. (1995). *Textbook of Carp biology and culture*. Praxis publishing, Chichester, U.K.

Bjørgen, H., & Koppang, E. O. (2022). Anatomy of teleost fish immune structures and organs. In K. Buchmann, & C. J. Secombes (Eds.), *Principles of fish immunology: From cells and molecules to host protection* (pp. 1-30). Springer, Cham.

Bongers, A. B. J., Sukkel, M., Gort, G., Komen, J., & Richter, C. J. J. (1998). Development and use of genetically uniform strains of common carp in experimental animal research. *Laboratory Animals*, *32*(4), 349-363. https://doi.org/10.1258/002367798780599749.

Canosa, L. F., & Bertucci, J. I. (2023). The effect of environmental stressors on growth in fish and its endocrine control. *Frontiers in Endocrinology*, *14*, Article 1109461. https://doi.org/10.3389/fendo.2023.1109461.

Chen, C., Zhou, P., Choi, Y. A., Huang, S., & Gmitter, F. G. (2006). Mining and characterizing microsatellites from citrus ESTs. *Theoretical and Applied Genetics*, *112*, 1248-1257. https://doi.org/10.1007/s00122-006-0226-1.

Chen, S., Zhou, Y., Chen, Y., & Gu, J. (2018). Fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics*, *34*(17), i884-i890. https://doi.org/10.1093/bioinformatics/bty560.

Ewels, P., Magnusson, M., Lundin, S., & Käller, M. (2016). MultiQC: summarize analysis results for multiple tools and samples in a single report. *Bioinformatics*, *32*(19), 3047-3048. https://doi.org/10.1093/bioinformatics/btw354.

Eze, F. (2019). Marker-assisted selection in fish: A review. *Asian Journal of Fisheries and Aquatic Research*, *3*(4), 1-11. https://doi.org/10.9734/ajfar/2019/v3i430038.

Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q., Chen, Z., Mauceli, E., Hacoben, N., Gnirke, A., Rhind, N., di Palma, F., Birren, B. W., Nusbaum, C., Lindblad-Toh, K., Friedman, N., & Regev, A. (2011). Trinity: reconstructing a full-length transcriptome without a genome from RNA-Seq data. *Nature Biotechnology*, *29*(7), 644-652. https://doi.org/10.1038/nbt.1883.

Huang, Y., Niu, B., Gao, Y., Fu, L., & Li, W. (2010). CD-HIT Suite: a web server for clustering and comparing biological sequences. *Bioinformatics*, *26*(5), 680-682. https://doi.org/10.1093/bioinformatics/btq003.

Ji, P., Liu, G., Xu, J., Wang, X., Li, J., Zhao, Z., Zhang, X., Zhang, Y., Xu, P., & Sun, X. (2012). Characterization of common carp transcriptome: sequencing, de novo assembly, annotation and comparative genomics. *PloS One*, *7*(4), Article e35152. https://doi.org/10.1371/journal.pone.0035152.

Jiang, R., Tang, J., Chen, Y., Deng, L., Ji, J., Xie, Y., Wang, K., Jia, W., Chu, W. M., & Sun, B. (2017). The long noncoding RNA lnc-EGFR stimulates T-regulatory cells differentiation thus promoting hepatocellular carcinoma immune evasion. *Nature Communications*, *8*(1), Article 15129. https://doi.org/10.1038/ncomms15129.

Kantety, R. V., La Rota, M., Matthews, D. E., & Sorrells, M. E. (2002). Data mining for simple sequence repeats in expressed sequence tags from barley, maize, rice, sorghum and wheat. *Plant Molecular Biology*, *48*, 501-510. https://doi.org/10.1023/A:1014875206165.

Kolder, I. C. R. M., van der Plas-Duivesteijn, S. J., Tan, G., Wiegertjes, G. F., Forlenza, M., Guler, A. T., Travin, D. Y., Nakao, M., Moritomo, T., Irnazarow, I., den Dunnen, J. T., Anvar, S. Y., Jansen, H. J., Dirks, R. P., Palmblad, M., Lenhard, B., Henkel, C. V., & Spaink, H. P. (2016). A full-body transcriptome and proteome resource for the European common carp. *BMC Genomics*, *17*, Article 701. https://doi.org/10.1186/s12864-016-3038-y.

Kongchum, P., Palti, Y., Hallerman, E. M., Hulata, G., & David, L. (2010). SNP discovery and development of genetic markers for mapping innate immune response genes in common carp (*Cyprinus carpio*). *Fish & Shellfish Immunology*, *29*(2), 356-361. https://doi.org/10.1016/j.fsi.2010.04.013.

Kumar, I. P., Sarkar, P., Raju, V. S., Manikandan, V., Guru, A., Arshad, A., Elumalai, P., & Arockiaraj, J. (2020). Pathogenicity and pathobiology of epizootic ulcerative syndrome (EUS) causing fungus *Aphanomyces invadans* and its immunological response in Fish. *Reviews in Fisheries Science & Aquaculture*, *28*(3), 358-375. https://doi.org/10.1080/23308249.2020.1753167.

Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nature Methods*, *9*(4), 357-359. https://doi.org/10.1038/nmeth.1923.

Li, W., & Godzik, A. (2006). Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics*, *22*(13), 1658–1659. https://doi.org/10.1093/bioinformatics/btl158.

Li, G., Zhao, Y., Liu, Z., Gao, C., Yan, F., Liu, B., & Feng, J. (2015). De novo assembly and characterization of the spleen transcriptome of common carp (*Cyprinus carpio*) using Illumina paired-end sequencing. *Fish & Shellfish Immunology*, *44*(2), 420-429. https://doi.org/10.1016/j.fsi.2015.03.014.

Mohindra, V., Chowdhury, L. M., Chauhan, N., Maurya, R. K., & Jena, J. K. (2022). Transcriptome analysis revealed hub genes for muscle growth in Indian major carp, *Catla catla* (Hamilton, 1822). *Genomics*, *114*(4), Article 110393. https://doi.org/10.1016/j.ygeno.2022.110393.

Montero, R., Chan, J. T. H., Köllner, B., Kuchta, R., Vyslouzil, J., Podhorec, P., Holzer, A. S., & Korytáø, T. (2022). The Acute immune responses of the common carp *Cyprinus carpio* to PLGA microparticles—the interactions of a teleost fish with a foreign material. *Biomolecules*, *12*(2), Article 326. https://doi.org/10.3390/biom12020326.

Moriya, Y., Itoh, M., Okuda, S., Yoshizawa, A. C., & Kanehisa, M. (2007). KAAS: an automatic genome annotation and pathway reconstruction server. *Nucleic Acids Research*, *35*(2), W182-W185. https://doi.org/10.1093/nar/gkm321.

Morozova, O., Hirst, M., & Marra, M. A. (2009). Applications of new sequencing technologies for transcriptome analysis. *Annual Review of Genomics and Human Genetics*, *10*(1), 135-151. https://doi.org/10.1146/

annurev-genom-082908-145957.

Moyerbrailean, G. A., Davis, G. O., Harvey, C. T., Watza, D., Wen, X., Pique-Regi, R., & Luca, F. (2015). A high-throughput RNA-seq approach to profile transcriptional responses. *Scientific Reports*, *5*(1), Article 14976. https://doi.org/10.1038/srep14976.

Nishimura, O., Hara, Y., & Kuraku, S., (2019). Evaluating genome assemblies and gene models using gVolante. In N. Kollmar (Ed.), *Gene Prediction: Methods and Protocols* (pp. 247-256). Humana, New York.

OIE. (2019). Infection with *Aphanomyces invadans* (Epizootic ulcerative syndrome). *Manual of diagnostic tests for aquatic animals*.

BioBam Bioinformatics. *(2019, March 3). OmicsBox – Bioinformatics made easy.* https://www.biobam.com/omicsbox.

Pietrzak, E., Mazurkiewicz, J., & Slawinska, A. (2020). Innate immune responses of skin mucosa in common carp (*Cyprinus carpio*) fed a diet supplemented with galactooligosaccharides. *Animals*, *10*(3), Article 438. https://doi.org/10.3390/ani10030438.

Qiao, D., Zhao, Y., Pei, C., Zhao, X., Jiang, X., Zhu, L., Zhang, J., Li, L., & Kong, X. (2022). Genome-wide identification, evolutionary analysis, and antimicrobial activity prediction of CC chemokines in allotetraploid common carp, *Cyprinus carpio*. *Fish & Shellfish Immunology*, *130*, 114-131. https://doi.org/10.1016/j.fsi.2022.09.002.

Rio, D. C., Ares, M., Hannon, G. J., & Nilsen, T. W. (2010). Purification of RNA using TRIzol (TRI reagent). *Cold Spring Harbor Protocols*, *2010*(6). https://doi.org/10.1101/pdb.prot5439.

Schreck, C. B., & Tort, L., (2016). The concept of stress in fish. In C. B. Schreck, L. Tort, A. P. Farrell, & C. J. Brauner (Eds.), *Fish physiology* (Vol. 35, pp. 1-34). Academic Press, London.

Shved, N., Berishvili, G., Häusermann, E., D'Cotta, H., Baroiller, J. F., & Eppler, E. (2009). Challenge with 17á-ethinylestradiol (EE2) during early development persistently impairs growth, differentiation, and local expression of IGF-I and IGF-II in immune organs of tilapia. *Fish & Shellfish Immunology*, *26*(3), 524-530. https://doi.org/10.1016/j.fsi.2009.02.003.

Stanke, M., & Morgenstern, B. (2005). AUGUSTUS: a web server for gene prediction in eukaryotes that allows user-defined constraints. *Nucleic Acids Research*, *33*(2), W465-W467. https://doi.org/10.1093/nar/gki458.

Sumana, S. L., Wang, P., Zhang, C., Jing, X., Zhu, J., Tang, Y., Liu, W., Su, S., & Liao, Y. (2024). Genetic diversity of the common carp black strain population based on mtDNA (D-loop and cytb). *Heliyon*, *10*(10), Article e30307. https://doi.org/10.1016/j.heliyon.2024.e30307.

Tong, C., Zhang, C., Zhang, R., & Zhao, K. (2015). Transcriptome profiling analysis of naked carp (*Gymnocypris przewalskii*) provides insights into the immune-related genes in highland fish. *Fish & Shellfish Immunology*, *46*(2), 366-377. https://doi.org/10.1016/j.fsi.2015.06.025.

Tóth, B., Khosravi, R., Ashrafzadeh, M. R., Bagi, Z., Fehér, M., Bársony, P., Kovács, G., & Kusza, S. (2020). Genetic diversity and structure of common carp (*Cyprinus carpio* L.) in the Centre of Carpathian Basin: Implications for conservation. *Genes*, *11*(11), Article 1268. https://doi.org/10.3390/genes11111268.

Wang, D., Liao, X., Cheng, L., Yu, X., & Tong, J. (2007). Development of novel EST-SSR markers in common carp by data mining from public EST sequences. *Aquaculture*, *271*(1-4), 558-574. https://doi.org/10.1016/j.aquaculture.2007.06.001.

Wang, M., Li, X., Wang, C., Zou, M., Yang, J., Li, X. D., & Guo, B. (2024). Asymmetric and parallel subgenome selection co-shape common carp domestication. *BMC Biology*, *22*(1), Article 4. https://doi.org/10.1186/s12915-023-01806-9.

Wang, J., Zhou, Q., & Jiang, Y. (2024). Genome-wide analysis of common carp (*Cyprinus carpio*) mucin genes and their roles in mucosal immune response following the *Aeromonas hydrophila* infection. *Comparative Immunology Reports*, *7*, Article 200167. https://doi.org/10.1016/j.cirep.2024.200167.

Xu, P., Wang, J., Wang, J. Cui, R., Li, Y., Zhao, Z., Ji, P., Zhang, Y., Li, J., & Sun, X. (2011). Generation of the first BAC-based physical map of the common carp genome. *BMC Genomics*, *12*, Article 537. https://doi.org/10.1186/1471-2164-12-537.

Xu, P., Zhang, X., Wang, X., Li, J., Liu, G., Kuang, Y., Xu, J., Zheng, X., Ren, L., Wang, G., … & Sun, X. (2014). Genome sequence and genetic diversity of the common carp, *Cyprinus carpio*. *Nature Genetics*, *46*(11), 1212-1219. https://doi.org/10.1038/ng.3098.

Zhang, X., Luan, P., Cao, D., & Hu, G. (2021). A high-density genetic linkage map and fine mapping of QTL for feed conversion efficiency in common carp (*Cyprinus carpio*). *Frontiers in Genetics*, *12*, Article 778487. https://doi.org/10.3389/fgene.2021.778487.

Zhao, Y., Zhu, X., Li, Z., Xu, W., Dong, J., Wei, H., Li, Y., & Li, X. (2019). Genetic diversity and structure of Chinese grass shrimp, *Palaemonetes sinensis*, inferred from transcriptome-derived microsatellite markers. *BMC Genetics*, *20*, Article 75. https://doi.org/10.1186/s12863-019-0779-z.

Zhu, R., Liu, X. X., Lv, X., Li, S. Y., Li, Y. D., Yu, X. J., & Wang, X. G. (2017). Deciphering transcriptome profile of the yellow catfish (*Pelteobagrus fulvidraco*) in response to *Edwardsiella ictaluri*. *Fish & Shellfish Immunology*, *70*, 593-608. https://doi.org/10.1016/j.fsi.2017.08.040.