BMC Genomics

RESEARCH

Open Access

Comparative genomic analysis of immune-related genes and chemosensory receptors provides insights into the evolution and adaptation of four major domesticated Asian carps

Ling Shi[1,2,3], Judy Kin-Wing Ng[1,2], Qing Xiong[1,2,4], Kelvin Fu-Kiu Ao[1,2], Soo-Kyung Shin[1,2], Cherie Tsz-Yiu Law[1,2], Weixue Mu[1,2], Guang-Ming Liu[5], Shitao Rao[6] and Stephen Kwok-Wing Tsui[1,2,7*]

Abstract

Background *Ctenopharyngodon idella* (grass carp), *Mylopharyngodon piceus* (black carp), *Hypophthalmichthys nobilis* (bighead carp), and *Hypophthalmichthys molitrix* (silver carp), collectively known as the four major domesticated Asian carp, are freshwater fish species from the family Cyprinidae and are widely consumed in China. Current studies on these species primarily focus on immune system regulation and the growth and development of individual species. However, in-depth genomic investigations and comprehensive comparative analysis remained limited.

Methods The complete genomes of *Ctenopharyngodon idella*, *Mylopharyngodon piceus* and *Hypophthalmichthys nobilis* were assembled using a hybrid approach that integrated both next- and third-generation sequencing reads, followed by annotation using the MAKER2 pipeline. Based on the high-quality genomes of *Ctenopharyngodon idella*, *Mylopharyngodon piceus Hypophthalmichthys nobilis*, and *Hypophthalmichthys molitrix*, a comparative genomic analysis was conducted using bioinformatic tools to investigate gene family evolution in these four domesticated Asian carp species.

Results High-quality genomes of *Ctenopharyngodon idella, Mylopharyngodon piceus*, and *Hypophthalmichthys nobilis* were assembled, achieving over 90% completeness. Immune-related gene families, including MHC class I and NLRC3-like genes, have undergone rapid evolution, with *Ctenopharyngodon idella* exhibiting significant expansion of NLRC3-like genes. Massive tandem duplication events were identified in trace amine-associated receptors (TAARs), and rapid expansion was observed in TAAR16 and TAAR29. Additionally, a novel TAAR gene cluster was identified in all four Asian carp species. Comparative genomic analysis revealed the expansion of type 1 taste receptor genes, particularly in *Ctenopharyngodon idella* and *Mylopharyngodon piceus*.

Conclusion This study has successfully constructed the high-quality genomes of *Ctenopharyngodon idella*, *Mylopharyngodon piceus*, and *Hypophthalmichthys nobilis*. The comparative genomic analysis revealed the evolution of immune-related genes and chemosensory receptors in the four major domesticated Asian carp species. These

*Correspondence:
Stephen Kwok-Wing Tsui
kwtsui@cuhk.edu.hk
Full list of author information is available at the end of the article

findings suggested the enhanced immunity and sensory perception in these species, providing valuable insights into their adaptation, survival and reproduction.

Keywords Four major domesticated Asian carp species, Comparative genomic analysis, Gene family evolution, Chemosensory receptors, Immune-related genes

Introduction

Ctenopharyngodon idella (grass carp), *Mylopharyngodon piceus* (black carp), *Hypophthalmichthys nobilis* (bighead carp) and *Hypophthalmichthys molitrix* (silver carp) are the four major domesticated Asian carp species that have been cultured for centuries. *C. idella* and *M. piceus* belong to the subfamily Squaliobarbinae, while *H. nobilis* and *H. molitrix* are classified under the genus, *Hypophthalmichthys*, within the subfamily Xenocyprinae. In China, these domesticated carp species play critical roles in freshwater fishery, serving as a significant source of high-quality proteins in traditional diets. However, when introduced to North America, these species have become invasive, posing a threat to the Great Lakes due to their rapid proliferation, which disrupts the local ecosystem and economy [1]. Previous studies have established the complete genomes of *C. idella*, providing insights into its genome evolution and vegetarian adaptation [2, 3]. Jian et al. [4] conducted a systematic genomic analysis of *H. nobilis* and *H. molitrix*, shedding light on their evolution and speciation. These studies suggested the construction of high-quality genomes as an efficient approach to gaining an understanding of the biology of Asian carp. Recent research on the four major domesticated Asian carp species studied immune system regulation, antiviral signaling, mitochondrial analysis, hypoxia adaptation, and the growth and development of individual species [5–9]. However, comprehensive genomic investigations and comparative analysis of all four Asian carp species remain limited. Wang et al. [10] studied the species differentiation of Asian carp by comparative genomic analysis, identifying positive selection in three genes (HHEX, PGK1, and WNT1) associated with body growth and energy regulation. Meanwhile, the PIM kinase family, involved in cell proliferation, survival, and body growth, has been reported to be under expansion in *C. idella* and *M. piceus*. These findings highlighted the importance of comparative genomic analysis in investigating the evolutionary dynamics and adaptations of the four major domesticated Asian carp species.

In this study, we have assembled and annotated the complete genomes of *C. idella*, *M. piceus,* and *H. nobilis* using a hybrid approach that integrated both short-read and long-read sequencing technologies. Utilizing these high-quality genomes, we performed a comparative genomic analysis to explore the evolution and adaptation of gene families in the four Asian carp species (*C. idella*, *M. piceus*, *H.* nobilis, and *H. molitrix*). The comparison of gene families in Cyprinid fish species revealed the rapid evolution of immune-related gene families, including MHC class I and NLRC3-like genes, driven by tandem duplication events, particularly in *C. idella*. Additionally, chemosensory receptor families, including trace amine-associated receptors, vomeronasal receptor type 1 genes, and type 1 taste receptors were undergoing expansion. This comparative analysis provided insights into the gene family evolution and adaptation of the four major domesticated Asian carp species.

Methods

Genomic DNA and RNA extraction

Three fish species, *C. idella*, *M. piceus*, and *H. nobilis*, were purchased from the Tai Po wet market in Hong Kong. These fish species were imported from a fish farm in Shunde, Guangdong Province, China. The genomic DNA of *C. idella*, *M. piceus* and *H. nobilis* were extracted from muscle tissue using the blood and cell culture DNA maxi kit (QIAGEN, Germany). Total RNA was isolated from 10 organ tissues (brain, muscle, liver, ovary/gonad, gall bladder, intestine, heart, skin, spleen, and kidney) of each carp species using TRIzol reagent (Invitrogen, USA), followed by phenol–chloroform extraction. The aqueous phase was then purified using PureLink RNA mini kit (QIAGEN, Germany). The concentrations of genomic DNA and RNA were measured using a Nanodrop 2000 spectrophotometer (Thermo Fisher Scientific, USA) and a Qubit Fluorometer (Thermo Fisher Scientific, USA). The integrities of RNA were evaluated by an Agilent 2100 Bioanalyzer (Agilent Technologies, USA).

Whole genome sequencing and de novo genome assembly

The genomic DNA of *C. idella*, *M. piceus*, and *H. nobilis* proceeded for long-read library construction using the genomic DNA ligation sequencing kit SQK-LSK110 (Oxford Nanopore Technologies, UK) according to manufacturer's instruction, and sequenced by GridION Mk1 platform (Oxford Nanopore Technologies, UK). For data processing, the raw Nanopore reads were trimmed and filtered by Porechop (v0.2.4) [11] and NanoFilt (v2.8.0) [12]. For short-read sequencing, the genomic DNA was sequenced on the BGI DNBseq platform (Groken Bioscience Limited, Hong Kong). The draft genomes of *C.*

idella, *M. piceus*, and *H. nobilis* were assembled by Flye (v2.8.3) [13] using clean Nanopore reads. Scaffolding was accomplished by SSPACE (v2.0) [14]. The short reads were aligned to the assemblies using BWA (v0.7.16a) [15] to produce bam files. Pilon (v1.23) [16] was then applied to polish the Nanopore assemblies with aligned short reads. The overall quality and completeness of genome assemblies were assessed by QUAST (v5.0.2) [17] and BUSCO (v3.0) [18].

Genome annotation

Total RNA sequencing of tissue samples from *C. idella*, *M. piceus* and *H. nobilis* were performed using the Illumina HiSeq 2500 platform (Groken Bioscience Limited, Hong Kong). Transcriptome mapping and the file format conversions were carried out using Hisat2 (v2.2.1) [19], StringTie (v2.2.0) [20], and SAMtools (v1.12) [21]. Repeat elements within the genomes of *C. idella*, *M. piceus*, and *H. nobilis* were identified and masked by RepeatModeler (v2.0.1) [22] and RepeatMasker (v4.1.1) [23]. A total of five rounds of genome annotation were conducted using the MAKER2 pipeline (v3.01.03), integrating ab initio gene prediction, protein homology prediction, and transcript evidence prediction. The initial round of annotation was performed using Exonerate (v2.2.0) based on the evidence from transcriptome and the proteome of a closely related reference species, common carp (*Cyprinus carpio*), available in the NCBI database (NCBI accession: GCF_018340385.1). The first round of annotation trained the ab initio gene prediction model of SNAP (v1.0) [24], which was then used for the second round of annotation. A third round of annotation further refined SNAP using results from the previous round. In the final two rounds of annotation, the gene prediction models of Augustus (v3.4.0) [25] and GeneMark-ES (v4.69) [26] were incorporated along with SNAP. Augustus was trained using the annotation results from the previous round, while GeneMark-ES employed a self-training approach. The completeness of annotation was assessed by BUSCO (v3.0).

Collection of genome, transcriptome, and protein sequences

The genome, transcriptome and protein sequences of *C. idella*, *M. piceus*, and *H. nobilis* were deposited in the NCBI database (BioProject accession: PRJNA890423, PRJNA891927, and PRJNA892279). The genome, transcriptome, and protein sequences of *H. molitrix* (BioProject accession: PRJNA631443) were based on the study by Zhou et al. [27] and downloaded from Figshare (https://doi.org/https://doi.org/10.6084/m9.figshare.12618884.v1). The sequences of *Cyprinus carpio* (NCBI accession: GCF_018340385.1), *Carassius auratus* (NCBI accession: GCF_003368295.1), *Danio rerio* (NCBI accession: GCF_000002035.6), *Labeo rohita* (NCBI accession: GCF_022985175.1), *Onychostoma macrolepis* (NCBI accession: GCA_012432095.1), *Puntigrus tetrazona* (NCBI accession: GCF_018831695.1), *Sinocyclocheilus grahami* (NCBI accession: GCF_001515645.1) and the outgroup species *Chanos chanos* (NCBI accession: GCF_902362185.1) were retrieved from the NCBI database.

Divergence time analysis of cyprinid fishes

The coding sequences of 301 overlapping single-copy BUSCO genes from 11 Cyprinidae species and the outgroup species *C. chanos* were collected based on the vertebrate_odb9 database. The coding sequences of the overlapping BUSCO genes were aligned using MAFFT (v7.310) [28] in fasta format. The sequence alignment of codons was then split using the online tool Sequence Manipulation Suite (v2.0) (https://www.bioinformatics.org/sms2) and subsequently converted into phylip format using the online tool Format Converter from HIV databases (https://www.hiv.lanl.gov/content/sequence/FORMAT_CONVERSION/form.html). The phylogenetic relationships of the 12 fish species were constructed from the multiple sequence alignments generated by MAFFT, using maximum likelihood phylogeny inference and 100 bootstrap replicates in RAxML (v8.2.12) [29], with the option "-m PROTGAMMALGX -f -# 100". Based on the phylogenetic tree established by RAxML (v8.2.12), the divergence time analysis was performed using MCMCtree in PAML (v4.9j) [30]. Two calibration points were employed: the divergence times between *C. auratus* and *C. carpio*, and between *C. idella* and *M. piceus*. The 95% confidence intervals of the divergence time used for calibration were retrieved from TimeTree database (http://www.timetree.org/). The ultrametric tree was edited using the online software Interactive Tree of Life (iTOL v6) [31].

Phylogenetic orthology inference

The protein sequences of 11 Cyprinidae species and the outgroup species *C. chanos* were filtered by removing duplicated copies and selecting the longest isoforms. The inference of orthologous groups were analysed by OrthoFinder (v2.5.4) [32] based on the filtered protein sequences of the 12 fish species.

Comparison of gene families

The comparative analysis of gene families was performed based on the results of OrthoFinder and the divergence time tree established by MCMCtree. Large orthogroups were filtered by Python scripts (cafetutorial_clade_and_size_filter.py) available on the GitHub repository (https://

github.com/hahnlab/cafe_tutorial/tree/main/python_scripts). The expansion and contraction of gene families within 12 fish species were analyzed using CAFE (v4.0) [33]. The significance of changes in gene family size was examined using computed *p*-values from CAFE based on a birth and death model, with $p < 0.05$ as the threshold for identifying rapidly evolving gene families. The sequences of rapidly evolving gene families in the four Asian carp species were subjected to BLASTP search with an e-value cut-off of 1e-6 to retrieve UniProt accession identifiers. Gene enrichment analysis was performed based on Gene ontology (GO) terms using the online tool Database for Annotation, Visualization and Integrated Discovery (DAVID) (https://david.ncifcrf.gov/).

In silico identification of genes from different families

The reference protein sequences of taste receptors in *C. carpio, D. rerio,* and *C. idella*, trace-amine associated receptors (TAARs) in *D. rerio* and *C. auratus*, and vomeronasal receptor type 1 (V1R/ORA) in *D. rerio* were obtained from previous studies [34–36]. The reference protein sequences of MHC class I genes and NLRC3/NLRC3-like genes were retrieved from the NCBI and UniProt databases. All the genes of the chemosensory receptor family and MHC class I family in the four annotated Asian carp genomes were identified by BLASTP (v2.11.0) [37]. The analysis was conducted against the corresponding reference sequences using the option "-evalue 1e-6 -outfmt 6", with the sequence identity no less than 30% as the threshold. HMMER (v 3.3.2) [38] search was further employed to validate all the NLRC3/NLRC3-like genes identified through BLASTP. The profile hidden Markov model (HMM) was established using the hmmbuild command, based on the protein sequence alignment of the reference protein sequences. The homologs from different gene families were detected using the hmmsearch command against the profile HMM model with the default parameters. Based on the gene synteny information from the genome annotation gff files, the tandemly or proximally arrayed genes (T/PAGs) were defined as zero spacer genes, and fewer than ten spacer genes, respectively.

Phylogenetic analysis of gene families

The multiple sequence alignment of the rapidly evolving gene families was performed using MUSCLE (v5.1) [39]. The phylogenetic trees of the rapidly evolving gene families were constructed using FastTree (v2.1.11) [40] and validated with MEGA (v10.1.8) [41] using the maximum likelihood algorithm with 100 bootstrap replications. The phylogenetic trees were edited using the online software Interactive Tree of Life (iTOL v6) [31].

Positive selection analysis

The sequences of all the single-copy orthogroups from the 12 species were collected and aligned using MAFFT (v7.0), with gene trees derived from the previous OrthoFinder analysis. Positive selection analysis was conducted using the branch-site model in PAML (v4.9j), where the alternative model was specified as model = 2 and NSites = 2. To identify the positively selected genes, a likelihood ratio test (LRT) was performed to compare the alternative model (fix_omega = 0) to the null model (fix_omega = 1) using codeml program in PAML (v4.9j). The *p*-value from the LRT was examined using chi-square statistics, with the significant threshold set at $p < 0.05$. All the positively selected genes were annotated using the BLASTP search program against the NCBI database.

Motif enrichment analysis

For TAARs, the protein sequences of cluster A TAAR genes were applied for motif enrichment analysis with the use of other subfamily TAAR gene sequences as controls. Regarding MHC class I genes, the protein sequences of the atypical Z2 lineage were collected for analysis using Z1 lineage sequences as controls. Motif enrichment analysis was performed using the simple enrichment analysis (SEA) program from the online software MEME Suite (v5.5.5) [42], with the PROSITE fixed-length motif database and an e-value threshold of < 0.01. For TAS1R2 genes, multiple sequence alignment was conducted for herbivorous and carnivorous fish species using Clustal Omega [43]. The domain structures of TAS1R2 genes were identified by InterProScan [44], available on the EMBL-EBI website. The motif analysis of TAS1R2 genes in the six examined fish species was conducted using the STREME program with a *p*-value threshold of < 0.05.

Results

Genome assembly and annotation

The high-quality genomes of three Asian carp species (*C. idella*, *M. piceus* and *H. nobilis*) were constructed by a hybrid assembly approach that combined long-read and short-read sequencing technologies (Table S1). The genome of *H. molitrix* was obtained from the NCBI database (BioProject accession: PRJNA631443). The assembled genome sizes for *C. idella, M. piceus and H. nobilis* were 880.8 Mb, 877.2 Mb and 861.0 Mb, respectively (Table 1). The genome of *H. molitrix* was assembled at the chromosome level, achieving a size of 856.6 Mb (Table 1). The completeness of the four Asian carp genomes ranged from 92.9% (*C. idella*) to 95.8% (*M. piceus* and *H. nobilis*), and their contig N50 ranged from 3,164 Kb (*M. piceus*) to 6,603 Kb (*H. molitrix*) (Table 1).

Table 1 Overview statistics of genome assemblies and annotations of four Asian carp species

	C. idella	*M. piceus*	*H. nobilis*	*H. molitrix*
Assembly features				
Genome size (Mb)	880.8	877.2	861.0	856.6
Scaffold number	1,949	3,348	2,001	25
Contig number	3056	3,398	2,051	214
Contig N50 (Kb)	4,807	3,164	5,617	6,603
BUSCO completeness (%)	92.9	95.8	95.8	95.0
BUSCO duplication (%)	3.5	3.8	3.6	3.8
Annotation features				
Number of genes	30,366	31,656	34,968	29,279
BUSCO completeness (%)	87.9	90.5	88.9	95.0
BUSCO duplication (%)	5.6	5.9	5.9	8.9
Repeat content (%)	50	46	48	46

Based on the assembled genomes, transcriptomes and the protein sequences of the closely related species *C. carpio* (common carp), genome annotation of *C. idella, M. piceus and H. nobilis* was performed using the MAKER2 pipeline. The genome annotation dataset of *H. molitrix* was downloaded from Figshare. The BUSCO completeness of the four Asian carp proteomes ranged from 87.9% to 95.0% (Table 1). The predicted number of protein-coding genes was comparable across the four species, with 30,366, 31,656, 34,968, and 29,729 genes identified in *C. idella*, *M. piceus*, *H. nobilis* and *H. molitrix*, respectively (Table 1). The repeat content annotation revealed a similar percentage of total repeat sequences in the four Asian carp genomes, ranging from 48.37% (*H. molitrix*) to 50.63% (*M. piceus*) (Table S2). In addition, the features of the de novo assembled Asian carp genomes were compared with previously published genomes (Table S3).

Phylogenomic analysis

Phylogenomic analysis was conducted to infer the evolutionary relationships among Asian carp and other Cyprinidae species. A total of 12 fish genomes were included in the analysis, comprising 11 species from Cyprinidae and one outgroup species, *C. chanos*, from Chanidae (Table S4). *M. piceus* and *C. idella* shared 18,648 overlapping orthogroups, while *H. nobilis* shared 21,836 orthogroups with *H. molitrix* (Fig. S1 A). All four Asian carp species exhibited over 80% overlap in orthogroups with each other (Fig. S1B). Based on 3,612 conserved amino acid sequences from 301 single-copy and complete BUSCO orthologs, the divergence time of the 12 species was estimated. The results demonstrated that *M. piceus* and *C. idella* diverged approximately 7.51 million years ago (Fig. S2). The divergence time between *H. nobilis* and *H. molitrix* was estimated to be 7.18 million years (Fig. S2). The Asian carp family was found to have diverged from *C. carpio* (common carp) and *D. rerio* (zebrafish) around 39.94 and 50.87 million years ago, respectively (Fig. S2).

Gene family evolution

To investigate the evolution of the four Asian carp species, a comparative analysis of gene families was performed using the annotated protein sequences of the 12 species with CAFE. In *C. idella*, 1,066 gene families were found to be expanded, while 3,635 gene families were contracted (Fig. 1A). A total of 1,521 gene families were expanded, and 2,891 gene families contracted in *M. piceus* (Fig. 1A). For *H. nobilis*, gene family comparative analysis identified 1,933 expanded gene families and 1,729 contracted gene families (Fig. 1A). In *H. molitrix*, 1,294 gene families were expanded, while 6,410 gene families were contracted (Fig. 1A). Meanwhile, 251 and 253 gene families were undergoing rapid evolution in *C. idella* and *M. piceus*, respectively (Fig. 1A). For *H. nobilis* and *H. molitrix*, the number of rapidly evolving gene families was 290 and 374, respectively (Fig. 1A).

The rapidly evolving gene families were further annotated through Gene Ontology (GO) enrichment analysis. In *C. idella*, genes related to the detection of chemical stimuli involved in smell perception were identified as rapidly evolving (Fig. S3A). For *M. piceus*, genes associated with synaptic transmission were undergoing rapid evolution (Fig. S3B). Gene families related to muscle contraction were identified as rapidly evolving in *H. nobilis* (Fig. S3C). In *H. molitrix*, genes families that positively regulate kinase activity and are involved in axon guidance were found to be rapidly evolving (Fig. S3D). Furthermore, the results indicated that genes associated with calcium ion transport were rapidly expanding across all four Asian carp species (Fig. S3E).

The unique gene families in each species were annotated using GO enrichment analysis (Fig. S4), and the genes undergoing positive selection were identified through PAML analysis (Table S5). The selection pressure analysis revealed that three genes—CAPN10, SVIL, and ENKD1—associated with calcium ion signaling and muscle function, were under positive selection in all four Asian carp species (Table S5A). The Venn diagram of species-specific orthogroups revealed that *H. molitrix* had the highest number of species-specific orthogroups (N= 921), while *M. piceus* exhibited the lowest number (N= 326) (Fig. 1B).

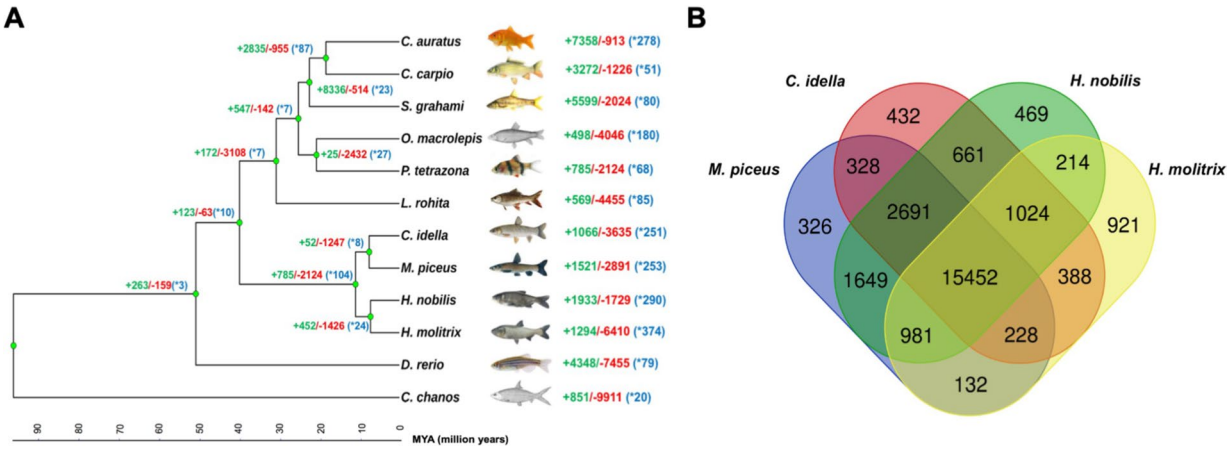

Fig. 1 The phylogenomic analysis of Asian carp and other Cyprinidae species. **A** The gene family expansion and contraction of 11 Cyprinidae speices: *Carassius auratus*, *Cyprio carpio*, *Sinocyclocheilus grahami*, *Onychostoma macrolepis*, *Puntigrus tetrazona*, *Labeo rohita*, *Ctenopharyngodon idella*, *Mylopharyngodon piceus*, *Hypophthalmichthys nobilis* and *Hypophthalmichthys molitrix*, *Danio rerio* and outgroup species *Chanos chanos* were analyzed by CAFE. The numbers indicate the number of orthogroups under expansion (+ and green) or contraction (- and red) and rapidly evolving orthogroups (* and blue). **B** Venn diagram of orthogroups of the group Asian carp species. The proteomes were assigned into orthogroups using OrthoFinder, and the overlapped orthogroups of the four Asian carp species were then presented using a Venn diagram

Immune-related gene families

Asian carp exhibit a complex and intricate immune system that enables them to defend against pathogens and maintain their overall health. The immunity of Asian carp involves both innate and adaptive immune responses. In this study, we investigated two specific immune-related gene families in the four major domesticated Asian carp species to understand their evolution.

Major Histocompatibility Complex (MHC) class I genes

The classical MHC class I molecules are critical for initiating immune responses against intracellular pathogens such as viruses. In teleost fish, five MHC class I lineages have been identified: U, Z, S, L, and P. Each lineage encodes transmembrane heavy chain molecules with a canonical organization of three extracellular domains [45]. In this study, we identified all MHC class I genes in the four Asian carp species using reported MHC class I genes from *D. rerio* as reference sequences via HMMER search. The results suggested that MHC class I genes in Asian carp could be classified into U, Z and L lineages, with the Z lineage further subdivided into a typical Z1 lineage and an atypical Z2 lineage (Fig. 2A). The number of MHC class I genes in *M. piceus*, *H. nobilis*, and *H. molitrix* was 58, 69 and 50, respectively, whereas *C. idella* possessed a significantly higher number of MHC class I genes (N= 97) (Fig. S5A). The phylogenetic analysis of U lineage MHC class I genes revealed massive tandem duplication events. Among the four species, *C. idella* had the highest number of U lineage genes (N= 24), with 18 of them as T/PAGs (Table 2, Fig. 2B). In contrast, *M. piceus* had the fewest U lineage genes (N= 10), with 5 of them tandemly or proximally arrayed (Fig. 2B). *H. nobilis* and *H. molitrix* contained 16 and 12 U lineage MHC class I genes, respectively, with 12 and 11 identified as T/PAGs (Table 2, Fig. 2B).

The Z lineage is divided into typical and atypical clusters, with the typical Z lineage representing a more ancestral cluster found not only in teleost, but also in spotted gar, bichir and lungfish [46]. The atypical Z lineage genes are highly differentiated and can only be found in certain teleost species [47, 48]. In this study, we identified the Z1 cluster as the typical Z lineage, while the Z2 cluster as the atypical Z lineage which is suggested to be prevalent in carp [48]. Among the four Asian carp species, *C. idella* had the highest number of Z1 sub-lineage genes (N= 26), with 23 of them tandemly or proximally arrayed (Table 2, Fig. 2C). The total number of Z1 sub-lineage genes in *M. piceus*, *H. nobilis* and *H. molitrix* was 11, 9 and 16, respectively, with the number of T/PAGs ranging from 5 in *H. nobilis* to 14 in *H. molitrix* (Fig. 2C). Phylogenetic analysis of the atypical Z2 sub-lineage genes revealed that *H. nobilis* had the highest number (N = 18), comparable to *C. idella* (N = 17) (Table 2, Fig. 2D). Massive tandem gene duplication events of Z2 sub-lineage genes were observed in *H. nobilis* and *C. idella*, and the number of T/PAGs were 15 and 12 in *H. nobilis* and *C. idella*, respectively (Fig. 2D). In contrast, *M. piceus* and *H. molitrix* possessed relatively fewer Z2 sub-lineage genes, with only 6 and 3 identified, respectively. Notably, all Z2 sub-lineage genes in both species were either tandemly or proximally arrayed (Fig. 2D). In addition, the

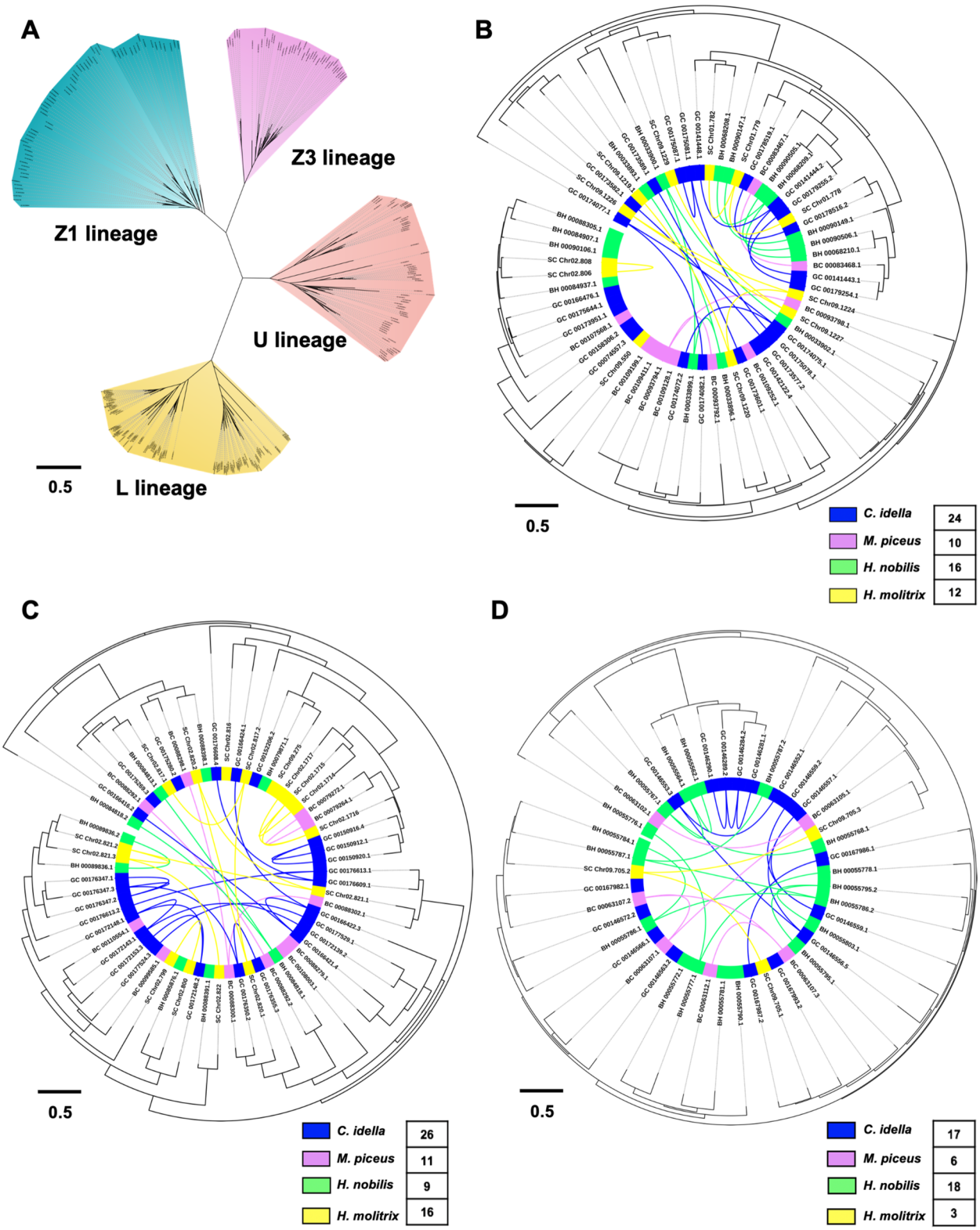

Fig. 2 The phylogenetic relationships of MHC class I gene family. **A** All the MHC class I genes in the four Asian carp species were classified into U, L and Z (Z1 and Z2) lineages based on the phylogenetic relationships. The phylogenetic analysis of MHC class I (**B**) U lineage genes, **C** Z1 lineage genes and (**D**) Z2 lineage genes in the four Asian carp species were examined. The number of MHC class I genes from different lineages in each Asian carp species was denoted, and all the T/PAGs were connected by curved lines

Table 2 The overview of expanding gene families in four Asian carp species

Functional category	Gene family	Gene name/ description	Representative	Length	Best hit in NCBI NR database				Copy number			
					Accession	Identity	E-value	Species	Grass carp	Black carp	Bighead carp	Silver carp
Immune related genes	MHC class I	L lineage	GC_00172322.1	358	XP_051758566.1	100.0%	0.0	Grass carp	30	31	26	19
		U lineage	GC_00173577.2	313	XP_051729032.1	96.81%	0.0	Grass carp	24	10	16	12
		Z1 lineage	GC_00177524.3	338	XP_051745366.1	95.15%	0.0	Grass carp	26	11	9	16
		Z2 lineage	GC_00146290.1	421	XP_051729407.1	98.34%	0.0	Grass carp	17	6	18	3
	Nucleotide-binding leucine-rich-repeat (NLR)	NLRC3/NLRC3-like	GC_00176041.4	840	XP_048036676.1	89.76%	0.0	Wuchang bream (*Megalobrama amblycephala*)	191	58	117	129
Chemosensory receptor	Taste receptor	TAS1R1	GC_00038307.2	821	XP_051736025.1	100.00%	0.0	Grass carp	1	1	1	1
		TAS1R2	GC_00140630.2	824	APG29583.1	99.64%	0.0	Grass carp	6	6	3	3
		TAS1R3	GC_00065666.2	845	XP_051769379.1	100.00%	0.0	Grass carp	1	1	1	1
		TAS2R	GC_00112906.1	345	XP_051760441.1	98.43%	0.0	Grass carp	3	3	3	3
	Olfactory receptor	TAAR1	GC_00021423.2	334	XP_051731363.1	100.00%	0.0	Grass carp	1	1	1	1
		TAAR10	GC_00021422.1	336	XP_051731730.1	99.70%	0.0	Grass carp	4	4	4	4
		TAAR11	GC_00021422.5	311	XP_051731364.1	100.00%	0.0	Grass carp	1	1	1	1
		TAAR12	GC_00021426.2	324	XP_051731356.1	100.00%	0.0	Grass carp	9	9	9	9
		TAAR13	GC_00138512.2	341	XP_051731456.1	99.71%	0.0	Grass carp	4	3	3	4
		TAAR14	GC_00138530.4	326	XP_051732302.1	99.69%	0.0	Grass carp	12	10	11	11
		TAAR15	GC_00138516.2	332	XP_051729896.1	100.00%	0.0	Grass carp	1	1	1	1
		TAAR16	GC_00017218.1	328	XP_051766486.1	99.70%	0.0	Grass carp	17	14	16	14
		TAAR17	GC_00017219.2	321	XP_048037774.1	97.50%	0.0	Wuchang bream (*Megalobrama amblycephala*)	2	2	2	2
		TAAR18	GC_00017188.3	337	XP_048037756.1	96.44%	0.0	Wuchang bream (*Megalobrama amblycephala*)	19	18	16	13
		TAAR29	GC_00017147.3	333	XP_048037739.1	93.09%	0.0	Wuchang bream (*Megalobrama amblycephala*)	15	16	10	15
		TAAR cluster A	GC_00017172.3	338	XP_051764283.1	95.22%	0.0	Grass carp	7	5	8	8
		ORA1	GC_00157407.2	316	XP_051766534.1	100.00%	0.0	Grass carp	1	1	1	1
		ORA2	GC_00157407.1	310	XP_051764039.1	99.68%	0.0	Grass carp	1	1	1	1
		ORA3	GC_00157467.2	329	XP_051766532.1	99.39%	0.0	Grass carp	2	2	2	2
		ORA4	GC_00157464.1	344	XP_051766557.1	99.42%	0.0	Grass carp	1	1	1	1
		ORA5	GC_00123591.1	325	XP_048037788.1	98.46%	0.0	Wuchang bream (*Megalobrama amblycephala*)	2	2	2	2
		ORA6	GC_0128612.1	302	XP_051730292.1	100.0%	0.0	Grass carp	1	1	1	1

motif enrichment analysis suggested that the signatures associated with casein kinase II phosphorylation site and leucine zipper pattern were relatively enriched in Z2 sublineage genes compared to MHC class I genes from other lineages (Table S6). Moreover, the evolution of the L lineage MHC class I genes was investigated through phylogenetic analysis. The results revealed that the four Asian carp species possessed a similar number of MHC class I genes from the L lineage, which were under rapid expansion through tandem duplication (Fig. S5B).

NLRC3/NLRC3-like genes

NLRC3/NLRC3-like genes belong to the nucleotide-binding domain and leucine-rich repeat-containing (NLR) protein family, which plays a crucial role in innate immune responses in teleost fish [49]. In *C. idella*, NLRC3/NLRC3-like proteins have been classified into four categories according to their domain architecture: (i) having PYD, FISNA, NACHT and LRR domains, (ii) possessing a FISNA-NACHT-LRR structure, (iii) containing FISNA, NACHT, LRR and SPRY-PRY (B30.2) domains and (iv) consisting of PYD, FISNA and NACHT domains, with no LRRs [50]. To identify all the NLRC3/NLRC3-like genes in the four Asian carp species, we conducted an HMMER search using the reported NLRC3/NLRC3-like protein sequences from *D. rerio* as a reference. All the genes were classified into three clusters (C1-3) based on their phylogenetic relationships and were further characterized by their domain structures (Fig. S6). Among the four species, *C. idella* had the largest number of NLRC3/NLRC3-like genes (N= 191), followed by *H. molitrix* (N= 129) and *H. nobilis* (N= 117), whereas *M. piceus* had the fewest NLRC3/NLRC3-like genes (N= 58) (Table 2). We further investigated the phylogenetic relationships of the C1 cluster, which includes a large number of NLRC3/NLRC3-like genes with a FISNA-NACHT-LRR-SPRY (B30.2) structure. The phylogenetic analysis identified multiple T/PAGs, with 22 in *C. idella*, 6 in *M. piceus*, 15 in *H. nobilis* and 18 in *H. molitrix*, suggesting their rapid expansion (Fig. 3A). Likewise, the C3 cluster contained a substantial number of NLRC3-like genes with FISNA-NACHT-LRR structure. The number of NLRC3-like genes with a FISNA-NACHT-LRR structure in *C. idella* (N =93) was significantly higher than that in the other three species, with 33 of them generated through tandem duplication (Fig. 3B). In contrast, *M. piceus* harboured 24 NLRC3-like genes with a FISNA-NACHT-LRR structure, of which only 2 were tandemly or proximally arrayed (Fig. 3B). These findings suggested that NLRC3-like genes with a FISNA-NACHT-LRR-SPRY (NLR-B30.2) and a FISNA-NACHT-LRR structure underwent rapid expansion in *C. idella*, *H. nobilis*, and *H. molitrix* through tandem duplication, while only a limited number of duplication events occurred in *M. piceus*.

Chemosensory receptors

Chemosensory receptors are sensory receptors used to detect and respond to chemical stimuli, including odors, flavors, and pheromones [51]. In vertebrates, these receptors are typically classified into six families: olfactory receptor (OR), trace amine-associated receptor (TAAR), vomeronasal receptor type 1 and 2 (V1R and V2R), and taste receptor type 1 and 2 (TAS1R and TAS2R) [52]. The OR, TAAR, V1R, and V2R serve as olfactory or pheromone receptors, while TAS1R and TAS2R are taste receptors [52]. In teleost fish, the chemosensory system is highly developed and plays an essential role in their survival and behavior. In this study, using the high-quality genomes of four Asian carp species, we investigated the evolution of several chemosensory receptor families.

Trace amine-associated receptors (TAARs)

TAARs are a distinct family of G-protein-coupled receptors responsible for detecting both endogenous and exogenous biogenic amines [34]. In general, classical TAARs are mainly classified into three clades, with clade I and II as tetrapod TAARs and clade III as teleost-specific [53]. In fish, it has been revealed that many TAARs were expressed in the olfactory organ and highly diversified, highlighting their pivotal role in the chemosensory system [54]. TAAR genes were initially divided into 28 existing subfamilies (TAAR1-28), with three additional subfamilies (TAAR29,30 and 31) identified in teleost [34, 53]. In this study, we conducted a homology search using both BLASTP and HMMER, based on TAAR protein sequences from two Cyprinidae species, *D. rerio* (zebrafish) and *C. auratus* (goldfish), as references to identify TAAR genes in the four Asian carp species. Based on the phylogenetic relationships, TAARs from the four Asian carp species, along with those of *D. rerio* and *C. auratus,* were classified into different subfamilies (Fig. 4A, Fig. S7). The four Asian carp species shared a similar number of TAAR genes, ranging from 83 in *H. nobilis* and *H. molitrix* to 91 in *C. idella*. Notably, these species exhibited a higher number of TAAR genes than *C. auratus*, but lower than that of *D. rerio* (Fig. S7). Regarding the classification of TAAR subfamilies, TAAR19 and TAAR20 were found exclusively in *D. rerio*, whereas TAAR29 was identified in all four Asian carp species and *C. auratus*, but not in *D. rerio* (Fig. S7). Interestingly, a distinct cluster of TAARs (Cluster A) was identified exclusively in the four Asian carp species (Fig. 4A). The number of cluster A genes varied across the four Asian carp species, ranging from 5 in *M. piceus* to 8 in *H. nobilis* and *H. molitrix*. Phylogenetic analysis

revealed multiple tandem duplication events of cluster A genes, which gave rise to tandemly or proximally arrayed genes, with 6 in *C. idella*, 4 in *M. piceus*, and 7 in both *H. nobilis* and *H. molitrix* (Table 2, Fig. 4B). Protein sequence analysis revealed enriched motifs in cluster A TAAR genes, including those identified in interleukin-1, peroxidases proximal heme-ligand and ubiquitin-activating enzyme E1 (Table S7). These findings suggested that cluster A TAAR genes might represent a novel subfamily of TAARs undergoing rapid expansion in the four Asian carp species. Moreover, Asian carp species harbored a significantly higher number of TAAR16 genes, ranging from 14 to 17, compared to fewer than 10 genes identified in *D. rerio* and *C. auratus* (Table 2, Fig. 4C). Extensive tandem duplication events of TAAR16 genes were observed in the four Asian carp species, with 11 to 15 T/PAGs identified (Fig. 4C). In the TAAR29 subfamily, *C. idella*, *M. piceus*, and *H. molitrix* possessed a similar number of TAAR29 genes, whereas *H. nobilis* contained fewer (Fig. 4D). Likewise, multiple T/PAGs of TAAR29 were identified in the four Asian carp species, ranging from 7 in *H. nobilis* to 15 in *M. piceus* (Fig. 4D). These findings suggested that TAAR16 and TAAR29 genes were rapidly expanding through tandem duplication in the four Asian carp species.

Vomeronasal receptor type 1 (V1R/ORA)

The vomeronasal receptor type 1 (V1R/ORA) refers to a specific class of receptors found in the vomeronasal organ. In teleost species, the ORA gene family consists of six genes (ORA1–ORA6) that encode chemosensory receptors, which are expected to detect pheromones. Unlike the rapidly evolving V1R gene family in mammals, the ORA genes in teleost are remarkably conserved [55]. The conserved nature of V1R/ORA receptors suggests their importance in mediating fundamental chemosensory processes and in regulating species-specific behaviours such as social recognition, mating, and territoriality. In this study, we identified all the ORA genes in the four Asian carp species using BLASTP search, followed by phylogenetic analysis to investigate their evolution. Notably, we identified an additional expansion of the ORA3 gene in the four Asian carp species, as well as in *D. rerio* (Fig. 5A). Furthermore, a tandem duplication event in the ORA5 genes was observed in the four Asian carp species but not in *D. rerio* (Fig. 5A and B). Interestingly, sequence alignment analysis revealed that the sequence identity between ORA5a and ORA5b within each Asian carp species was relatively low (< 75%). In contrast, the sequence identity of ORA5a or ORA5b across different species was notably high (Fig. 5C).

Taste receptor

In vertebrates, taste receptors consist of two G-protein-coupled receptor families, TAS1R and TAS2R, which combine in different ways to generate sweet, umami, and bitter taste reception [56]. The TAS1R family is composed of three genes: TAS1R1, TAS1R2 and TAS1R3, mediating sweet and umami taste perception, while the TAS2R gene is responsible for bitter compound detection [56]. Most vertebrates have three TAS1R genes, while some teleost fishes have been reported to exhibit an expansion of TAS1R2 genes, potentially associated with their food habit formation and adaptation to environmental change [57, 58]. To investigate the evolution of the taste receptor family in the four major domesticated Asian carp species, we identified all the TAS1R genes in their annotated genomes using the BLASTP search, with the reported protein sequences of TAS1R in *C. idella* as a reference. A total of 8 TAS1R genes were identified in *C. idella* and *M. piceu*s (Table 2, Fig. 6A). In contrast, *H. nobilis* and *H. molitrix* were found to possess 5 TAS1R genes each (Table 2, Fig. 6A). Tandem duplication events of TAS1R2 genes were observed in all four Asian carp species. Specifically, *C. idella and M. piceus* each contained 6 TAS1R2 genes, while *H. nobilis* and *H. molitrix* each possessed 3 TAS1R2 genes (Table 2, Fig. 6A). Compared to *D. rerio*, which contains two duplicated TAS1R2 genes, the four Asian carp species exhibited an additional tandem duplicated cluster of TAS1R2 genes, indicating an expansion of the TAS1R2 gene family (Fig. 6B). The TAS2R genes of Asian carp species were also identified and analyzed, revealing that all four Asian carp species harbored 3 TAS2R genes each (Table 2, Fig. S8).

Discussion

In this study, we constructed the high-quality genomes of three major domesticated Asian carp species (*C. idella*, *M. piceus* and *H. nobilis*), originating from cultured populations maintained in fish farms. Notably, previous

(See figure on next page.)

Fig. 3 The phylogenetic relationships of NLRC3/NLRC3-like gene in C1 and C3 cluster. All the NLRC3/NLRC3-like genes in the four Asian carp species were classified into three clusters (C1-C3) based on four categories of domain structures and the phylogenetic relationships. C1 cluster is consisted of genes with NACHT-LRR-SPRY structure, whereas C3 cluster is composed of genes with mainly NACHT-LRR structure. The phylogenetic relationships of (**A**) C1 cluster genes and (**B**) C3 cluster genes were examined, respectively. The number of NLRC3/NLRC3-like genes from C1 and C3 cluster in each Asian carp species was denoted, and all the T/PAGs were connected by curved lines

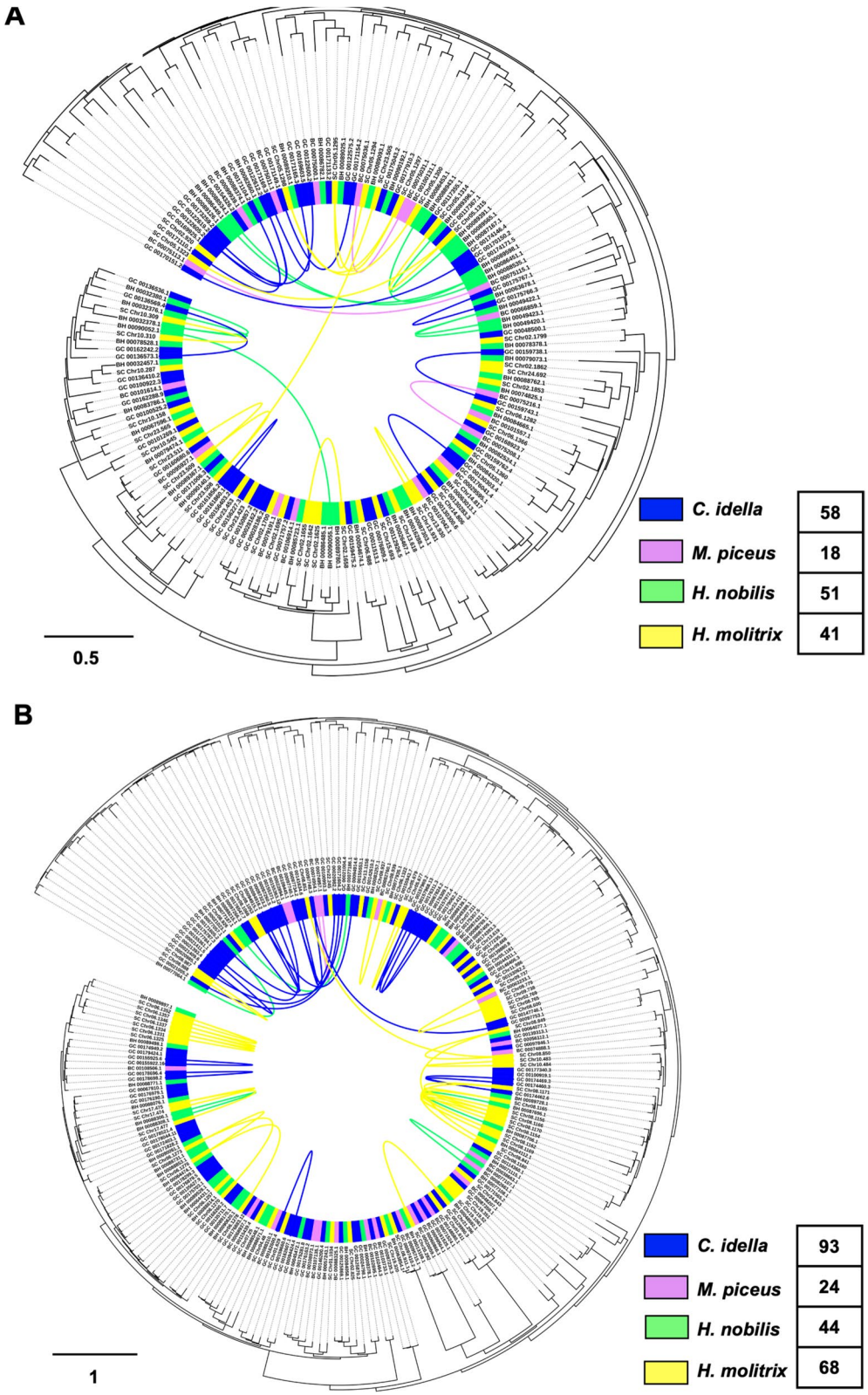

Fig. 3 (See legend on previous page.)

studies have demonstrated a reduced genetic diversity in cultured *C. idella* and *M. piceus* populations compared to wild populations [59, 60], suggesting the degeneration in these major domesticated Asian carp species. A comparative genomic analysis of the four major domesticated Asian carp species *C. idella*, *M. piceus*, *H. nobilis*, and *H. molitrix* was conducted using high-quality genome assemblies. Our findings provided insights into the evolution of immune-related gene families and chemosensory receptors, thereby highlighting their roles in the adaptation of the four Asian carp species.

Expansion of Immune-related Gene Families

Asian carp have been characterized as carriers of various pathogens while exhibiting strong disease resistance [61]. Previous studies have shown that immune-related gene families were under rapid expansion in Asian carp, suggesting the beneficial role of these genes in defending against pathogen invasion [2, 4, 62]. In this study, we reported the expansion of MHC class I and NLRC3/NLRC3-like genes in the four domesticated Asian carp species, suggesting their potentially enhanced immunity to combat various pathogen challenges. Previous studies have demonstrated that the increased variation of immunity in invasive species could help them respond to diverse pathogens rapidly, thereby supporting their survival and successful establishment in new environments [63, 64]. Consequently, the expansion of immune-related genes in the four domesticated Asian carp species might facilitate their adaptation and success in colonizing new environments as invasive species.

Major Histocompatibility Complex (MHC) class I genes

Major histocompatibility complex (MHC) genes are essential components of the adaptive immune system in vertebrates, as they enable the presentation of foreign antigens to immune cells. MHC class I molecules play a critical role in presenting peptide antigens on the cell surface. To date, five lineages of MHC class I genes (U, L, Z, S, and P) have been reported in teleost fish based on phylogenetic clustering [48]. In this study, we identified the presence of U, L, and Z lineage MHC class I genes in the four Asian carp species. They exhibited a significantly higher number of MHC class I genes compared to other species within the Cyprinidae family, such as zebrafish (N= 39) and goldfish (N= 22) [65, 66]. The expansion of U lineage MHC class I genes was observed in the four Asian carp species, especially in *C. idella*. In teleost, U lineage MHC class I genes have been demonstrated to exhibit classical functions of MHC class I molecules, presenting antigens to activate the immune response of $CD8^+$ T cells [65]. The activation of $CD8^+$ T cells will induce cytotoxicity to combat pathogenic infections, including those caused by viruses, parasites, and bacteria [67–69]. For example, Chen et al. [69] revealed that the U lineage MHC class I molecules of *C. idella* could recognize antigenic peptides from grass carp hemorrhagic virus, thereby triggering $CD8^+$ T cell response [70]. Therefore, the expansion of U lineage genes suggested a more diverse repertoire of MHC class I genes in the four domesticated Asian carp species. This diversification might enhance their ability to recognize and present a broader range of antigens, thereby strengthening their adaptive immunity to combat pathogens. In addition, multiple duplication events of Z lineage MHC class I genes were identified within the four Asian carp species. Notably, *C. idella* had the highest number of typical Z1 sub-lineage genes, while *H. nobilis* had the highest number of atypical Z2 sub-lineage genes. The typical Z1 sub-lineage genes are known to be conserved in teleost, whereas carp species have divergent atypical Z2 sub-lineage gene sequences [48, 71]. Previous studies indicated that Z lineage MHC class I genes were involved in fish immune responses, with altered expression levels in response to different immune stimuli and disease states, as well as sub-functionalization of duplicates [48, 72]. However, it remains unclear whether Z lineage genes function as classical MHC class I molecules involved in antigen presentation. Collectively, our findings suggest that the expansion of Z lineage genes might enhance the immune adaptations of the four domesticated Asian carp species to various immune stimuli and disease conditions. The observed expansion of atypical Z2 sub-lineage genes indicated the species-specific diversification of MHC class I gene repertoire in Asian carp, potentially contributing to their distinct immune response.

NLRC3/NLRC3-like genes

In addition to adaptive immunity, the innate immune system plays a critical role in recognizing different pathogens, primarily through interactions between pathogen-associated molecular patterns (PAMPs) and pattern recognition receptors (PRRs) [73]. NOD-like receptors (NLRs) represent a key family of intracellular PRRs that respond to pathogen infections. The NLRC3/NLRC3-like gene family has been demonstrated as a distinct subfamily of NLRs unique to teleost fish, characterized by extensive repertoires and rapid expansion in various species [74]. In this study, we identified all the NLRC3/NLRC3-like genes in the four Asian carp species and classified them into different clusters based on their domain structures. The SPRY/PRY (B30.2) domain, characterized by a β-barrel like structure, is involved in pathogen recognition and regulation of cytokine signaling in innate immunity [75]. We discovered that NLR-B30.2 proteins underwent tandem duplication in the four Asian

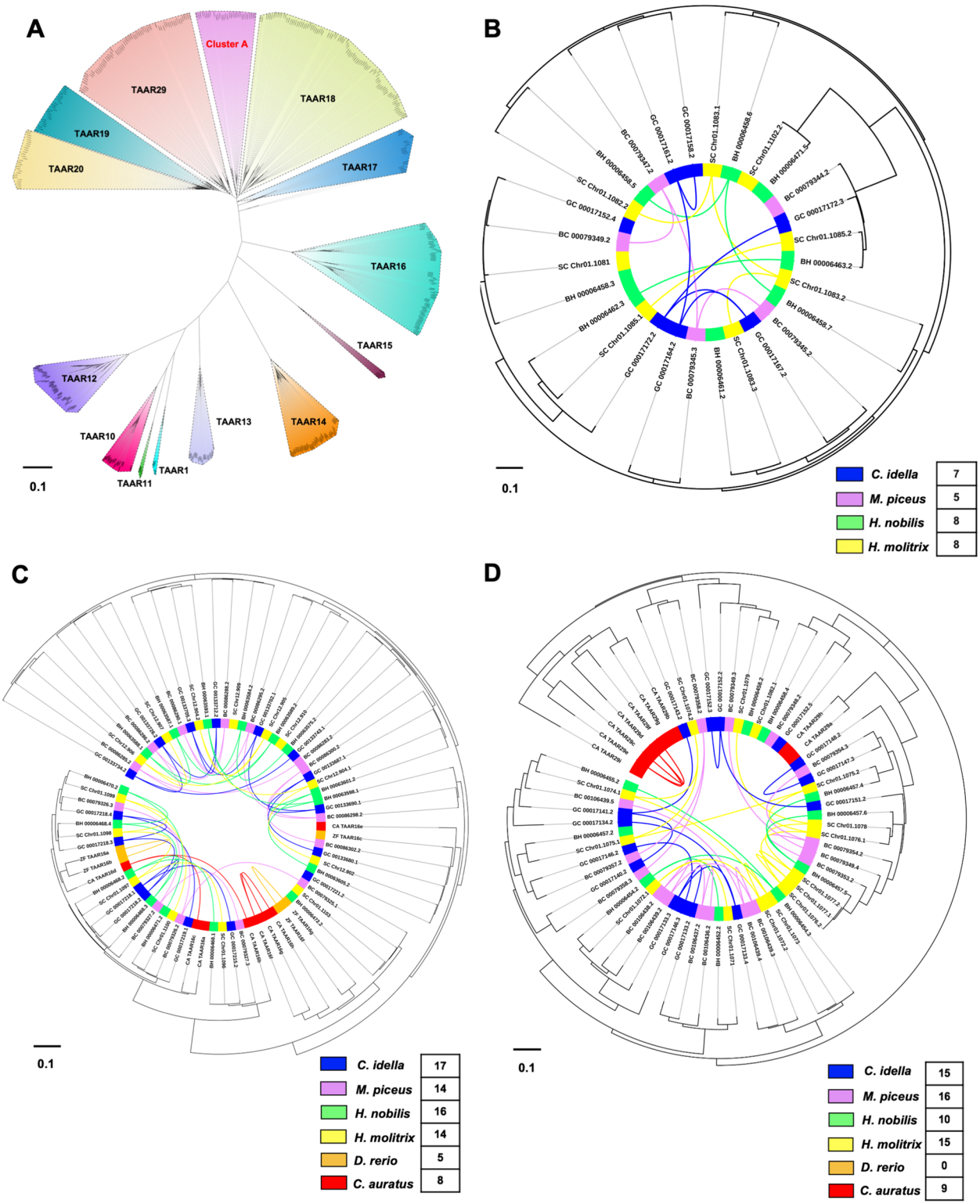

Fig. 4 The phylogenetic relationships of trace amine-associated receptors (TAARs). **A** All the TAAR genes in the four Asian carp species (*C. idella, M. piceus, H. nobilis and H. molitrix*) together with *D. rerio* and *C. auratus* were classified into 14 different TAAR subfamilies based on phylogenetic relationships. The phylogenetic relationships of (**B**) cluster A TAAR genes, **C** TAAR16 subfamily, and **D** TAAR29 subfamily were analyzed. Cluster A TAARs were exclusively observed in the four Asian carp species, while TAAR16 and TAAR29 genes can also be identified in *D. rerio and C. auratus*. The number of TAARs in different species was denoted, and all the T/PAGs were connected by curved lines

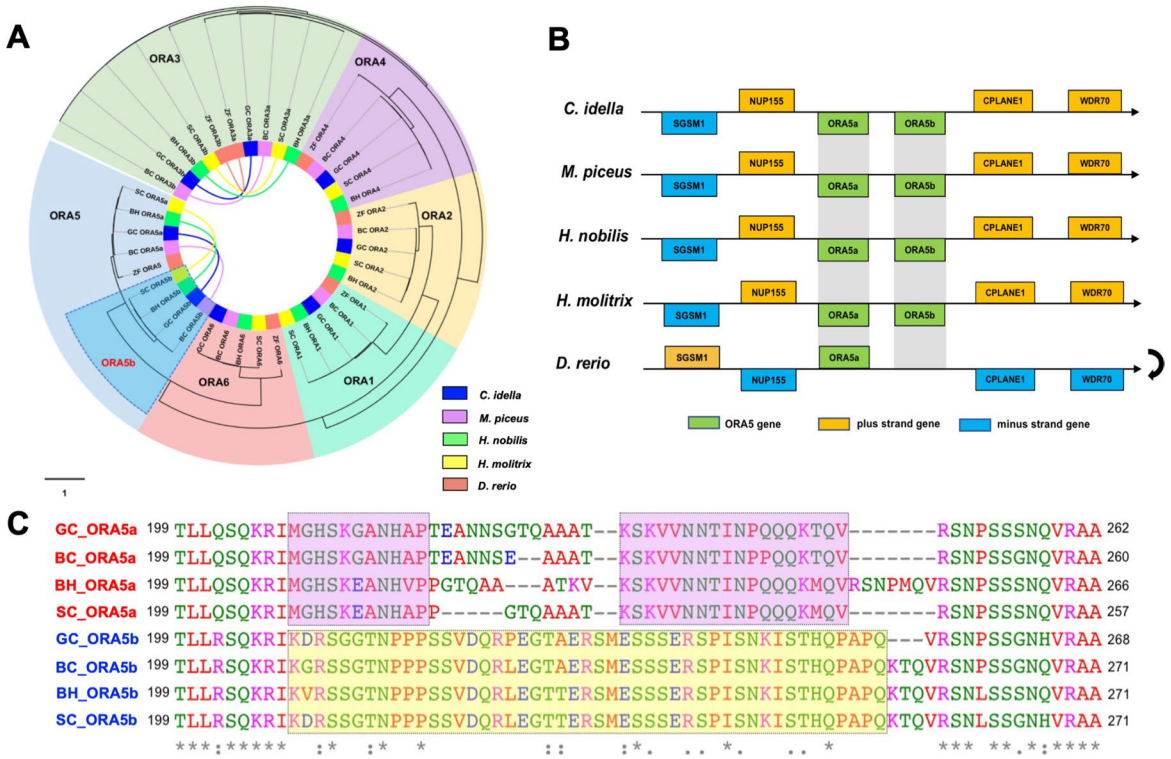

Fig. 5 The expansion of fish vomeronasal type-1 receptors (V1R/ORA). **A** The phylogenetic relationships of V1R/ORA family in the four Asian carp species and *D. rerio* were analyzed and classified into six clusters (ORA1-ORA6). The T/PAGs were connected by curved lines. **B** The gene synteny alignment of ORA5 genes in the four Asian carp species and *D. rerio* were compared, with an additional ORA5 gene identified in Asian carp. The black arrow indicated reverse complement. **C** The sequence alignment of ORA5 genes in the four Asian carp species, with "GC" as *C. idella* (grass carp), "BC" as *M. piceus* (black carp), BH as *H. nobilis* (bighead carp), and SC as *H. molitrix* (silver carp)

carp species. Notably, *C. idella*, *H. nobilis*, and *H. molitrix* exhibited a similar number of duplicated NLR-B30.2 genes, whereas *M. piceus* contained fewer copies of these genes. NLR-B30.2 proteins have been demonstrated to exhibit diversity in their specific pathogen recognition modules among different fish species [27]. Furthermore, NLR-30.2 genes have been reported to exhibit greater sensitivity to viral infections compared to bacterial pathogens, suggesting their important roles in antiviral defense [76]. Therefore, the expansion of NLR-B30.2 genes in the four Asian carp species might facilitate their species-specific antiviral innate immunity. Meanwhile, NLRC3-like genes, characterized by NACHT and LRR domains, were also rapidly expanding through tandem duplication in the four Asian carp species. Our analysis revealed that *C. idella* possessed a significantly higher number of these genes compared to the other three species, while *M. piceus* exhibited the lowest gene count. The leucine-rich repeat (LRR) domain in NLRC3-like genes has been reported to play a critical role in the recognition of PAMPs from a wide range of pathogens, including bacteria, viruses, and fungi [77]. Thus, the expansion of NLRC3-like genes, characterized by NACHT and LRR domains, might increase the diversity of pathogen recognition modules within the four Asian carp species, thereby enhancing their innate immunity against various pathogen invasions. Overall, these findings shed light on the potential species-specific variations in the abundance of NLRC3/NLRC3-like genes among the four Asian carp species. The extensive repertoire of NLRC3/NLRC3-like genes suggests their crucial role in innate immune adaptation and defense mechanisms against pathogens in these domesticated Asian carp species.

Evolution of Chemosensory Receptors

The four major domesticated Asian carp species are known for their rapid growth, high reproductive potential, and great adaptability to varied environments [10, 78, 79]. Chemosensory receptors play an essential role in a wide range of fish behaviors, including feeding, foraging, pheromone sensation, and predator avoidance, by mediating the response to chemical stimuli [80]. These receptors are critical for supporting the survival, growth, and reproduction in fish. Recognizing their functional importance, we investigated the evolution of different

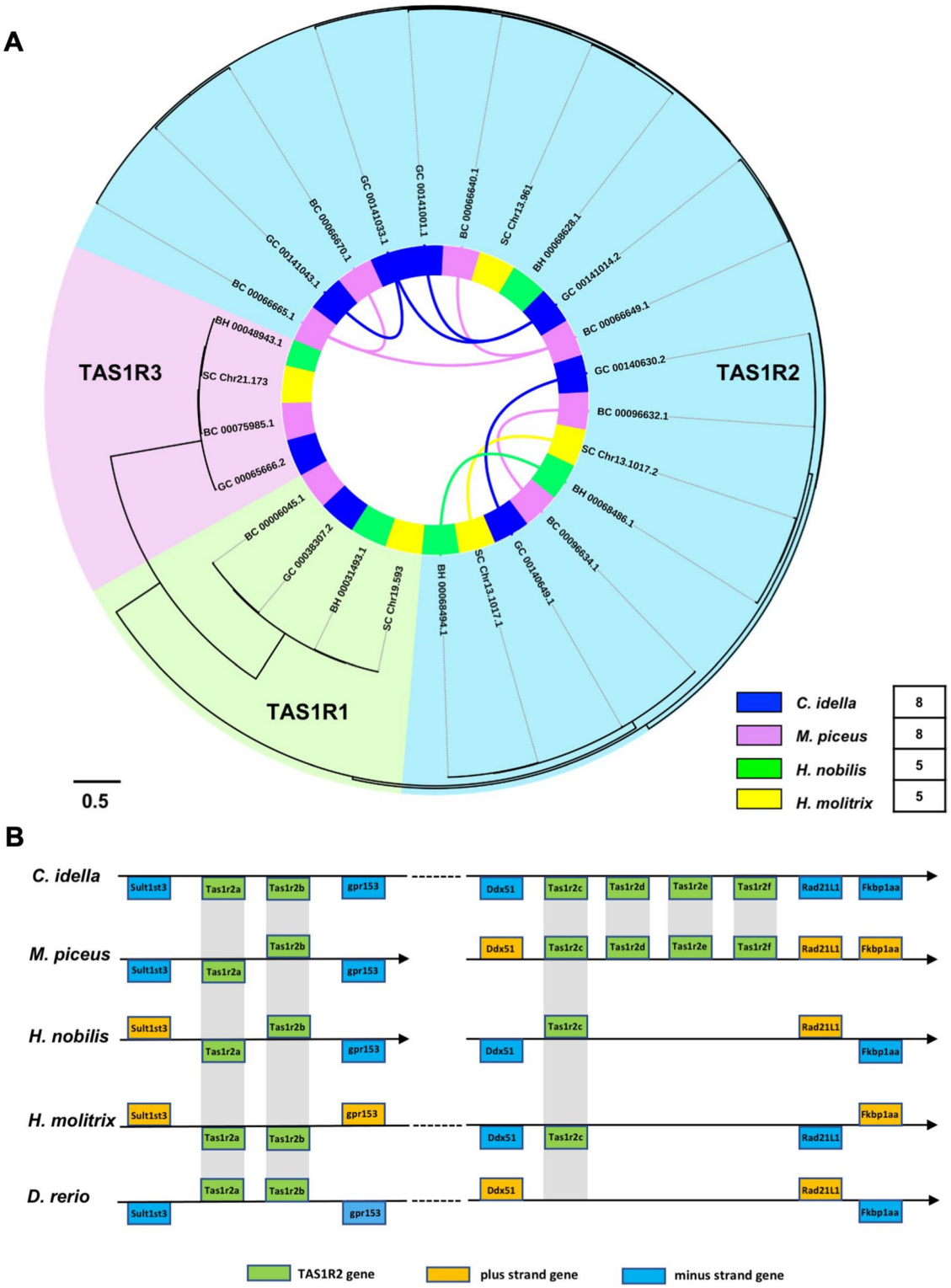

Fig. 6 The expansion of taste receptor type 1 genes (TAS1R). **A** The phylogenetic relationships of TAS1R genes were revealed in the four Asian carp species. All the TAS1R genes were classified into TAS1R1, TAS1R2 and TAS1R3 gene clusters. The number of TAS1R genes in each species was denoted and all the T/PAGs were connected by curved lines. **B** The gene synteny alignment of TAS1R2 genes in the four Asian carp species and *D. rerio* were analyzed. TAS1R2 A and TAS1R2B gene were tandemly arrayed in the five species examined. TAS1R2C, TAS1R2D, TAS1R2E and TAS1R2F gene were generated through tandem duplications in *C. idella* and *M. piceus*

chemosensory receptor families in the four domesticated Asian carp species.

TAARs

TAARs serve as an essential type of olfactory receptors in teleost fish, primarily involved in the detection of biogenic amines. These receptors have been classified into three clades, with clade III being specific to teleost [53]. In ray-finned fishes, the number of TAARs can range from 3 to 497, driven by species-specific gene expansion and loss events of TAARs [81]. In this study, we conducted a genome-wide analysis of the four Asian carp species to identify and classify all the TAAR genes into different subfamilies. The four Asian carp species demonstrated a similar number of TAAR genes; however, notable duplication events were predominantly observed in the TAAR16 and TAAR29 subfamilies when compared to zebrafish and goldfish. TAAR16 has been shown to recognize several biogenic amine ligands, including isoamylamine, N-methylpiperidine and N-methylpyrrolidine in fish [82]. Isoamylamine, in particular, has been found to act as a pheromone that can accelerate the onset of puberty in female mice [83]. Given the high fecundity and rapid growth of Asian carp species [84], the expansion of TAAR16 genes may enhance pheromone sensation, potentially contributing to the sexual maturation and reproductive success of these species. Meanwhile, isoamylamine, N-methylpiperidine and N-methylpyrrolidine are biogenic amines commonly found in decaying meat, fermented foodstuffs, and plants [85]. These compounds have been implicated in the regulation of gastrointestinal peristalsis and stimulation of gastric and pancreatic secretions, highlighting their potential role in digestive processes [86, 87]. The expansion of TAAR16 genes in the four Asian carp species is therefore suggested to enhance their intestinal function and digestive regulation, potentially supporting their feeding behaviors. Furthermore, TAAR29, a newly discovered teleost TAAR subfamily [16], was identified in Asian carp and goldfish but not in zebrafish. In addition, a distinct cluster of TAARs has been found exclusively in the four Asian carp species. These findings suggest potential species-specific adaptations and functional divergence in trace amine sensing within these domesticated Asian carp. Except for TAAR1, all the other TAARs function as olfactory receptors that bind to various biogenic amines, including volatile and highly aversive amines [88]. TAAR-mediated detection of different amines has been indicated to play an essential role in intra- and interspecific communication, including food source detection, predator proximity, and conspecific sex signalling [89]. Meanwhile, TAARs have been shown to elicit innate behaviours in fish. In zebrafish, TAAR13 is activated by foul-smelling amines such as cadaverine and putrescine, which triggers aversive responses to facilitate the avoidance of potential threats [90]. Therefore, the expansion of TAAR29 and the emergence of the novel TAAR cluster genes might contribute to the variations in innate behaviours and intra- and interspecific communication in the four Asian carp species, likely through differential recognition of specific amine ligands. Further investigation is required to identify the ligands and functional roles of the expanded and unique TAARs, as well as to elucidate their contributions to the sensory functions and behavioral adaptations of Asian carp.

ORA

The fish ORA gene family, homologous to mammalian V1Rs, is crucial for pheromone detection and highly conserved in teleost [36]. Typically, most of the fish species possess six canonical ORA genes (ORA1-ORA6), with rare duplication and loss events occurring [36]. Previous studies have revealed the species-specific gene duplication events of the ORA3 genes in zebrafish, salmon and medaka [91]. In this study, we analyzed all the ORA genes in the four domesticated Asian carp species and identified an additional copy of ORA5 gene, denoted as ORA5b. The sequence alignment revealed that ORA5b gene exhibited higher sequence identity among the four Asian carp species than between ORA5a and ORA5b. This result aligns with the previous studies indicating that the ORA orthologs were more closely related than paralogs in teleost [36]. These findings suggest that the ORA gene family has an ancient evolutionary origin, with the emergence of the ORA genes predating the speciation events. Meanwhile, the high conservation of the ORA gene in teleost fish indicates its co-evolution with the conserved ligands, enabling important pheromone functions. Previous studies have shown that the ORA genes exhibited high affinity for various bile acids in fish [55, 92]. In zebrafish, ORA5 gene has been reported to interact with the bile acid lithocholic acid and cholanic acid, functioning in pheromone recognition [55]. Although the function of ORA5 has not been clearly elucidated yet, bile acids have been demonstrated to play essential roles in various behaviors in fish, including spawning, conspecific detection, and migration [93]. Therefore, the emergence of the ORA5b gene through tandem duplication in the four Asian carp species is suggested to enable the binding to a similar type of ligands, thereby supporting behaviors related to reproduction, communication, and migration.

Taste receptors

Most vertebrates harbor three TAS1R genes, classified into TAS1R1, TAS1R2 and TASR3. In mammals,

the TAS1R1/TAS1R3 heterodimer functions as umami taste receptor, while the TAS1R2/TAS1R3 heterodimer acts as sweet sensor [94]. It has been proposed that the evolution of TAS1R genes in mammals was associated with diet specializations, with the umami taste receptor TAS1R1/TAS1R3 being more related to a carnivorous diet [95, 96]. However, a recent study investigated the evolution of different chemosensory receptors across various vertebrates and demonstrated that: 1) the number of TAS1R genes remains relatively stable compared to other chemosensory receptor families, and 2) no significant association was found between the loss of TAS1R1 or TAS1R3 genes and dietary preferences in mammals [81]. Fish TAS1R1 and TAS1R3 genes have been found to share a high degree of sequence identity with mammalian TAS1R1 and TAS1R3 genes, respectively [58]. Previous studies revealed that most ray-finned fish species typically possess a single ortholog of TAS1R1 and TAS1R3 gene (Table S8), with no significant association identified between TAS1R gene copy number and dietary preferences [56, 81]. These findings indicate the evolutionary conservation of TAS1R1 and TAS1R3 in fish and provide a plausible explanation for the identical copy number of these genes in the four Asian carp species, despite their different diets. Although the copy number of TAS1R1 and TAS1R3 genes is consistent across many fish species, their expression profiles and regulatory mechanisms may vary, potentially reflecting adaptations to diverse dietary preferences. Cai et al. [97] demonstrated that *C. idella* undergoes a transition from carnivore to herbivore throughout their development. This dietary shift is accompanied by a significant decrease in TAS1R1 expression, driven by elevated DNA methylation, highlighting the role of epigenetic modifications in dietary adaptation. Therefore, the four Asian carp species may exhibit different expression profiles of TAS1R1/TAS1R3 genes to support diet adaptation, despite having the same gene copy number.

While TAS1R1 and TAS1R3 genes are typically present as a single copy in fish, the copy number of TAS1R2 gene varies among different species, contributing to the diversification of the TAS1R family expression profile in fish taste buds [56]. The heterodimeric TAS1R2/TAS1R3 is involved in mediating glucose perception. Recent studies further indicated that a higher abundance of TAS1R2 is linked to increased carbohydrate intake in fish species including rainbow trout (*Oncorhynchus mykiss*) and Japanese flounder (*Paralichthys olivaceus*), highlighting the important role of the TAS1R2 genes in sweet taste perception [98, 99]. Yuan et al. [35] performed a genome-wide analysis of the TAS1R gene family in various teleost fish species and showed that the majority of omnivorous and carnivorous fish species examined possessed two or three TAS1R2 genes. Furthermore, their findings revealed the expansion of TAS1R2 genes in *C. idella* and suggested that the expansion and increased expression of TAS1R2 genes could mediate carbohydrate-induced signaling and contribute to the dietary shift of *C. idella* from carnivory to herbivory [35]. These results align with our findings that the omnivorous filter-feeders *H. nobilis* and *H. molitrix* harbored three copies of TAS1R2 genes, while the herbivorous *C. idella* exhibited an expansion of TAS1R2 genes, with six copies identified. Previous studies in mammals have reported that carnivores are more prone to the loss of TAS1R2 genes compared to herbivores and omnivores [95, 100]. Therefore, it is plausible to expect that carnivorous fish species may have fewer copies of TAS1R2 genes. Nevertheless, Policarpio et al. [81] analyzed the evolution of different chemosensory receptors in various vertebrates and found no significant association between the number of TAS1R family genes and dietary preferences within Actinopterygii (ray-finned fishes). Based on their findings, we investigated the correlation between the copy number of TAS1R2 genes and dietary preference in various fish species but similarly found no significant association (Fig. S9, Table S9). Thus, variations in the copy number of TAS1R2 genes alone may not be able to fully explain the dietary specializations among different fish species. Previous studies have reported the expansion of TAS1R2 genes in carnivorous fish species such as *Sparus aurata* and *Gasterosteus aculeatus* [101, 102]. Consistent with these findings, we also observed the expansion of TAS1R2 genes in the carnivorous species *M. piceus*. Furthermore, the TAS1R2 genes have been shown to respond to L-amino acids in omnivorous species such as *D. rerio* (zebrafish) and *Oryzias latipes* (medaka), as well as in the carnivorous species *Sparus aurata* [103]. In contrast, in the herbivorous species *C. idella*, the TAS1R2 genes were reported to preferentially respond to plant-specific fructose [35]. These findings suggest that the expansion of TAS1R2 genes might contribute to species-specific taste adaptation to diverse diets in fish. The close phylogenetic relationship between *M. piceus* and *C. idella*, coupled with their dietary divergence, prompted us to investigate the functional roles of the expanded TAS1R2 genes. Our results revealed that all the TAS1R2 genes in *C. idella* and *M. piceus* retain intact taste receptor domains, suggesting the preservation of TAS1R2 functional structures in both species (Fig S10). Meanwhile, sequence alignment of TAS1R2 proteins revealed amino acid residue variation at the ligand binding sites and glycosylation sites among *C. idella*, *M. piceus* and other fish species (Fig. S10). Motif analysis further showed the enriched motif patterns in the herbivorous *C. idella* and the carnivorous *M. piceus*, respectively (Table S10). These findings might shed light on the understanding of TAS1R2 gene expansion and its role in dietary adaptation

in *C. idella* and *M. piceus*. The expansion of TAS1R2 genes in certain carnivorous and omnivorous fish species is associated with the recognition of amino acids rather than sugar ligands. Therefore, it is plausible that the expansion of TAS1R2 genes in *M. piceus* may facilitate responses to amino acid stimuli, whereas in *C. idella*, it is primarily associated with sugar sensation. Further investigation is needed to characterize the specific ligands and the expression profiles of TAS1R2 genes in *M. piceus* and *C. idella* to elucidate their functions in species-specific taste adaptation.

Taste receptor 2 (TAS2R) genes function as bitter taste receptors, essential for detecting potentially toxic compounds [104]. In this study, we identified all the TAS2R genes in the four Asian carp species and found that each species contained an identical number of TAS2R genes (N= 3), indicating a small gene repertoire. This finding aligns with the previous research that ray-finned fishes have a relatively small repertoire of TAS2R genes compared to other vertebrates [105]. Specifically, Cypriniformes, the order to which the four Asian carp species belong, have an average of fewer than 10 TAS2R genes [106]. For vertebrates, although it has been proposed that herbivores possessed a larger TAS2R gene repertoire than carnivores, enabling them to detect toxic compounds in plants, recent findings revealed that omnivores had the highest number of TAS2R genes, followed by herbivores and carnivores [81]. Interestingly, this correlation between dietary preferences and TAS2R gene repertoire does not apply to ray-finned fishes, birds, or crocodiles [81]. This might explain why the four Asian carp species exhibited an identical number of TAS2R genes, despite having different diets.

Potential significance of calcium ion transport genes

Comparative analysis of gene families revealed both species-specific and shared rapidly evolving gene families among the four Asian carp species. Notably, genes associated with calcium ion transport were found to be rapidly expanding across all four species (Fig. S3E). Selection pressure analysis further identified three genes (CAPN10, SVIL, and ENKD1) involved in calcium ion signaling and muscle function [107–109], under positive selection across all four Asian carp species (Table S5A). The four Asian carp species are recognized as rapidly growing cyprinids, characterized by their large body size and fast skeletal muscle development [10]. Calcium ion signaling is fundamental to muscle development, growth, maintenance, and regeneration across diverse taxa, including both invertebrates and vertebrates [110]. Wang et al. [111] analyzed the genomic features of *H. nobilis* and *H. molitrix* populations in the Mississippi River Basin, revealing significant expansion of gene families related to calcium ion homeostasis, binding and transport. These gene expansions are implicated in facilitating muscle contraction and myosin filament assembly. These findings are consistent with our observations, suggesting the critical role of calcium ion regulation in muscle development of the four Asian carp species, which likely supports their rapid growth.

Conclusion

This study utilized high-quality genome assemblies to perform a comprehensive comparative genomic analysis of the four major domesticated Asian carp species (*C. idella*, *M. piceus*, *H. nobilis*, and *H. molitrix*). Our findings revealed the rapid expansion of MHC class I genes in the four Asian carp species, particularly within the U and Z lineages. This suggests the enhanced adaptive immunity of the four major domesticated Asian carp species, potentially improving their ability to combat various pathogens. Massive duplication events of the NLRC3/NLRC3-like gene family in the four Asian carp species indicated their elevated innate immunity. The species-specific variation of NLRC3/NLRC3-like gene repertoire might reflect their differences in innate immune responses. The expansion of these immune-related genes may contribute to the ability of the four domesticated Asian carp species to successfully colonize new habitats as invasive species. The comparative analysis of chemosensory receptor genes revealed a notable expansion of TAAR genes in the four Asian carp species, particularly within the TAAR16 and TAAR29 subfamilies. This expansion likely enhances the olfactory function and pheromone detection in the four Asian carp species, supporting their social communication, innate behaviors and reproductive success. In addition, the expansion of ORA5 and TAS1R2 genes in the four domesticated Asian carp species may play a critical role in pheromone sensation and taste adaptation, providing insights into their reproductive behavior and dietary specialization. Overall, the evolution of immune-related genes and chemosensory receptors suggests the enhanced immunity and sensory perception of the four domesticated Asian carp species, shedding light on their adaptation, survival and reproduction.

Abbreviations

OR	Olfactory receptor
TAAR	Trace amine-associated receptor
V1R/ORA	Vomeronasal receptor type 1
TAS1R	Taste receptor type 1
TAS2R	Taste receptor type 2
MHC	Major histocompatibility complex

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12864-025-11719-2.

Supplementary Material 1.

Supplementary Material 2.

Supplementary Material 3.

Acknowledgements
Not applicable.

Authors' contributions
L.S., Q.X. and S.T. contributed to the conception and design of the study. L.S., J.N., Q.X., K.A., S.S. and C.L. were responsible for laboratory study. L.S. and J.N. performed data curation and bioinformatic analysis. L.S., W.M. and S.R. validated the results and L.S. wrote the first draft of the manuscript. L.S., Q.X., K.A., G.L. and S.T. finalized the manuscript. All authors contributed to the manuscript and approved the final submitted version.

Funding
This project was supported by the General Research Funds (Ref. No.: 14119420 and 14121222) and Theme-based Research Scheme (T11-709/21-N, T12-716/22-R and T11-702/24-N) of the Hong Kong Research Grants Council.

Data availability
The genomes of *Ctenopharyngodon idella*, *Mylopharyngodon piceus* and *Hypophthalmichthys nobilis* were available on NCBI database (BioProject accession: PRJNA890423, PRJNA891927 and PRJNA892279). The genome sequencing data of *Hypophthalmichthys molitrix* were available on NCBI database (BioProject accession: PRJNA631443) and the assembled genome and annotation files were deposited in FigShare (https://doi.org/10.6084/m9.figshare.12618884.v1).

Declarations

Ethics approval and consent to participate
Ethical approval does not apply to this study since live animals were not involved. The carp samples were sacrificed at the wet market before purchase.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Author details
[1]School of Biomedical Sciences, The Chinese University of Hong Kong, Hong Kong, China. [2]Hong Kong Bioinformatics Centre, The Chinese University of Hong Kong, Hong Kong, China. [3]Department of Applied Science, School of Science and Technology, Hong Kong Metropolitan University, Hong Kong, China. [4]Department of Health Technology and Informatics, The Hong Kong Polytechnic University, Hong Kong, China. [5]College of Ocean Food and Biological Engineering, Xiamen Key Laboratory of Marine Functional Food, Fujian Provincial Engineering Technology Research Center of Marine Functional Food, Jimei University, Xiamen 361021, China. [6]Department of Bioinformatics, Fujian Key Laboratory of Medical Bioinformatics, School of Medical Technology and Engineering, Fujian Medical University, Fuzhou 350122, China. [7]Centre for Microbial Genomics and Proteomics, The Chinese University of Hong Kong, Hong Kong, China.

Received: 14 October 2024 Accepted: 15 May 2025
Published online: 26 May 2025

References

1. Ivan LN, Mason DM, Zhang H, Rutherford ES, Hunter T, Sable S, et al. Potential establishment and ecological effects of bighead and silver carp in a productive embayment of the Laurentian Great Lakes. Biol Invasions. 2020;22(8):2473–95.
2. Wu CS, Ma ZY, Zheng GD, Zou SM, Zhang XJ, Zhang YA. Chromosome-level genome assembly of grass carp (*Ctenopharyngodon idella*) provides insights into its genome evolution. BMC Genomics. 2022;23(1):271.
3. Wang Y, Lu Y, Zhang Y, Ning Z, Li Y, Zhao Q, et al. The draft genome of the grass carp (*Ctenopharyngodon idellus*) provides insights into its evolution and vegetarian adaptation. Nat Genet. 2015;47(6):625–31.
4. Jian J, Yang L, Gan X, Wu B, Gao L, Zeng H, et al. Whole genome sequencing of silver carp (*Hypophthalmichthys molitrix*) and bighead carp (*Hypophthalmichthys nobilis*) provide novel insights into their evolution and speciation. Mol Ecol Resour. 2021;21(3):912–23.
5. Chen G, Pang MX, Yu XM, Wang JR, Tong JO. Transcriptome sequencing provides insights into the mechanism of hypoxia adaption in bighead carp (*Hypophthalmichthys nobilis*). Comp Biochem Physiol Part D Genomics Proteomics. 2021;40:100891.
6. George AE, Chapman DC. Aspects of embryonic and larval development in bighead carp *Hypophthalmichthys nobilis* and silver carp *Hypophthalmichthys molitrix*. PLoS One. 2013;8(8):e73829.
7. Li XM, Liu L, Zhu YJ, Zhu TB, Wu XB, Yang DG. Microbial community structure and its driving environmental factors in black carp (*Mylopharyngodon piceus*) aquaculture pond. Water (Basel). 2021;13(21):3089.
8. Lu L, Wang X, Wu S, Song X, Zou Z, Xie X, et al. Black carp STING functions importantly in innate immune defense against RNA virus. Fish Shellfish Immunol. 2017;70:13–24.
9. Wang C, Li J, Yang X, Wang Q, Zhong H, Liu Y, et al. Black carp IKKε collaborates with IRF3 in the antiviral signaling. Fish Shellfish Immunol. 2021;118:160–8.
10. Wang C, Yang L, Lu Y, Fang C, Gan X, Chen Y, et al. Genomic features for adaptation and evolutionary dynamics of four major Asian domestic carps. Sci China Life Sci. 2024;67(6):1308–10.
11. Wick RR, Judd LM, Gorrie CL, Holt KE. Completing bacterial genome assemblies with multiplex MinION sequencing. Microb Genom. 2017;3(10): e000132.
12. De Coster W, D'Hert S, Schultz DT, Cruts M, Van Broeckhoven C. NanoPack: visualizing and processing long-read sequencing data. Bioinformatics. 2018;34(15):2666–9.
13. Kolmogorov M, Yuan J, Lin Y, Pevzner PA. Assembly of long, error-prone reads using repeat graphs. Nat Biotechnol. 2019;37(5):540–6.
14. Boetzer M, Pirovano W. SSPACE-LongRead: scaffolding bacterial draft genomes using long read sequence information. BMC Bioinformatics. 2014;15:211.
15. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics. 2009;25(14):1754–60.
16. Walker BJ, Abeel T, Shea T, Priest M, Abouelliel A, Sakthikumar S, et al. Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. PLoS ONE. 2014;9(11): e112963.
17. Gurevich A, Saveliev V, Vyahhi N, Tesler G. QUAST: quality assessment tool for genome assemblies. Bioinformatics. 2013;29(8):1072–5.
18. Simao FA, Waterhouse RM, Ioannidis P, Kriventseva EV, Zdobnov EM. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. Bioinformatics. 2015;31(19):3210–2.
19. Kim D, Paggi JM, Park C, Bennett C, Salzberg SL. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. Nat Biotechnol. 2019;37(8):907–15.
20. Pertea M, Pertea GM, Antonescu CM, Chang TC, Mendell JT, Salzberg SL. StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. Nat Biotechnol. 2015;33(3):290–5.
21. Danecek P, Bonfield JK, Liddle J, Marshall J, Ohan V, Pollard MO, et al. Twelve years of SAMtools and BCFtools. Gigascience. 2021;10(2):giab008.
22. Flynn JM, Hubley R, Goubert C, Rosen J, Clark AG, Feschotte C, et al. RepeatModeler2 for automated genomic discovery of transposable element families. Proc Natl Acad Sci U S A. 2020;117(17):9451–7.

23. Tarailo-Graovac M, Chen N. Using RepeatMasker to identify repetitive elements in genomic sequences. Curr Protoc Bioinformatics. 2009;Chapter 4:4 10 1-4 10 14.
24. Johnson AD, Handsaker RE, Pulit SL, Nizzari MM, O'Donnell CJ, de Bakker PI. SNAP: a web-based tool for identification and annotation of proxy SNPs using HapMap. Bioinformatics. 2008;24(24):2938–9.
25. Stanke M, Keller O, Gunduz I, Hayes A, Waack S, Morgenstern B. AUGUSTUS: *ab initio* prediction of alternative transcripts. Nucleic Acids Res. 2006;34(Web Server issue):W435-9.
26. Borodovsky M, Lomsadze A. Eukaryotic gene prediction using GeneMark.hmm-E and GeneMark-ES. Curr Protoc Bioinformatics. 2011;Chapter 4:4 6 1-4 6 10.
27. Zhou Y, Qin W, Zhong H, Zhang H, Zhou L. Chromosome-level assembly of the *Hypophthalmichthys molitrix* (Cypriniformes: Cyprinidae) genome provides insights into its ecological adaptation. Genomics. 2021;113(5):2944–52.
28. Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol. 2013;30(4):772–80.
29. Stamatakis A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics. 2014;30(9):1312–3.
30. Yang Z. PAML 4: phylogenetic analysis by maximum likelihood. Mol Biol Evol. 2007;24(8):1586–91.
31. Letunic I, Bork P. Interactive Tree of Life (iToL): an online tool for phylogenetic tree display and annotation. Bioinformatics. 2007;23(1):127–8.
32. Emms DM, Kelly S. Orthofinder: Solving fundamental biases in whole genome comparisons dramatically improves orthogroup inference accuracy. Genome Biol. 2015;16(1):157.
33. De Bie T, Cristianini N, Demuth JP, Hahn MW. CAFE: a computational tool for the study of gene family evolution. Bioinformatics. 2006;22(10):1269–71.
34. Guo L, Dai W, Xu Z, Liang Q, Miller ET, Li S, et al. Evolution of brain-expressed biogenic amine receptors into olfactory trace amine-associated receptors. Mol Biol Evol. 2022;39(3):msac006.
35. Yuan XC, Liang XF, Cai WJ, He S, Guo WJ, Mai KS. Expansion of sweet taste receptor genes in grass carp (*Ctenopharyngodon idellus*) coincided with vegetarian adaptation. BMC Evol Biol. 2020;20(1):25.
36. Saraiva LR, Korsching SI. A novel olfactory receptor gene family in teleost fish. Genome Res. 2007;17(10):1448–57.
37. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, et al. BLAST+: architecture and applications. BMC Bioinformatics. 2009;10:421.
38. Eddy SR. Accelerated profile HMM searches. PLoS Comput Biol. 2011;7(10): e1002195.
39. Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 2004;32(5):1792–7.
40. Price MN, Dehal PS, Arkin AP. FastTree: computing large minimum evolution trees with profiles instead of a distance matrix. Mol Biol Evol. 2009;26(7):1641–50.
41. Kumar S, Stecher G, Li M, Knyaz C, Tamura K. MEGA X: molecular evolutionary genetics analysis across computing platforms. Mol Biol Evol. 2018;35(6):1547–9.
42. Bailey TL, Boden M, Buske FA, Frith M, Grant CE, Clementi L, et al. MEME SUITE: tools for motif discovery and searching. Nucleic Acids Res. 2009;37(Web Server issue):W202-8.
43. Sievers F, Higgins DG. Clustal Omega for making accurate alignments of many protein sequences. Protein Sci. 2018;27(1):135–45.
44. Quevillon E, Silventoinen V, Pillai S, Harte N, Mulder N, Apweiler R, et al. InterProScan: protein domains identifier. Nucleic Acids Res. 2005;33(Web Server issue):W116-20.
45. Grimholt U, Tsukamoto K, Hashimoto K, Dijkstra JM. Discovery of a novel MHC class I lineage in teleost fish which shows unprecedented levels of ectodomain deterioration while possessing an impressive cytoplasmic tail motif. Cells. 2019;8(9):1056.
46. Yamaguchi T, Dijkstra JM. Major histocompatibility complex (MHC) genes and disease resistance in fish. Cells. 2019;8(4):378.
47. Hashimoto K, Nakanishi T, Kurosawa Y. Isolation of carp genes encoding major histocompatibility complex antigens. Proc Natl Acad Sci U S A. 1990;87(17):6863–7.
48. Grimholt U, Tsukamoto K, Azuma T, Leong J, Koop BF, Dijkstra JM. A comprehensive analysis of teleost MHC class I sequences. BMC Evol Biol. 2015;15:32.
49. Chang MX, Xiong F, Wu XM, Hu YW. The expanding and function of NLRC3 or NLRC3-like in teleost fish: recent advances and novel insights. Dev Comp Immunol. 2021;114: 103859.
50. Xu T, Liao Z, Su J. Pattern recognition receptors in grass carp *Ctenopharyngodon idella*: II. organization and expression analysis of NOD-like receptors. Dev Comp Immunol. 2020;110:103734.
51. Ferkey DM, Sengupta P, L'Etoile ND. Chemosensory signal transduction in *Caenorhabditis elegans*. Genetics. 2021;217(3):iyab004.
52. Nei M, Niimura Y, Nozawa M. The evolution of animal chemosensory receptor gene repertoires: roles of chance and necessity. Nat Rev Genet. 2008;9(12):951–63.
53. Hussain A, Saraiva LR, Korsching SI. Positive Darwinian selection and the birth of an olfactory receptor clade in teleosts. Proc Natl Acad Sci U S A. 2009;106(11):4313–8.
54. Hashiguchi Y, Nishida M. Evolution of trace amine associated receptor (TAAR) gene family in vertebrates: lineage-specific expansions and degradations of a second class of vertebrate chemosensory receptors expressed in the olfactory epithelium. Mol Biol Evol. 2007;24(9):2099–107.
55. Cong X, Zheng Q, Ren W, Cheron JB, Fiorucci S, Wen T, et al. Zebrafish olfactory receptors ORAs differentially detect bile acids and bile salts. J Biol Chem. 2019;294(17):6762–71.
56. Matsumoto I, Ohmoto M, Abe K. Functional diversification of taste cells in vertebrates. Semin Cell Dev Biol. 2013;24(3):210–4.
57. Picone B, Hesse U, Panji S, Van Heusden P, Jonas M, Christoffels A. Taste and odorant receptors of the coelacanth–a gene repertoire in transition. J Exp Zool B Mol Dev Evol. 2014;322(6):403–14.
58. Ishimaru Y, Okada S, Naito H, Nagai T, Yasuoka A, Matsumoto I, et al. Two families of candidate taste receptors in fishes. Mech Dev. 2005;122(12):1310–21.
59. Tang H, Wang Y, Mao S, Guo J, Bao S, Xu X, et al. The complete mitogenome reveals genetic diversity and differentiation among wild and farmed black carp (*Mylopharyngodon piceus*) populations. Aquac Fish. 2024. https://doi.org/10.1016/j.aaf.2024.07.001.
60. Wang Y, Liu W, Li Z, Qiu B, Li J, Geng G, et al. Improvement and application of genetic resources of grass carp (*Ctenopharyngodon idella*). Reprod Breed. 2024;4(3):126–33.
61. Peng J, Lu X, Fan R, Ren Y, Sun H, Yu Y, et al. Analysis and evaluation of muscle quality in different parts of the bighead carp (*Aristichthys nobilis*). Foods. 2023;12(24):4430.
62. Lu Y, Xia H, Zhai W, Liu J, Zhou Y, Xu X, et al. Genome survey sequence of black carp provides insights into development-related gene duplications. 2022;53(6):1197–214.
63. Pruter H, Franz M, Twietmeyer S, Bohm N, Middendorff G, Portas R, et al. Increased immune marker variance in a population of invasive birds. Sci Rep. 2020;10(1):21764.
64. Biedrzycka A, Popiolek M, Zalewski A. Host-parasite interactions in non-native invasive species are dependent on the levels of standing genetic variation at the immune locus. BMC Evol Biol. 2020;20(1):43.
65. Dirscherl H, McConnell SC, Yoder JA, de Jong JL. The MHC class I genes of zebrafish. Dev Comp Immunol. 2014;46(1):11–23.
66. Xu C, Xue M, Jiang N, Li Y, Meng Y, Liu W, et al. Characteristics and expression profiles of MHC class I molecules in *Carassius auratus*. Fish Shellfish Immunol. 2023;137: 108794.
67. Sukeda M, Shiota K, Kondo M, Nagasawa T, Nakao M, Somamoto T. Innate cell-mediated cytotoxicity of $CD8^+$ T cells against the protozoan parasite *Ichthyophthirius multifiliis* in the ginbuna crucian carp. Carassius auratus langsdorfii Dev Comp Immunol. 2021;115: 103886.
68. Rakus KL, Irnazarow I, Adamek M, Palmeira L, Kawana Y, Hirono I, et al. Gene expression analysis of common carp (*Cyprinus carpio L.*) lines during Cyprinid herpesvirus 3 infection yields insights into differential immune responses. Dev Comp Immunol. 2012;37(1):65–76.
69. Nayak SK, Nakanishi T. Direct antibacterial activity of CD8+/CD4+ T-cells in ginbuna crucian carp. Carassius auratus langsdorfii Fish Shellfish Immunol. 2013;34(1):136–41.
70. Chen W, Jia Z, Zhang T, Zhang N, Lin C, Gao F, et al. MHC class I presentation and regulation by IFN in bony fish determined by molecular analysis of the class I locus in grass carp. J Immunol. 2010;185(4):2209–21.

71. Okamura K, Nakanishi T, Kurosawa Y, Hashimoto K. Expansion of genes that encode MHC class I molecules in cyprinid fishes. J Immunol. 1993;151(1):188–200.
72. Dirscherl H, Yoder JA. Characterization of the Z lineage major histocompatability complex class I genes in zebrafish. Immunogenetics. 2014;66(3):185–98.
73. Chuphal B, Rai U, Roy B. Teleost NOD-like receptors and their downstream signaling pathways: a brief review. Fish Shellfish Immunol Rep. 2022;3: 100056.
74. Howe K, Schiffer PH, Zielinski J, Wiehe T, Laird GK, Marioni JC, et al. Structure and evolutionary history of a large family of NLR proteins in the zebrafish. Open Biol. 2016;6(4): 160009.
75. Woo JS, Suh HY, Park SY, Oh BH. Structural basis for protein recognition by B30.2/SPRY domains. Mol Cell. 2006;24(6):967–76.
76. Liu Y, Sheng X, Tang X, Xing J, Chi H, Zhan W. Genome-wide identification, phylogenetic relationships and expression patterns of the NOD-like receptor (NLR) gene family in flounder (*Paralichthys olivaceus*). Fish Shellfish Immunol. 2023;141: 109083.
77. Sahoo BR. Structure of fish Toll-like receptors (TLR) and NOD-like receptors (NLR). Int J Biol Macromol. 2020;161:1602–17.
78. Zhao J, Xu D, Zhao K, Diogo R, Yang J, Peng Z. The origin and divergence of Gobioninae fishes (Teleostei: Cyprinidae) based on complete mitochondrial genome sequences. J Appl Ichthyol. 2016;32(1):32–9.
79. Shen YB, Wang L, Fu JJ, Xu XY, Yue GH, Li JL. Population structure, demographic history and local adaptation of the grass carp. BMC Genomics. 2019;20(1):467.
80. Daghfous G, Green WW, Zielinski BS, Dubuc R. Chemosensory-induced motor behaviors in fish. Curr Opin Neurobiol. 2012;22(2):223–30.
81. Policarpo M, Baldwin MW, Casane D, Salzburger W. Diversity and evolution of the vertebrate chemoreceptor gene repertoire. Nat Commun. 2024;15(1):1421.
82. Li Q, Tachie-Baffour Y, Liu Z, Baldwin MW, Kruse AC, Liberles SD. Non-classical amine recognition evolved in a large clade of olfactory receptors. Elife. 2015;4: e10441.
83. Price MA, Vandenbergh JG. Analysis of puberty-accelerating pheromones. J Exp Zool. 1992;264(1):42–5.
84. Jawdhari A, Mihailescu DF, Fendrihan S, Jujea V, Stoilov-Linu V, Negrea BM. Silver carp (*Hypophthalmichthys molitrix*) (Asian silver carp) presence in Danube Delta and Romania-a review with data on natural reproduction. Life (Basel). 2022;12(10):1582.
85. Gainetdinov RR, Hoener MC, Berry MD. Trace amines and their receptors. Pharmacol Rev. 2018;70(3):549–620.
86. Dial EJ, Cooper LC, Lichtenberger LM. Amino acid- and amine-induced gastrin release from isolated rat endocrine granules. Am J Physiol. 1991;260(2 Pt 1):G175–81.
87. Tsutsumi E, Kanai S, Ohta M, Suwa Y, Miyasaka K. Stimulatory effect of N-methyltyramine, a congener of beer, on pancreatic secretion in conscious rats. Alcohol Clin Exp Res. 2010;34(Suppl 1):S14–7.
88. Johnson MA, Tsai L, Roy DS, Valenzuela DH, Mosley C, Magklara A, et al. Neurons expressing trace amine-associated receptors project to discrete glomeruli and constitute an olfactory subsystem. Proc Natl Acad Sci U S A. 2012;109(33):13410–5.
89. Dewan A. Olfactory signaling via trace amine-associated receptors. Cell Tissue Res. 2021;383(1):395–407.
90. Hussain A, Saraiva LR, Ferrero DM, Ahuja G, Krishna VS, Liberles SD, et al. High-affinity olfactory receptor for the death-associated odor cadaverine. Proc Natl Acad Sci U S A. 2013;110(48):19579–84.
91. Zapilko V, Korsching SI. Tetrapod V1R-like ora genes in an early-diverging ray-finned fish species: The canonical six ora gene repertoire of teleost fish resulted from gene loss in a larger ancestral repertoire. BMC Genomics. 2016;17:83.
92. Dong X, Lv M, Zeng M, Chen X, Wang J, Liang XF. Genome-wide identification, characterization of the ora (olfactory receptor class a) gene family, and potential roles in bile acid and pheromone recognition in mandarin fish (*Siniperca chuatsi*). Cells. 2025;14(3):189.
93. Buchinger TJ, Li W, Johnson NS. Bile salts as semiochemicals in fish. Chem Senses. 2014;39(8):647–54.
94. Zhao GQ, Zhang Y, Hoon MA, Chandrashekar J, Erlenbach I, Ryba NJ, et al. The receptors for mammalian sweet and umami taste. Cell. 2003;115(3):255–66.
95. Jiang P, Josue J, Li X, Glaser D, Li W, Brand JG, et al. Major taste loss in carnivorous mammals. Proc Natl Acad Sci U S A. 2012;109(13):4956–61.
96. Zhao H, Yang JR, Xu H, Zhang J. Pseudogenization of the umami taste receptor gene Tas1r1 in the giant panda coincided with its dietary switch to bamboo. Mol Biol Evol. 2010;27(12):2669–73.
97. Cai W, He S, Liang XF, Yuan X. DNA methylation of T1R1 gene in the vegetarian adaptation of grass carp *Ctenopharyngodon idella*. Sci Rep. 2018;8(1):6934.
98. Balasubramanian MN, Panserat S, Dupont-Nivet M, Quillet E, Montfort J, Le Cam A, et al. Molecular pathways associated with the nutritional programming of plant-based diet acceptance in rainbow trout following an early feeding exposure. BMC Genomics. 2016;17:449.
99. Liu D, Deng KY, Sampath WWHA, Gu ZX, Pan MZ, Zhang Y, et al. Responses of glucosensing system to glucose in japanese flounder fed diets with different carbohydrate content. Comp Biochem Physiol B Biochem Mol Biol. 2019;232:72–8.
100. Wolsan M, Sato JJ. Role of feeding specialization in taste receptor loss: insights from sweet and umami receptor evolution in carnivora. Chem Senses. 2022;47:bjac033.
101. Hashiguchi Y, Furuta Y, Kawahara R, Nishida M. Diversification and adaptive evolution of putative sweet taste receptors in threespine stickleback. Gene. 2007;396(1):170–9.
102. Angotzi AR, Puchol S, Cerda-Reverter JM, Morais S. Insights into the function and evolution of taste 1 receptor gene family in the carnivore fish gilthead seabream (*Sparus aurata*). Int J Mol Sci. 2020;21(20):7732.
103. Oike H, Nagai T, Furuyama A, Okada S, Aihara Y, Ishimaru Y, et al. Characterization of ligands for fish taste receptors. J Neurosci. 2007;27(21):5584–92.
104. Chandrashekar J, Mueller KL, Hoon MA, Adler E, Feng L, Guo W, et al. T2Rs function as bitter taste receptors. Cell. 2000;100(6):703–11.
105. Behrens M, Di Pizio A, Redel U, Meyerhof W, Korsching SI. At the root of T2R gene evolution: recognition profiles of coelacanth and zebrafish bitter receptors. Genome Biol Evol. 2021;13(1):evaa264.
106. Itoigawa A, Nakagita T, Toda Y. The remarkable diversity of vertebrate bitter taste receptors: recent advances in genomic and functional studies. Int J Mol Sci. 2024;25(23):12654.
107. Metwally E, Al-Abbadi HA, Hussain T, Murtaza G, Abdellatif AM, Ahmed MF. Calpain signaling: from biology to therapeutic opportunities in neurodegenerative disorders. Front Vet Sci. 2023;10:1235163.
108. Crowley JL, Smith TC, Fang Z, Takizawa N, Luna EJ. Supervillin reorganizes the actin cytoskeleton and increases invadopodial efficiency. Mol Biol Cell. 2009;20(3):948–62.
109. Tiryaki F, Deretic J, Firat-Karalar EN. ENKD1 is a centrosomal and ciliary microtubule-associated protein important for primary cilium content regulation. FEBS J. 2022;289(13):3789–812.
110. Tu MK, Levin JB, Hamilton AM, Borodinsky LN. Calcium signaling in skeletal muscle development, maintenance and regeneration. Cell Calcium. 2016;59(2–3):91–7.
111. Wang J, Gaughan S, Lamer JT, Deng C, Hu W, Wachholtz M, et al. Resolving the genetic paradox of invasions: preadapted genomes and postintroduction hybridization of bigheaded carps in the Mississippi river basin. Evol Appl. 2020;13(2):263–77.

Publisher's Note