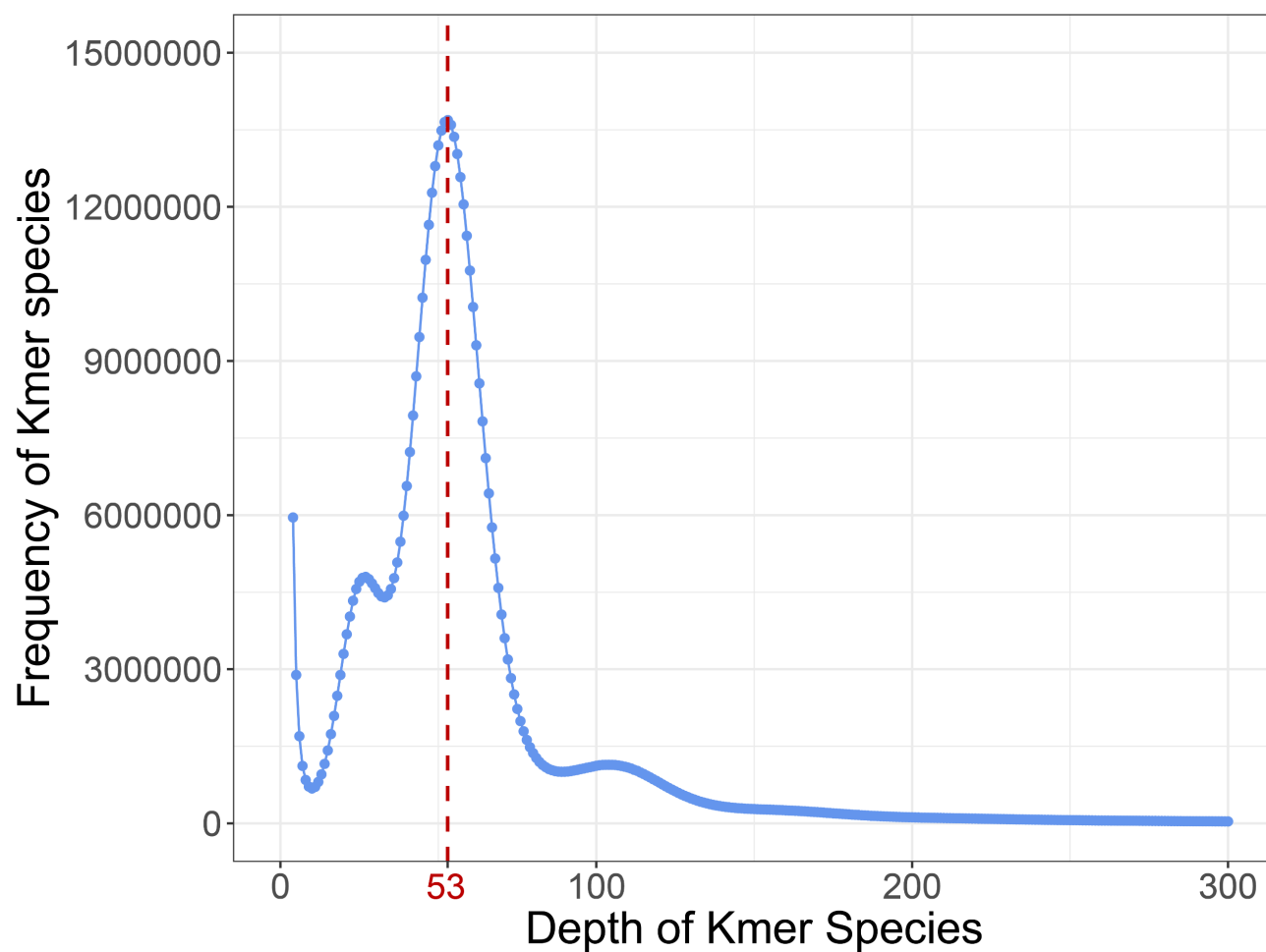


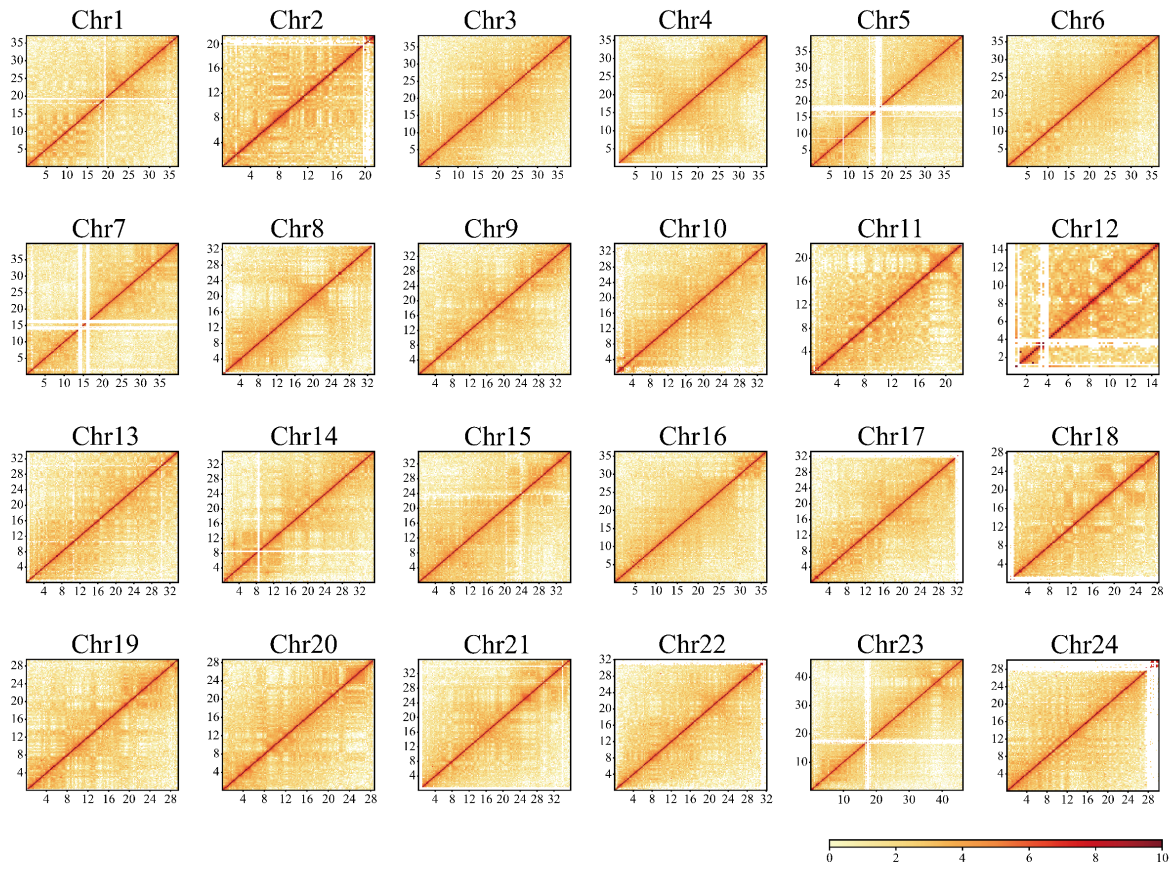
Supplementary Material

1 Supplementary Figures and Tables

1.1 Supplementary Figures



Supplementary Figure 1. Distribution of K-mers of length 17 from the Illumina data. The red dashed line represents the main peak of K-mer frequency.



Supplementary Figure 2. Hi-C matrix for each of the 24 chromosomes. The color of each pixel represents the count of Hi-C links between 200kb windows on each chromosome on a logarithmic scale. Darker red pixels represent higher inter-chromosomal interactions.

1.2 Supplementary Table

Supplementary Table 1. Statistics of sequencing data.

Platform	Reads	Total bases (bp)	N50 length (bp)	Average length (bp)
HiFi	1553426	21067117955	13869	13562
Illumina	337644822	50646723300		150
Hi-C	665871558	99880733700		150
RNA-seq	196675134	29501270100		150

Supplementary Table 2. Estimated genome size of *Scartelaos histophorus* from k-mer analysis.

K-mer	K-mer number	K-mer Depth	Genome Size (Mbp)	Heterozygous Ratio (%)	Repeat (%)
17	44350933428	53	817.12	0.62	47.54

Supplementary Table 3. Statistics of assembled 24 chromosomes.

Chromosome name	Total length (bp)	GC content(%)	Contig number
Chr1	37060754	39.58	32
Chr2	21348399	40.11	27
Chr3	38359349	39.56	33
Chr4	36246144	39.6	25
Chr5	39964816	39.61	36
Chr6	36692761	39.41	11
Chr7	39851156	39.95	29
Chr8	33576242	39.26	26
Chr9	35519128	39.68	21
Chr10	34051215	39.32	18
Chr11	22418847	39.8	26
Chr12	14652100	39.49	13
Chr13	33880487	39.75	52
Chr14	35592069	39.44	17
Chr15	35283103	39.36	12
Chr16	36355260	39.86	29
Chr17	33206237	39.53	14
Chr18	28216037	39.66	15

Chr19	29453772	39.43	40
Chr20	28651772	39.32	33
Chr21	35875693	39.84	36
Chr22	32163828	40.04	37
Chr23	46379213	39.63	29
Chr24	30077440	39.27	5
Linked Total	794875822	39.6	616
Unlinked Total	74608786	45.47	578
ALL	869484608	40.1	1194

Supplementary Table 4. BUSCO assessment of the genome of *Scartelaos histophorus*.

genome	BUSCO assessment	Complete BUSCOs	Complete and single-copy BUSCOs	Complete and duplicated BUSCOs	Fragmented BUSCOs	Missing BUSCOs
This study	vertebrata_odb10	3243(96.7%)	3168(94.5%)	75(2.2%)	48(1.4%)	63(1.9%)
	actinopterygii_odb10	3546(97.4%)	3477(95.5%)	69(1.9%)	32(0.9%)	62(1.7%)
The Previous (You et al., 2014)	vertebrata_odb10	2492(74.3%)	2466(73.5%)	26(0.8%)	542(16.2%)	320(9.5%)
	actinopterygii_odb10	2746(75.5%)	2715(74.6%)	31(0.9%)	251(6.9%)	643(17.6%)

Supplementary Table 5. The genome coverage statistics of Illumina sequence reads.

	Percentage
Mapping rate(%)	99.05%
Coverage(%)	99.87%
Coverage at least 4X (%)	99.64%
Coverage at least 10X (%)	98.97%
Coverage at least 30X (%)	92.44%

Supplementary Table 6. Statistic of repeat content in the *Scartelaos histophorus* genome.

Repeat type		Classification	Number	Masked(bp)	Masked(%)
Class I: Retrotransposon	LTR- Retrotransposon	Copia	5014	1073677	0.12
		Gypsy	83658	27366377	3.15
		Others	798156	156039770	17.95
	Non-LTR Retrotransposon	LINE	186641	37028081	4.26
		SINE	31640	3142879	0.36
Class II: DNA Transposon		hobo-Activator	421090	68852219	7.92
		Tc1-IS630-Pogo	33012	7882457	0.91
		PiggyBac	208	43902	0.01
		Tourist/Harbinger	7791	1699376	0.20
		Other	49559	13286416	1.53
Tandem Repeat		Simple_repeat	317552	30548809	3.51
		Satellite	3272	1202900	0.14
Unknown		-	-	78485443	9.03
Total		-	-	426652306	49.07

Supplementary Table 7. Statistics of predicted protein-coding genes.

	Software		Number of genes	Average Transcript Length (bp)	Average CDS Length(bp)	Average Exons per gene	Average Exon Length(bp)	Average Intron Length(bp)
<i>Ab initio</i>	Augustus、 STAR、 genemake	/	52110	7363.85	1095.65	6.26	174.98	1191.35
Homology- based	genomeThreader	<i>D.rerio</i>	10343	15073.33	1493.72	9.78	152.81	1547.50
		<i>A.testudineus</i>	14122	17083.67	1740.79	10.93	159.33	1545.83
		<i>L.crocea</i>	14814	16550.36	1695.38	10.65	159.15	1538.96
		<i>O.niloticus</i>	14239	16673.51	1701.62	10.75	158.31	1535.78
		<i>P.magnuspinnatus</i>	21444	15170.35	1749.70	10.52	166.32	1409.76
		<i>B.pectinirostris</i>	22869	14505.08	1691.35	10.26	164.89	1384.12
RNA-seq	hisat2、 PASA、 TransDecoder、 stringtie、 TransDecoder	/	35751	15722.65	1069.40	5.90	181.15	2988.34
Integration	EVM	/	24223	12364.79	1613.28	9.59	168.20	1251.43

Supplementary Table 8. BUSCO assessment of gene prediction of *Scartelaos histophorus*.

BUSCO assessment	Complete BUSCOs	Complete and single-copy BUSCOs	Complete and duplicated BUSCOs	Fragmented BUSCOs	Missing BUSCOs
vertebrata_odb10	3050(90.9%)	2969(88.5%)	81(2.4%)	114(3.4%)	190(5.7%)
actinopterygii_odb10	3300(90.7%)	3238(89.0%)	62(1.7%)	109(3.0%)	231(6.3%)

Supplementary Table 9. The statistical results of gene function annotation.

		Number	Percent (%)
Swiss-Prot		17084	70.53%
NR		22416	92.54%
KEGG		15745	65.00%
	all	20674	85.35%
InterPro	Pfam	20674	85.35%
	GO	13128	54.20%
Annotated		22702	93.72%
Total		24223	

Supplementary Table 10. Data address and taxonomy of species used in comparative genomics analysis.

Species	Order/Family	data
<i>Scartelaos histophorus</i>	Gobiiformes / Oxudercidae	This study
<i>Periophthalmus magnuspinnatus</i>	Gobiiformes / Oxudercidae	https://www.ncbi.nlm.nih.gov/labs/data-hub/taxonomy/409849/
<i>Periophthalmus modestus</i>	Gobiiformes / Oxudercidae	http://gigadb.org/dataset/100957
<i>Boleophthalmus pectinirostris</i>	Gobiiformes / Oxudercidae	https://www.ncbi.nlm.nih.gov/labs/data-hub/taxonomy/150288/
<i>Danio rerio</i>	Cypriniformes / Cyprinidae	https://www.ncbi.nlm.nih.gov/labs/data-hub/taxonomy/7955/
<i>Lepisosteus oculatus</i>	Lepisosteiformes / Lepisosteidae	https://www.ncbi.nlm.nih.gov/labs/data-hub/taxonomy/7918/
<i>Esox lucius</i>	Esociformes / Esocidae	https://www.ncbi.nlm.nih.gov/labs/data-hub/taxonomy/8010/
<i>Cheilinus undulatus</i>	Labriiformes / Labridae	https://www.ncbi.nlm.nih.gov/data-hub/taxonomy/241271/
<i>Larimichthys crocea</i>	Acanthuriformes / Sciaenidae	https://www.ncbi.nlm.nih.gov/labs/data-hub/taxonomy/215358/
<i>Acanthopagrus latus</i>	Spariformes / Sparidae	https://www.ncbi.nlm.nih.gov/data-hub/taxonomy/8177/
<i>Thunnus maccoyii</i>	Scombriformes / Scombridae	https://www.ncbi.nlm.nih.gov/data-hub/assembly/GCF_910596095.1/
<i>Xiphias gladius</i>	Istiophoriformes / Xiphiidae	https://www.ncbi.nlm.nih.gov/labs/data-hub/taxonomy/8245/

<i>Echeneis naucrates</i>	Carangiformes / Echeneidae	https://www.ncbi.nlm.nih.gov/data-hub/taxonomy/173247/
<i>Anabas testudineus</i>	Anabantiformes / Anabantidae	https://www.ncbi.nlm.nih.gov/labs/data-hub/taxonomy/64144/
<i>Epinephelus lanceolatus</i>	Perciformes / Serranidae	https://www.ncbi.nlm.nih.gov/data-hub/taxonomy/310571/
<i>Sebastes umbrosus</i>	Perciformes / Sebastidae	https://www.ncbi.nlm.nih.gov/data-hub/assembly/GCF_015220745.1/
<i>Trematomus bernacchii</i>	Perciformes / Nototheniidae	https://www.ncbi.nlm.nih.gov/data-hub/taxonomy/40690/
<i>Salarias fasciatus</i>	Blenniiformes / Blenniidae	https://www.ncbi.nlm.nih.gov/data-hub/taxonomy/181472/
<i>Oreochromis niloticus</i>	Cichliformes / Cichlidae	https://www.ncbi.nlm.nih.gov/labs/data-hub/taxonomy/8128/
<i>Sphaeramia orbicularis</i>	Kurtiformes / Apogonidae	https://www.ncbi.nlm.nih.gov/data-hub/taxonomy/375764/

Supplementary Table 11. A list of software and parameters used for this study.

Softwares	Version	Parameters/Commands
Genome survey:		
kmerfreq	4.0	-k 17
GCE	1.0.2	-g 44350933428 -f *.freq.stat.2column
Genome assembly:		
hifiasm	0.16.1-r375	default
Juicer	1.6	-s MboI
3D-DNA	180922	default
Genome annotation:		
LTR_FINDER	v1.07	-D 20000 -d 1000 -L 700 -l 100 -p 20 -C -M 0.9
RepeatModeler	2.0.3	default
RepeatMasker	4.1.2-p1	-e rmbblast -pa 20 -lib Repeat.fasta -html -gff
Genewise	wise2-4-1	-both -sum -genesf -gff
Trinity	v2.13.2	--genome_guided_max_intron 10000
PASA	v2.5.2	-C -R --ALIGNER gmap
TransDecoder	v5.5.0	default
GeneMark	4	default
Augustus	3.4.0	gff2gbSmallDNA.pl 1000/etraining -- stopCodonExcludedFromCDS=false/randomSplit.pl 500/optimize_augustus.pl --rounds=3/Other processes use default parameters --genome genome.fasta --weights `pwd`/weights.txt -- gene_predictions gene_predictions.gff3 -- protein_alignments protein_alignments.gff3 -- transcript_alignments transcript_alignments.gff3 -- output_file_name evm.out --partitions partitions_list.out
EVidenceModeler	v1.1.1	-appl Pfam -f TSV -i pep/genome.pep.fa -cpu 8 -b out -goterms -iprlookup -pa
InterProScan	5.55-88.0	
Comparative Phylogenomics:		
OrthoFinder	2.5.4	-M msa -S diamond -T fasttree -t 20
ProtTest	3.4.2	-all-distributions -F -DT -tc 0.5 -threads 20 -T 60 -f a -x 12345 -p 12345 -N 100 -m PROTGAMMAIVTF -o Lepisosteus_oculatus.pep -k
RAxML	8.2.12	
PAML	4.9f	default
CAFÉ	4.0.2	-p 0.05 lambda -s -t