Parallel evolution in Crassostrea oysters along the latitudinal gradient is associated with variation in multiple genes involved in adipogenesis

Article *in* Molecular Ecology · August 2023

DOI: 10.1111/mec.17108

9 authors, including:

Wen Teng
Ocean University of China
5 PUBLICATIONS **13** CITATIONS
SEE PROFILE

Huiru Fu
Ocean University of China
15 PUBLICATIONS **112** CITATIONS
SEE PROFILE

Zhuanzhuan Li
Ocean University of China
10 PUBLICATIONS **50** CITATIONS
SEE PROFILE

Chengxun Xu
Ocean University of China
56 PUBLICATIONS **282** CITATIONS
SEE PROFILE

Received: 8 May 2023 | Revised: 5 August 2023 | Accepted: 10 August 2023

DOI: 10.1111/mec.17108

ORIGINAL ARTICLE

MOLECULAR ECOLOGY WILEY

Parallel evolution in *Crassostrea* oysters along the latitudinal gradient is associated with variation in multiple genes involved in adipogenesis

Wen Teng[1] | Huiru Fu[1] | Zhuanzhuan Li[1] | Qian Zhang[2] | Chengxun Xu[1] | Hong Yu[1] | Lingfeng Kong[1] | Shikai Liu[1] | Qi Li[1,3]

[1]Key Laboratory of Mariculture, Ministry of Education, Ocean University of China, Qingdao, China

[2]Public Technology Service Center, Institute of Oceanology, Chinese Academy of Sciences, Qingdao, China

[3]Laboratory for Marine Fisheries Science and Food Production Processes, Qingdao National Laboratory for Marine Science and Technology, Qingdao, China

Correspondence
Qi Li, Key Laboratory of Mariculture, Ministry of Education, Ocean University of China, Qingdao 266003, China.
Email: qili66@ouc.edu.cn

Funding information
China Agriculture Research System Project, Grant/Award Number: CARS-49; Earmarked Fund for Agriculture Seed Improvement Project of Shandong Province, Grant/Award Number: 2021ZLGX03 and 2022LZGCQY010; the National Key R&D Program of China, Grant/Award Number: 2022YFD2400305

Handling Editor: J. A. H. Benzie

Abstract

Parallel diversification provides a proper framework for studying the role of natural selection in evolution. Yet, empirical studies from ecological 'non-model' species of invertebrates are limited at the whole genome level. Here, we presented a chromosome-scale genome assembly for *Crassostrea angulata* and investigated the parallel genomic evolution in oysters. Specifically, we used population genomics approaches to compare two southern–northern oyster species pairs (*C. angulata*–*C. gigas* and southern–northern *C. ariakensis*) along the coast of China. The estimated divergence time of *C. angulata* and *C. gigas* is earlier than that of southern and northern *C. ariakensis*, which aligns with the overall elevated genome-wide divergence. However, the southern–northern *C. ariakensis* F_{ST} profile represented more extremely divergent "islands". Combined with recent reciprocal hybridization studies, we proposed that they are currently at an early stage of speciation. These two southern–northern oyster species pairs exhibited significant repeatability in patterns of genome-wide differentiation, especially in genomic regions with extremely high and low divergence. This suggested that divergent and purifying selection has contributed to the genomic parallelism between southern and northern latitudes. Top differentiated genomic regions shared in these two oyster species pairs contained candidate genes enriched for functions in energy metabolism, especially adipogenesis, which are closely related to reproductive behaviours. These genes might be good candidates for further investigation in vivo. In conclusion, our results suggest that similar divergent selection and shared genomic features could predictably transform standing genetic variation within one species pair into differences in another.

KEYWORDS
adaptation, invertebrates, molluscs, population ecology, population genetics, speciation

1 | INTRODUCTION

It has been suggested that most species appear to evolve by natural selection ever since Darwin (Coyne & Orr, 2004; Schluter, 2009). Closely related populations or species in similar environments are expected to have parallel genetic bases to parallel phenotypes. This supplies a proper framework to infer the deterministic role natural selection plays (Elmer & Meyer, 2011). Candidate gene studies have provided excellent examples of parallel animal genotype and phenotype changes. For instance, frequent mutations targeting the melanocortin-1 receptor (*Mc1r*) gene underlie brown phenotypes in geographically separate cavefish (Gross et al., 2009). Multiple

marine stickleback populations showed independent deletions in a *Pitx1* enhancer that were found to be associated with the parallel reduction in the pelvic girdle (Chan et al., 2010). A derived allele of the *Ectodysplasin* locus has been fixed repeatedly in the parallel evolution of stickleback low-plated phenotypes at most freshwater locations around the world (Barrett et al., 2008; Colosimo et al., 2005). However, the candidate gene approach only focused on a single locus and inevitably introduced the ascertainment bias, which may exaggerate the genetic parallelism (Elmer & Meyer, 2011). The genome-wide approaches overcome that limitation and help to assess the entire genome to identify many genetic bases for phenotypic parallelisms (Elmer & Meyer, 2011). For example, marine threespine sticklebacks have independently colonized many freshwater habitats and showed globally shared marine–freshwater divergence, including 0.5% of its genome regions and three chromosomal inversions, underlying parallel armour, and body shape evolution (Jones et al., 2012). However, most of the experimental and comparative work has been conducted between pairs of populations or races rather than between pairs of different species.

Many vicariant events in marine animals between southern and northern latitudes along the coast of China have been associated with Pleistocene sea level changes (Ni et al., 2014; Wang, 1999). Recent population genomic studies have integrated diverse selection measures, such as the fixation index (F_{ST}) and genetic diversity ratio ($\theta\pi$-Ratio), to explore the divergence between southern and northern shellfish (Li et al., 2018; Li, Li et al., 2021; Wu et al., 2022). However, these studies have neglected to investigate parallel genomic differentiation by utilizing multiple species pairs. This approach may offer valuable insights into the adaptive divergence between southern and northern latitudes.

Oysters are one of the widely distributed marine bivalves, harbouring more than 20 species in China seas (Guo et al., 2018; Yang et al., 2021). Notably, *C. angulata* and *C. gigas* are naturally distributed in the southern and northern waters, respectively (Figure 1). They presented similar morphology but different thermotolerance and reproductive cycles (Ghaffari et al., 2019; Shi et al., 2019). Estuarine oyster *C. ariakensis* is another native distributed in the estuarine, intertidal environment along the coastline of China (Figure 1). The southern and northern *C. ariakensis* represented distinct genome-wide variations (Li, Li et al., 2021; Wu et al., 2022). Physiology experiments also revealed that the southern *C. ariakensis* had evolved a higher thermotolerance than the northern *C. ariakensis* (Li et al., 2020). Interpopulation hybrids showed no severe reproductive isolation but a significantly low hatching rate between southern and northern populations (Qin et al., 2022). These results implied that they have evolved into different species. Given all that, these two oyster groups (*C. angulata*–*C. gigas* and southern–northern *C. ariakensis*) diverged along a similar latitudinal gradient and appeared to have adapted to similar changes in temperature. They would serve as a proper framework to study parallel genomics.

In this study, we begin by constructing a high-quality genome for *C. angulata*. Subsequently, we integrate whole-genome resequencing data from *C. angulata*, *C. gigas*, and southern and northern *C. ariakensis* to explore the genomic patterns of divergence and parallel speciation. Our findings provide valuable insights into the genomics of vicariant speciation in *Crassostrea* oysters, contributing to a more comprehensive understanding of this process.

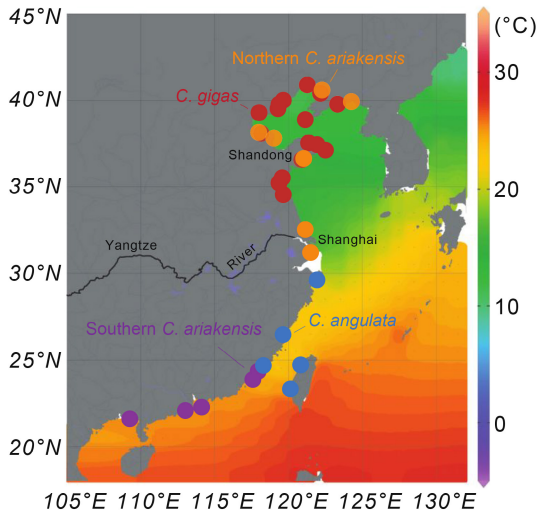

FIGURE 1 The oyster species range (modified and redrawn from various sources; Li et al., 2018; Li, Dai et al., 2021, Li, Li et al., 2021; Vaschenko et al., 2013; Wang, 2004; Wu et al., 2022). Blue dots represent *Crassostrea angulata*; Red dots represent *Crassostrea gigas*; Purple dots represent southern *Crassostrea ariakensis*; Orange dots represent northern *Crassostrea ariakensis*. The annual SST data for the year 2018 was downloaded from Ocean Data View (Schlitzer, 2018).

2 | MATERIALS AND METHODS

2.1 | Genome assembly and annotation

A wild female Fujian oyster (Portuguese oyster), *C. angulata*, was sampled from Ningde, China. Genomic DNA was extracted from adductor muscle using the TIANamp Marine Animals DNA Kit. Total RNA was extracted from the adductor muscle, mantle, gill, and digestive gland using the Trizol reagent (Thermo Fisher Scientific). Subsequently, an Iso-Seq library was constructed from the pooled RNA using the SMRTbell prep kit 3.0 for sequencing on the PacBio Sequel II System. To estimate the genome size and complexity, we constructed a 150 bp paired-end library using the NEBNext Ultra DNA Library Prep Kit and subsequently sequenced it on the Illumina Hiseq Xten platform. Raw reads were filtered using fastp v0.20.1 (Chen et al., 2018). Bases with a mean quality lower than 20 in the sliding window were removed. Additionally, a minimum read length of 36 was required, and five bases at the front of both read1 and read2 were also discarded. Genome size and complexity were estimated with Jellyfish (Marçais & Kingsford, 2011) and GenomeScope (Vurture et al., 2017). To perform a chromosome-level genome assembly, we first constructed a SMRTbell library using the SMRTbell prep kit 3.0 and sequenced it on the PacBio Sequel II System.

Additionally, a Hi-C library was prepared using the Dovetail™ Hi-C Library Preparation Kit and sequenced on the Illumina Hiseq Xten platform. The primary genome was initially performed with hifiasm v0.16 using default parameters (Cheng et al., 2021). Subsequently, the haploid sub-assembly was reconstructed using HaploMerger2 (Huang et al., 2017). Finally, the chromosome-level assembly for *C. angulata* was accomplished using Juicer v1.6 with default parameters (Durand et al., 2016). We accessed the genome assembly completeness using BUSCO v5.1.3 (Simão et al., 2015) with the "mollusca_odb10" BUSCO gene set collection (Waterhouse et al., 2018).

Transposable elements de novo prediction and homology searching were performed by RepeatModeler v2.0.3 and RepeatMasker v4.1.1 pipeline (Chen, 2004). Alignment assembly was first completed using transcriptome data by PASA pipeline v2.5.1 (Haas et al., 2003). The genome was also annotated by MAKER v3.01.03 (Cantarel et al., 2008). Specifically, the initial MAKER analysis employed the empirical transcript of *C. angulata* and protein models from bivalves, including *C. gigas*, *Crassostrea virginica*, *Saccostrea glomerata*, *Pinctada fucata*, and *Mizuhopecten yessoensis*. Gene models were trained by AUGUSTUS v 3.4.0 (Stanke et al., 2008) and SNAP (Johnson et al., 2008). We finally ran three rounds of MAKER annotation with 92.2% of the gene models having an annotation edit distance (AED) of 0.5, and the returns started diminishing thereafter. Predicted genes were functionally annotated with Nr, Uniport, Swiss-Prot, and COG databases using BLAST (Camacho et al., 2009) with an E-value cutoff of 1e-10.

2.2 | Phylogenetic analysis of southern–northern oyster species

Crassostrea angulata settled in southern China with an annual average sea surface temperature (SST) above 20°C (Figure 1). Its closest species, *C. gigas*, settled in northern China with an annual average SST lower than 20°C (Figure 1). The estuarine oysters, *C. ariakensis*, settled in southern and northern China, including the Yangtze River mouth (Figure 1). Here, whole-genome sequencing (WGS) data of *C. angulata* (21 samples) and *C. gigas* (472 samples) were downloaded from the Sequence Read Archive (SRA) database under accession numbers PRJNA394055 (Li et al., 2018). WGS data of *C. ariakensis* (69 southern samples and 192 northern samples) was downloaded from the SRA database under accession numbers PRJNA715058 (Wu et al., 2022).

Since the southern *C. ariakensis* genome assembly is unavailable, we utilized mitochondrial DNA sequences to estimate the divergence time, commonly employed in oyster molecular phylogenetics (Guo et al., 2018). Specifically, WGS reads of 69 southern and 192 northern *C. ariakensis* individuals were mapped to *C. ariakensis* genome assembly with accession GCA_020458035.1 (Wu et al., 2022), comprising only 10 primary super scaffolds. Unmapped reads were extracted for de novo assembly by SPAdes v3.15.4 (Prjibelski et al., 2020). Then, a BLASTN-algorithm (Camacho et al., 2009) based search using a cytochrome c oxidase subunit 1 (COX1) nucleotide sequences, downloaded from NCBI with accession number FJ841964.1, as a query was conducted. The aligned subject sequence was finally extracted and translated into protein sequences using our custom Python scripts (https://github.com/tengwen2018/Parallel_evol_oysters_Asia/blob/main/Part2.Phylogenetic_analysis/cds2prot.py). Similarly, WGS data of 21 *C. angulata* and 472 *C. gigas* were mapped to our *C. angulata* reference genome, which comprises only 10 primary scaffolds. Unmapped reads were extracted for the de novo assembly. The COX1 nucleotide sequence, downloaded from NCBI with accession number NC_012648.1, was used as a query to extract the COX1 sequence for each sample. Next, multiple alignments were performed in each southern and northern oyster species pair with MUSCLE v3.8.31 (Edgar, 2004). Pairwise Kimura 2-parameter (K2P) distances were calculated with the DNADIST program (Felsenstein, 1988). Two samples from *C. angulata* and *C. gigas* with median K2P values were randomly selected. Two samples from northern and southern *C. ariakensis* with median K2P values were chosen randomly. One COX1 gene sequence of *C. virginica* was downloaded from NCBI with accession number MN817966.1 and served as an outgroup. To construct phylogenetic relationships, COX1 gene sequences from all five species were used for multiple alignment analysis with MUSCLE v3.8.31 (Edgar, 2004). Then the phylogenetic tree was estimated using RaxML v.8.2.12 (Stamatakis, 2014). Finally, the divergence time was estimated using MCMCTree from PAML v.4.9 (Yang, 2007) with main parameters of 'RootAge=<90 model=HKY85 alpha=0.969 clock=2' and calibration time between *C. angulata* and *C. virginica*: minimum = 63 Ma and soft maximum = 83 Ma.

The estimation of the divergence time was also performed using SNAPP (SNP and AFLP Package for Phylogenetic analysis; Bryant et al., 2012) with SNP data. Specifically, one sample per species was employed, including *C. angulata*, *C. gigas*, southern *C. ariakensis*, northern *C. ariakensis*, and *C. virginica* (Puritz et al., 2023). The calibration divergence of *C. virginica* and *C. angulata* was estimated at 73 Ma (Plazzi & Passamonti, 2010; Ren et al., 2010).

2.3 | Genome-wide divergence of southern–northern oyster species

Variation calling was performed on 21 *C. angulata* and 472 *C. gigas* WGS data to study the genome-wide divergence using our *C. angulata* reference genome. Variation calling of *C. ariakensis* (69 southern samples and 192 northern samples) was performed with the *C. ariakensis* reference genome with the GenBank assembly accession number GCA_020458035.1 (Wu et al., 2022).

Specifically, WGS raw reads were filtered and trimmed using fastp v0.20.1 (Chen et al., 2018). Bases with a mean quality lower than 20 in the sliding window were removed. Additionally, a minimum read length of 36 was required, and five bases at the front of both read1 and read2 were also discarded. Clean reads were mapped to genome reference using BWA v0.7.17 with default parameters (Li & Durbin, 2009). SNP calling was performed using VCFtools v0.1.16 (Danecek et al., 2011).

SNPs retained in the following analysis were filtered by the following criterion: called in more than 90% of samples; with minor allele frequency ≥ 0.1; a quality score of more than 30; and a mean depth of more than 10 but less than 50. To clarify the population structure, principal components analysis (PCA) was performed using PLINK v1.9 with the main parameter (–indep-pairwise 50 10 0.1; Purcell et al., 2007). And individual ancestries were estimated using ADMIXTURE (Alexander et al., 2009). Population admixture graphs within *C. ariakensis* were inferred using TreeMix (Pickrell & Pritchard, 2012).

To exclude the effect of sampling size in population genomic analysis, we randomly selected 20 specimens of *C. angulata* and 20 specimens of *C. gigas* to calculate F_{ST}, repeating this process 100 times. Similarly, we randomly sampled 20 southern *C. ariakensis* specimens and 20 northern *C. ariakensis* specimens to calculate F_{ST}, repeating the procedure 100 times. Specifically, per-site F_{ST} and sliding-window F_{ST} were calculated using VCFtools v0.1.16 (Danecek et al., 2011) and custom scripts written by Simon Martin (Martin, 2022), respectively. Sliding window F_{ST} was measured with a window size of 2 kb that was sliding by 2 kb and requested these windows to have a minimum number of 10 sites covered. To evaluate the role of linkage disequilibrium in a selective sweep, the integrated haplotype homozygosity score (iHS; Voight et al., 2006) was calculated within these two southern–northern oyster pairs using *rehh* version 3.2.2 (Gautier & Vitalis, 2012). Meanwhile, we estimated the synonymous substitution rate (d_s) as a proxy for the mutation rate (Baer et al., 2007) using SNPGenie (Nelson et al., 2015) with whole-genome variants for each species's sample.

2.4 | Repeatability in genomic patterns: Quantile approach

To investigate the parallel genomic evolution along the similar latitudinal gradient, we first selected two southern–northern oyster species pairs as the foreground. Specifically, we utilized WGS data from 21 *C. angulata*, 21 *C. gigas*, 69 southern *C. ariakensis*, and 69 northern *C. ariakensis*, and mapped them to our *C. angulata* genome assembly. F_{ST} values were calculated using sliding window (2 kb windows) as described earlier. Subsequently, we calculated the Pearson correlation coefficient of F_{ST} values between the *C. angulata*–*C. gigas* pair and the southern–northern *C. ariakensis* pair. Secondly, we selected two subgroups of the northern *C. ariakensis* as the background. We named one of the subgroups of northern *C. ariakensis* Shandong *C. ariakensis* and the other one Shanghai *C. ariakensis*. They were distributed in and above the northern regions of the Yangtze River mouth and experienced similar temperature conditions (Figure 1). Specifically, WGS data of 29 Shandong *C. ariakensis* and 29 Shanghai *C. ariakensis* were mapped to our *C. angulata* genome, and F_{ST} was calculated within 2 kb windows, subsequently. Subsequently, we calculated the Pearson correlation coefficient of F_{ST} values between the *C. angulata*–*C. gigas* pair and the Shandong–Shanghai *C. ariakensis* pair.

The quantile approach was performed as described in the literature (Renaut et al., 2014). Specifically, genomic regions were ranked according to F_{ST} for each species pair and split into 20 quantiles. The first quantile comprises the most divergent regions, while the 20th quantile includes the least divergent regions. Next, observed shared windows per quantile were compared using BEDTools v2.29.2 (Quinlan & Hall, 2010) between oyster species pairs. Expected shared regions per quantile were calculated with the formula of 5% × 5% × the total number of polymorphic genomic regions. Genes closest to the shared genomic regions, along with a 3 kb region upstream and downstream, within the first quantile, were utilized for a Gene Ontology (GO) analysis using the clusterProfiler R package (Yu et al., 2012).

3 | RESULTS

3.1 | Genome assembly of *C. angulata* and annotation

Three sequencing and assembly technologies were integrated to perform the de novo assembly. Firstly, the genome of *C. angulata* was sequenced with 64-fold coverage of 150 bp Illumina paired-end sequencing reads to estimate the genome size and complexity. The *C. angulata* genome has a high level of heterozygosity (2.6%), and the genome size was estimated at 548 megabases (Mb). Then, 76.3-fold coverage of HiFi reads from the Pacific Biosciences Sequel II system was assembled into 372 contigs with a contig N50 of 12 Mb. Lastly, 97.7-fold coverage of the Hi-C library was employed to help generate a chromosome-level assembly, and the final genome assembly of *C. angulata* comprised 624 Mb with scaffold N50 of 60 Mb (Figure S1a). Genome assembly completeness and sequence accuracy assessment of the genome in this study showed that >99.0% of the 5295 genes were complete, and 98.1% occurred as single-copy genes.

Repeat elements constituted 49.67% of the *C. angulata* genome and were distributed unevenly along the chromosomes (Figure S1b), where DNA transposons (19.04%) and rolling-circle transposons (*Helitrons*) (13.2%) predominate. In contrast, retroelements [including long interspersed nuclear element (LINE), long terminal repeat (LTR), and short interspersed nuclear element (SINE)] constitute a minor portion of the genome (5.81%) (Table S1). Gene annotation was performed based on a strategy combining homology searches, de novo gene prediction, and isoform sequencing (Iso-Seq). A total of 29,608 protein-coding gene models were predicted (Table S2).

3.2 | Phylogenetic divergence of southern–northern oyster species pairs

Based on the mitochondrial COX1 sequence, the *C. angulata*–*C. gigas* species pair has a divergence of 1.9%–2.7% with a mean value of 2.3% [measured by Kimura 2-parameter (K2P) distances; Figure S2a]. However, the average K2P distance of the *C. ariakensis* south–north

population pair is 0%–2.4% with a mean value of 0.6% (Figure S2b). The point of divergence between the *C. gigas*–*C. angulata* pair and the southern–northern *C. ariakensis* pair was estimated to be 29 Mya (Figure 2). The divergence time estimation between *C. gigas* and *C. angulata* is 3.6 Mya, earlier than that of 0.98 Mya between southern and northern *C. ariakensis* (Figure 2). The estimation of divergence time between these two southern–northern oyster species pairs using SNP data was much shorter than that using the COX1 sequences (Figure S3). Despite this, the divergence of southern and northern *C. ariakensis* was still later than that of *C. gigas* and *C. angulata*.

3.3 | Genome-wide islands of speciation of southern–northern oyster species

Principal component analysis (PCA) revealed a pronounced population structure between southern and northern oyster species (Tracy-Widom statistics: $p < 1 \times 10$–200; Tracy & Widom, 1994). In the southern–northern *C. ariakensis* pair, southern and northern individuals were clearly separated by the first eigenvector (Figure 3a). Similarly, individuals from *C. angulata* and *C. gigas* formed distinct clusters (Figure 3b). ADMIXTURE analysis showed genetic clusters are similar to the result of PCA analysis (Figure S4). Considering the *C. angulata* and *C. gigas* pairs, the ADMIXTURE cross-validation approach chose the K parameter of 3 or 4 with the lowest error (Figure S4b). There is no admixture between *C. angulata* and *C. gigas* in these situations (Figure S4b). However, recent hybrids seem to be ubiquitous within *C. gigas* (Figure S4a). Considering the southern–northern *C. ariakensis* pair, the ADMIXTURE cross-validation approach chose the *K* parameter of 3 with the lowest error (Figure S4d). The northern *C. ariakensis* has diverged not only from the southern *C. ariakensis* but has also divided into two distinct subgroups (Figure S4c). One includes samples from Shanghai. The other includes samples from Shandong (Figure 1). A little gene flow from the southern *C. ariakensis* into the Shanghai *C. ariakensis* is observed when $K = 2$. This may be due to an admixture event in history (Figure S5).

Genome-wide F_{ST} profiles clearly display demarcated "islands" of high differentiation in both the south–north oyster species pairs (Figure 3c and Table S3). Comparisons between *C. angulata* and southern *C. ariakensis* and between *C. gigas* and northern *C. ariakensis* revealed consistently high F_{ST} patterns (Figure S6). After performing 100 repetitions of F_{ST} calculations with randomly selected samples, we observed that the standard deviations (SD) of the median, the 95th, and the 99th percentiles of F_{ST} were less than 0.003. Based on this analysis, the median F_{ST} in the *C. angulata*–*C. gigas* pair is 0.06, higher than that of 0.04 in the southern–northern *C. ariakensis* pair (Table S3; $p < .05$). However, the 95th and 99th percentiles of F_{ST} in the *C. angulata*–*C. gigas* pair are 0.28 and 0.44, lower than the 95th and 99th percentiles of F_{ST} of 0.48 and 0.72 in the southern–northern *C. ariakensis* pair (Table S3; $p < .05$). For single sites, the maximum F_{ST} value of the *C. angulata*–*C. gigas* species pair is 5.4 SD above the mean, and 49,495 sites have an F_{ST} over 3 SD above the mean. The maximum F_{ST} of the southern–northern *C. ariakensis* pair is 4 SD above the mean, and an even greater number of sites, 131,082, have an F_{ST} over 3 SD above the mean ($p < .05$). Overall, the southern–northern *C. ariakensis* pair exhibits fewer genomic regions with moderate F_{ST} but a greater number of genomic regions with extremely low and high F_{ST} values compared to the *C. angulata*–*C. gigas* species pair (Figure 3d). Besides, SNPs with high iHS in one species but low iHS in another are likely to show elevated F_{ST} in both the southern–northern oyster species pairs (Figure S7). Furthermore, the rate of synonymous mutation per synonymous site (d_s) in *C. ariakensis* species is 0.008, which is higher than that of *C. angulata* or *C. gigas*, which is 0.006 ($p < .05$). In the southern–northern *C. ariakensis* species pair, the Pearson correlation coefficient between d_s and F_{ST} is 0.2 ($p < .05$). In the *C. angulata*–*C. gigas* species pairs, the Pearson correlation coefficient between d_s and F_{ST} is 0.055 ($p < .05$; Figure S8).

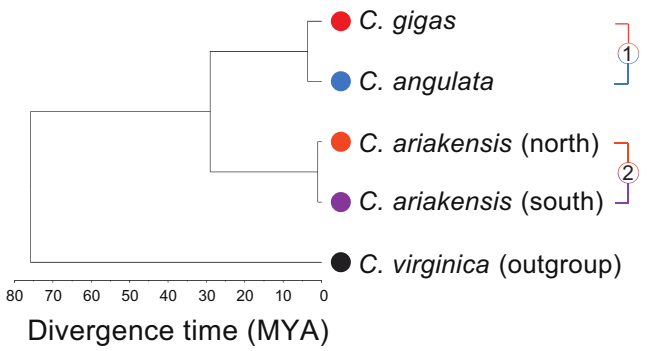

FIGURE 2 A phylogenetic relationship of five *Crassostrea* oyster species. Divergence time was estimated with a relaxed molecular clock Bayesian method. *Crassostrea virginica* served as an outgroup.

3.4 | Repeatability in genomic patterns and genes in response to parallel selection

Because of the genome-wide differentiation and inhabiting the northern waters with similar temperatures, the Shandong–Shanghai *C. ariakensis* pair served as a background to investigate parallel genomic divergence between southern and northern oyster species. Based on strict criteria, 2,847,615, 1,219,101, and 1,086,786 SNPs were detected for *C. angulata*–*C. gigas*, southern–northern *C. ariakensis*, and Shandong–Shanghai *C. ariakensis*, respectively. We observed the Pearson's correlation coefficient of 0.11 ($p < 2.2e\text{-}16$) for *C. angulata*–*C. giga* versus Shandong–Shanghai *C. ariakensis* (Figure S9) and 0.25 ($p < 2.2e\text{-}16$) for *C. angulata*–*C. giga* versus southern–northern *C. ariakensis* (Figure 4a). Genomic regions with extremely high and low average F_{ST} estimates were more likely to be shared than genomic regions with modest average F_{ST} estimates (Figure 4b). Significant quadratic relationships between the 20 quantiles and the ratio of observed versus expected shared genomic regions were identified in *C. angulata*–*C. gigas* versus southern–northern *C. ariakensis* (Figure 4b, $p = 7.8e\text{-}07$). Additionally, the most divergent regions exhibited high iHS in the southern species but low iHS in the northern species (Figure S10). Sixty-eight genes closest to these most divergent shared regions (top

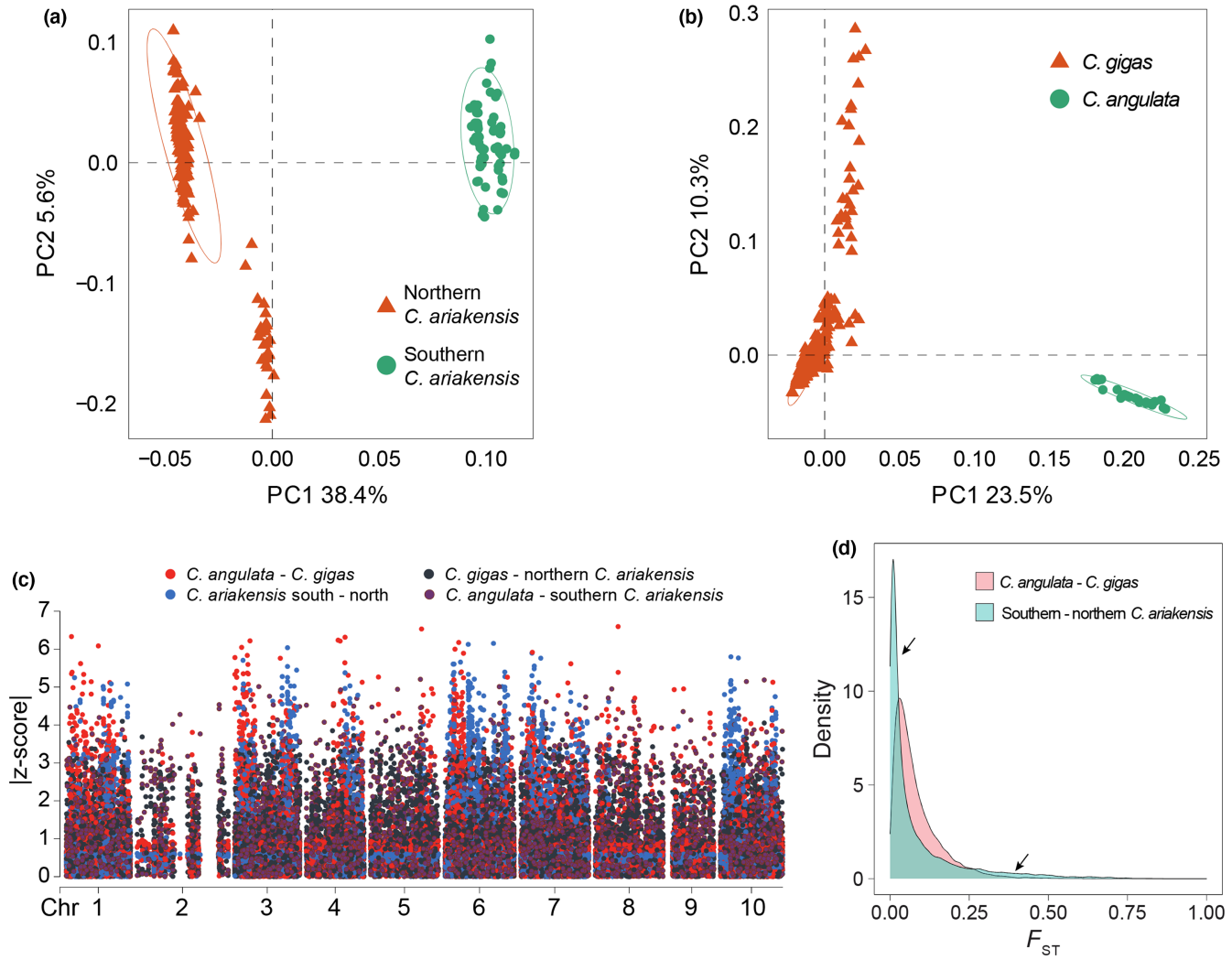

FIGURE 3 Genome-wide divergence of southern and northern oysters (a) PCA of genetic variation within the southern–northern *C. ariakensis* pair. (b) PCA of genetic variation within the *C. angulata*–*C. gigas* species pair. (c) Absolute standard scores of southern–northern *C. ariakensis* divergence (blue) and *C. angulata*–*C. gigas* divergence (red) overlaid on the divergence between southern and northern *C. angulata* (purple), and northern *C. ariakensis* and *C. gigas* (black). (d) Density plot of sliding-window-based F_{ST} for *C. angulata*–*C. gigas* (pink) and southern–northern *C. ariakensis* (cyan) divergence. The *C. angulata*–*C. gigas* species pair has more genomic regions with moderate F_{ST} values, whereas the southern–northern *C. ariakensis* exhibits a greater number of genomic regions with extremely low and high F_{ST} values, which are highlighted with black arrows.

5%) were utilized for a GO analysis. Among these, 21 genes were significantly enriched in 21 GO terms ($Q < 0.05$; Figure 4c and Table S4). And most of these enriched terms were related to four clusters: 'glycolytic process', 'lipid metabolic process', 'fat cell differentiation', and 'circadian rhythm' (Figure 4c and Table S4).

4 | DISCUSSION

Crassostrea oysters constitute a collection of recent divergence and vicariant speciation events along the coast of China (Guo et al., 2018). Among these events, two southern–northern species pairs (*C. angulata*–*C. gigas* and southern–northern *C. ariakensis*) exhibit local adaptation to a similar climate gradient. The speciation status remains controversial, especially in the southern–northern *C. ariakensis* pair. Moreover, the availability of large-scale whole-genome sequencing data and genome scans permits the investigation of the parallel genomic changes underlying local adaptation to southern and northern latitudes. In this study, we first present a viewpoint that suggests the southern and northern *C. ariakensis* are in an early stage of speciation. Secondly, we demonstrate the genomic parallelism and its potential causes between the two southern–northern oyster species pairs.

4.1 | The early stage of allopatric speciation in the southern–northern oyster species

The Pacific oyster *C. gigas* and Fujian oyster *C. angulata* are native species in northern and southern China's coastal waters (Figure 1). They used to be considered the same species (Huvet et al., 2002;

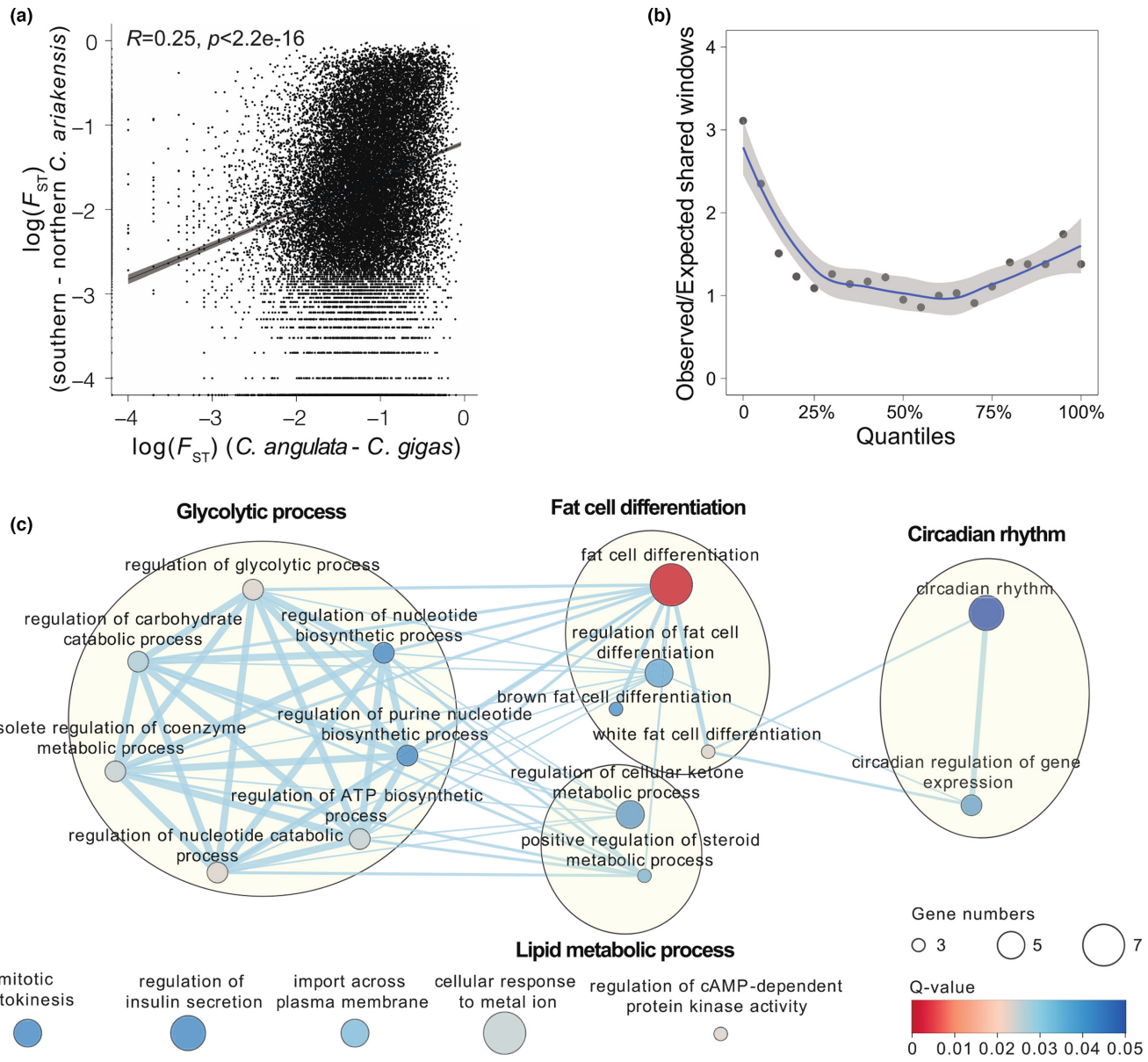

FIGURE 4 Parallel genomic divergence. (a) Scatterplot displaying correlation of genomic divergence between the *C. angulata*–*C. gigas* species pair and the southern–northern *C. ariakensis* species pair. (b) Genomic repeatability patterns quantified with respect to F_{ST} at the genomic level (2 kb windows). The left end part of the quadratic relationship represents those genomic regions with extremely high divergence in two southern–northern oyster species pairs that are likely to be shared. The right end of the quadratic relationship represents those genomic regions with extremely low divergence in these two southern–northern oyster species pairs that are also likely to be shared. (c) Enrichment Map for significantly enriched GO terms ($Q < 0.05$). The level of overlap between GO enriched terms is indicated by the thickness of the edge between them. The size of the node indicates the gene numbers for the term, and the colour of the node indicates the enrichment *Q*-value of the Go term.

Menzel, 1974). However, the differing physiological characteristics (Ghaffari et al., 2019; Haure et al., 2003), the mitochondrial DNA differentiation (Figure S2a), the genome-wide differentiation (Figure 3b and Figure S4a), and ecological reproductive isolation (Li, Dai et al., 2021) tend to divide they into two sister species (Guo et al., 2018). The Suminoe oyster, *C. ariakensis*, is another native occurring along China's coast (Figure 1). Estimation of divergence time revealed that the southern and northern *C. ariakensis* pair diverged at a more recent date compared to the *C. gigas* and *C. angulata* pair (Figure 2 and Figure S3). Their taxonomic status was previously clarified as the same species (Li et al., 2020). However, reciprocal transplant experiments have demonstrated ecological reproductive isolation between the southern and northern *C. ariakensis* (Li et al., 2020). Moreover, the *C. angulata*–*C. gigas* species pair showed no significant differences ($p > .05$) in the fertilization rate and hatching rate between interspecific hybrids and intraspecific inbreeds (Huvet et al., 2002; Jiang et al., 2021). In contrast, the fertilization and hatching rate of interspecific hybrids is significantly lower

($p < .05$) than that of intraspecific inbreeds in the southern–northern *C. ariakensis* pair (Qin et al., 2022). To some extent, these results indicate that the southern and northern *C. ariakensis* have undergone speciation accompanied by the development of intrinsic postzygotic barriers.

In this study, genome scans revealed "islands" of differentiation across the genome in southern - northern oyster species pairs (Figure 3c and Figure S4), like many other observed cases (Hohenlohe et al., 2010; Jones et al., 2012; Lawniczak et al., 2010; Malinsky et al., 2015). Interestingly, our data uncovered that the *C. angulata–C. gigas* pair exhibits more genomic regions with moderate F_{ST} values, while the southern–northern *C. ariakensis* pair exhibits a greater number of genomic regions with extremely low and high F_{ST} values (Figure 3d and Table S3). According to the genome hitchhiking (GH) theory, genomic islands under selection from the start of the speciation process remained highly differentiated, whereas other putatively neutral ones have steadily increased in divergence (Feder et al., 2012). The higher occurrence of genomic regions with moderate differentiation (F_{ST}) in the early differentiated *C. angulata–C. gigas* pair could be attributed to GH effects. On the other hand, the greater number of genomic regions with extremely high differentiation (F_{ST}) in the southern–northern *C. ariakensis* species pair may be influenced by the increased mutation rate and its positive correlation with genomic divergence (F_{ST}) (Figure S8). Additionally, it is essential to explore other potential contributing factors, such as divergent selection introduced by estuarine environments. Moreover, the relationship between these extremely divergent "islands" and intrinsic postzygotic barriers in the southern–northern *C. ariakensis* species pair is worth further study.

Above all, we suggested that the southern–northern *C. ariakensis* species pair may have evolved at a relatively later stage of speciation and could have a higher speciation rate than the *C. angulata–C. gigas* species pair.

4.2 | Parallel genomic divergence between southern–northern oyster species pairs

Closely related lineages in a common landscape often face similar selective pressure, which may lead to parallel phenotypic and genetic changes (Cresko et al., 2004; Elmer & Meyer, 2011; Jones et al., 2012; Pearse et al., 2014; Pearse & Pogson, 2000; Renaut et al., 2014; Waples et al., 2004). The *C. angulata–C. gigas* and southern–northern *C. ariakensis* species pairs have diverged independently along a similar latitudinal gradient (Figures 1 and 2). As such, they have presumably experienced similar selective pressure introduced by the gradient of solar energy, temperature, and nutrients (Rohde, 1992). Here, we predicted and confirmed that the pattern of repeatability was higher between the oyster species pairs that have diverged along the similar latitudinal cline than that in the background (Figure 4a and Figure S9). Moreover, the GO analyses also confirmed that biological processes related to circadian rhythm were overrepresented among the genes closest to the highly divergent shared regions in *C. angulata–C. gigas* versus southern–northern *C. ariakensis* (Figure 4c). The regions with low divergence shared by the species pairs have been interpreted as being affected by purifying selection (Renaut et al., 2014). While highly divergent regions are often associated with divergent selection, shared ancestral polymorphism or recombination rate variation could also contribute to similar patterns (Nosil et al., 2009; Renaut et al., 2013, 2014; Sodeland et al., 2016). As haplotypes are expected to be longer in regions of low recombination than in regions of high recombination, we analysed the iHS statistic between the southern–northern oyster species pairs to assess the role of linkage disequilibrium (Voight et al., 2006). Subsequently, a certain relationship between iHS and F_{ST} is observed both in the *C. angulata–C. gigas* and the southern–northern *C. ariakensis* species pairs (Figure S7). Moreover, the genomic regions showing high parallel divergence tend to have contrasting iHS values between southern and northern oyster species (Figure S10). Therefore, the patterns of genomic repeatability between the southern and northern latitudes in oysters may also be attributed to the shared recombination heterogeneity. Besides, the elevated mutation rate is also believed to play a crucial role in adaptive evolution, as it increases the supply of beneficial mutations (Losos, 2010; Raynes & Sniegowski, 2014). The southern and northern oyster species have evolved with reduced gene flow due to geographic isolation. As a result, the variability in mutation rate is expected to have influenced their divergence. Indeed, in the southern–northern *C. ariakensis* species pair, the rate of synonymous divergence (d_s), a proxy for the neutral mutation rate (Baer et al., 2007), displayed a positive correlation with the genomic divergence (F_{ST}). However, this correlation was extremely weak in the *C. angulata–C. gigas* species pairs (Figure S8). It implied that mutation rate may play a relatively minor role in the genomic parallelism, but further investigations are warranted.

Oysters are broadcast-spawning marine organisms with environment-dependent reproductive behaviours. Northern oyster species, like *C. gigas*, naturally have one reproductive cycle per year (Liu et al., 2020). While some southern oyster species, like *C. angulata*, naturally have multiple reproductive cycles per year with the prolonged spawning season and rapid gonad maturation and recovery (Shi et al., 2019). Energy metabolism in marine bivalves is closely associated with reproductive activities (Hassan et al., 2018; Li et al., 2006). In the process of oyster oogenesis, the triglyceride increases with the course of sexual maturation and has been regarded as a source of energy and a temporary polyunsaturated fatty acid reservoir (Li et al., 2000). While glycogen decreases with the course of sexual maturation, and some of it might be converted to triglyceride as a yolk lipid in the ovary (Li et al., 2000).

Here, we found that numerous genes involved in adipogenesis and glycolysis are under parallel selection in oysters between northern and southern latitudes (Figure 4c). Among these genes involved in adipogenesis, 5′-AMP-activated protein kinase catalytic subunit alpha-2 (PRKAA2) (Figure S11l), stearoyl-CoA desaturase (SCD) (Figure S11i), and elongation of very long chain fatty acids protein 5 (ELOVL5) (Figure S11d) are three key genes in lipid anabolic pathway. Specifically, PRKAA2 is the catalytic subunit of

AMP-activated protein kinase (AMPK), an energy sensor protein kinase that negatively modulates the first committed step of the fatty acid synthesis by phosphorylating acetyl-CoA carboxylase (ACC) (Davis et al., 2000; Hardie, 2007; Munday & Hemingway, 1999). SCD catalyses the rate-limiting step for converting saturated fatty acids (SFA) into monounsaturated fatty acids (MUFA) (Paton & Ntambi, 2009). ELOVL5 catalyses the first and rate-limiting reaction of the four reactions in the elongation cycle of long-chain fatty acids (Liu et al., 2013; Sassa & Kihara, 2014; Zhang et al., 2018). Among these genes under parallel selection and involved in glycolysis, PRKAA2 activates catabolic processes by phosphorylating phosphofructo-2-kinase, a rate-limiting enzyme in glycolysis (Bando et al., 2005). Midnolin (MIDN) (Figure S11i) plays a role in inhibiting the activity of glucokinase, which controls the first step of glycolysis (Hofmeister-Brix et al., 2013). Besides, other genes involved in insulin secretion and circadian rhythm indirectly regulate lipid and glucose metabolisms (Acosta-Rodriguez et al., 2021; Saltiel & Kahn, 2001). The present results suggested that natural selection and local genomic landscape tend to act on multiple genes involved in the same pathway, leading to polygenic adaptation (Barghi et al., 2020). These differentiated genes associated with cellular energy metabolism, especially adipogenesis, might play roles in the distinct reproductive behaviours adapted to southern or northern waters. These genes stand out as promising targets for further investigation in vivo.

5 | CONCLUSIONS

We integrated genomic data and previous reciprocal hybridization studies to elucidate the speciation status in two southern–northern oyster species pairs. Grounded in these observations, we provided evidence that parallel genomic changes could occur within well-differentiated species. This genomic parallelism may be attributed to similar natural selection as well as shared genomic features. Numerous genes involved in cellular energy metabolisms, especially adipogenesis, emerge as potential candidates under parallel selection. The hypothetical effects of these differentiated loci on reproductive behaviour differentiation between southern and northern oyster species are worth further investigation.

AUTHOR CONTRIBUTIONS

QL conceived the project and supervised this study. WT led the bioinformatic analyses and wrote the paper. HF and ZL prepared the materials for genome sequencing. QZ, CX, HY, LK, and SL participated in the figures' design and visualization. All authors discussed the results and commented on the manuscript. The authors read and approved the final manuscript.

ACKNOWLEDGEMENTS

We would like to express our gratitude to Tu Kang from the Putian Institute of Aquaculture Science of Fujian Province for his valuable contributions to this research. We are also grateful to the anonymous reviewers for helpful comments on an earlier version of the manuscript.

FUNDING INFORMATION

This research was supported by grants from the National Key R&D Program of China (2022YFD2400305), Earmarked Fund for Agriculture Seed Improvement Project of Shandong Province (2021ZLGX03 and 2022LZGCQY010), and China Agriculture Research System Project (CARS-49).

CONFLICT OF INTEREST STATEMENT

Authors declare no competing interest.

OPEN RESEARCH BADGES

This article has earned Open Data, Open Materials and Preregistered Research Design badges. Data, materials and the preregistered design and analysis plan are available at https://www.ncbi.nlm.nih.gov/bioproject/PRJNA857752 and https://github.com/tengwen2018/Parallel_evol_oysters_Asia.

DATA AVAILABILITY STATEMENT

This Whole Genome Shotgun project has been deposited at GenBank under the accession JANFPH000000000. The genome assembly of *C. angulata* is available at NCBI with the accession ASM2561291v2. We have uploaded all code, scripts, and outputs related to this manuscript on GitHub (https://github.com/tengwen2018/Parallel_evol_oysters_Asia).

ORCID

Huiru Fu https://orcid.org/0000-0002-3152-8743
Qi Li https://orcid.org/0000-0002-3937-9324

REFERENCES

Acosta-Rodriguez, V. A., Rijo-Ferreira, F., Green, C. B., & Takahashi, J. S. (2021). Importance of circadian timing for aging and longevity. *Nature Communications*, *12*, 2862. https://doi.org/10.1038/s41467-021-22922-6

Alexander, D. H., Novembre, J., & Lange, K. (2009). Fast model-based estimation of ancestry in unrelated individuals. *Genome Research*, *19*, 1655–1664. https://doi.org/10.1101/gr.094052.109

Baer, C. F., Miyamoto, M. M., & Denver, D. R. (2007). Mutation rate variation in multicellular eukaryotes: Causes and consequences. *Nature Reviews. Genetics*, *8*, 619–631. https://doi.org/10.1038/nrg2158

Bando, H., Atsumi, T., Nishio, T., Niwa, H., Mishima, S., Shimizu, C., Yoshioka, N., Bucala, R., & Koike, T. (2005). Phosphorylation of the 6-phosphofructo-2-kinase/fructose 2,6-bisphosphatase/PFKFB3 family of glycolytic regulators in human cancer. *Clinical Cancer Research*, *11*, 5784–5792. https://doi.org/10.1158/1078-0432.Ccr-05-0149

Barghi, N., Hermisson, J., & Schloetterer, C. (2020). Polygenic adaptation: A unifying framework to understand positive selection. *Nature Reviews Genetics*, *21*, 769–781. https://doi.org/10.1038/s41576-020-0250-z

Barrett, R. D. H., Rogers, S. M., & Schluter, D. (2008). Natural selection on a major armor gene in threespine stickleback. *Science*, *322*, 255–257. https://doi.org/10.1126/science.1159978

Bryant, D., Bouckaert, R., Felsenstein, J., Rosenberg, N. A., & RoyChoudhury, A. (2012). Inferring species trees directly from biallelic genetic markers: Bypassing gene trees in a full coalescent analysis. *Molecular Biology and Evolution*, *29*, 1917–1932. https://doi.org/10.1093/molbev/mss086

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., & Madden, T. L. (2009). BLAST+: Architecture and applications. *BMC Bioinformatics*, *10*, 1–9. https://doi.org/10.1186/1471-2105-10-421

Cantarel, B. L., Korf, I., Robb, S. M., Parra, G., Ross, E., Moore, B., Holt, C., Sánchez Alvarado, A., & Yandell, M. (2008). MAKER: An easy-to-use annotation pipeline designed for emerging model organism genomes. *Genome Research*, *18*, 188–196. https://doi.org/10.1101/gr.6743907

Chan, Y. F., Marks, M. E., Jones, F. C., Villarreal, G., Jr., Shapiro, M. D., Brady, S. D., Southwick, A. M., Absher, D. M., Grimwood, J., Schmutz, J., Myers, R. M., Petrov, D., Jónsson, B., Schluter, D., Bell, M. A., & Kingsley, D. M. (2010). Adaptive evolution of pelvic reduction in sticklebacks by recurrent deletion of a Pitx1 enhancer. *Science*, *327*, 302–305. https://doi.org/10.1126/science.1182213

Chen, N. (2004). Using Repeat Masker to identify repetitive elements in genomic sequences. *Current Protocols in Bioinformatics*, *5*, 4.10. 11–14.10. 14. https://doi.org/10.1002/0471250953.bi0410s05

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

Cheng, H., Concepcion, G. T., Feng, X., Zhang, H., & Li, H. (2021). Haplotype-resolved de novo assembly using phased assembly graphs with hifiasm. *Nature Methods*, *18*, 170–175. https://doi.org/10.1038/s41592-020-01056-5

Colosimo, P. F., Hosemann, K. E., Balabhadra, S., Villarreal, G., Jr., Dickson, M., Grimwood, J., Schmutz, J., Myers, R. M., Schluter, D., & Kingsley, D. M. (2005). Widespread parallel evolution in sticklebacks by repeated fixation of ectodysplasin alleles. *Science*, *307*, 1928–1933. https://doi.org/10.1126/science.1107239

Coyne, J. A., & Orr, H. A. (2004). *Speciation* (Vol. *37*). Sinauer Associates.

Cresko, W. A., Amores, A., Wilson, C., Murphy, J., Currey, M., Phillips, P., Bell, M. A., Kimmel, C. B., & Postlethwait, J. H. (2004). Parallel genetic basis for repeated evolution of armor loss in Alaskan threespine stickleback populations. *Proceedings of the National Academy of Sciences of the United States of America*, *101*, 6050–6055. https://doi.org/10.1073/pnas.0308479101

Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., Handsaker, R. E., Lunter, G., Marth, G. T., Sherry, S. T., McVean, G., Durbin, R., & 1000 Genomes Project Analysis Group. (2011). The variant call format and VCFtools. *Bioinformatics*, *27*, 2156–2158. https://doi.org/10.1093/bioinformatics/btr330

Davis, M. S., Solbiati, J., & Cronan, J. E. (2000). Overproduction of acetyl-CoA carboxylase activity increases the rate of fatty acid biosynthesis in *Escherichia coli*. *Journal of Biological Chemistry*, *275*, 28593–28598. https://doi.org/10.1074/jbc.M004756200

Durand, N. C., Robinson, J. T., Shamim, M. S., Machol, I., Mesirov, J. P., Lander, E. S., & Aiden, E. L. (2016). Juicebox provides a visualization system for Hi-C contact maps with unlimited zoom. *Cell Systems*, *3*, 99–101. https://doi.org/10.1016/j.cels.2015.07.012

Edgar, R. C. (2004). MUSCLE: A multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics*, *5*, 1–19. https://doi.org/10.1186/1471-2105-5-113

Elmer, K. R., & Meyer, A. (2011). Adaptation in the age of ecological genomics: Insights from parallelism and convergence. *Trends in Ecology & Evolution*, *26*, 298–306. https://doi.org/10.1016/j.tree.2011.02.008

Feder, J. L., Egan, S. P., & Nosil, P. (2012). The genomics of speciation-with-gene-flow. *Trends in Genetics*, *28*, 342–350. https://doi.org/10.1016/j.tig.2012.03.009

Felsenstein, J. (1988). Phylogenies from molecular sequences: Inference and reliability. *Annual Review of Genetics*, *22*, 521–565. https://doi.org/10.1146/annurev.ge.22.120188.002513

Gautier, M., & Vitalis, R. (2012). rehh: An R package to detect footprints of selection in genome-wide SNP data from haplotype structure. *Bioinformatics*, *28*, 1176–1177. https://doi.org/10.1093/bioinformatics/bts115

Ghaffari, H., Wang, W., Li, A., Zhang, G., & Li, L. (2019). Thermotolerance divergence revealed by the physiological and molecular responses in two oyster subspecies of *Crassostrea gigas* in China. *Frontiers in Physiology*, *10*, 1137. https://doi.org/10.3389/fphys.2019.01137

Gross, J. B., Borowsky, R., & Tabin, C. J. (2009). A novel role for Mc1r in the parallel evolution of depigmentation in independent populations of the cavefish *Astyanax mexicanus*. *PLoS Genetics*, *5*, e1000326. https://doi.org/10.1371/journal.pgen.1000326

Guo, X., Li, C., Wang, H., & Xu, Z. (2018). Diversity and evolution of living oysters. *Journal of Shellfish Research*, *37*, 755–771. https://doi.org/10.2983/035.037.0407

Haas, B. J., Delcher, A. L., Mount, S. M., Wortman, J. R., Smith, R. K., Jr., Hannick, L. I., Maiti, R., Ronning, C. M., Rusch, D. B., Town, C. D., Salzberg, S. L., White, O., & Town, C. D. (2003). Improving the Arabidopsis genome annotation using maximal transcript alignment assemblies. *Nucleic Acids Research*, *31*, 5654–5666. https://doi.org/10.1093/nar/gkg770

Hardie, D. G. (2007). AMP-activated/SNF1 protein kinases: Conserved guardians of cellular energy. *Nature Reviews Molecular Cell Biology*, *8*, 774–785. https://doi.org/10.1038/nrm2249

Hassan, M. M., Qin, J. G., & Li, X. (2018). Gametogenesis, sex ratio and energy metabolism in *Ostrea angasi*: Implications for the reproductive strategy of spermcasting marine bivalves. *Journal of Molluscan Studies*, *84*, 38–45. https://doi.org/10.1093/mollus/eyx041

Haure, J., Huvet, A., Palvadeau, H., Nourry, M., Penisson, C., Martin, J. L. Y., & Boudry, P. (2003). Feeding and respiratory time activities in the cupped oysters *Crassostrea gigas*, *Crassostrea angulata* and their hybrids. *Aquaculture*, *218*, 539–551. https://doi.org/10.1016/s0044-8486(02)00493-3

Hofmeister-Brix, A., Kollmann, K., Langer, S., Schultz, J., Lenzen, S., & Baltrusch, S. (2013). Identification of the ubiquitin-like domain of midnolin as a new glucokinase interaction partner. *Journal of Biological Chemistry*, *288*, 35824–35839. https://doi.org/10.1074/jbc.M113.526632

Hohenlohe, P. A., Bassham, S., Etter, P. D., Stiffler, N., Johnson, E. A., & Cresko, W. A. (2010). Population genomics of parallel adaptation in threespine stickleback using sequenced RAD tags. *PLoS Genetics*, *6*, e1000862. https://doi.org/10.1371/journal.pgen.1000862

Huang, S., Kang, M., & Xu, A. (2017). HaploMerger2: Rebuilding both haploid sub-assemblies from high-heterozygosity diploid genome assembly. *Bioinformatics*, *33*, 2577–2579. https://doi.org/10.1093/bioinformatics/btx220

Huvet, A., Gerard, A., Ledu, C., Phelipot, P., Heurtebise, S., & Boudry, P. (2002). Is fertility of hybrids enough to conclude that the two oysters *Crassostrea gigas* and *Crassostrea angulata* are the same species? *Aquatic Living Resources*, *15*, 45–52. https://doi.org/10.1016/s0990-7440(01)01148-2

Jiang, G., Li, Q., Xu, C., Liu, S., Kong, L., & Yu, H. (2021). Reciprocal hybrids derived from *Crassostrea gigas* and *C. angulata* exhibit high heterosis in growth, survival and thermotolerance in northern China. *Aquaculture*, *545*, 737173. https://doi.org/10.1016/j.aquaculture.2021.737173

Johnson, A. D., Handsaker, R. E., Pulit, S. L., Nizzari, M. M., O'Donnell, C. J., & de Bakker, P. I. W. (2008). SNAP: A web-based tool for identification and annotation of proxy SNPs using HapMap. *Bioinformatics*, *24*, 2938–2939. https://doi.org/10.1093/bioinformatics/btn564

Jones, F. C., Grabherr, M. G., Chan, Y. F., Russell, P., Mauceli, E., Johnson, J., Swofford, R., Pirun, M., Zody, M. C., White, S., Birney, E., Searle,

S., Schmutz, J., Grimwood, J., Dickson, M. C., Myers, R. M., Miller, C. T., Summers, B. R., Knecht, A. K., … Amemiya, C. (2012). The genomic basis of adaptive evolution in threespine sticklebacks. *Nature*, *484*, 55–61. https://doi.org/10.1038/nature10944

Lawniczak, M. K. N., Emrich, S. J., Holloway, A. K., Regier, A. P., Olson, M., White, B., Redmond, S., Fulton, L., Appelbaum, E., Godfrey, J., Farmer, C., Chinwalla, A., Yang, S. P., Minx, P., Nelson, J., Kyung, K., Walenz, B. P., Garcia-Hernandez, E., Aguiar, M., … Besansky, N. J. (2010). Widespread divergence between incipient *Anopheles gambiae* species revealed by whole genome sequences. *Science*, *330*, 512–514. https://doi.org/10.1126/science.1195755

Li, A., Dai, H., Guo, X., Zhang, Z., Zhang, K., Wang, C., Wang, X., Wang, W., Chen, H., Li, X., Zheng, H., Li, L., & Zhang, G. (2021). Genome of the estuarine oyster provides insights into climate impact and adaptive plasticity. *Communications Biology*, *4*, 1287. https://doi.org/10.1038/s42003-021-02823-6

Li, A., Li, L., Zhang, Z., Li, S., Wang, W., Guo, X., & Zhang, G. (2021). Noncoding variation and transcriptional plasticity promote thermal adaptation in oysters by altering energy metabolism. *Molecular Biology and Evolution*, *38*, 5144–5155. https://doi.org/10.1093/molbev/msab241

Li, A., Wang, C., Wang, W., Zhang, Z., Liu, M., She, Z., Jia, Z., Zhang, G., & Li, L. (2020). Molecular and fitness data reveal local adaptation of southern and northern estuarine oysters (*Crassostrea ariakensis*). *Frontiers in Marine Science*, *7*, 589099. https://doi.org/10.3389/fmars.2020.589099

Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with burrows-wheeler transform. *Bioinformatics*, *25*, 1754–1760. https://doi.org/10.1093/bioinformatics/btp324

Li, L., Li, A., Song, K., Meng, J., Guo, X., Li, S., Li, C., De Wit, P., Que, H., Wu, F., Wang, W., Qi, H., Xu, F., Cong, R., Huang, B., Li, Y., Wang, T., Tang, X., Liu, S., … Zhang, G. (2018). Divergence and plasticity shape adaptive potential of the Pacific oyster. *Nature Ecology & Evolution*, *2*, 1751–1760. https://doi.org/10.1038/s41559-018-0668-2

Li, Q., Liu, W., Shirasu, K., Chen, W., & Jiang, S. (2006). Reproductive cycle and biochemical composition of the Zhe oyster *Crassostrea plicatula* Gmelin in an eastern coastal bay of China. *Aquaculture*, *261*, 752–759. https://doi.org/10.1016/j.aquaculture.2006.08.023

Li, Q., Osada, M., & Mori, K. (2000). Seasonal biochemical variations in Pacific oyster gonadal tissue during sexual maturation. *Fisheries Science*, *66*, 502–508. https://doi.org/10.1046/j.1444-2906.2000.00067.x

Liu, H., Zheng, H., Wang, S., Wang, Y., Li, S., Liu, W., & Zhang, G. (2013). Cloning and functional characterization of a polyunsaturated fatty acid elongase in a marine bivalve noble scallop *Chlamys nobilis* Reeve. *Aquaculture*, *416*, 146–151. https://doi.org/10.1016/j.aquaculture.2013.09.015

Liu, S., Li, L., Wang, W., Li, B., & Zhang, G. (2020). Characterization, fluctuation and tissue differences in nutrient content in the Pacific oyster (*Crassostrea gigas*) in Qingdao, northern China. *Aquaculture Research*, *51*, 1353–1364. https://doi.org/10.1111/are.14463

Losos, J. B. (2010). Adaptive radiation, ecological opportunity, and evolutionary determinism. American Society of Naturalists E. O. Wilson award address. *The American Naturalist*, *175*, 623–639. https://doi.org/10.1086/652433

Malinsky, M., Challis, R. J., Tyers, A. M., Schiffels, S., Terai, Y., Ngatunga, B. P., Miska, E. A., Durbin, R., Genner, M. J., & Turner, G. F. (2015). Genomic islands of speciation separate cichlid ecomorphs in an east African crater lake. *Science*, *350*, 1493–1498. https://doi.org/10.1126/science.aac9927

Marçais, G., & Kingsford, C. (2011). A fast, lock-free approach for efficient parallel counting of occurrences of k-mers. *Bioinformatics*, *27*, 764–770. https://doi.org/10.1093/bioinformatics/btr011

Martin, S. (2022). Github. https://github.com/simonhmartin

Menzel, R. W. (1974). Portuguese and Japanese oysters are the same species. *Journal of the Fisheries Research Board of Canada*, *31*, 453–456. https://doi.org/10.1139/f74-074

Munday, M. R., & Hemingway, C. J. (1999). The regulation of acetyl-CoA carboxylase – A potential target for the action of hypolipidemic agents. In G. Weber (Ed.), *Advances in enzyme regulation* (Vol. *39*, pp. 205–234). Elseiver.

Nelson, C. W., Moncla, L. H., & Hughes, A. L. (2015). SNPGenie: Estimating evolutionary parameters to detect natural selection using pooled next-generation sequencing data. *Bioinformatics*, *31*, 3709–3711. https://doi.org/10.1093/bioinformatics/btv449

Ni, G., Li, Q., Kong, L., & Yu, H. (2014). Comparative phylogeography in marginal seas of the northwestern Pacific. *Molecular Ecology*, *23*, 534–548. https://doi.org/10.1111/mec.12620

Nosil, P., Funk, D. J., & Ortiz-Barrientos, D. (2009). Divergent selection and heterogeneous genomic divergence. *Molecular Ecology*, *18*, 375–402. https://doi.org/10.1111/j.1365-294X.2008.03946.x

Paton, C. M., & Ntambi, J. M. (2009). Biochemical and physiological function of stearoyl-CoA desaturase. *American Journal of Physiology-Endocrinology and Metabolism*, *297*, E28–E37. https://doi.org/10.1152/ajpendo.90897.2008

Pearse, D. E., Miller, M. R., Abadia-Cardoso, A., & Garza, J. C. (2014). Rapid parallel evolution of standing variation in a single, complex, genomic region is associated with life history in steelhead/rainbow trout. *Proceedings of the Royal Society B-Biological Sciences*, *281*, 20140012. https://doi.org/10.1098/rspb.2014.0012

Pearse, D. E., & Pogson, G. H. (2000). Parallel evolution of the melanic form of the California legless lizard, *Anniella pulchra*, inferred from mitochondrial DNA sequence variation. *Evolution*, *54*, 1041–1046. https://doi.org/10.1111/j.0014-3820.2000.tb00104.x

Pickrell, J. K., & Pritchard, J. K. (2012). Inference of population splits and mixtures from genome-wide allele frequency data. *PLoS Genetics*, *8*, e1002967. https://doi.org/10.1371/journal.pgen.1002967

Plazzi, F., & Passamonti, M. (2010). Towards a molecular phylogeny of Mollusks: Bivalves' early evolution as revealed by mitochondrial genes. *Molecular Phylogenetics and Evolution*, *57*, 641–657. https://doi.org/10.1016/j.ympev.2010.08.032

Prjibelski, A., Antipov, D., Meleshko, D., Lapidus, A., & Korobeynikov, A. (2020). Using SPAdes de novo assembler. *Current Protocols in Bioinformatics*, *70*, e102. https://doi.org/10.1002/cpbi.102

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A. R., Bender, D., Maller, J., Sklar, P., de Bakker, P. I., Daly, M. J., & Sham, P. C. (2007). PLINK: A tool set for whole-genome association and population-based linkage analyses. *American Journal of Human Genetics*, *81*, 559–575. https://doi.org/10.1086/519795

Puritz, J. B., Guo, X., Hare, M., He, Y., Hillier, L. W., Jin, S., Liu, M., Lotterhos, K. E., Minx, P., Modak, T., Proestou, D., Rice, E. S., Tomlinson, C., Warren, W. C., Witkop, E., Zhao, H., & Gomez-Chiarri, M. (2023). A second unveiling: Haplotig masking of the eastern oyster genome improves population-level inference. *Molecular Ecology Resources*. Epub ahead of print. https://doi.org/10.1111/1755-0998.13801

Qin, Y., Liao, Q., Shi, G., Yang, Y., Zhou, Y., Li, J., Ma, H., Zhang, Y., & Yu, Z. (2022). Comparison of growth, survival and fertility of the southern and northern populations of *Crassostrea ariakensis* and their hybrids in southern China. *Aquaculture*, *549*, 737744. https://doi.org/10.1016/j.aquaculture.2021.737744

Quinlan, A. R., & Hall, I. M. (2010). BEDTools: A flexible suite of utilities for comparing genomic features. *Bioinformatics*, *26*, 841–842. https://doi.org/10.1093/bioinformatics/btq033

Raynes, Y., & Sniegowski, P. D. (2014). Experimental evolution and the dynamics of genomic mutation rate modifiers. *Heredity*, *113*, 375–380. https://doi.org/10.1038/hdy.2014.49

Ren, J., Liu, X., Jiang, F., Guo, X., & Liu, B. (2010). Unusual conservation of mitochondrial gene order in *Crassostrea oysters*: Evidence for recent speciation in Asia. *BMC Evolutionary Biology*, *10*, 1–14. https://doi.org/10.1186/1471-2148-10-394

Renaut, S., Grassa, C. J., Yeaman, S., Moyers, B. T., Lai, Z., Kane, N. C., Bowers, J. E., Burke, J. M., & Rieseberg, L. H. (2013). Genomic islands of divergence are not affected by geography of

speciation in sunflowers. *Nature Communications*, *4*, 1827. https://doi.org/10.1038/ncomms2833

Renaut, S., Owens, G. L., & Rieseberg, L. H. (2014). Shared selective pressure and local genomic landscape lead to repeatable patterns of genomic divergence in sunflowers. *Molecular Ecology*, *23*, 311–324. https://doi.org/10.1111/mec.12600

Rohde, K. (1992). Latitudinal gradients in species diversity: The search for the primary cause. *Oikos*, *65*, 514–527. https://doi.org/10.2307/3545569

Saltiel, A. R., & Kahn, C. R. (2001). Insulin signalling and the regulation of glucose and lipid metabolism. *Nature*, *414*, 799–806. https://doi.org/10.1038/414799a

Sassa, T., & Kihara, A. (2014). Metabolism of very long-chain fatty acids: Genes and pathophysiology. *Biomolecules & Therapeutics*, *22*, 83–92. https://doi.org/10.4062/biomolther.2014.017

Schlitzer, R. (2018). *Ocean data view*. https://odv.awi.de

Schluter, D. (2009). Evidence for ecological speciation and its alternative. *Science*, *323*, 737–741. https://doi.org/10.1126/science.1160006

Shi, B., Wu, Y., Zhou, L., You, W., & Ke, C. (2019). Influence of gonadal development on metal accumulation in the Portuguese oyster, *Crassostrea angulata*, in subtropical areas. *Aquaculture Research*, *50*, 1142–1152. https://doi.org/10.1111/are.13988

Simão, F. A., Waterhouse, R. M., Ioannidis, P., Kriventseva, E. V., & Zdobnov, E. M. (2015). BUSCO: Assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics*, *31*, 3210–3212. https://doi.org/10.1093/bioinformatics/btv351

Sodeland, M., Jorde, P. E., Lien, S., Jentoft, S., Berg, P. R., Grove, H., Kent, M. P., Arnyasi, M., Olsen, E. M., & Knutsen, H. (2016). "Islands of Divergence" in the Atlantic cod genome represent polymorphic chromosomal rearrangements. *Genome Biology and Evolution*, *8*, 1012–1022. https://doi.org/10.1093/gbe/evw057

Stamatakis, A. (2014). RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, *30*, 1312–1313. https://doi.org/10.1093/bioinformatics/btu033

Stanke, M., Diekhans, M., Baertsch, R., & Haussler, D. (2008). Using native and syntenically mapped cDNA alignments to improve *de novo* gene finding. *Bioinformatics*, *24*, 637–644. https://doi.org/10.1093/bioinformatics/btn013

Tracy, C. A., & Widom, H. (1994). Level-spacing distributions and the Airy kernel. *Communications in Mathematical Physics*, *159*, 151–174. https://doi.org/10.1007/bf02100489

Vaschenko, M. A., Hsieh, H.-L., & Radashevsky, V. I. (2013). Gonadal state of the oyster *Crassostrea angulata* cultivated in Taiwan. *Journal of Shellfish Research*, *32*, 471–482. https://doi.org/10.2983/035.032.0227

Voight, B. F., Kudaravalli, S., Wen, X. Q., & Pritchard, J. K. (2006). A map of recent positive selection in the human genome. *PLoS Biology*, *4*, 446–458. https://doi.org/10.1371/journal.pbio.0040072

Vurture, G. W., Sedlazeck, F. J., Nattestad, M., Underwood, C. J., Fang, H., Gurtowski, J., & Schatz, M. C. (2017). GenomeScope: Fast reference-free genome profiling from short reads. *Bioinformatics*, *33*, 2202–2204. https://doi.org/10.1093/bioinformatics/btx153

Wang, H. (2004). Studies on the molecular phylogeny and taxonomy of common oysters in China seas. (Doctoral thesis). Institute of Oceanology, Chinese Academy of Sciences.

Wang, P. X. (1999). Response of western Pacific marginal seas to glacial cycles: Paleoceanographic and sedimentological features. *Marine Geology*, *156*, 5–39. https://doi.org/10.1016/s0025-3227(98)00172-8

Waples, R. S., Teel, D. J., Myers, J. M., & Marshall, A. R. (2004). Life-history divergence in Chinook salmon: Historic contingency and parallel evolution. *Evolution*, *58*, 386–403. https://doi.org/10.1111/j.0014-3820.2004.tb01654.x

Waterhouse, R. M., Seppey, M., Simão, F. A., Manni, M., Ioannidis, P., Klioutchnikov, G., … Zdobnov, E. M. (2018). BUSCO applications from quality assessments to gene prediction and phylogenomics. *Molecular Biology and Evolution*, *35*, 543–548. https://doi.org/10.1093/molbev/msx319

Wu, B., Chen, X., Yu, M. J., Ren, J. F., Hu, J., Shao, C. W., Zhou, L., Sun, X., Yu, T., Zheng, Y., Wang, Y., Wang, Z., Zhang, H., Fan, G., & Liu, Z. H. (2022). Chromosome-level genome and population genomic analysis provide insights into the evolution and environmental adaptation of Jinjiang oyster *Crassostrea ariakensis*. *Molecular Ecology Resources*, *22*, 1529–1544. https://doi.org/10.1111/1755-0998.13556

Yang, H., Ma, M., Thompson, J. R., & Flower, R. J. (2021). Protect and regulate China's oyster resources. *Science*, *371*, 790. https://doi.org/10.1126/science.abg8656

Yang, Z. (2007). PAML 4: Phylogenetic analysis by maximum likelihood. *Molecular Biology and Evolution*, *24*, 1586–1591. https://doi.org/10.1093/molbev/msm088

Yu, G., Wang, L. G., Han, Y., & He, Q. Y. (2012). clusterProfiler: An R package for comparing biological themes among gene clusters. *Omics*, *16*, 284–287. https://doi.org/10.1089/omi.2011.0118

Zhang, H., Liu, H., Cheng, D., Liu, H., & Zheng, H. (2018). Molecular cloning and functional characterisation of a polyunsaturated fatty acid elongase in a marine bivalve *Crassostrea angulata*. *Journal of Food Nutrition Research*, *6*, 89–95. https://doi.org/10.12691/jfnr-6-2-4

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Teng, W., Fu, H., Li, Z., Zhang, Q., Xu, C., Yu, H., Kong, L., Liu, S., & Li, Q. (2023). Parallel evolution in *Crassostrea* oysters along the latitudinal gradient is associated with variation in multiple genes involved in adipogenesis. *Molecular Ecology*, *00*, 1–12. https://doi.org/10.1111/mec.17108