Received: 12 March 2020 | Revised: 3 September 2020 | Accepted: 7 September 2020

DOI: 10.1111/1755-0998.13261

RESOURCE ARTICLE

MOLECULAR ECOLOGY RESOURCES WILEY

Giant African snail genomes provide insights into molluscan whole-genome duplication and aquatic–terrestrial transition

Conghui Liu | Yuwei Ren | Zaiyuan Li | Qi Hu | Lijuan Yin | Hengchao Wang | Xi Qiao | Yan Zhang | Longsheng Xing | Yu Xi | Fan Jiang | Sen Wang | Cong Huang | Bo Liu | Hangwei Liu | Fanghao Wan | Wanqiang Qian | Wei Fan

Shenzhen Branch, Guangdong Laboratory for Lingnan Modern Agriculture, Agricultural Genomics Institute at Shenzhen, Chinese Academy of Agricultural Sciences, Shenzhen, China

Correspondence
Wanqiang Qian and Wei Fan, Shenzhen Branch, Guangdong Laboratory for Lingnan Modern Agriculture, Agricultural Genomics Institute at Shenzhen, Chinese Academy of Agricultural Sciences, Shenzhen, 518120, China.
Email: qianwanqiang@caas.cn; fanwei@caas.cn.

Funding information
National Natural Science Foundation of China, Grant/Award Number: 31901950; the Agricultural Science and Technology Innovation Program && The Elite Young Scientists Program of CAAS, Fundamental Research Funds for Central Non-profit Scientific Institution, Grant/Award Number: Y2017JC01; the National key research and development program of China, Grant/Award Number: 2016YFC1200600; Natural Science Foundation of Shenzhen, Grant/Award Number: JCYJ20190813120401662; the Agricultural Science and Technology Innovation Program Cooperation and Innovation Mission, Grant/Award Number: CAAS-XTCX2016; Fund of Key Laboratory of Shenzhen, Grant/Award Number: ZDSYS20141118170111640

Abstract

Whole-genome duplication (WGD), contributing to evolutionary diversity and environmental adaptability, has been observed across a wide variety of eukaryotic groups, but not in molluscs. Molluscs are the second largest animal phylum in terms of species numbers, and among the organisms that have successfully adapted to the nonmarine realm through aquatic–terrestrial (A-T) transition. We assembled a chromosome-level reference genome for *Achatina immaculata*, a globally invasive species, and compared the genomes of two giant African snails (*A. immaculata* and *Achatina fulica*) to other available mollusc genomes. Macrosynteny, colinearity blocks, Ks peak and Hox gene clusters collectively suggested a WGD event in the two snails. The estimated WGD timing (~70 million years ago) was close to the speciation age of the Sigmurethra–Orthurethra (within Stylommatophora) lineage and the Cretaceous–Tertiary (K-T) mass extinction, indicating that the WGD may have been a common event shared by all Sigmurethra–Orthurethra species and conferred ecological adaptability allowing survival after the K-T extinction event. Furthermore, the adaptive mechanism of WGD in terrestrial ecosystems was confirmed by the presence of gene families related to the respiration, aestivation and immune defence. Several mucus-related gene families expanded early in the Stylommatophora lineage, and the haemocyanin and phosphoenolpyruvate carboxykinase families doubled during WGD, and zinc metalloproteinase genes were highly tandemly duplicated after WGD. This evidence suggests that although WGD may not have been the direct driver of the A-T transition, it played an important part in the terrestrial adaptation of giant African snails.

KEYWORDS
aquatic–terrestrial transition, giant African snails, molluscs, whole-genome duplication

1 | INTRODUCTION

Whole-genome duplication (WGD) has been proposed to be a key evolutionary event that drives phenotypic complexity, functional novelty and ecological adaptation (Mable et al., 2011; Van de Peer et al., 2009). With the recent advances in high-throughput DNA sequencing technologies, almost all fundamental lineages of land plants have been characterized as having undergone WGD, while far fewer WGDs have

Liu and Ren are contributed equally to this work.

been reported among animal lineages, especially in protostome invertebrates (Clark & Donoghue, 2018; Orr, 1990). Based on chromosome counts and genome size, WGD events are thought to have occurred in molluscs, the second largest animal phylum in terms of species numbers, which have survived major extinction events and become some of the most successful members of marine, freshwater and terrestrial environments (Guo et al., 2015; Hallinan & Lindberg, 2011). Giant African snails, a group of species in the family Achatinidae that possess unusually large body sizes, evolved from aquatic ancestors and have achieved a pulmonate terrestriality through an aquatic–terrestrial (A-T) transition (Jörger et al., 2010). Because of their remarkable ecological adaptability, some species of giant African snails are considered to be globally invasive species that cause serious damage to agricultural crops and ornamental plants (Cowie et al., 2009). Their success in colonizing terrestrial ecosystems and adaptation to diverse environments suggests that giant African snails possess advanced genetic plasticity and evolutionary novelties, although the key drivers and underlying mechanisms remain poorly understood.

In contrast to plants, only a few ancient WGDs have been well documented in animals, because polyploidy is usually an evolutionary dead end that results from associated meiotic difficulties (Comai, 2005; Van de Peer et al., 2017). In vertebrates, several ancient WGD events that occurred at the origin of vertebrates and teleost fishes are proposed to have shaped genetic diversity and adaptive radiation (Furlong & Holland, 2002; Hoegg et al., 2004). In animals other than vertebrates, less conclusive evidence of ancient WGD events has been reported due to a deficiency in the continuous production of high-quality genomic data. Nevertheless, there are several known cases of large-scale genome duplication in certain groups, such as rotifers, chelicerates and insects. The genomes of bdelloid rotifers are tetraploid, and the scaffolds are rearranged during asexual reproduction (Flot et al., 2013). A possible ancestral WGD in chelicerates has been identified in the horseshoe crab and was dated to at least 135 million years ago (Ma) in a phylogenetic analysis of Hox genes (Kenny et al., 2016), and subsequent surveys in chelicerates revealed that the spider and scorpion lineages underwent another separate WGD event before 430 Ma (Schwager et al., 2017). Because of the fragmented genome assemblies and uncertain accuracy of traditional analysis methods, the validation of WGD in invertebrates remains controversial. Macrosyteny analysis in lepidopterans has suggested that WGD events should be verified by chromosome-scale genome assembly and macrosynteny analysis (Li et al., 2018; Nakatani & McLysaght, 2019).

Regarding the A-T transition, the conquest of land by organisms that evolved from aquatic ancestors represents one of the most astonishing events in the history of life on Earth (Lillywhite et al., 2012). The dramatic environmental changes that occur when transitioning from homogeneous water habitats to the heterogenous land environment brought about a significant revolution in species evolution, leading to a diversity of radiation and complexity of life (Vermeij & Dudley, 2000; Volkmann & Baluška, 2006). This step was achieved in multiple phyla independently via a set of adaptations such as water balance, air breathing, nitrogen excretion, neural–immune system interactions and certain behaviours (Little, 1990). Within Mollusca, the clade Pulmonata includes several lineages that invaded the terrestrial zone and nonaquatic ecosystems, especially the Stylommatophora (Romero et al., 2016). The giant African snail, a species in the order Stylommatophora, is a representative of the group of land snails that have also developed fundamental morphological and physiological changes and behaviours, such as a pulmonate lung, complex immune system, mucus production and aestivation, that make it highly adapted to the terrestrial environment (Hiong et al., 2005; Mukherjee et al., 2017; Romero et al., 2016). Several reports on amphibious species have shown that the expansion and positive selection of genes related to innate immunity, nitrogen excretion, hormonal regulation and pulmonary surfactants were closely connected to the A-T transition (Biscotti et al., 2016); however, the genomic features and evolutionary characteristics of terrestrial invertebrates remain poorly described. The giant African snail, a representative of the Sigmurethra–Orthurethra lineage (snails and slugs) in the Stylommatophora, is a promising model for terrestrial molluscs, which will facilitate our understanding of the invertebrate A-T transition.

Mollusca is the second largest animal phylum, accounting for 7% of all species of living animals on Earth (Bouchet, 2006), and it comprises numerous species of evolutionary and economic importance (Ferreira et al., 2012; Thiengo et al., 2007). However, the available genomic resources for Mollusca remain low in comparison with those for other large phyla such as Arthropoda and Nematoda (Faure & Casanova, 2006; Serb & Lydeard, 2003). The genomes of aquatic molluscs such as the California sea hare (Aplysia Genome Project, 2009), Pacific oyster (Zhang, Fang et al., 2012; Zhang, Xiao et al., 2012), pearl oyster (Du et al., 2017; Takeuchi et al., 2012), owl limpet (Simakov et al., 2013), California two-spot octopus (Albertin et al., 2015), golden mussel (Uliano-Silva et al., 2018), *Biomphalaria glabrata* (Adema et al., 2017) and golden apple snail (Liu et al., 2018) have been sequenced, whereas few terrestrial species in Mollusca have well-documented genomic information. Recently, genomic data for the invasive land snail *Achatina fulica* were released, but without in-depth studies of biological issues related to the terrestrial lifestyle (Guo et al., 2019). To address the genetic and evolutionary characteristics of terrestrial molluscs, we report the genome of another similarly shaped giant African snail species, *Achatina immaculata* (also known as *Achatina panthera* and *Lissachatina immaculata*), with a larger body size and greater invasiveness (Figure 1a–c) (Sow-Yan, 2019), and provide insights into the WGD, A-T transition and invasion mechanisms through comparative genomic analyses between the two giant African snails and other closely related molluscs.

2 | MATERIALS AND METHODS

2.1 | Sample collection and genome sequencing

Adults *Achatina immaculata* were collected from a local farm in Yangjiang, Guangdong province, China, and maintained at

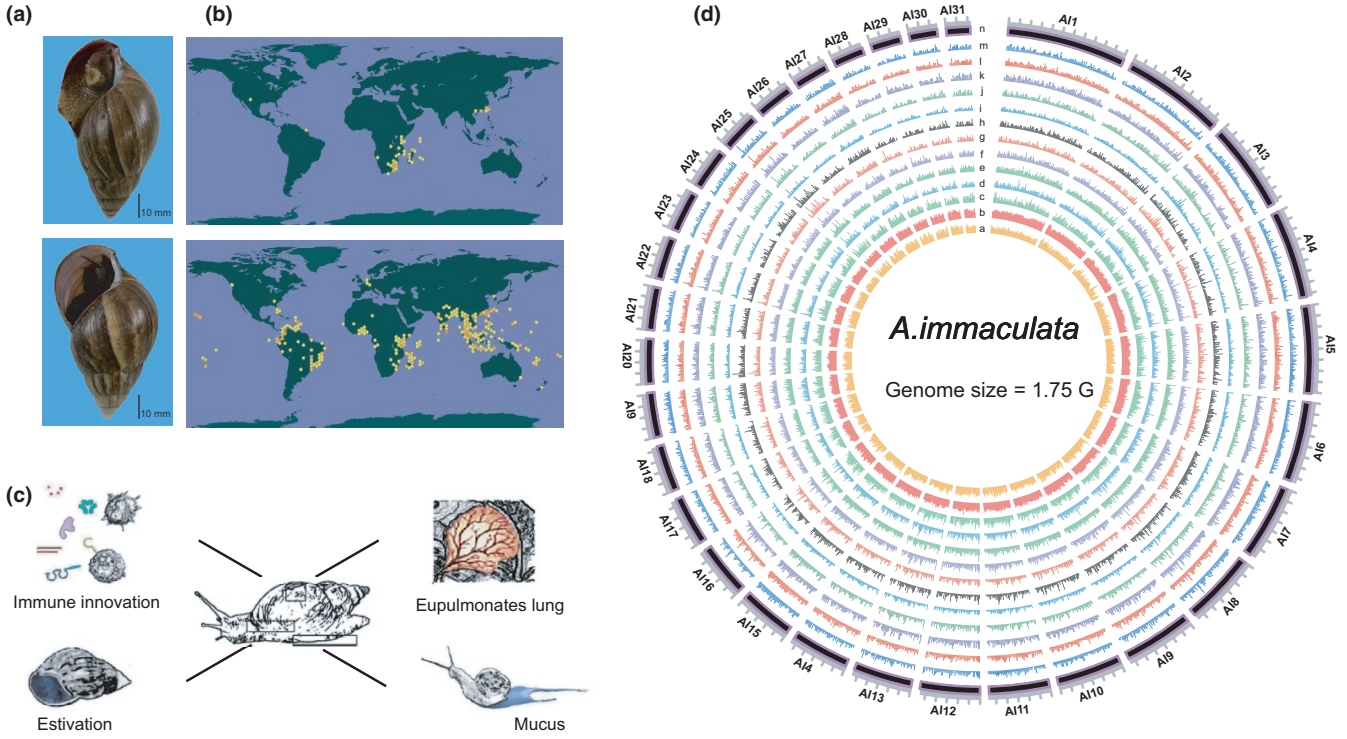

FIGURE 1 The general characteristics of *Achatina immaculata*. (a) Morphological differences: *A. immaculata* has a relatively longer shell with pink or purple columella, while *Achatina fulica* has a shorter shell with white columella. Scale of 10 mm is indicated at the bottom. (b) Global invasion regions for *A. immaculata* (top) and *A. fulica* (bottom). Currently, *A. fulica* is more widely distributed, but *A. immaculata* has a faster invasion speed. (c) Sketch map showing the major physiological specialty of *A. immaculata*, including immune innovation, eupulmonates lung, aestivation and mucus. (d) Genomic features shown by a CIRCOS plot. Track n: 31 linkage groups of the genome; Tracks d–m: expression profile of brain, egg, eye, haemocytes, hepatopancreas, kidney, lungs, muscle, ovary and testis tissues; Track c: distribution of transposon elements; Track b: distribution of gene density; Track a: distribution of GC content

25 ± 2°C for a week before processing. The average weight of these adult snails is about 75 g, while the average size of adults is about 7 cm in length and 4.5 cm in width. Genomic DNA was extracted from the foot muscles of a single snail using DNeasy Blood & Tissue Kit (Catalogue no. 69504, Qiagen) for constructing PCR-free Illumina 350-bp insert libraries and PacBio 20-kb insert library, and sequenced on Illumina HiSeq-X and PacBio SMRT platforms, respectively. The Hi-C library was prepared using the muscle tissue of another single snail by the following methods: nuclear DNA was cross-linked *in situ*, extracted and then digested with a restriction enzyme. The sticky ends of the digested fragments were biotinylated, diluted and then ligated to each other randomly. Biotinylated DNA fragments were enriched and sheared again for preparing the sequencing library, which was then sequenced on a HiSeq-X platform (Illumina).

2.2 | RNA sample preparation and transcriptome sequencing

Ten tissue types, brains, eggs (24 hr post-oviposition), eyes, haemocytes, hepatopancreas, kidneys, lungs, muscles, ovaries and testes, from six animals were collected as replicates. Eighty snails were employed for the immune elicitor challenge experiment, and they were divided equally into a control group, LPS (lipopolysaccharide) group, PGN (peptidoglycan) group, GLU (β-glucan) group and IC (poly I:C) group. The five groups of snails received injections of 100 μl phosphate-buffered saline (PBS, 0.14 m NaCl, 3 mm KCl, 8 mm $NaH_2PO_4.12H_2O$, 1.5 mm K_2HPO_4, pH 7.4), LPS from *Escherichia coli* 0111:B4 (Sigma-Aldrich, 0.5 mg/ml in PBS), PGN from *Staphylococcus aureus* (Sigma-Aldrich, 0.8 mg/ml in PBS), GLU from *Saccharomyces cerevisiae* (Sigma-Aldrich, 1.0 mg/ml in PBS) and poly I:C (Sigma-Aldrich, 1.0 mg/ml in PBS), respectively. These treated snails were maintained after injection, and 15 individuals from each group were randomly sampled at 12 hr post-injection. Nonaestivating snails were fed with enough food and water, whereas aestivating snails were fasted and treated at high temperature, and the whole blood of these two groups were collected around 10 days after epiphragm formation. Whole blood samples of all the treated group were collected from five individuals and pooled. There were three replicates for each sample. Total RNAs were extracted from the stored tissues using an E.Z.N.A. Mollusc RNA Kit (Catalogue no. R6875-01, Omega), and then mRNAs were pulled out by beads with poly-T to construct cDNA libraries (insert 350 bp) using a VAHTS TM mRNA-seq V3 Library Prep Kit (NR611), and sequenced on an Illumina HiSeq-X sequencer.

2.3 | Genome assembly

The Illumina raw reads were filtered by trimming the adapter sequence and low-quality regions (https://github.com/fanagislab/DBG_assembly/tree/master/clean_illumina), resulting in clean and high-quality reads with an average error rate <0.001. For the PacBio raw data, the short subreads (<2 kb) and low-quality (error rate >0.2) subreads were filtered out, and only one representative subread was retained for each PacBio read. The clean PacBio reads were corrected and trimmed by CANU 1.8 (Koren et al., 2017) and then assembled by WTDBG2 (Ruan & Li, 2020). Default parameters were employed by CANU-correct and CANU-trim during the SMRT reads correction, and by WTDBG2 during assembly. The PacBio reads were used to polish the raw contigs by a module within WTDBG2, after which Illumina reads were aligned to the contigs by BWA-MEM, and single base errors in the contigs were corrected by PILON 2.10 (Walker et al., 2014) with the parameters "-fix bases, -nonpf, -minqual 20." Next, Hi-C sequencing data were aligned to the haploid reference contigs by HIC-PRO 2.11.1 (Servant et al., 2015), and these contigs were then clustered, ordered and orientated into chromosomes with LACH-ESIS (Burton et al., 2013).

2.4 | Genome annotation

A de novo repeat library for *A. immaculata* was constructed by REPEATMODELER (version 1.0.4; http://www.repeatmasker.org/RepeatModeler.html). Transposable elements (TEs) in the *A. immaculata* genome were identified by REPEATMASKER (version 4.0.6; http://www.repeatmasker.org/) using both the Repbase library and the *de novo* library. Tandem repeats in the *A. immaculata* genome were predicted using TANDEM REPEATS FINDER version 4.07b (Benson, 1999). The divergence rates of TEs were calculated between the identified TEs in the genome and their consensus sequence at the TE family level.

The gene models in the *A. immaculata* genome were predicted by EVM version 1.1.1 (Haas et al., 2008) integrating evidences from *ab initio* predictions, homology-based searches and RNA-sequencing (RNA-seq) alignments. In annotation of RNA-seq alignments, the corrected RNA-seq reads were mapped to the genome with TOPHAT version 2.1.0 and the gene regions were extracted with CUFFLINKS version 2.2.1. Then, these gene models were annotated by RNA-seq data, the UniProt database and INTERPROSCAN software (Jones et al., 2014). Finally, the gene models were retained if they had at least one piece of supporting evidence from the UniProt database, INTERPROSCAN domain and RNA-seq data (mapping rate >60%). Gene functional annotation was performed by aligning the protein sequences to the NCBI NR, UniProt, COG and KEGG (Kyoto Encyclopedia of Genes and Genomes) databases with BLASTP version 2.3.0+ under an E-value cutoff of 10^{-5} and choosing the best hit. Pathway analysis and functional classification were conducted based on the KEGG database (Aoki & Kanehisa, 2005). INTERPROSCAN was used to assign IPR domains and GO terms to the gene models. BUSCO 3.0.2 (https://github.com/guyleonard/busco_hmms_for_anvio/) and metazoan core genes from OrthoDB (http://www.orthodb.org/) were adopted to evaluate the integrity of the genome assembly and the gene models.

2.5 | Evolutionary analysis

Orthologous and paralogous groups were assigned from 11 species (*A. immaculata*, *A. fulica*, *Aplysia californica*, *Biomphalaria glabrata*, *Pomacea canaliculata*, *Lottia gigantea*, *Crassostrea gigas*, *Pinctada fucata*, *Limnoperna fortunei*, *Octopus bimaculoides* and *Lingula anatina*) by ORTHOFINDER (Emms & Kelly, 2015) with default parameters. Orthologous groups (OGs) that contained only one gene for each species were selected to construct the phylogenetic tree. The protein sequences of each gene family were independently aligned by MUSCLE version 3.8.31 (Edgar, 2004) and then concatenated into one super-sequence. The phylogenetic tree was constructed by maximum likelihood (ML) using RAXML 8.2.12 (Stamatakis, 2014), with the best-fit model (LG + I+G + F) estimated by PROTTEST3 (Darriba et al., 2011). The absolute rates of molecular evolution and divergence times were estimated by R8S (Sanderson, 2003). The tree was calibrated with the following time frames to constrain the age of the nodes between the species: minimum = 260 Ma and maximum = 290 Ma for *Pinctada fucata* and *C. gigas* (Sun et al., 2017); minimum = 450 Ma and maximum = 480 Ma for *Aplysia californica* (or *B. glabrata*) and *Lottia gigantea* (Benton et al., 2009). The calibration time (fossil record time) interval (550–610 Ma) of *O. bimaculoides* was adopted from previous results (Zapata et al., 2014). Subsequently, CAFÉ (Computational Analysis of gene Family Evolution) (De Bie et al., 2006) was used to identify gene family expansion and contraction with default parameters.

2.6 | Transcriptome data analysis

Transcriptome reads were trimmed using the same method as for genomic reads (https://github.com/fanagislab/DBG_assembly/tree/master/clean_illumina), and then mapped to the reference genome of *A. immaculata* using TOPHAT (version 2.1.0) with default settings. The expression level of each reference gene in terms of FPKM (fragments per kilobase of transcript per million mapped reads) was computed by CUFFLINKS version 2.2.1. A gene was considered to be expressed if FPKM> 0. Analysis of differentially expressed genes (DEGs) was conducted using CUFFDIFF version 2.2.1, requiring fold-change over 2 or less than 0.5, and q-value (false discovery rate [FDR]) less than 0.05.

2.7 | WGD verification

The analyses incorporating chromosome-level macrosynteny analysis, colinearity blocks, Ks peak and Hox gene clusters were used to verify the WGD in giant African snails. Based on four chromosome-level molluscan genomes, macrosynteny was identified using

homologous gene sets. Genes from different species would be identified as homologous gene pairs when they had mutual best BLASTP hits with each other. The conserved macrosynteny between species with chromosome-level assemblies was displayed in dot plot. Each dot in the dot plot comparison represents a one-to-one homologous gene pair mentioned above. Based on the dot plot, we inferred the circus plot and dual synteny plot. Then, MCSCANX was also used with default parameters to identify the colinearity blocks in *A. immaculata* and *A. fulica* (Wang et al., 2012). The Ks distribution of gene pairs in colinearity blocks was calculated by PARAAT (Zhang, Fang et al., 2012; Zhang, Xiao et al., 2012) and KAKS _ CALCULATOR 2.0 (Wang et al., 2010). Homeobox genes were identified in the giant African snails using BLAST ($E < e^{-5}$) against all homeodomain sequences from the HomeoDB database (http://homeodb.zoo.ox.ac.uk/), and were further confirmed by comparing to the Conserved Domains Database (http://www.ncbi.nlm.nih.gov/cdd).

2.8 | Gene family analysis

Gene families involved in pathways of respiration, aestivation and immune functions for closely related species were found by literature review, and the protein sequence data were download from public databases. The identification of homologous genes in *A. immaculata* was performed in three steps: initially, we aligned known genes of other species closely related to the *A. immaculata* genome by BLASTP with best hits ($E < e^{-5}$), and followed by the analysis of paralogous genes performed by ORTHOFINDER. Then, the obtained genes were used to perform phylogeny analysis by ML with MEGA7 (Kumar et al., 2016), to further validate the accuracy and reveal the phylogenetic relationship of these genes.

3 | RESULTS

3.1 | The *A. immaculata* genome provides a high-quality assembly for giant African snails

We generated 200 Gb (121×) of PacBio SMRT raw reads with an average read length of 14 kb, and 145 Gb (82×) of Illumina HiSeq paired-end reads using DNA extracted from a single adult of *Achatina immaculata* (Table S1). After quality filtering, 199 Gb (120×) of clean PacBio SMRT reads were corrected with CANU (Koren et al., 2017), assembled with WTDBG2 (Ruan & Li, 2020), and polished with PILON (Walker et al., 2014), resulting in an assembly of 563 raw contigs with a total length of 1,653 Mb and an N50 length of 3.80 Mb (Table S2). Based on the Hi-C data, 1,648 Mb (99.7%) of the final contigs was anchored and arranged into 31 linkage groups, each possibly corresponding to a natural chromosome (Sun, 1995; Thiriot-Quiévreux, 2003), where the longest was 111.19 Mb and the shortest was 34.32 Mb (Figure 1d; Figure S1). We estimated the genome size of *A. immaculata* to be 1.75 Gb, using GCE software based on the distribution of the k-mer frequency (Liu et al., 2013) (Figure S2), which means that ~ 95% of the *A. immaculata* genome is included in the assembly. To further confirm the accuracy and completeness of the assembly, we mapped the Illumina reads to the assembled reference genome. Significantly, 99.11% of the genome-derived reads could be aligned to the reference genome with a coverage rate of 96.50%, suggesting no obvious bias in the sequencing and assembly. Besides, 93.2% of the metazoan core genes (Complete 92.0% and Fragmentary 1.2%) from OrthoDB (http://www.orthodb.org/) were identified in the reference genome assembly using BUSCO. Furthermore, the rate of heterozygosity was estimated from the percentage of heterozygous single nucleotide polymorphisms (SNPs) as 0.24%, and this low rate ensures that the assembly is accurate. The high-quality reference genome provides a good foundation for gene annotation.

In total, 30,194 gene models were predicted as the reference gene set, with coding regions spanning ~ 39.1 Mb (2.37%) of the genome (Tables S3 and S4). The distribution of CDS (coding sequence) length in *A. immaculata* is similar to that in closely related species (Figures S3 and S4). Overall, 87.7% of the reference genes were supported by transcriptome data, and 97.9% of the metazoan core genes (complete 89.6% and fragmentary 8.3%) from OrthoDB (http://www.orthodb.org/) were identified in the reference gene set using BUSCO. The difference in BUSCO scores between the genome assembly and the gene models is mostly due to the different gene annotation methods adopted by evaluating the reference genome assembly (BUSCO integrated augustus prediction method) and the reference gene models (our gene annotation method, see section 2.4); additional explanations and examples can be found in the BUSCO method paper (Simão et al., 2015). These results are comparable to those obtained from other published molluscan genomes (Table S5). In the functional annotation, a total of 27,349 (90.58%) reference genes were annotated based on at least one functional database (Figure S5).

The quality of this genome assembly is comparable to, or better than, that of other published molluscan genomes (Table S2). In particular, the coverage rate and sequence continuity were greatly improved compared with the most recently published *A. fulica* genome. The estimated genome size is 2.12 Gb, and the *A. fulica* genome assembly has a total length of 1.86 Gb (Guo et al., 2019). The coverage rate of the *A. immaculata* genome (95%) is 8% higher than that of *A. fulica* (87%). The N50 and N90 lengths of the *A. immaculata* contigs were increased by 5.3- and 4.9-fold compared to *A. fulica*, respectively. With the better assembly quality, ~6,500 additional gene models were annotated in *A. immaculata* compared to *A. fulica*. The high quality of the assembly and annotation of *A. immaculata* provide an effective resource for genomic research on giant African snails.

3.2 | Signs of adaptive evolution in giant African snails

To gain insights into the evolution of giant African snails (*A. immaculata* and *A. fulica*), a total of 292,034 reference genes from 10

molluscan genomes were clustered into 17,949 OGs containing at least two genes each. A phylogenetic tree was then constructed based on 229 high-confidence single-copy orthologous genes. The results showed that *A. immaculata* diverged from *A. fulica* 21 Ma, from *B. glabrata* (Panpulmonata) 174 Ma, from *Aplysia californica* (Euopisthobranchia) 205 Ma and from *Pomacea canaliculata* (Caenogastropoda) 416 Ma (Figure 2a). Using CAFE (De Bie et al., 2006), we identified 2,225 expanded OGs in giant African snails (both *A. immaculata* and *A. fulica*). The functions of these OGs are related mainly to signal transduction; the endocrine, immune, and nervous systems; and reproduction (Figure S6). Additionally, 836 OGs were found exclusively in the lineage of giant African snails, and these mainly function in neurohormonal regulation and mucus synthesis and include genes for proteins such as acetylcholine receptors, pannexin, tenascin, adrenocorticotropic hormone receptors, neuropeptide receptors, the 5-hydroxytryptamine receptor, mucin, heparan sulphate glucosamine 3-*O*-sulfotransferase and proteophosphoglycan (Figure 2b; Data S1). The expanded and lineage-specific OGs are enriched in functions which may enhance terrestrial living, suggesting that they may play a role in the adaptation and invasion of giant African snails.

Our high-quality genome assembly enables a comprehensive analysis of TEs, which play multiple roles in driving genome evolution in eukaryotes (Feschotte & Wessler, 2002). We identified a total of 954 Mb of repetitive sequences in the assembled *A. immaculata* genome and 1,366 Mb in *A. fulica* (Figure 2c). We next analysed the divergence rate of each class of TE among the available molluscan genomes. Notably, the TE class of DNA transposons showed a specific peak at a divergence rate of 4%–6% for four invasive species, *A. immaculata*, *A. fulica*, *Pomacea canaliculata* (Carlsson et al., 2004) and *C. gigas* (Troost, 2010), indicating a recent rapid expansion of DNA transposons (Figure 2d). We identified 9,291 genes in regions that contain DNA transposons distributed within the specific divergence peak. Based on KEGG annotation, these genes were found to be mainly enriched in signal transduction; the endocrine, immune and nervous systems; and reproduction (Figure S7). TEs are powerful facilitators of evolution that generate the "evolutionary potential" to introduce small adaptive changes within a lineage, and the importance of TEs in stress responses and adaptation has been reported in numerous studies (Hua-Van et al., 2011; Werren, 2011). The recent explosion of DNA TEs in the giant African snail genomes could have played important roles in promoting their potential plasticity in stress adaptation.

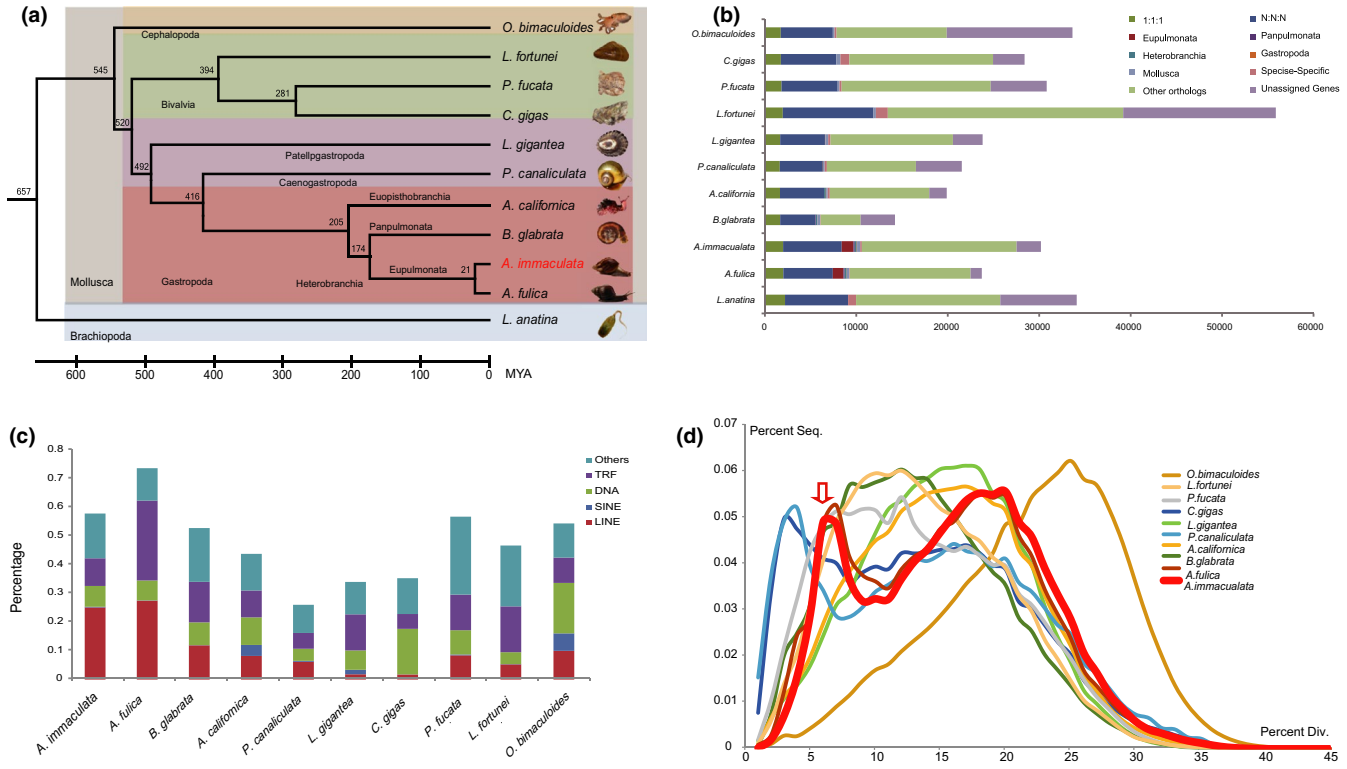

FIGURE 2 Evolutionary analysis of *Achatina immaculata*. (a) Phylogenetic placement of *A. immaculata* within the dated tree of molluscs. The estimated divergence time is shown at each branch point, and each lineage is shown in different colour blocks. *A. immaculata* is highlighted in red. The scale of the dated tree is indicated at the bottom in million years ago. (b) Categories of orthologous and paralogous genes in various phyla and species. (c) Classification and contents of repetitive sequences in *A. immaculata* and other molluscs. (d) Distribution of divergence rate for the class of DNA transposons in mollusc genomes. The divergence rate was calculated by comparing all TE sequences identified in the genome to the corresponding consensus sequence in each TE subfamily. The arrow indicates that both *A. immaculata* and *Achatina fulica* had a recent explosion of TEs at a divergence rate of ~5%

3.3 | Whole genome duplication in the Sigmurethra–Orthurethra branch

Using the chromosome-level assemblies, we searched for macrosynteny based on homologous gene pairs among four molluscan genomes: *A. immaculata*, *A. fulica*, *Pomacea canaliculata* and *Pinctada fucata*. Our chromosome-level macrosynteny revealed a WGD event shared by the two giant African snails, *A. immaculata* and *A. fulica*. The 31 chromosomes of *A. immaculata* could be divided into 14 groups with the preservation of correspondence (Figures 3a and 4a; Figure S8), and the same situation was found in *A. fulica* (Figures 3b and 4b; Figure S8). In *A. immaculata*, 2,092 homologous gene pairs with mutual best BLASTP hits were located on the corresponding chromosomes, and the number was 2,364 for *A. fulica* (Table S6). The genome comparison of the giant African snails to *Pomacea canaliculata* and *Pinctada fucata*, both of which have 14 haploid chromosomes, revealed that the karyotype doubled in the lineage leading to giant African snails (Figure 3c–e). The chromosomes of *A. immaculata* and *A. fulica* showed a 1-to-1 correspondence relationship (Figure 3c; Figure S8), while most chromosomes from *Pomacea canaliculata* (Figure 3d; Figure S8) and *Pinctada fucata* (Figure 3e; Figure S8) shared macrosynteny with two corresponding chromosomes from giant African snails. The colinearity blocks identified by MCSCANX based on BLASTP hits in the corresponding chromosome groups also suggested that WGD events have occurred in the two giant African snail species (Figures S9 and S10). In the gene age distribution plot of homologous gene pairs, a specific Ks peak shared by *A. immaculata* and *A. fulica* was observed, which is consistent with a WGD (Figure 3f). The duplication of Hox gene clusters is a powerful clue that led to the discovery of ancient WGDs in vertebrates (Amores et al., 1998), so we further compared the giant African snail genomes to those (*Pomacea canaliculata*, *Lottia gigantea*, *C. gigas*, *Pinctada fucata* and *O. bimaculoides*) with a single Hox cluster and no evidence of WGD. Duplicated Hox gene clusters with specific rearrangements were found in the two giant African snail genomes (Figure 3g). In conclusion, the chromosomal macrosynteny, colinearity blocks, Ks peak and Hox gene clusters collectively suggest that a WGD event occurred before the speciation of giant African snails.

The timing of WGD events has been reported to show a significant correlation with specific geological and global climatic change (Sessa, 2019). Based on the gene age distribution of homologous gene pairs and the diverge time of *B. glabrata* and the giant African snails (~174 Ma), we deduced that the timing of the WGD event was ~ 70 Ma (Figure 3f). An earlier estimation suggested the occurrence of a WGD event at the Sigmurethra–Orthurethra branch within Stylommatophora (Figure S11) by comparing chromosome numbers among closely related molluscs (Hallinan & Lindberg, 2011). The speciation age of the Sigmurethra–Orthurethra branch was reported to be 65 Ma (Hallinan & Lindberg, 2011), which is after our deduced timing of the WGD (~70 Ma), indicating that the WGD reported here arose before the Sigmurethra–Orthurethra divergence and it may be a common signature shared by all other Sigmurethra and Orthurethra species (Figure 3f; Figure S11). The timing of the WGD of the Sigmurethra–Orthurethra branch was also close to the Cretaceous–Tertiary (K-T) mass extinction (~66 Ma), in which global climate change caused the extinction of 60%–70% of all plant and animal life, including most molluscs (Petersen et al., 2016). In plants, the K-T mass extinction is considered to be a shared common causal factor in the genome-wide doubling in diverse angiosperm lineages (Fawcett et al., 2009). It has also been previously suggested that polyploidization in animals is correlated with periods of unstable environments (Mable et al., 2011). The WGD in the Sigmurethra–Orthurethra branch is proposed to have provided the ecological adaptability and genome plasticity that allowed these taxa to survive the K-T mass extinction.

As the terrestrial area expanded during the K-T mass extinction due to Maastrichtian sea-level regression (Marshall & Ward, 1996), the WGD is thought to have both promoted the adaptability of land snails to terrestrial ecosystems and promoted speciation diversity. The functions of WGD-derived homologous gene pairs are significantly enriched in biological regulation, signal transduction, energy generation and the response to a stimulus (Figure S12). These functions are closely related to terrestrial living, indicating that the retained WGD-associated genes increased the terrestrial ecological tolerance of giant African snails.

3.4 | Expansion of haemocyanins and zinc metalloproteinases improves terrestrial respiratory function

The innovation of respiratory gas exchange is the signature of the A–T transition, which was one of the most conspicuous evolutionary events to have occurred on Earth (Ashley-Ross et al., 2013). The evolution of the oxygen transportation system allowed land snails to utilize O_2 from the air far more efficiently than aquatic molluscs (Hsia et al., 2013). Additionally, an advanced system is needed to eliminate the accompanying oxidative stress to maintain O_2 homeostasis. However, the underlying mechanisms of O_2 transport and antioxidation are less well understood.

Within the six species of Gastropoda (*A. immucualta*, *A. fulica*, *B. glabrata*, *Aplysia californica*, *Pomacea canaliculata*) and Cephalopoda (*O. bimaculoides*) we analysed, most are dependent on haemocyanin (Markl, 2013), with the exception of *B. glabrata*, a species that uses haemoglobin for O_2 transport (Lieb et al., 2006). Based on an analysis of orthologous genes and phylogenetics, haemocyanin genes were observed in all six species (including *B. glabrata*), indicating that *B. glabrata* not only possesses haemoglobin genes but also possesses haemocyanin genes. However, the *B. glabrata* haemocyanin is composed of only six functional units, compared to eight functional units in the other five species, indicating some level of function loss (Lieb et al., 2006; Markl, 2013; Peña & Adema, 2016). Notably, there are four haemocyanin genes in both *A. immaculata* and *A. fulica*, which is twice as many as in any other Gastropoda species (Figure 5a). Moreover, the four homologous genes from *A. immaculata* and *A. fulica* are located on

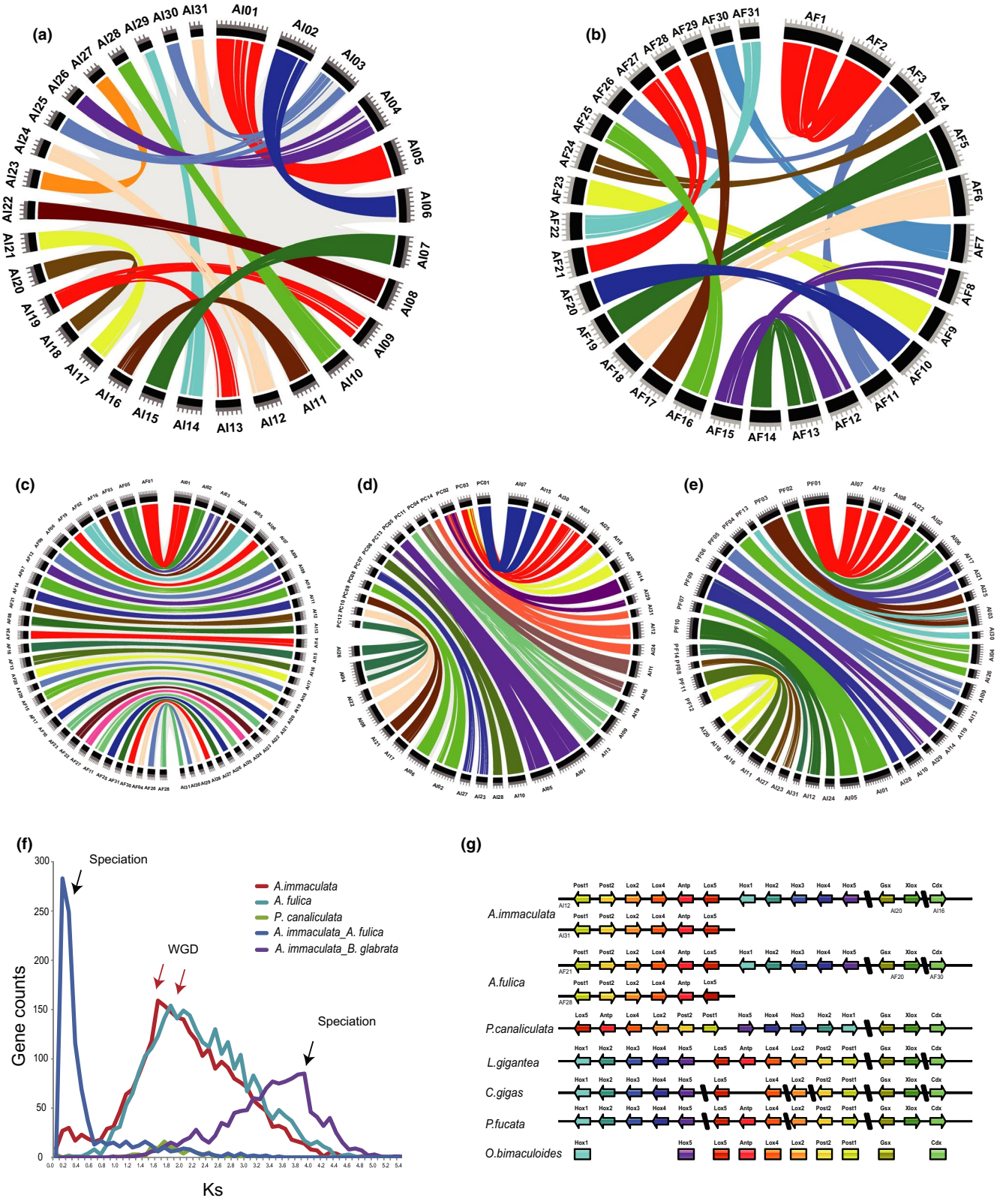

FIGURE 3 Whole genome duplication shared by *Achatina immaculata* and *Achatina fulica*. CIRCOS plots showing the homologous gene pairs in chromosome-level microsynteny among *A. immaculata* (a) and *A. fulica* (b) individually, as well as chromosomes between *A. immaculata* and *A. fulica* (c), *A. immaculata* and *Pomacea canaliculata* (d), *A. immaculata* and *Pinctada fucata* (e). (F) Gene age distribution of Ks values calculated from orthologous gene pairs, among *A. immaculata*, *A. fulica*, and *Pomacea canaliculata* individually, as well as between *A. immaculata* and *A. fulica*, *A. immaculata* and *Biomphalaria glabrata*. Black arrows indicate speciation, and red arrows indicate WGD. (g) Comparison of the catalogue of Hox genes. *A. immaculata* and *A. fulica* have more Hox genes than the other molluscs, because of the retained genes on the duplicated chromosome. This evidence collectively supports a whole genome duplication event shared by *A. immaculata* and *A. fulica*, which is not found in other molluscs

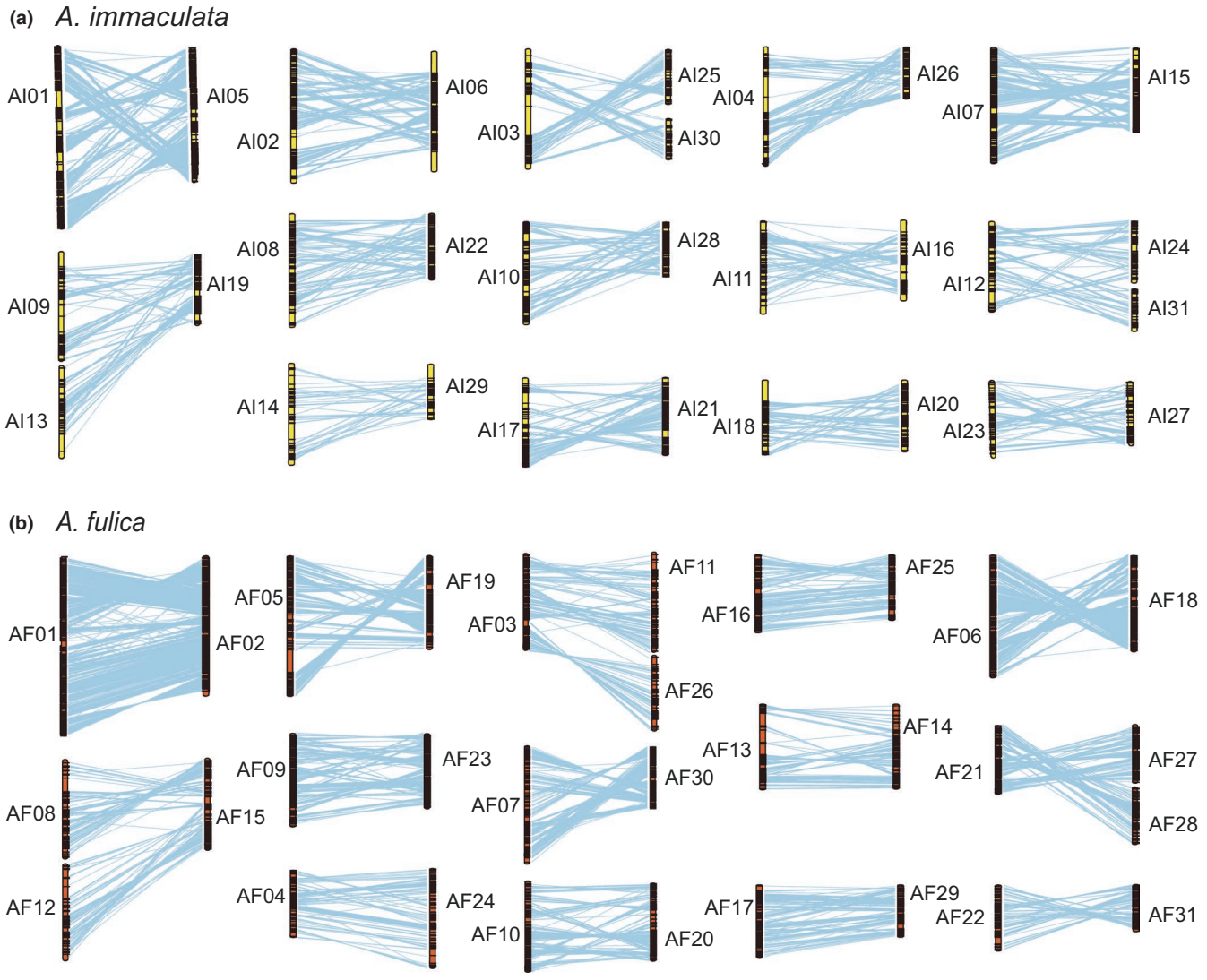

FIGURE 4 Chromosome relationship and karyotype inference. Clustering of chromosomes for *Achatina immaculata* (a) and *Achatina fulica* (b). Each group corresponds to an orthologous chromosome pair derived from WGD, with the links between chromosome pairs representing the mutual best-hit gene pairs

two chromosomes that were derived from the WGD, whereas the haemocyanin genes in the other species without WGD are located on only one chromosome or scaffold (Figure 5b; Figures S13 and S14). Thus, the doubling of the haemocyanin gene number through WGD may have increased the ability of giant African snails to transport O_2, which allowed them to adapt to land living.

Reactive oxygen species (ROS) are generated during O_2 metabolism, and excessive ROS cause oxidative stress and trigger inflammation (Shen et al., 2010). However, antioxidant enzymes protect the host from excess oxidative damage. Superoxide dismutase (SOD), acid phosphatase (AP) and glutathione *S* transferase (GST) genes were identified in giant African snails, with gene numbers comparable to those of other molluscan species (Table S7). In addition to antioxidant enzymes, metalloproteinases have the ability to hydrolyse fibronectin to reduce the injuries caused by ROS (Faa et al., 2008; Kruse et al., 2004). The number of zinc metalloproteinase genes in the genomes of both giant African snail species was largely expanded, to 11 in *A. immaculata* and nine in *A. fulica*, compared to only one or two in the other species (Figure 5c; Figure S15). Importantly, the zinc metalloproteinase genes of the two species are located in two syntenic clusters on homologous chromosomes, while no homologous genes were found on the corresponding duplicated chromosomes resulting from WGD. The finding that the same location pattern was shared by the two species indicates that these genes were tandemly duplicated after the WGD event but before the speciation of *A. immaculata* and *A. fulica* (Figure 5d). Furthermore, the expression levels of all zinc metalloproteinases in the hepatopancreas and blood were higher than those in other tissues (Figure S16), which is consistent with the importance of the antioxidative functions of these two tissues (Faa et al., 2008). Therefore, the expansion of zinc metalloproteinases genes after WGD might have played an important role in the defence of giant African snails against the damage resulting from oxidative stress and inflammation.

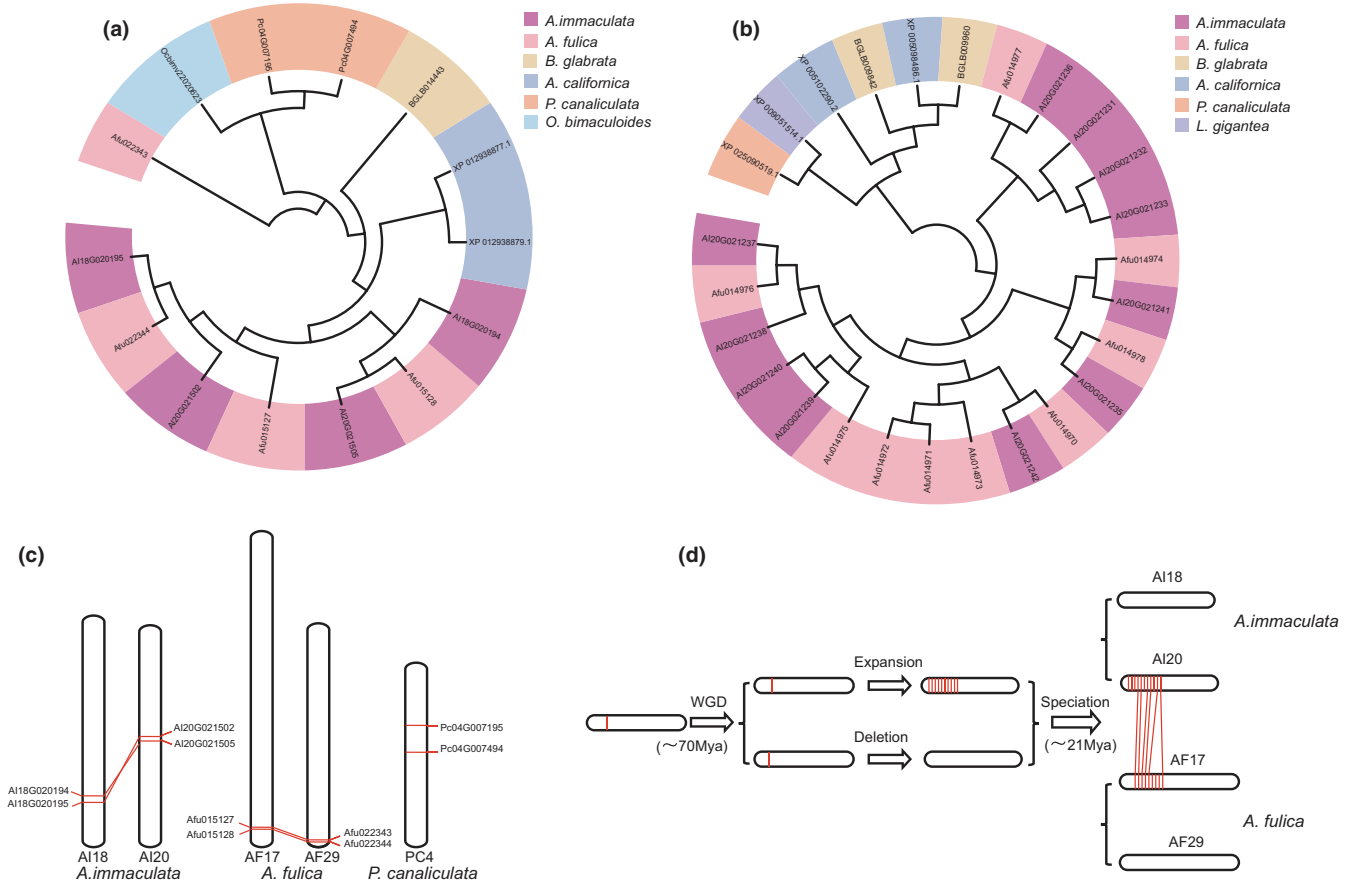

FIGURE 5 Oxygen transport and anti-oxidization in respiratory function. (a) The phylogenetic relationship of haemocyanin genes among molluscs. Protein IDs with different colours indicate different species. (b) The phylogenetic relationship of zinc metalloproteinases among gastropoda, with same style to that of (a). (c) Comparison of haemocyanin locations in the chromosomes of *Achatina immaculata*, *Achatina fulica* and *Pomacea canaliculata*. Haemocyanin genes of *A. immaculata* and *A. fulica* were distributed in two chromosomes, whereas those of *Pomacea canaliculata* were only in one contig. (d) The evolution process of zinc metalloproteinase genes in *A. immaculata* and *A. fulica*. The shared ancient ancestor might have only one zinc metalloproteinase gene located on one chromosome, a WGD (~70 Ma) event doubles the chromosome and gene number, followed by gene expansion on one chromosome and deletion on the corresponding chromosome, resulting in a tandem cluster of 11 genes and nine genes in *A. immaculata* and *A. fulica*, respectively

3.5 | Glucose homeostasis and ureagenesis benefit survival in aestivation

Aestivation in land animals is a particular long-term torpid state that occurs in response to the extreme environmental conditions of summer that include desiccation, high temperatures and starvation (Storey, 2002). During aestivation, giant African snails seal their shell aperture with epiphragm, and their body remains within the solid shell for several months, isolated from feeding and excretion (Hiong et al., 2005). How the snails regulate blood glucose homeostasis and eliminate toxins during aestivation is still not well understood.

Because blood sugar is the only directed energy source for aestivating gastropods (Rahman & Raut, 2012), without sugar intake during aestivation, blood glucose homeostasis is mainly achieved via the exploitation of endogenous resources and a reduction in energy expenditure (Ip & Chew, 2010). In this study, we analysed the expression level changes in the genes for all 10 rate-limiting enzymes in the metabolism of glucose homeostasis (Table S8). Among them, genes that encode pyruvate carboxylase (PC, EC:6.4.1.1) and phosphoenolpyruvate carboxykinase (PCK, EC:4.1.1.31) are DEGs involved in gluconeogenesis. PC is encoded by one gene that shows significantly increased expression during aestivation ($\log_2$FC > 1, q-value < 0.01), while PCKs are encoded by two homologous genes derived from WGD, and the expression level of one of these homologous genes is significantly increased during aestivation ($\log_2$FC > 1, q-value < 0.01). However, the consumption of glucose through glycolysis and the tricarboxylic acid (TCA) cycle is minimized during aestivation (Bell et al., 2012). Genes that encode hexokinase (HK, EC:2.7.1.1) and pyruvate kinase (PK, EC:2.7.1.40) are DEGs involved in glycolysis, and the gene for the oxoglutarate dehydrogenase complex (OGDC, EC:2.3.1.61) is a DEG involved in the TCA cycle (Bulutoglu et al., 2016) (Table S8). HK, PK and OGDC are each encoded by single genes with significantly decreased expression levels ($\log_2$FC < −1, q-value < 0.01). These results indicate that blood glucose homeostasis during aestivation is achieved via both the upregulation of gluconeogenesis and the downregulation of glycolysis and

the TCA cycle (Figure 6a), which is consistent with previous reports for animals in aestivation (Bell et al., 2012).

The oxidative deamination of amino acids generates toxic ammonium (NH_3, NH_4^+), which can cause immunosuppression and neurotoxic effects (Hermenegildo et al., 2000; Oja et al., 2017). During aestivation, giant African snails convert toxic ammonium into nontoxic urea (Hiong et al., 2005; Onadeko & Egonmwan, 2017), and the accumulation of nontoxic urea in the body has also been proposed to play roles in the reduction of evaporative water loss and the reutilization of nitrogen resources (Ip & Chew, 2010). In this study, we analysed the expression profiles of genes for three important enzymes: carbamoyl phosphate synthetase (CPS, EC:6.3.5.5), argininosuccinate synthetase (ASS, EC:6.3.4.5) and arginase (EC:3.5.3.1) (Diez-Fernandez et al., 2015). All of the genes encoding these enzymes were significantly upregulated in the aestivation group compared with the normal group ($\log_2$FC > 1, q-value < 0.01) (Table S8), providing insight into the mechanism by which ammonia is transformed into urea during the aestivation of giant African snails (Figure 6b).

3.6 | Expanded mucus-related gene families reinforce immune defence

As one of the most successful colonizers of terrestrial environments within Mollusca, giant African snails may possess advanced molecular mechanisms for host defence against biotic and abiotic stresses (Begum et al., 2000). Despite the fact that they lack the canonical vertebrate immunoglobulins, these snails have developed diverse repertoires of receptors, regulators and effectors (Mukherjee et al., 2017). Remarkably, the mucus system was reported to be one of the unique evolutionary modalities (Begum et al., 2000). To investigate the genes and pathways involved in the immune and stress responses, we characterized the mucus-related gene families in the genome of *A. immaculata* (Figure 7a; Table S9). The transcriptomes of the haemocytes after exposure to different stimuli (LPS, PGN, poly I:C and GLU) were also analysed to address the potential roles of these genes (Figure S17).

The mucus of giant African snails is mainly composed of glycosaminoglycans and glycoproteins, and its antimicrobial properties were first reported in the 1980s (Iguchi et al., 1982). To investigate the glycosaminoglycan synthesis pathway and the glycoproteins in the mucus of *A. immaculata*, we identified genes for 11 heparan sulphate glucosamine 3-*O*-sulfotransferases (OSTs) that function in glycosaminoglycan synthesis and 99 mucin genes that encode the major glycoprotein in mucus. OST gene numbers were found to be expanded in the Panpulmonata (10 in *A. immaculata*, 11 in *A. fulica*, 10 in *B. glabrata*) and the Cephalopoda (13 in *O. bimaculoides*), in which they were twice as abundant as than in other molluscs (Table S10). The expansion of OSTs in the Panpulmonata occurred in the subfamilies OST-1 and OST-3B (Figure 7b). Acharan sulphate, a product of OSTs, is a primary constituent of mucus that functions in phenotypic and functional maturation of immune cells (Kim et al., 2007). Thus, the expansion of the OST genes is therefore closely related to the GAT/GALT (galactosyltransferase) abundant surface mucus. The increased expression levels of OST genes observed after LPS stimulus (FPKM fold change = 82.41, q-value = 0.02) suggest a possible immune role in the response to Gram-negative bacteria (Figure 7a; Table S11). The other

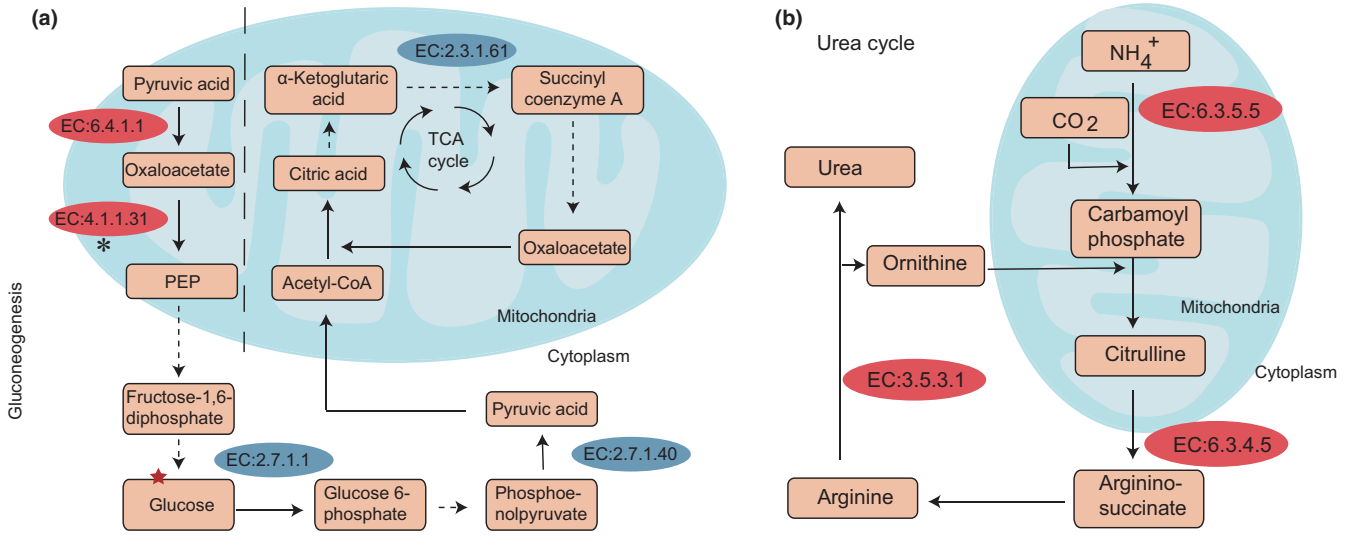

FIGURE 6 Flowchart of glucose homeostasis and urea-cycle in aestivation of *Achatina immaculata*. (a) The key processes and related gene expression change of gluconeogenesis, glycolysis and TCA cycle in the aestivation compared with normal group. The left part represents gluconeogenesis, the bottom right area in the cytoplasm represents glycolysis, while the right part in the mitochondria represents the TCA cycle. A dotted-line arrow indicates a number of steps, while a solid arrow indicates one step. A rectangle indicates production, while an ellipse indicates enzymes, with red and blue representing up-regulated and down-regulated respectively. Rectangle with five-pointed red star represents glucose. Ellipse with an asterisk indicates that this enzyme had two genes derived from WGD. (b) The key processes and related gene expression change of the ornithine–urea cycle in the aestivation compared with normal groups; style is as in (a)

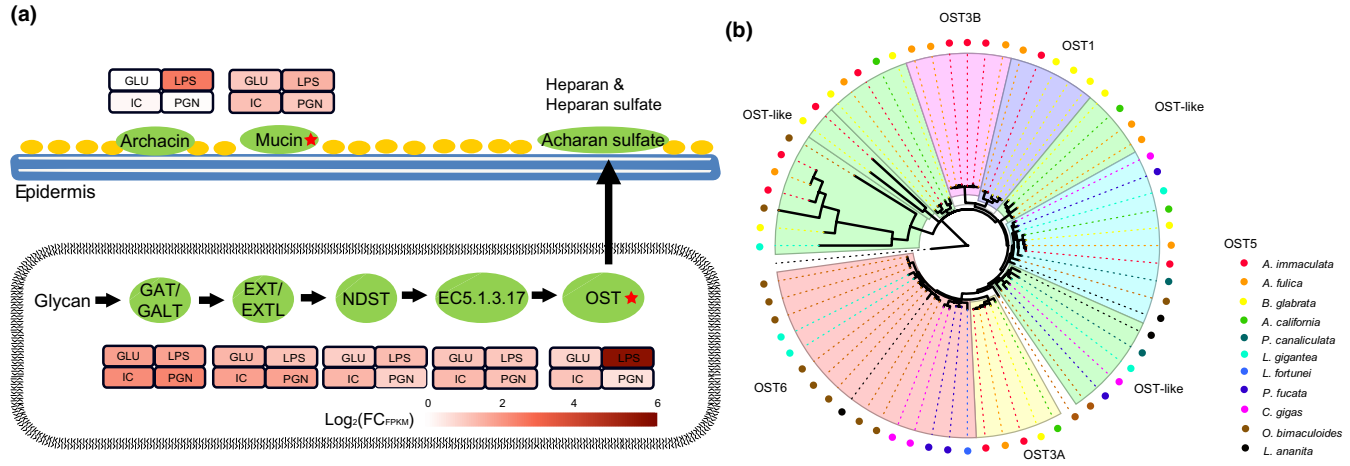

FIGURE 7 The mucus-related gene repertoire of *Achatina immaculata* and the expansion of OST. (A) The immunity of *A. immaculata* is presented in the mucus system. The mucus system covers epidermis, mainly including archacin, mucin and acharan sulphate secreted through the heparin and heparin sulphate synthesis pathway, presented with a green ellipse. The heparin and heparin sulphate synthesis pathway includes EXT/EXTL (acetylglucosaminyltransferase), GAT/GALT (galactosyltransferase), GLCE (heparosan-*N*-sulfate-glucuronate 5-epimerase), NDST (*N*-deacetylase and *N*-sulphotransferase) and OST (glucosamine 3-*O*-sulfotransferases). The FPKM fold change post-lipopolysaccharide (LPS), peptidoglycan (PGN), poly (I:C) (IC) and β-glucan (GLU) compared with the control group is indicated in a heatmap next to the gene ellipse. The scale at the bottom shows a $\log_2$ transform. (b) The expansion of OST genes. Phylogenetic analyses were performed using MEGA7 through maximum likelihood. OST genes from different species are indicated in CIRCUS plots filled with different colours

enzymes involved in the biosynthesis of acharan sulphate also showed increased expression patterns after challenge with, for example, GAT/GALT and EXT/EXTL (acetylglucosaminyltransferase) (Table S11). The mucin genes were also found to be expanded on the Stylommatophora branch (99 for *A. immaculata*, 71 for *A. fulica*), showing a number approximately three times greater than in other molluscs. Snail mucins are reported to function in the antibacterial process (Portela et al., 2013), so the expansion of mucins indicates that the mucus of Stylommatophora species contains more mucin proteins to defend against microorganism infection (Figure S18, Table S10). Increased expression patterns of mucin genes were mainly detected in the LPS (FPKM fold-change = 1.27, *q*-value = 0.02) and IC (FPKM fold-change = 1.72, *q*-value = 0.04) treatment groups (Figure 7a; Table S11), implying their immune roles against Gram-negative bacteria (LPS), and viruses and fungi (IC). In addition, a significant increase in the level of mRNA of genes involved in achacin synthesis, an antibacterial glycoprotein, was found after LPS stimulation (FPKM fold change = 4.66, *q*-value = 0.01, Table S11).

A total of 34 mucins from 19 OGs and three OSTs from two OGs were found exclusively on the Stylommatophora branch (Data S1). Their homologous pseudogenes were located on the corresponding duplicated chromosomes, indicating that these genes were generated before WGD, doubled in number through WGD and were transformed into pseudogenes after WGD (Tables S12 and S13). In addition to their roles in immunity, the functions of mucins and the products of OSTs in wound healing, locomotion and other terrestrial-living-related processes have been observed (Adikwu & Alozie, 2007; Denny, 1979), indicating that their expansions before WGD may have played important roles in the A-T transition.

4 | DISCUSSION

Whole genome duplication provides evolutionary novelties that increase environmental adaptation and species diversity, although this process is regarded as being rare in animals because it is hampered by sex determination (Wertheim et al., 2013). As one of the best-known simultaneous hermaphrodites (Barrows, 2011), the giant African snail possesses the ability for autofecundation under certain circumstances, thereby overcoming the evolutionary dead end resulting from WGD. Based on high-quality genome assemblies from two species of giant African snails, we report a high-credibility WGD on the Sigmurethra–Orthurethra branch of molluscs (order Stylommatophora) with collective evidence including macrosynteny data, colinearity blocks, the gene-age distribution and Hox gene clusters. In particular, chromosome-level macrosynteny was used for identification of WGD, a method that was highly recommended in a recent paper to avoid possible false-positive results (Nakatani & McLysaght, 2019). To the best of our knowledge, this WGD analysis based on chromosome-level macrosynteny is the first to be reported in Mollusca. Compared with other reported WGDs in protostome invertebrates at the scaffold level, such as chelicerates, this chromosome-level WGD analysis can provide more detailed information and a deeper understanding of the process. In contrast to the two WGDs found in chelicerates scaffolds (~135 and ~430 Ma), the molluscan WGD was a relatively more recent event with an estimated timing of ~70 Ma. In contrast to ancient WGDs, the identification of this recent event will help to reveal the adaptive mechanisms and consequences of WGD, especially regarding subgenome divergence, providing more traceable genomic clues (Soltis et al., 2009). In most cases, WGD occurs only under remarkably uncommon

circumstances and provides a driving force during subsequent evolution. As an example, the WGDs identified in a cluster of angiosperm plants share a common causal factor corresponding with the K-T mass extinction (Vanneste et al., 2014). The timing of the WGD that occurred in chelicerates at ~430 Ma and that in Sigmurethra–Orthurethra at ~70 Ma is close to the Ordovician–Silurian (O-S) extinction and the K-T mass extinction, respectively, indicating that the invertebrate WGD was also connected to mass extinction events. Furthermore, it has been reported that WGD is followed by a substantial increase in morphological complexity and species numbers (Van de Peer et al., 2009). Therefore, the species richness and wide range of adaptations of invertebrates imply that more undiscovered WGDs exist in this clade and might be revealed as available genomic resources increase.

The A-T transition from water-living to land-living was a milestone in the evolutionary history on Earth (Volkmann & Baluška, 2006). The Stylommatophora lineage, which originated from a marine ancestor, and breathes through permanently air-filled lungs, successfully completed terrestrial adaptation ~100–150 Ma (Kameda & Kato, 2011). With all of the associated anatomical and physiological changes required for terrestrial living (Barker, 2001), Stylommatophora is considered to provide good study material for the investigation of the A-T transition. The WGD event and sea-level regression that occurred during the K-T boundary are expected to have resulted in greater adaptability to terrestrial ecosystems (Marshall & Ward, 1996). To reveal the mechanisms underlying the A-T transition and the influence of WGD, we investigated gene families related to respiration, aestivation and immune defence in the giant African snail. The terrestrial adaptation of Stylommatophora species was initiated before the WGD event that occurred ~70 Ma, and we found that several mucus-related gene families expanded early in the Stylommatophora lineage (100–150 Ma); these genes include the mucin and OST families, which function in water retention, immune defence and wound healing (Adikwu & Alozie, 2007; Denny, 1979). WGD has been proposed to provide functional redundancy and mutational robustness to increase the rates of evolution and adaptation (Crow & Wagner, 2006). The genes encoding haemocyanins, proteins involved in O_2 transport, and PCK (phosphoenolpyruvate carboxykinase), an enzyme involved in gluconeogenesis, were doubled and retained after WGD, enhancing the capacity for gas exchange and glucose homeostasis during aestivation. The extra chromosome copy resulting from WGD might limit the risk of genetic variation occurring in one copy of the duplicated chromosome. In the post-WGD period, zinc metalloproteinase genes were highly tandemly duplicated, facilitating the protection of tissue against damage by ROS that arise from respiration. This evidence indicates that although the A-T transition of the giant African snail was not initially driven by WGD, WGD could have facilitated its terrestrial adaptation by providing additional genomic resources, thus increasing the survival rate in the dramatic transition from water to land.

Biological invasion has become an increasingly serious problem worldwide, as the international flow of people and goods has increased significantly over the years (Wan et al., 2017). With its rich species diversity and strong environmental adaptability, Mollusca is among the phyla with the greatest numbers of invasive species, such as the golden apple snail (*Pomacea canaliculata*) and giant African snail (*A. fulica* and *A. immaculata*), which are listed among the top 100 global invasive species, although the mechanisms that facilitate invasion by molluscs are not yet clear. During genetic bottlenecks, invasive species can still adapt to new habitats and expand their populations, which is referred to as the "genetic paradox of invasive species" (Frankham, 2005). Previous reports have shown that TEs are powerful facilitators of rapid adaptation that generate "evolutionary potential" by introducing stress-induced changes in invasive species (Stapley et al., 2015). In this study, recent TE explosions were observed in all four invasive molluscs but were absent in the other non-invasive molluscs. Genes located near recently emerged TEs were enriched in functions such as stress responses, which is consistent with our previous findings in the golden apple snail (Liu et al., 2018), indicating that the recent TE explosion might be a common genetic force contributing to biological invasion. In addition, WGD has been proposed to provide robustness of genetic variation and redundant gene resources for the rapid evolution of novel functions, driving phenotypic complexity and ecological adaptation. Therefore, WGD may be another explanation for the genetic paradox of biological invasion, and species exhibiting WGD may be more invasive.

In summary, our results have revealed a WGD that occurred on the Sigmurethra–Orthurethra branch within the Stylommatophora, providing genomic evidence for the A-T transition in Mollusca, and we propose WGD as a potential mechanism that contributes to biological invasion. On the other hand, the genome sequence assembly for the giant African snail *A. immaculata* will enable us to develop more environmentally friendly and efficient control measures using species-specific gene targets, benefiting the protection of agricultural crops and the ecological environment as well as the prevention of human disease caused by zoonotic parasites. However, giant African snails are considered to be a high-protein food source in some parts of the world, especially in Africa and Asia, and the invasive characteristics of giant African snails, such as their rapid growth, high production rate and ability to survive harsh conditions, give them the potential to be cultivated as an economic species. Thus, the genome sequence of the giant African snail *A. immaculata* provides a powerful platform to enable the genetic improvement of this species through breeding, turning "waste" into wealth.

ACKNOWLEDGEMENTS

The work was funded by the National key research and development programme of China (2016YFC1200600), the Agricultural Science and Technology Innovation Program & The Elite Young Scientists Program of CAAS, Fundamental Research Funds for Central Non-profit Scientific Institution (No. Y2017JC01), the Agricultural Science and Technology Innovation Program Cooperation and Innovation Mission (CAAS-XTCX2016), Fund of Key Laboratory of Shenzhen (ZDSYS20141118170111640), Natural Science Foundation of Shenzhen (No. JCYJ20190813120401662), as well as National Natural Science Foundation of China (Grant No. 31901950).

CONFLICT OF INTEREST

The authors declare no competing interests.

AUTHOR CONTRIBUTIONS

W.F. and W.Q.Q. conceived and led the project; C.H.L. and Y.W.R. prepared DNA and RNA for sequencing; C.H.L. performed genome assembly, annotation, evolution, whole genome duplication and immune analysis; Y.W.R. performed respiration and aestivation analysis; W. F., W.Q.Q., C.H.L. and Y.W.R. wrote and revised the manuscript.

DATA AVAILABILITY STATEMENT

Data relating to the findings of this study are available within the paper and the Appendix S1. A summary for this article is available in the Appendix S1. Source data are provided as Data S1. All the raw sequencing data generated during this study have been deposited at NCBI as a BioProject under accession PRJNA561271. Genomic and transcriptome sequence reads have been deposited in the SRA database with BioSample: SAMN12612888. The Whole Genome Shotgun project of *A. immaculata* has been deposited at DDBJ/ENA/GenBank under accession WNKJ00000000. The version described in this paper is version WNKJ00000000. The genome assemblies and annotation files are available at ftp://ftp.agis.org.cn/~fanwei/Achatina_immaculata_genome/.

Data citations: *Achatina fulica*: http://dx.doi.org/10.5524/100647; *Aplysia californica*: https://www.ncbi.nlm.nih.gov/assembly/?term=Aplysia+californica; *Biomphalaria glabrata*: https://www.ncbi.nlm.nih.gov/assembly/GCF_000457365.1/; *Limnoperna fortunei*: https://www.ncbi.nlm.nih.gov/assembly/GCA_003130415.1/; *Lingula anatina*: https://www.ncbi.nlm.nih.gov/assembly/?term=Lingula+anatina; *Lottia gigantea*: https://www.ncbi.nlm.nih.gov/assembly/GCF_000327385.1/; *Octopus bimaculoides*: https://www.ncbi.nlm.nih.gov/assembly/GCF_001194135.1/; *Crassostrea gigas*: https://www.ncbi.nlm.nih.gov/assembly/GCA_000297895.2; *Pinctada fucata*: https://www.ncbi.nlm.nih.gov/assembly/GCA_002216045.1/; *Pomacea canaliculata*: https://www.ncbi.nlm.nih.gov/assembly/GCF_003073045.1/

ORCID

Wei Fan https://orcid.org/0000-0001-5036-8733

REFERENCES

Adema, C. M., Hillier, L. W., Jones, C. S., Loker, E. S., Knight, M., Minx, P., & Wilson, R. K. (2017). Whole genome analysis of a schistosomiasis-transmitting freshwater snail. *Nature Communications*, *8*, 15451.

Adikwu, M. U., & Alozie, B. U. (2007). Application of snail mucin dispersed in detarium gum gel in wound healing. *Scientific Research and Essays*, *2*(6), 195–198.

Albertin, C. B., Simakov, O., Mitros, T., Wang, Z. Y., Pungor, J. R., Edsinger-Gonzales, E., & Rokhsar, D. S. (2015). The octopus genome and the evolution of cephalopod neural and morphological novelties. *Nature*, *524*(7564), 220–224.

Amores, A., Force, A., Yan, Y.-L., Joly, L., Amemiya, C., Fritz, A., & Wang, Y.-L. (1998). Zebrafish hox clusters and vertebrate genome evolution. *Science*, *282*(5394), 1711–1714.

Aoki, K. F., & Kanehisa, M. (2005). Using the KEGG database resource. *Current Protocols in Bioinformatics*, *11*(1), 1.12.1–1.12.43. https://doi.org/10.1002/0471250953.bi0112s11

Aplysia Genome Project (2009). Broad Institute. Vertebrate Biology Group. https://www.broadinstitute.org/aplysia/aplysia-genome-project

Ashley-Ross, M. A., Hsieh, S. T., Gibb, A. C., & Blob, R. W. (2013). Vertebrate land invasions-past, present, and future: an introduction to the symposium. *Integrative and Comparative Biology*, *53*(2), 192–196.

Barker, G. M. (2001). Gastropods on land: phylogeny, diversity and adaptive morphology. In G. M. Barker (Ed.), *The Biology of terrestrial molluscs* (pp. 1–146). Wallingford, NY: CABI. https://doi.org/10.1079/9780851993188.0001

Barrows, E. M. (2011). Organisms. In E. M. Barrows (Ed.), *Animal behavior desk reference: a dictionary of animal behavior, ecology, and evolution*, 3rd ed. (pp. 680–735). Boca Raton, FL: CRC Press.

Begum, N., Matsumoto, M., Tsuji, S., Toyoshima, K., & Seya, T. (2000). The primary host defense system across humans, flies and plants. *Current Trends in Immunology*, *3*, 59–74.

Bell, R. A., Dawson, N. J., & Storey, K. B. (2012). Insights into the in vivo regulation of glutamate dehydrogenase from the foot muscle of an estivating land snail. *Enzyme Research*, *2012*, 317314.

Benson, G. (1999). Tandem repeats finder: a program to analyze DNA sequences. *Nucleic Acids Research*, *27*(2), 573–580.

Benton, M., Donoghue, P., & Asher, R. (2009). Calibrating and constraining molecular clocks. *The Timetree of Life*, *35*, 86.

Biscotti, M. A., Gerdol, M., Canapa, A., Forconi, M., Olmo, E., Pallavicini, A., & Schartl, M. (2016). The lungfish transcriptome: a glimpse into molecular evolution events at the transition from water to land. *Scientific Reports*, *6*, 21571.

Bouchet, P. (2006). The magnitude of marine biodiversity. In C. M. Duarte (Ed.), *The exploration of marine biodiversity: scientific and technological challenges* (pp. 33–64). France: Fundación BBVA. www.fbbva.es

Bulutoglu, B., Garcia, K. E., Wu, F., Minteer, S. D., & Banta, S. (2016). Direct evidence for metabolon formation and substrate channeling in recombinant tca cycle enzymes. *ACS Chemical Biology*, *11*(10), 2847–2853.

Burton, J. N., Adey, A., Patwardhan, R. P., Qiu, R., Kitzman, J. O., & Shendure, J. (2013). Chromosome-scale scaffolding of de novo genome assemblies based on chromatin interactions. *Nature Biotechnology*, *31*(12), 1119.

Carlsson, N. O. L., Brönmark, C., & Hansson, L. A. (2004). Invading herbivory: the golden apple snail alters ecosystem functioning in Asian wetlands. *Ecology*, *85*(6), 1575–1580.

Clark, J. W., & Donoghue, P. C. J. (2018). Whole-genome duplication and plant macroevolution. *Trends in Plant Science*, *23*(10), 933–945.

Comai, L. (2005). The advantages and disadvantages of being polyploid. *Nature Reviews Genetics*, *6*(11), 836.

Cowie, R. H., Dillon, R. T., Robinson, D. G., & Smith, J. W. (2009). Alien non-marine snails and slugs of priority quarantine importance in the United States: A preliminary risk assessment. *American Malacological Bulletin*, *27*(1/2), 113–133.

Crow, K. D., & Wagner, G. P. (2006). What is the role of genome duplication in the evolution of complexity and diversity? *Molecular Biology and Evolution*, *23*(5), 887–892.

Darriba, D., Taboada, G. L., Doallo, R., & Posada, D. (2011). ProtTest 3: fast selection of best-fit models of protein evolution. *Bioinformatics*, *27*(8), 1164–1165.

De Bie, T., Cristianini, N., Demuth, J. P., & Hahn, M. W. (2006). CAFE: a computational tool for the study of gene family evolution. *Bioinformatics*, *22*(10), 1269–1271. https://doi.org/10.1093/bioinformatics/btl097

Denny, M. W. (1979). Introduction. In M. W. Denny (Ed.), *The role of mucus in the locomotion and adhesion of the pulmonate slug* (pp. 1–4). The

Faculty of Graduate Studies Department of Zoology, University of British Columbia.

Diez-Fernandez, C., Gallego, J., Haberle, J., Cervera, J., & Rubio, V. (2015). The study of carbamoyl phosphate synthetase 1 deficiency sheds light on the mechanism for switching on/off the urea cycle. *Journal of Genetics and Genomics*, *42*(5), 249–260.

Du, X., Fan, G., Jiao, Y., Zhang, H., Guo, X., Huang, R., & Liu, X. (2017). The pearl oyster Pinctada fucata martensii genome and multi-omic analyses provide insights into biomineralization. *Gigascience*, *6*(8), 1–12.

Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research*, *32*(5), 1792–1797.

Emms, D. M., & Kelly, S. (2015). OrthoFinder: solving fundamental biases in whole genome comparisons dramatically improves orthogroup inference accuracy. *Genome Biology*, *16*(1), 157.

Faa, G., Nurchi, V. M., Ravarino, A., Fanni, D., Nemolato, S., Gerosa, C., & Geboes, K. (2008). Zinc in gastrointestinal and liver disease. *Coordination Chemistry Reviews*, *252*(10–11), 1257–1269.

Faure, E., & Casanova, J. P. (2006). Comparison of chaetognath mitochondrial genomes and phylogenetical implications. *Mitochondrion*, *6*(5), 258–262.

Fawcett, J. A., Maere, S., & Van De Peer, Y. (2009). Plants with double genomes might have had a better chance to survive the Cretaceous-Tertiary extinction event. *Proceedings of the National Academy of Sciences*, *106*(14), 5737–5742.

Ferreira, F. S., Albuquerque, U. P., Coutinho, H. D., Almeida Wde, O., & Alves, R. R. (2012). The trade in medicinal animals in northeastern Brazil. *Evidence-based Complementary and Alternative Medicine*, *2012*, 126938.

Feschotte, C., & Wessler, S. R. (2002). Mariner-like transposases are widespread and diverse in flowering plants. *Proceedings of the National Academy of Sciences*, *99*(1), 280–285.

Flot, J.-F., Hespeels, B., Li, X., Noel, B., Arkhipova, I., Danchin, E. G., & Aury, J.-M. (2013). Genomic evidence for ameiotic evolution in the bdelloid rotifer Adineta vaga. *Nature*, *500*(7463), 453–457.

Frankham, R. (2005). Resolving the genetic paradox in invasive species. *Heredity*, *94*(4), 385.

Furlong, R. F., & Holland, P. W. (2002). Were vertebrates octoploid? *Philosophical Transactions of the Royal Society of London Series B: Biological Sciences*, *357*(1420), 531–544.

Guo, X., He, Y., Zhang, L., Lelong, C., & Jouaux, A. (2015). Immune and stress responses in oysters with insights on adaptation. *Fish & Shellfish Immunology*, *46*(1), 107–119.

Guo, Y., Zhang, Y., Liu, Q., Huang, Y., Mao, G., Yue, Z., & Xiao, N. (2019). A chromosomal-level genome assembly for the giant African snail *Achatina fulica*. *Gigascience*, *8*(10), 1–8.

Haas, B. J., Salzberg, S. L., Zhu, W., Pertea, M., Allen, J. E., Orvis, J., & Wortman, J. R. (2008). Automated eukaryotic gene structure annotation using evidencemodeler and the program to assemble spliced alignments. *Genome Biology*, *9*(1), R7.

Hallinan, N. M., & Lindberg, D. R. (2011). Comparative analysis of chromosome counts infers three paleopolyploidies in the mollusca. *Genome Biology and Evolution*, *3*, 1150–1163.

Hermenegildo, C., Monfort, P., & Felipo, V. (2000). Activation of N-methyl-D-aspartate receptors in rat brain in vivo following acute ammonia intoxication: characterization by in vivo brain microdialysis. *Hepatology*, *31*(3), 709–715.

Hiong, K. C., Loong, A. M., Chew, S. F., & Ip, Y. K. (2005). Increases in urea synthesis and the ornithine–urea cycle capacity in the Giant African Snail, *Achatina fulica*, during fasting or aestivation, or after the Injection with ammonium chloride. *Journal of Experimental Zoology Part A: Comparative Experimental Biology*, *303*(12), 1040–1053.

Hoegg, S., Brinkmann, H., Taylor, J. S., & Meyer, A. (2004). Phylogenetic timing of the fish-specific genome duplication correlates with the diversification of teleost fish. *Journal of Molecular Evolution*, *59*(2), 190–203.

Hsia, C. C., Schmitz, A., Lambertz, M., Perry, S. F., & Maina, J. N. (2013). Evolution of air breathing: oxygen homeostasis and the transitions from water to land and sky. *Comprehensive Physiology*, *3*(2), 849–915.

Hua-Van, A., Le Rouzic, A., Boutin, T. S., Filée, J., & Capy, P. (2011). The struggle for life of the genome's selfish architects. *Biology Direct*, *6*(1), 19.

Iguchi, S. M., Aikawa, T., & Matsumoto, J. J. (1982). Antibacterial activity of snail mucus mucin. *Comparative Biochemistry and Physiology Part A: Physiology*, *72*(3), 571–574.

Ip, Y. K., & Chew, S. F. (2010). Nitrogen metabolism and excretion during aestivation. In A. N. Carlos & E. C. José (Eds.), *Aestivation*, Progress in molecular and subcellular biology, Vol. *49* (pp. 63–94). Berlin, Heidelberg: Springer. https://doi.org/10.1007/978-3-642-02421-4_4

Jones, P., Binns, D., Chang, H.-Y., Fraser, M., Li, W., McAnulla, C., & Nuka, G. (2014). InterProScan 5: genome-scale protein function classification. *Bioinformatics*, *30*(9), 1236–1240.

Jörger, K. M., Stöger, I., Kano, Y., Fukuda, H., Knebelsberger, T., & Schrödl, M. (2010). On the origin of Acochlidia and other enigmatic euthyneuran gastropods, with implications for the systematics of Heterobranchia. *BMC Evolutionary Biology*, *10*(1), 323.

Kameda, Y., & Kato, M. (2011). Terrestrial invasion of pomatiopsid gastropods in the heavy-snow region of the Japanese Archipelago. *BMC Evolutionary Biology*, *11*, 118.

Kenny, N. J., Chan, K. W., Nong, W., Qu, Z., Maeso, I., Yip, H. Y., & Chu, K. H. (2016). Ancestral whole-genome duplication in the marine chelicerate horseshoe crabs. *Heredity*, *116*(2), 190–199.

Kim, H. S., Lee, Y. H., Lee, Y. R., Im, S. A., Lee, J. K., Kim, Y. S., & Lee, C. K. (2007). Activation of professional antigen presenting cells by acharan sulfate isolated from giant african snail, achatina fulica. *Archives of Pharmacal Research*, *30*(7), 866–870.

Koren, S., Walenz, B. P., Berlin, K., Miller, J. R., Bergman, N. H., & Phillippy, A. M. (2017). Canu: scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. *Genome Research*, *27*(5), 722–736.

Kruse, M. N., Becker, C., Lottaz, D., Köhler, D., Yiallouros, I., Krell, H. W., & Stöcker, W. (2004). Human meprin α and β homo-oligomers: cleavage of basement membrane proteins and sensitivity to metalloprotease inhibitors. *Biochemical Journal*, *378*, 383–389.

Kumar, S., Stecher, G., & Tamura, K. (2016). MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Molecular Biology and Evolution*, *33*(7), 1870–1874.

Li, Z., Tiley, G. P., Galuska, S. R., Reardon, C. R., Kidder, T. I., Rundell, R. J., & Barker, M. S. (2018). Multiple large-scale gene and genome duplications during the evolution of hexapods. *Proceedings of the National Academy of Sciences*, *115*(18), 4713–4718.

Lieb, B., Dimitrova, K., Kang, H. S., Braun, S., Gebauer, W., Martin, A., & Markl, J. (2006). Red blood with blue-blood ancestry: Intriguing structure of a snail hemoglobin. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(32), 12011–12016.

Lillywhite, H. B., Albert, J. S., Sheehy, C. M. III, & Seymour, R. S. (2012). Gravity and the evolution of cardiopulmonary morphology in snakes. *Comparative Biochemistry and Physiology Part A: Molecular & Integrative Physiology*, *161*(2), 230–242.

Little, C. (1990). Part III: life on land. In C. Little (Ed.), *The terrestrial invasion: an ecophysiological approach to the origins of land animals*, 1st ed. (pp. 201–203). Cambridge, UK: Cambridge University Press.

Liu, B., Shi, Y., Yuan, J., Hu, X., Zhang, H., Li, N., Fan, W. (2013). Estimation of genomic characteristics by analyzing k-mer frequency in de novo genome projects. *arXiv Preprint* arXiv:1308.2012. https://arxiv.org/abs/1308.2012

Liu, C., Zhang, Y., Ren, Y., Wang, H., Li, S., Jiang, F., Yin, L., Qiao, X. I., Zhang, G., Qian, W., Liu, B. O., & Fan, W. (2018). The genome of the golden apple snail Pomacea canaliculata provides insight into stress

tolerance and invasive adaptation. *Gigascience*, *7*(9), 1–13. https://academic.oup.com/gigascience/article/7/9/giy101/5069392

Mable, B., Alexandrou, M., & Taylor, M. (2011). Genome duplication in amphibians and fish: an extended synthesis. *Journal of Zoology*, *284*(3), 151–182.

Markl, J. (2013). Evolution of molluscan hemocyanin structures. *Biochimica Et Biophysica Acta (BBA)-Proteins and Proteomics*, *1834*(9), 1840–1852.

Marshall, C. R., & Ward, P. D. (1996). Sudden and gradual molluscan extinctions in the latest Cretaceous of western European Tethys. *Science*, *274*(5291), 1360–1363.

Mukherjee, S., Sarkar, S., Munshi, C., & Bhattacharya, S. (2017). The uniqueness of Achatina Fulica in its evolutionary success. In S. Ray (Ed.), *Organismal* (pp. 219–232). IntechOpen. https://doi.org/10.5772/68134

Nakatani, Y., & McLysaght, A. (2019). Macrosynteny analysis shows the absence of ancient whole-genome duplication in lepidopteran insects. *Proceedings of the National Academy of Sciences*, *116*(6), 1816–1818.

Oja, S. S., Saransaari, P., & Korpi, E. R. (2017). Neurotoxicity of ammonia. *Neurochemical Research*, *42*, 713.

Onadeko, A. B., & Egonmwan, R. I. (2017). Water relations, weight loss and nitrogenous waste products in the giant African land snail Archachatina marginata (swainson) - pulmonata: achatinidae. *Journal of Scientific Research and Development*, *17*(1), 40–47.

Orr, H. A. (1990). " Why polyploidy is rarer in animals than in plants" revisited. *The American Naturalist*, *136*(6), 759–770.

Peña, J. J., & Adema, C. M. (2016). The Planorbid Snail Biomphalaria glabrata Expresses a Hemocyanin-Like Sequence in the Albumen Gland. *PLoS One*, *11*(12), e0168665. https://doi.org/10.1371/journal.pone.0168665

Petersen, S. V., Dutton, A., & Lohmann, K. C. (2016). End-cretaceous extinction in antarctica linked to both deccan volcanism and meteorite impact via climate change. *Nature Communications*, *7*, 12079.

Portela, J., Duval, D., Rognon, A., Galinier, R., Boissier, J., Coustau, C., & Gourbal, B. (2013). Evidence for specific genotype-dependent immune priming in the lophotrochozoan Biomphalaria glabrata snail. *Journal of Innate Immunity*, *5*(3), 261–276.

Rahman, M. S., & Raut, S. K. (2012). Consequences of aestivation in the giant African land snail Achatina fulica Bowdich (Gastropoda: Achatinidae). *Proceedings of the Zoological Society*, *65*(2), 95–104.

Romero, P. E., Pfenninger, M., Kano, Y., & Klussmann-Kolb, A. (2016). Molecular phylogeny of the Ellobiidae (Gastropoda: Panpulmonata) supports independent terrestrial invasions. *Molecular Phylogenetics and Evolution*, *97*, 43–54.

Ruan, J., & Li, H. (2020). Fast and accurate long-read assembly with wtdbg2. *Nature Methods*, *17*, 155–158. https://doi.org/10.1038/s41592-019-0669-3

Sanderson, M. J. (2003). r8s: inferring absolute rates of molecular evolution and divergence times in the absence of a molecular clock. *Bioinformatics*, *19*(2), 301–302.

Schwager, E. E., Sharma, P. P., Clarke, T., Leite, D. J., Wierschin, T., Pechmann, M., & Bilde, T. (2017). The house spider genome reveals an ancient whole-genome duplication during arachnid evolution. *BMC Biology*, *15*(1), 62.

Serb, J. M., & Lydeard, C. (2003). Complete mtDNA sequence of the North American freshwater mussel, Lampsilis ornata (Unionidae): an examination of the evolution and phylogenetic utility of mitochondrial genome organization in Bivalvia (Mollusca). *Molecular Biology and Evolution*, *20*(11), 1854–1866.

Servant, N., Varoquaux, N., Lajoie, B. R., Viara, E., Chen, C.-J., Vert, J.-P., & Barillot, E. (2015). HiC-Pro: an optimized and flexible pipeline for Hi-C data processing. *Genome Biology*, *16*(1), 259.

Sessa, E. B. (2019). Polyploidy as a mechanism for surviving global change. *New Phytologist*, *221*(1), 5–6.

Shen, S., Callaghan, D., Juzwik, C., Xiong, H., Huang, P., & Zhang, W. (2010). ABCG2 reduces ROS-mediated toxicity and inflammation: a potential role in Alzheimer's disease. *Journal of Neurochemistry*, *114*(6), 1590–1604.

Simakov, O., Marletaz, F., Cho, S. J., Edsinger-Gonzales, E., Havlak, P., Hellsten, U., & Rokhsar, D. S. (2013). Insights into bilaterian evolution from three spiralian genomes. *Nature*, *493*(7433), 526–531.

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*(19), 3210–3212.

Soltis, D. E., Albert, V. A., Leebens-Mack, J., Bell, C. D., Paterson, A. H., Zheng, C., & Soltis, P. S. (2009). Polyploidy and angiosperm diversification. *American Journal of Botany*, *96*(1), 336–348.

Sow-Yan, C. (2019). Shell of a pink-lipped agate snail, Achatina immaculata, found in Singapore. *Singapore Biodiversity Records*, 19. https://lkcnhm.nus.edu.sg/app/uploads/2019/01/sbr2019-019.pdf

Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, *30*(9), 1312–1313.

Stapley, J., Santure, A. W., & Dennis, S. R. (2015). Transposable elements as agents of rapid adaptation may explain the genetic paradox of invasive species. *Molecular Ecology*, *24*(9), 2241–2252.

Storey, K. B. (2002). Life in the slow lane: Molecular mechanisms of estivation. *Comparative Biochemistry and Physiology Part A: Molecular & Integrative Physiology*, *133*(3), 733–754.

Sun, J., Zhang, Y., Xu, T., Zhang, Y., Mu, H., Zhang, Y., & Zhang, W. (2017). Adaptation to deep-sea chemosynthetic environments as revealed by mussel genomes. *Nature Ecology & Evolution*, *1*(5), 1–7.

Sun, T. (1995). Chromosomal studies in three land snails. *Sinozoologia*, *12*, 154–162.

Takeuchi, T., Kawashima, T., Koyanagi, R., Gyoja, F., Tanaka, M., Ikuta, T., & Satoh, N. (2012). Draft genome of the pearl oyster Pinctada fucata: a platform for understanding bivalve biology. *DNA Research*, *19*(2), 117–130.

Thiengo, S. C., Faracas, F. A., Salgado, N. C., Cowie, R. H., & Fernandez, M. A. (2007). Rapid spread of an invasive snail in South America: the giant African snail, *Achatina fulica*, in Brasil. *Biological Invasions*, *9*, 693–702.

Thiriot-Quiévreux, C. (2003). Advances in chromosomal studies of gastropod molluscs. *Journal of Molluscan Studies*, *69*(3), 187–202.

Troost, K. (2010). Causes and effects of a highly successful marine invasion: Case-study of the introduced Pacific oyster Crassostrea gigas in continental NW European estuaries. *Journal of Sea Research*, *64*(2010), 145–165.

Uliano-Silva, M., Dondero, F., Dan Otto, T., Costa, I., Lima, N. C. B., Americo, J. A., & Rebelo, M. F. (2018). A hybrid-hierarchical genome assembly strategy to sequence the invasive golden mussel, Limnoperna fortunei. *Gigascience*, *7*(2), 1–10. https://academic.oup.com/gigascience/article-abstract/7/2/gix128/4750781

Van de Peer, Y., Maere, S., & Meyer, A. (2009). The evolutionary significance of ancient genome duplications. *Nature Reviews Genetics*, *10*(10), 725.

Van de Peer, Y., Mizrachi, E., & Marchal, K. (2017). The evolutionary significance of polyploidy. *Nature Reviews Genetics*, *18*(7), 411.

Vanneste, K., Baele, G., Maere, S., & Van de Peer, Y. (2014). Analysis of 41 plant genomes supports a wave of successful genome duplications in association with the Cretaceous-Paleogene boundary. *Genome Research*, *24*(8), 1334–1347.

Vermeij, G. J., & Dudley, R. (2000). Why are there so few evolutionary transitions between aquatic and terrestrial ecosystems? *Biological Journal of the Linnean Society*, *70*(4), 541–554.

Volkmann, D., & Baluška, F. (2006). Gravity: one of the driving forces for evolution. *Protoplasma*, *229*(2–4), 143–148.

Walker, B. J., Abeel, T., Shea, T., Priest, M., Abouelliel, A., Sakthikumar, S., & Young, S. K. (2014). Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. *PLoS One*, *9*(11), e112963.

Wan, F., Jiang, M., & Zhan, A. (2017). Biological invasions in China. In F. Wan, M. Jiang, A. Zhan (Eds.), *Biological invasions and its management in china*, Vol. *1* (pp. 3–147). Springer. https://doi.org/10.1007/978-94-024-0948-2

Wang, D., Zhang, Y., Zhang, Z., Zhu, J., & Yu, J. (2010). KaKs_Calculator 2.0: a toolkit incorporating gamma-series methods and sliding window strategies. *Genomics, Proteomics & Bioinformatics*, *8*(1), 77–80.

Wang, Y., Tang, H., DeBarry, J. D., Tan, X., Li, J., Wang, X., & Guo, H. (2012). MCScanX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity. *Nucleic Acids Research*, *40*(7), e49.

Werren, J. H. (2011). Selfish genetic elements, genetic conflict, and evolutionary innovation. *Proceedings of the National Academy of Sciences*, *108*(Supplement 2), 10863–10870.

Wertheim, B., Beukeboom, L., & Van de Zande, L. (2013). Polyploidy in animals: effects of gene expression on sex determination, evolution and ecology. *Cytogenetic and Genome Research*, *140*(2–4), 256–269.

Zapata, F., Wilson, N. G., Howison, M., Andrade, S. C., Jörger, K. M., Schrödl, M., & Dunn, C. W. (2014). Phylogenomic analyses of deep gastropod relationships reject Orthogastropoda. *Proceedings of the Royal Society B: Biological Sciences*, *281*(1794), 20141739.

Zhang, G., Fang, X., Guo, X., Li, L. I., Luo, R., Xu, F., Yang, P., Zhang, L., Wang, X., Qi, H., Xiong, Z., Que, H., Xie, Y., Holland, P. W. H., Paps, J., Zhu, Y., Wu, F., Chen, Y., Wang, J., … Wang, J. (2012). The oyster genome reveals stress adaptation and complexity of shell formation. *Nature*, *490*(7418), 49–54.

Zhang, Z., Xiao, J., Wu, J., Zhang, H., Liu, G., Wang, X., & Dai, L. (2012). ParaAT: a parallel tool for constructing multiple protein-coding DNA alignments. *Biochemical and Biophysical Research Communications*, *419*(4), 779–781.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Liu C, Ren Y, Li Z, et al. Giant African snail genomes provide insights into molluscan whole-genome duplication and aquatic–terrestrial transition. *Mol Ecol Resour*. 2020;00:1–17. https://doi.org/10.1111/1755-0998.13261