Received: 31 March 2021 | Revised: 2 August 2021 | Accepted: 31 August 2021

DOI: 10.1111/1755-0998.13504

MOLECULAR ECOLOGY RESOURCES WILEY

RESOURCE ARTICLE

Chromosome-scale assembly and whole-genome sequencing of 266 giant panda roundworms provide insights into their evolution, adaptation and potential drug targets

Lei Han[1,2] | Tianming Lan[3,4] | Desheng Li[5] | Haimeng Li[3,6] | Linhua Deng[5] | Zhiwei Peng[1] | Shaowen He[7] | Yanqiang Zhou[1] | Ruobing Han[1] | Lingling Li[1] | Yaxian Lu[1] | Haorong Lu[8,9] | Qing Wang[10] | Shangchen Yang[11] | Yixin Zhu[10] | Yunting Huang[8,9] | Xiaofang Cheng[12] | Jieyao Yu[8,9] | Yulong Wang[1] | Heting Sun[13] | Hongliang Chai[1] | Huanming Yang[3] | Xun Xu[3,8] | Michael Lisby[4] | Quan Liu[1,14] | Karsten Kristiansen[4,15] | Huan Liu[3,4] | Zhijun Hou[1,2]

[1]College of Wildlife and Protected Area, Northeast Forestry University, Harbin, China

[2]Key Laboratory of Wildlife Conservation, China State Forestry Administration, Harbin, China

[3]State Key Laboratory of Agricultural Genomics, BGI-Shenzhen, Shenzhen, China

[4]Laboratory of Genomics and Molecular Biomedicine, Department of Biology, University of Copenhagen, Copenhagen, Denmark

[5]Key Laboratory of SFGA on Conservation Biology of Rare Animals in the Giant Panda National Park (CCRCGP), Sichuan, China

[6]School of Future Technology, University of Chinese Academy of Sciences, Beijing, China

[7]Foping National Nature Reserve, Hanzhong, China

[8]Guangdong Provincial Key Laboratory of Genome Read and Write, BGI-Shenzhen, Shenzhen, China

[9]China National GeneBank, BGI-Shenzhen, Shenzhen, China

[10]BGI Education Center, University of Chinese Academy of Sciences, Shenzhen, China

[11]College of Life Sciences, Zhejiang University, Hangzhou, China

[12]MGI, BGI-Shenzhen, Shenzhen, China

[13]General Station for Surveillance of Wildlife Diseases, National Forestry and Grassland Administration, Harbin, China

[14]School of Life Sciences and Engineering, Foshan University, Foshan, China

[15]Qingdao-Europe Advanced Institute for Life Sciences, Qingdao, China

Correspondence
Quan Liu and Zhijun Hou, College of Wildlife and Protected Area, Northeast Forestry University, Harbin, China.
Emails: houzhijundb@163.com (Z.H); liuquan1973@hotmail.com (Q.L)

Huan Liu, State Key Laboratory of Agricultural Genomics, BGI-Shenzhen, Shenzhen, China.
Email: liuhuan@genomics.cn

Karsten Kristiansen, Laboratory of Genomics and Molecular Biomedicine, Department of Biology, University of

Abstract

Helminth diseases have long been a threat to the health of humans and animals. Roundworms are important organisms for studying parasitic mechanisms, disease transmission and prevention. The study of parasites in the giant panda is of importance for understanding how roundworms adapt to the host. Here, we report a high-quality chromosome-scale genome of *Baylisascaris schroederi* with a genome size of 253.60 Mb and 19,262 predicted protein-coding genes. We found that gene families related to epidermal chitin synthesis and environmental information processes in the roundworm genome have expanded significantly. Furthermore, we demonstrated

Lei Han, Tianming Lan and Desheng Li contributed equally to this work.

Copenhagen, Copenhagen, Denmark.
Email: kk@bio.ku.dk

Funding information
Forestry science and technology research project, Grant/Award Number: 20180302; The Foundation of Key Laboratory of State Forestry and Grassland Administration (State Park Administration) on Conservation Biology of Rare Animals in the Giant Panda National Park, Grant/Award Number: KLSFGAGP2020.002; The Pearl River Talent Recruitment Program in Guangdong Province, Grant/Award Number: 2019CX01N111; The Guangdong Provincial Key Laboratory of Genome Read and Write, Grant/Award Number: 2017B030301011; National Key R&D Program, Grant/Award Number: 2017YFD0501702; Open Project of Key Laboratory of SFGA on Conservation Biology of Rare Animals in The Giant Panda National Park, Grant/Award Number: 2020004; Fundamental Research Funds for the Central Universities of China, Grant/Award Number: 2572020AA30

unique genes involved in essential amino acid metabolism in the *B. schroederi* genome, inferred to be essential for the adaptation to the giant panda-specific diet. In addition, under different deworming pressures, we found that four resistance-related genes (*glc-1*, *nrf-6*, *bre-4* and *ced-7*) were under strong positive selection in a captive population. Finally, 23 known drug targets and 47 potential drug target proteins were identified. The genome provides a unique reference for inferring the early evolution of roundworms and their adaptation to the host. Population genetic analysis and drug sensitivity prediction provide insights revealing the impact of deworming history on population genetic structure of importance for disease prevention.

KEYWORDS
adaptation, anthelmintics, *Baylisascaris schroederi*, genetic diversity, roundworms

1 | INTRODUCTION

Parasitic ascariasis has long been a threat to the health of humans, livestock and wildlife worldwide (Hotez et al., 2009). With the expansion of towns, cities, and the wild land-urban interface, geographic isolation is no longer an effective barrier for transmission of helminth infections. As a result, the risks for transmission of diseases once isolated in wildlife have never been greater (Kazacos & Boyce, 1989). Due to its wide distribution and long incubation period, soil-transmitted helminth eggs are easily transmitted between wildlife and livestock, and even to humans through contaminated faeces or soil. In-depth studies of helminths in wildlife can provide information of relevance for identifying and detecting pathogens and instigate appropriate actions to deal with possible risks with broad and far-reaching implications for wildlife and human health.

Baylisascaris schroederi, a parasitic nematode specific for the giant panda (*Ailuropoda melanoleuca*), is a soil-transmitted nematode and can directly infect the giant panda without passing through an intermediate host (Bethony et al., 2006; De Silva et al., 2003). Baylisascaris species also cause infection as patent or latent larva migrans (LM) in a variety of mammals (Kazacos & Boyce, 1989), birds (Wolf et al., 2007) and humans (Murray, 2002; Wise et al., 2005), and are therefore considered zoonotic parasites with potential public health and safety risks. Even though not all details of the life cycle of *B. schroederi* have been established, it is known that eggs secreted with faeces are very resistant and can survive for extended periods in the soil. Fertile eggs can become infective under suitable temperature conditions (12–22°C) and after being ingested by a panda, the eggs hatch in the small intestine, the larvae migrate to several organs, and eventually returns to the intestine where they mature into adult reproductive worms (Wang et al., 2018). Migration of larvae to different organs may cause serious damage to the organs, with different roundworm species being associate with different syndromes including visceral larva migrans (VLM), ocular larva migrans (OLM), neural larva migrans (NLM) and even severe pneumonia and hepatitis (Kazacos & Boyce, 1989; Papini et al., 1995; Wildt et al., 2006; Zhang et al., 2011). Infection by Baylisascaris species can in addition cause severe baylisascariasis, intestinal blockage, and even fatal bowel rupture (Schaul, 2006; Yang, 1998). Compared with other roundworms, *B. schroederi* is smaller in size and is mainly found in the small intestine of giant pandas. Giant pandas have typical carnivorous intestinal characteristics (short and thick small intestines), but eat bamboo, a diet with low digestibility and absorption. This challenges the nutrient absorption of *B. schroederi* for survival in the small intestine. Based on available epidemiological data of the giant panda, *B. schroederi* is the leading cause of death from primary and secondary infection in wild and captive populations (Hu et al., 2018; Li et al., 2014). Moreover, the problem of increased resistance to anthelmintics is likely to be seriously underestimated. Giant pandas in captivity are regularly dewormed (every 60 days). According to investigations, *B. schroederi* eggs can still be detected in the faeces 10–15 days after treatment with anthelmintics (Li et al., 2015), indicating that a development of drug-resistant subtypes had occurred in the *B. schroederi* population, and that *B. schroederi* variants with resistance to a variety of anthelmintics had survived. These variants may potentially become anthelmintic-resistant pathogens.

Although *B. schroederi* poses threats to both wild and captive giant pandas, current studies are limited to morphological and single or multiple gene analyses, preventing in-depth exploration of genetic mechanisms of adaptations and further prevention and control

of infections (Xie et al., 2011). Here, we report a chromosome-scale reference genome of *B. schroederi*, which is also the first chromosome-scale reference genome of ascaridoids. Based on the genome, we explored possible genetic mechanisms of the adaptation of *B. schroederi* to the intestinal environment, especially the specific diet of the giant panda, as well as the potential genetic basis of drug resistance. Finally, potential drug target proteins were identified, which provides new insights into the potential disease management of *Baylisascaris* and related roundworms.

2 | MATERIALS AND METHODS

2.1 | Samples

All specimens of Sichuan *B. schroederi* population were collected from the giant panda (*Ailuropoda melanoleuca*) with naturally acquired infections from the China Conservation and Research Centre for Giant Panda during the period August 2018 to July 2019. Samples of Qinling *B. schroederi* population were obtained from the intestines of a wild giant panda that died shortly after being found in the Foping National Nature Reserve (Qinling Mountains) in 2018. The captive and wild individuals differed by geography and the different pressure of deworming history. Roundworms were washed extensively in sterile physiological saline (37°C), sexes separated, snap-frozen and transported on dry ice and then stored at −80°C until use. Several specimens were stored in RNA preservation solution for transcriptome sequencing. All experimental designs and handling of nematodes were approved by the Institutional Animal Care and Use Committee of Northeast Forestry University.

2.2 | Library construction, sequencing and assembly

PacBio libraries were generated with a SMRTbell Template Prep Kit 1.0 (Pacific Biosciences) and the SMRTbell Damage Repair Kit (Pacific Biosciences) and were sequenced with one SMRT cell on the PacBio SMRTplatform. In addition, a 100-bp paired-end library was constructed and sequenced on a BGISEQ platform to assess the complexity of genome and to polish PacBio data (for details, see Supporting Information).

The *B. schroederi* genome was assembled using a "correct-then-assemble" strategy. First, NEXTDENOVO (v2.0-beta.1; https://github.com/Nextomics/NextDenovo) was used to correct and assemble a draft genome. The Arrow (v0.3.2) algorithm was then used to carry out a second round of correction for this assembly (Archibald, 2017). NEXTPOLISH (v1.0.5) (Hu et al., 2019) was further used for genome polishing by using the WGS data. To finally ligate the scaffolds to chromosomes, Hi-C technology (Lieberman-Aiden et al., 2009) was used to capture the chromosome conformations. 105 Gb (~400x) Hi-C sequencing data were generated from a single Hi-C library which was constructed as previously described. We next used JUICER (v1.5.7) to analyse the Hi-C data sets, and a three-dimensional de novo assembly (3D-DNA v180322) pipeline (Dudchenko et al., 2017) to scaffold spotted *B. schroederi* genome to chromosome-length scaffolds (for details, see Supporting Information). We used the DIAMOND (v0.9.10) software to blast the genome against the NCBI NR database, then deleted the scaffolds that blasted to bacteria (such as *Escherichia coli*, *Lactococcus lactis*) and generated the final genome file. Finally, we used the Purge_Dups pipeline (Guan et al., 2020) to remove haplotig sequences from the genome. The completeness of the genome was evaluated using sets of Benchmarking Universal Single-Copy Orthologues (BUSCO v3.0.1) with genome mode and lineage data from nematode odb9 and eukaryote odb9, respectively (Simão et al., 2015).

2.3 | Detection and classification of repetitive elements

For de novo identification of repeat elements, we constructed a transposable element (TE) library of the *B. schroederi* genome to identified types of repeat elements by using TANDEM REPEAT FINDER (TRF) (Benson, 1999), LTR_FINDER (Zhao & Hao, 2007) and REPEATMODELER (v1.0.8) (Smit et al., 2015). REPEATMASKER (Tarailo-Graovac & Chen, 2009) and REPEATPROTEINMASK (Tempel, 2012) were used to search the genome sequences for known repeat elements against the Repbase database (Jurka et al., 2005).

2.4 | Prediction of protein-coding genes and functional annotation

A combined strategy of de novo gene prediction, homology-based search and RNA sequencing-aided annotation was used to perform gene prediction. For homology-based annotation, we selected the protein-coding sequences of five homologous species (*Brugia malayi*, *C. elegans*, *Pristionchus pacificus*, *Steinernema carpocapsae* and *T. canis*) from NCBI (https://www.ncbi.nlm.nih.gov/). For RNA-based prediction, a male and a female transcriptome sequences were aligned to the genome for assembly using TOPHAT (v2.1.0) (Trapnell et al., 2009) plus TRINITY (v2.0.6) (Haas et al., 2013) strategy. PASAPIPELINE (v.2.1.0) was applied to predict gene structure after which the inferred gene structures were used in AUGUSTUS (v.3.2.3) (Mario et al., 2006) to train gene models based on transcript evidence. In addition, the genome sequence was analysed by the program GENEMARK (v1.0) (John & Mark, 2005) utilizing unsupervised training to build a hidden Markov model. The consistent gene sets were generated by combining all above evidence using MAKER (v.2.31.8) (Campbell et al., 2013). All gene evidence was merged to form a comprehensive and nonredundant gene set using EVIDENCEMODELER (v1.1.1, EVM) (Haas et al., 2008).

In order to perform gene functional annotation, we aligned the gene sets against several known databases, including SwissProt (Amos & Rolf, 2000), TrEMBL (Amos & Rolf, 2000), Kyoto

encyclopedia of genes and genomes (KEGG) (Pitk, 2006), clusters of orthologous groups of proteins (COG) (Tatusov et al., 2003) and NR database. Gene ontology (GO) information was obtained through BLAST2GO (v.2.5.0) (Conesa et al., 2005). TRNASCAN-SE (v1.3.1) (Lowe & Eddy, 1997) was used to predict tRNAs. We aligned the *B. schroederi* genome against RFAM (v12.0) (Kalvari et al., 2018) database and invertebrate rRNA database to predict snRNA, miRNA and rRNA, respectively. In addition, proteases, protease inhibitors (PIs), and excretory/secretory proteins (EPs) were annotated and identified (for details, see supplementary text).

2.5 | Reconstruction of phylogeny and evolutional analysis using genomes from six nematodes

Genomes and annotation files of three parasitic roundworms (*A. suum*, *T. canis* and *P. univalens*), one free-living nematode (*C. elegans*) and one parasitic plant root nematode (*M. hapla*) were download from NCBI or the WormBase database (Table S10) (Price & Hurd, 2019). The syntenic analyses both at whole-genome nucleotide-level and protein-level were performed by aligning the five nematodes genomes to our assembled *B. schroederi* genome by using MCSCANX (Wang et al., 2012) software. Orthogroups among the six nematodes were defined using TREEFAM (v4.0) (Li et al., 2006). Next, the protein sequences from each family were aligned using MUSCLE (v3.8.31) (Edgar, 2004) with default parameters. The conserved CDS alignments were extracted by Gblocks (Gerard & Jose, 2007). We selected single-copy gene families to construct phylogenetic trees based on maximum likelihood using RAXML (v8.2.4) (Alexandros, 2014) with PROTCATGTR nucleotide substitution model with 500 bootstrap replicates.

Using the divergence time between *C. elegans* and *A. suum*, which was calculated based on fossil evidence, as the reference time points (Mcgill et al., 2017), we estimated the divergence time between each species by MCMCtree from the PAML (v4.8) (Bo & Yang, 2013) package with default parameters.

2.6 | Expanded and contracted gene families

Based on the phylogenetic tree we constructed using the 2451 single-copy genes, we explored significantly expanded and contracted gene families in the six nematodes (*A. suum*, *B. schroederi*, *T. canis*, *P. univalens*, *C. elegans* and *M. hapla*). Gene family expansion and contraction analyses were performed using CAFÉ (v4.2) (Bie et al., 2006). We used a median gene number to estimate the changes in gene family size. The overall *p*-value of each gene family were calculated and the details of each branch and node with the exact *p*-values of each significant overall *p*-value (≤.01) gene family were also calculated. Here, we performed analyses in two levels: (1) Comparing gene families of roundworms to those of other nematodes to determine the expanded and contracted genes in roundworms. (2) Comparing the *B. schroederi* and other three roundworms (*A. suum*, *P. univalens* and *T. canis*) to identify gain and loss of genes in *B. schroederi*, to further determine their differentiation within the roundworm branch. In addition, seven detoxification-related gene families of four roundworms were identified by HMMER3 (for details, see Supporting Information text) (Jaina et al.,2013).

2.7 | Identification of positively selected genes (PSGs)

Based on gene families retrieved from TreeFam, we identify PSGs between *B. schroederi* and the other five species (*A. suum*, *T. canis*, *P. univalens*, *C. elegans* and *M. hapla*). The branch-site model of CodeML in PAML (v4.8) (Bo & Yang, 2013) was used to detect potential PSGs. The null hypothesis and the alternative hypothesis was used to estimate whether the dN/dS (ω) value of the foreground branch (Marked as *B. schroederi*) was larger than 1 or not followed by likelihood ratio test (LRT) using R "chisq.test()" function to calculate chi-squared distributions with 1 degree of freedom. A PSG was identified by meeting the requirements of a corrected *p*-value (<.05) and containing at least one positively selected site with a posterior probability ≥1.

2.8 | The change of effective population size as a function of time

We inferred the demographic history of the *B. schroederi* by using the WGS data generated by DNBSEQ-T1 from one individual. We simultaneously perform the same analysis on a giant panda by using resequencing short reads of an individual download from SRA database (accession SRA053353). For this analysis, we used BWA (v0.7.13-r1126) (Li & Durbin, 2009) to map the clean reads to each genome with the default parameters. Next, the PSMC method (Li & Durbin, 2011) was used to evaluate the dynamic changes of the effective population size (N_e) of *B. schroederi* and the giant panda. Following Li's procedure (Li & Durbin, 2011), we applied a bootstrapping approach, repeated sampling 100 times to estimate the variance of simulated results for both *B. schroederi* and giant panda. We used 0.17 and 12 years per generation (g) and a mutation rate (μ) of 9×10^{-9} and 1.29×10^{-8} for *B. schroederi* and giant panda, respectively (Cutter, 2008). Since fluctuations in the effective population size of giant pandas have been reported to closely reflect changes in climate and atmospheric dust (Zhao et al., 2013), we added the mass accumulation rate (MAR) of Chinese loess over the past 250,000 years for comparison. In addition, we implemented the MSMC2 (Schiffels & Durbin, 2014) which can infer the recent effective population size history. We phased all SNPs of each individual by using BEAGLE (v5.0) (Browning & Browning, 2007) using the following parameters: -i 20 -t 6 -p '10*1+15*2'. The mutation rate (μ) of *B. schroederi* for MSMC2 was the same as used for PSMC.

2.9 | Population structure of *B. schroederi* in Qinling and Sichuan

A total of 240 samples collected from individuals in captivity and 26 samples from individuals in wild were re-sequenced using the DNBSEQ-T1&T5 platform. High-quality reads were aligned to the reference genome using BWA-MEM (0.7.13-r1126) (Li & Durbin, 2009) with default parameters. SAMTOOLS (v0.1.19) (Li et al., 2009) and genome analysis toolkit (GATK v 4.0.3.0) (Depristo et al., 2011) were used to obtain the SNP set within the population. Hard filtering was applied to the raw variant set using "QD <2.0 || FS >60.0 || MQ <40.0 || MQRankSum < −12.5 || ReadPosRankSum < −8.0" --filter-name "snp_filter". SNPs with >0.5% missing data or <0.01 minor allele frequency (MAF) were filtered out using VCFTOOLS (v0.1.12a) (Danecek et al., 2011). Principal components analysis (PCA) analysis of SNPs was carried out using EIGENSOFT (Patterson et al., 2006) software, and the population clustering analysis was conducted in PLINK (Purcell et al., 2007). We used the whole-genome SNPs to construct the ML phylogenetic tree with 1000 bootstrap using IQTREE (v1.6.12) (Lam-Tung et al.,2015), and using an genome sequence information of *P. univalens* as an outgroup. Population structure of all was analysed using the ADMIXTURE (v1.3.0) program with a block-relaxation algorithm. To explore the convergence of individuals, we predefined the number of genetic clusters K from 2 to 5.

2.10 | Recent natural selection analysis

Extended haplotype homozygosity (EHH) and iHS methods were used for detecting SNPs under strong positive selection of the captive and wild population, respectively (Mathieu & Renaud, 2012). We use the data2haplohh (), scan_hh () and ihhh2ihs () functions of the R package rehh (v3.1.2) to calculate the iHS scores (Mathieu & Renaud, 2012). We took the top 0.5% iHS score as the loci with strong positive selection. We added accumulated iHS scores at intervals of genes. The XP-EHH method was used to detect selective sweeps using the REHH (v3.1.2). We then split the genome into nonoverlapping segments of 50 kb to use the maximum (positive) XP-EHH score of all SNPs. The regions with *p*-values <.01 were considered significant signals in the population of interest.

2.11 | Known and potential drug targets

All compound-related proteins were searched against target proteins from CHEMBL v26 (Gaulton et al., 2012) using BLASTP ($E \leq 1 \times 10^{-10}$). We screened out all types of single proteins in the ChEMBL database for blasting. Known drug targets were identified from available publications and by searching for "anthelmintics" in the DrugBank (Wishart et al., 2017) database. For potential drug targets, we screened out all the single protein targets as described (Coghlan et al., 2018) and made adjustments to evaluate each gene. We set a score of "0/1" considering six main factors: (1) Similarity with ChEMBL drug targets and a highly conserved alignment (>80%). (2) Lack of human homologues. (3) Related to lethal, L3 arrest, flaccid, molt defect and sterile phenotypes. Lethal phenotypes were identified in WormBase WS240. (4) Whether the protein was a predicted chokepoint enzyme (Tyagi et al., 2018). (5) Whether the protein was predicted as an excretory/secretory protein (EP). (6) Whether the protein had a structure in the PDBe (Sameer et al., 2016). In order to make the screening most efficient, we searched for commercially available compounds against the target protein in ZINC15 (Sterling & Irwin, 2015). Finally, we selected compounds approved in phase III or above as suggested chemical compounds.

3 | RESULTS

3.1 | Genome assembly, annotation and evaluation

A total of 29 Gb (110x) PacBio long reads were generated (Table S1). The genome size was estimated to be 266.8 Mb based on K-mer depth distribution analysis (Table S2 and Figure S1), and the size of the assembled *B. schroederi* genome reached 253.61 Mb, accounting for 95.05% of the estimated genome. This genome contained 75 scaffolds, 21 of which were superscaffolds ligated using 105 Gb (~386x) Hi-C sequencing data (Figure 1a; Figure S2). The total length of these 21 superscaffolds reached ~250.70 Mb, accounting for 98.86% of the whole genome (Table S3). The final scaffold N50 was 12.32 Mb (Table 1), which is significantly better than the published genome (Hu et al., 2020). The GC-depth distribution (Figure S3a,b) further showed that most genomic regions have a GC content narrowly centred around 37%, which is similar to that of other three roundworms (*Ascaris suum*, *Parascaris univalens* and *Toxocara canis*; Table 1) (Jex et al., 2011; Zhu et al., 2015). BUSCO scores against nematode and eukaryote databases were 91.7% and 92.8%, respectively (Figure S4), reflecting the highest genome completeness among published roundworm genomes (Table 1). The blobplot also shows the accuracy of the assembly (Figure S5). 19,262 protein-coding genes were predicted via ab initio, homology-based and RNA sequencing-aided methods (see Section 2). KEGG, COG, TrEMBL, GO, Swissprot and InterPro (Figures S6 and S7). The average length of coding sequences (CDS) was 1052 bp with an average of 6.87 exons per gene, which is similar to that of other related roundworms (Table S4). We identified the mitochondrial genome (see Supporting Information text and Table S5). To evaluate the completeness of the predicted protein-coding genes, we compared the length distributions of mRNA, CDS, exons and introns in *B. schroederi* with those in other five nematodes (Figures S9 and S10).

Total repeats (DNA transposons and RNA transposons) accounted for 12.66% of the genome (Tables S6–S8 and Figure S11). Huge variation in the proportion of repeat content is found among published nematode genomes (from 1% to 48%) (Coghlan, Tyagi, et al., 2018; Schiffer et al., 2013). Transposable elements (TEs) account for 10.24% of the *B. schroederi* genome (Table S8), while

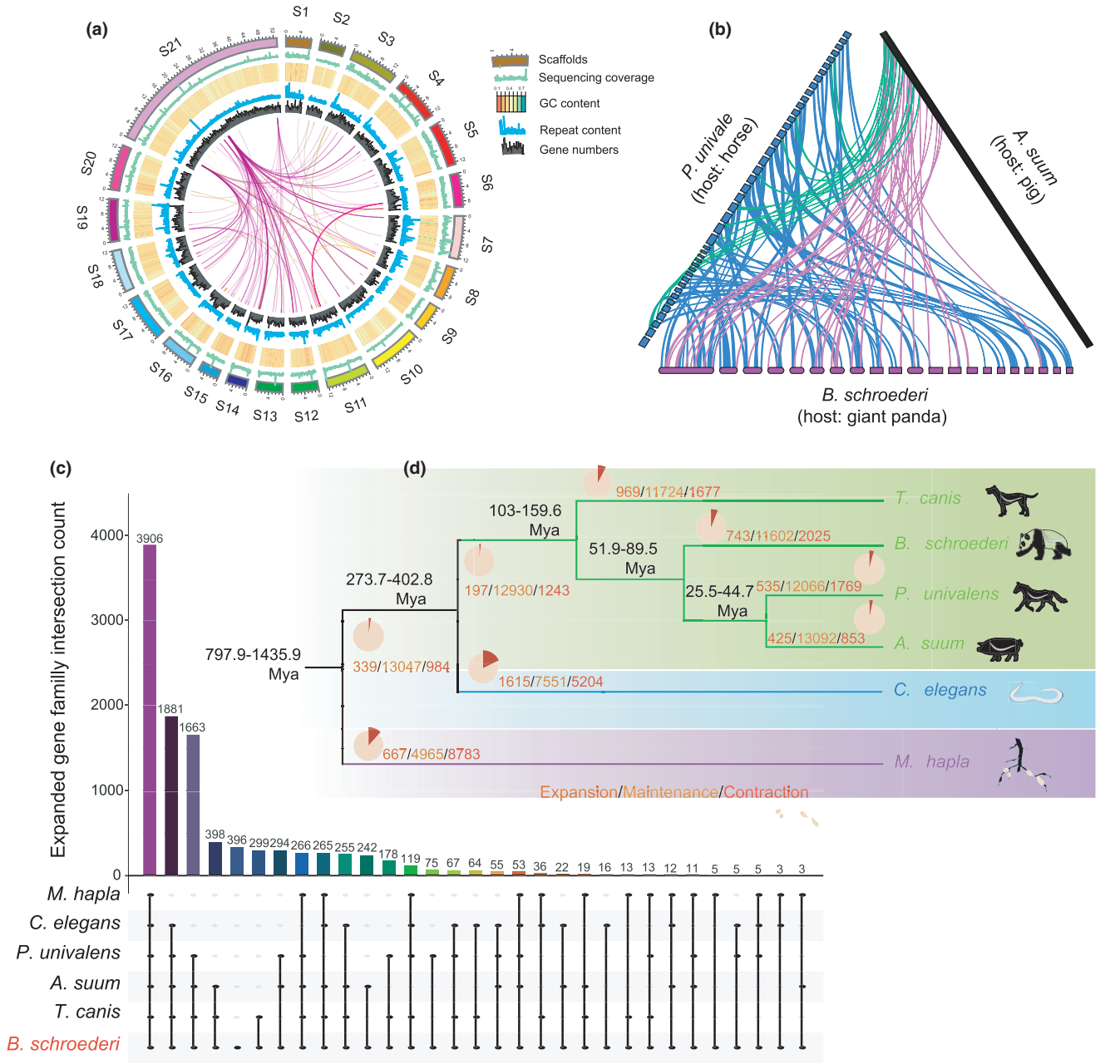

FIGURE 1 The phylogenetic relationships among six nematodes and genomic characteristics and synteny of *B. schroederi*. (a) Genomic characterization of *B. schroederi* genome. The figure shows the gene number, repeat content, GC content, sequencing coverage and scaffolds from the centre to the edge. (b) Synteny of *B. schroederi* with *P. univalens* and *T. canis* at the gene level. Different colours represent different synteny blocks. (c) Upset plot showing the intersection of gene family expansions in nematodes. Each row represents a nematode. Black circles and vertical lines between the rows represent the intersection of expanded families between species. The barplot indicates the total gene family count in each intersection. (d) Time-calibrated maximum likelihood phylogenetic tree of six nematodes. The estimated divergent times are shown at the bifurcations. The numbers below the nodes represent the number of gene families significantly expanded, maintained, and contracted, respectively [Colour figure can be viewed at wileyonlinelibrary.com]

TEs constitute 4.4% and 13.5% in the genomes of *A. suum* (Jex et al., 2011) and *T. canis* (Zhu et al., 2015), respectively. We identified a significant expansion of DNA transposons in *B. schroederi* compared to *T. canis* (Zhu et al., 2015) and *A. suum* (Data S1) (Jex et al., 2011). There are at least 64 DNA transposon families of which CMC-EnSpm, DNA and MULE-MuDR dominated the genome. We identified 17 long terminal repeats (LTRs) retrotransposon and 41 non-LTRs retrotransposon families (25 LINE and 16 SINE). Pao and Gypsy are the predominant LTRs, and CR1, RTE-RTE and L2 are the predominant non-LTRs. The number and size of the retrotransposon families are similar to those of other related roundworms (Ghedin et al., 2007; Jex et al., 2011; Zhu et al., 2015). In addition, four types of non-coding RNA (ncRNA; including tRNA, snRNA, miRNA, and rRNA) were predicted (Table S9).

TABLE 1 Summary of the features of the *B. schroederi* genome

Description	*B. schroederi* (this study)	*B. schroederi*[a]	*A. suum*	*P. univalens*	*T. canis*
Genome size (bp)	253,610,985	281,639,769	265,545,801	253,353,821	317,115,901
Number of scaffolds; contigs	75; 536	2778; 15,567 (>1000 bp)	31,538; 40,509	1274	22,857; 51,969
Average length of scaffolds; contigs (bp)	3,381,480; 436,058	–	8420; 6506	198,864; 6918	13,874; 5747
Gap length (bp; % of genome)	261,088 (0.10%)	13,257,555 (4.70%)	1,980,846 (0.7%)	966,507 (0.38%)	18,474,532 (5.82%)
N50 of scaffolds; contigs (bp)	12,324,682; 1,221,088	888,870; 42,126	290,558; 46,632	1,825,986; 20,376	375,067; 16,980
N90 of scaffolds; contigs (bp)	6,864,504; 192,063	104,281; 7439	48,674; 10,466	204,976; 2991	66,363; 3511
Genome GC content (%)	37.30	37.26	37.85	39.07	39.95
Repetitive sequences (%)	9.53	–	4.4	–	13.5

[a]The published genome information of *B. schroederi* (not released) (Hu et al., 2020).

3.2 | Development-related genes for key enzymes, ion channels, receptors and secretome

To understand the genetic basis for the adaptation of *B. schroederi* to the parasitic life, we assessed the abundance of several major protein classes in *B. schroederi*, *A. suum*, *P. univalens*, and *T. canis* (Table S10 and Figure S12b). A genome-wide search enabled us to identify 62 G-protein coupled receptors (GPCRs), 437 proteases and protease inhibitors (Data S2c). Some chemoreceptor families, especially those differentially expressed during the life cycle, were almost completely conserved among nematodes. For example, the homologues of *Caenorhabditis elegans daf-37* (GeneID: Baysch11898) and *daf-38* (GeneID: Baysch06944), which are known to mediate ascaroside signalling, are expressed during the transition from dauer larvae to infective larvae (Park et al., 2012). By comparing with the ligand-gated ion channel gene set collected from wormBase, 65 genes were identified, including members of the previously described nematode acetylcholine receptor classes (*deg-3*, *acr-16*, *unc-29* and *unc-38*) (Coghlan, Tyagi, et al., 2018). We predicted excretory/secretory proteins (EPs) of *B. schroederi*, and at least 1395 EPs (5.26% of total protein) with diverse functions were identified, including 1046 conventionally secreted proteins and 349 nonconventionally secreted proteins (Data S2d–e).

3.3 | Evolutionary and phylogenetic relationships among six nematodes

We clustered the *B. schroederi* gene models with the genes from five other nematode genomes (*A. suum*, *P. univalens*, *T. canis*, *C. elegans*, and *Meloidogyne hapla*; Table S11). We found that the six nematodes share 3906 homologous gene families (Figure 1c). In addition, four roundworms show high consistency in the number of single-copy and multicopy genes (Figure S11c). Collinearity results showed that although several roundworms are closely related, collinearities among the genomes are low (Figure 1b). The proportions of collinearity between *B. schroederi* and the other three roundworms in the genome species are 35.37% (vs. *T. canis*), 43.86% (vs. *A. suum*), and 55.12% (vs. *P. univalens*), respectively, which indicate that genetic differentiation among roundworms is considerable.

We used 2451 single-copy genes shared within the six nematode genomes to reconstruct a phylogenetic tree (Figure 1d). The relationships among the six nematodes in the phylogenetic tree are consistent with a previous study (Coghlan, Tyagi, et al., 2018). *B. schroederi* is closely related to *A. suum* and *P. univalens*. According to the TimeTree (Hedges, 2011) database and fossil evidence from *A. suum* and *C. elegans* (Mcgill et al., 2017), we estimated the divergence time to approximately 400–269 million years ago (Ma), and the divergence time between the four roundworms is approximately 160–26 Ma (Figure 1d). Among the four roundworms, *T. canis* was identified as the earliest branch to the other three roundworms (approximately 134 Ma).

3.4 | Expanded and contracted genes in Ascariasis

Compared with *C. elegans* and *M. hapla*, a large number of genes have been lost or contracted in the branch of roundworms (Figure 2a). Specifically, we found that 563 gene families are significantly contracted ($p < .05$), with an average loss of 1.61 genes in each family (Table S12). However, roundworms also show significant expansion in 29 gene families ($p < .05$). We focused on the changes in gene number related to free life and those related to parasitic life in nematodes. For both the expanded and contracted gene families, KEGG enrichment analysis showed that pathways related to tissue development, metabolism and environmental information processing had undergone significant changes (Figure S13a,b). Interestingly,

FIGURE 2 The expansion and contraction of roundworm gene families. (a) Significant increases and decreases in roundworm gene families. The solid circle and the solid triangle represent the top KEGG pathways that are enriched in the expanded gene families of Ascariasis compared with *C. elegans* or *M. hapla*, respectively. The open circle and the open triangle represent the top KEGG pathways that are enriched in the contracted gene families of Ascariasis compared with *C. elegans* or *M. hapla*, respectively. (b) GO function enrichment and gene copy number of the significantly expanded gene families in roundworms; (c) The proportion of GO functional genes in the gene family with significant expansion (or contraction) in roundworms compared to the total number of expansion (or contraction) genes. The red asterisk represents the *p*-value of statistical Sidak's multiple comparisons tests of expansion and contraction of genes comparing with *C. elegans* or *M. hapla* (one asterisk represent 10^{-1}) [Colour figure can be viewed at wileyonlinelibrary.com]

a similar expansion also appeared in the roundworm branch compared to *M. hapla* (Figure S13c–d). For tissue development, we found that the chitin-binding protein *CPG-2* gene family issignificantly expanded in the roundworm branch ($p < .01$). In addition, a significant expansion of the tight junctions (ko04530), phagosome pathway (ko04145) and Rap1 signalling pathway (ko04015) was observed ($p < .01$; Figure 2a). In relation to self-defense, consistent with a previous study, we observed an expansion of the chymotrypsin/elastase inhibitor gene family, which may be related to a protection of roundworms from host proteases (Coghlan, Tyagi, et al., 2018). In addition, according to the copy number statistics of expanded gene families, the actin family is significantly expanded among all four roundworms (Figure 2b). GO enrichment analysis showed that the genes related to nematode behavior and biological adhesion accounted for the most significant difference among all gene families exhibiting expansions in roundworms (Figure 2c).

3.5 | Expanded or contracted gene families related to the adaptation of *B. schroederi*

To better understand the adaptation to the specific habitat and intestinal environment of the giant panda, we analysed the expanded and contracted genes in *B. schroederi* compared with three roundworms (*A. suum*, *P. univalens* and *T. canis*).

We identified expanded gene families with functions involved in striated muscle contraction (GO:0006941), nematode larval development (GO:0002119), larval feeding behavior (GO:0030536), chitin metabolic process (GO:0006030) and actin cytoskeleton organization (GO:0030036). The most significantly enriched GO term was the straight muscle contraction, which was largely due to the highly significant expansion of the actin family (Figure 3a,d). KEGG enrichment showed that the number of genes involved in metabolic pathways, including drug metabolism (metabolism of xenobiotics by cytochrome P450, ko00980; $p < .01$) and autoimmunity (ko05130, ko05100; $p < .01$), exhibited significant changes (Figure 3a). We observed an expansion of the gene family involved in positive regulation of eating behaviour (Figure 3b). Finally, we found that *B. schroederi* has 654 unique annotated proteins, which were mainly enriched in the synthesis and recycling pathways of essential amino acids especially the degradation of valine, leucine and isoleucine (ko00280, $p < .01$; Figure 3c).

3.6 | The significant expansion of the actin family in Ascariasis and positive selection genes (PSGs) in *B. schroederi*

The migration of roundworms is the main cause of VLM (Figure 4a). Actin polymerization is controlled by intracellular signals that are mediated by small GTPases of the Rho family (*RhoA*, *Rac1*, and *Cdc42*; Figure 4c) (Ammassari-Teule & Segal, 2017). We observed a very significant expansion of the actin family in the *B. schroederi* genome (Figure 3d). Surprisingly, the three upstream regulators of actin (*Rac1*, *ROCK* and *MLCK*), were under strong positive selection ($p < .01$). *ROCK* phosphorylates the LIM-kinase which then phosphorylates cofilin to promote rho-induced actin cytoskeleton reorganization (Maekawa et al.,1999). These finding suggest a possible effect on myosin-actin interaction. Using the branch-site model implemented in PAML, 475 genes in the *B. schroederi* genome were found to be under strong positive selection compared with other three roundworms (Data S3a). Compared with three roundworms, the acetylcholine receptor subunit alpha-type deg-3 ($p = 6.5 \times 10^{-3}$), which is an important drug target (Jones et al., 2007), was shown to be under strong positive selection in the *B. schroederi* genome (Data S3a).

3.7 | Demographic history of *B. schroederi*

To reconstruct the demographic history of *B. schroederi*, and explore the relationship with the giant panda, we used a pairwise sequentially Markovian coalescent (PSMC) model to estimate the changes of the effective population size of both *B. schroederi* and the giant panda within the last one million years (Li & Durbin, 2011). PSMC analysis showed that the effective population sizes (N_e) of *B. schroederi* and the host giant panda have almost exactly the same trend from 300 kya (thousand years ago) to 10 kya (Figure 5a), but the change of effective population size of the roundworms showed a slight lag. The effective population size of the giant panda declined significantly during the last two Pleistocene glacial periods (300–130 kya and 50–10 kya), and the effective population size of the roundworm also reached a historical low level. The most obvious change occurred in the Greatest Lake Period (50–30 kya), where the effective population size of the two species reached their pinnacle (Jinchu & Wei, 2004). In addition, we applied multiple sequentially Markovian coalescent (MSMC) method (Schiffels & Durbin, 2014) to infer the recent demographic events of *B. schroederi* and performed five repetitions (each repetition used four individuals per population). The effective population size of roundworms showed a sharp decline in the last 10,000 years, which also may be related to the host population dynamics. Studies have indicated that human activities may have caused the decline of the giant panda population in the past few thousand years, and the roundworm population may be affected by this.

3.8 | Population structure of *B. schroederi* populations

We carried out whole-genome resequencing of 266 samples, including 240 from captive pandas of the Sichuan subspecies and 26 samples from individuals of the wild Qinling panda subspecies (Figure 6a). The average sequencing coverage and sequencing depth reached 97.91% and 41-fold, respectively (Tables S1 and S13). A total of 6.32 million SNPs were obtained after filtering (see Section 2). PCA supported the clear separation between *B. schroederi* from the captive Sichuan and the wild Qinling panda subspecies (Figure 6b), with PC1 separating the Qinling and Sichuan populations and PC2 separating the Sichuan population into two clusters ($p < .05$). We constructed a phylogenetic tree using the maximum likelihood (ML) method, which showed two distinct clusters in the whole population, with all Qinling individuals forming a single cluster and all Sichuan individuals forming another single cluster (Figure 6c). In addition, the results of structure analyses also indicated that there were almost no shared ancestral components between the Qinling and Sichuan populations, supporting the results from the phylogenetic tree and the PCA (Figure 6d). Although the two populations showed extremely similar genetic diversity (Figures S14 and S15), our results consistently supported two distinct groups corresponding to the Sichuan and Qinling populations.

3.9 | Recent positive selections in the *B. schroederi* populations

We used integrated haplotype score (iHS) to detect genes under recent natural selection in the captive and wild populations. A total

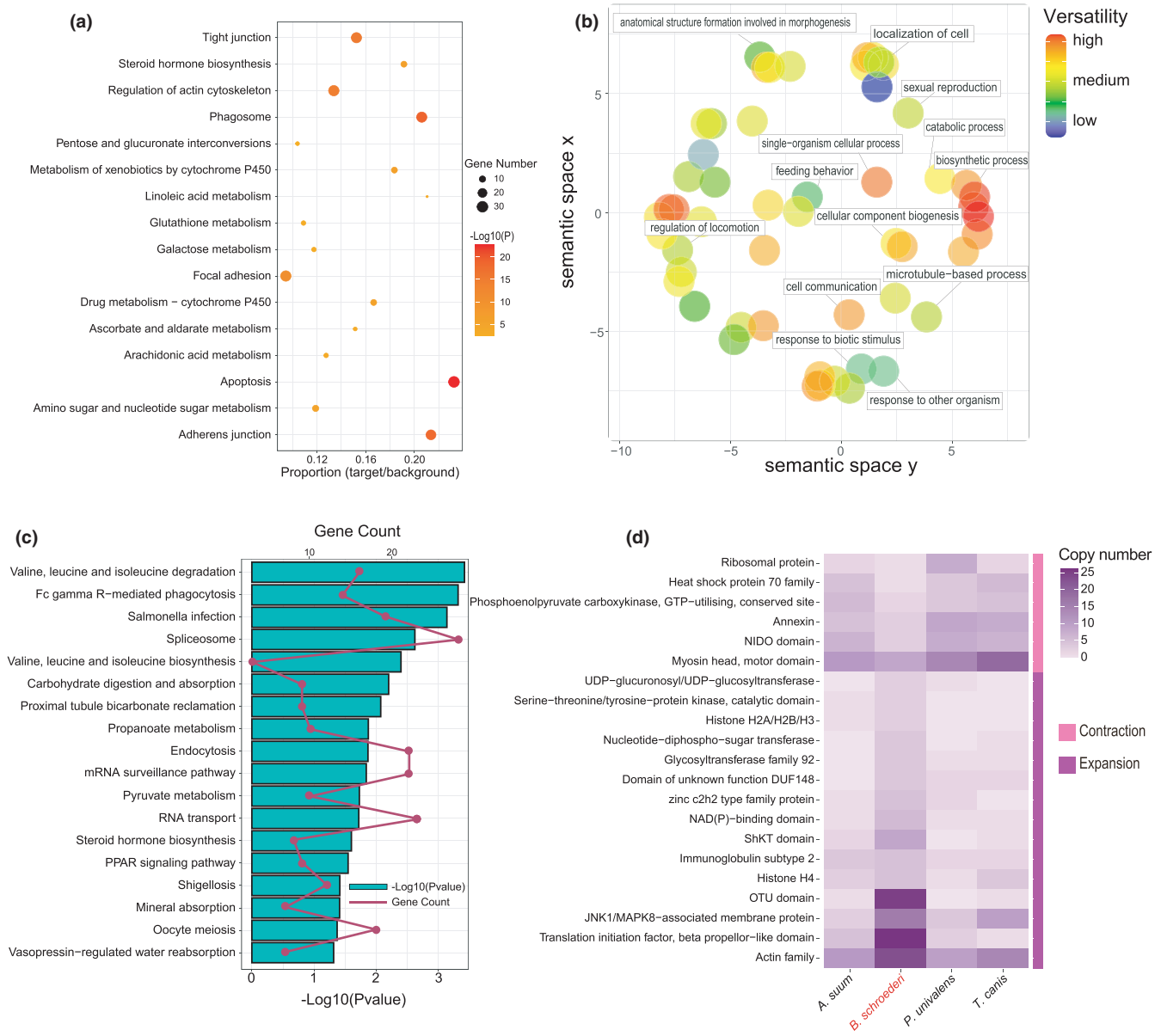

FIGURE 3 Expansion and contraction of *B. schroederi* gene families compared with three roundworms (*A. suum P. univalens* and *T. canis*). (a) Enrichment of the KEGG pathway in some significantly expanded gene families of *B. schroederi*. The proportion represents the ratio of the number of expanded genes located in the pathway (target genes) to all genes in the pathway (background genes). (b) REVIGO clusters of significantly overrepresented GO items for significantly expanded gene families in *B. schroederi*. The position of the bubbles is based on semantic similarity of GO terms. (c) Enrichment of KEGG pathways in *B. schroederi's* unique gene families. (d) Heatmap showing the gene families of *B. schroederi* that are significantly expanded or contracted ($p < .01$). The *x*-axis represents the four roundworms of Ascariasis, whereas the *y*-axis represents the families [Colour figure can be viewed at wileyonlinelibrary.com]

of 29,553 SNPs in captive and 18,953 SNPs in wild were identified within the top 1% iHS scores. By extending the 25 kb distance around the top 1% SNPs, filtering out SNPs in the nongene regions, a total of 518 and 370 genes were located in the positively selected regions in captive and wild populations, respectively (Data S3b,c). We further calculated the distribution of nucleotide diversity on the 21 superscaffolds (Figure S15), and found that the genetic diversity in some regions was significantly lower than in the flanking genome regions such as glutamate-gated chloride channel alpha (*glc-1*, a receptor for anthelmintic ivermectin [Cook et al., 2006]), nose resistant to fluoxetine protein 6 *nrf-6*, a fluoxetine (Prozac) resistance gene (Choy et al., 2006; Fares & Grant, 2002), ABC transporter ced-7 (*ced-7*, phagocyte corpse) (Wu & Horvitz, 1998) and β-1,4-N-acetylgalactosaminyltransferase bre-4 (*bre-4*, resistance to pore-forming toxin [Griffitts et al., 2003]) genes (Figure 7b). Interestingly, the four genes were only found under positive selection in the captive population, but not in the wild population. In addition, we also used the cross-population extended haplotype homozygosity (XP-EHH) method (Sabeti et al., 2007) to screen for genes that might have been positively selected by different deworming selection pressures by comparing the captive and wild populations (Figures S16 and S17, Data S3d). Similarly, the genes encoding the multidrug

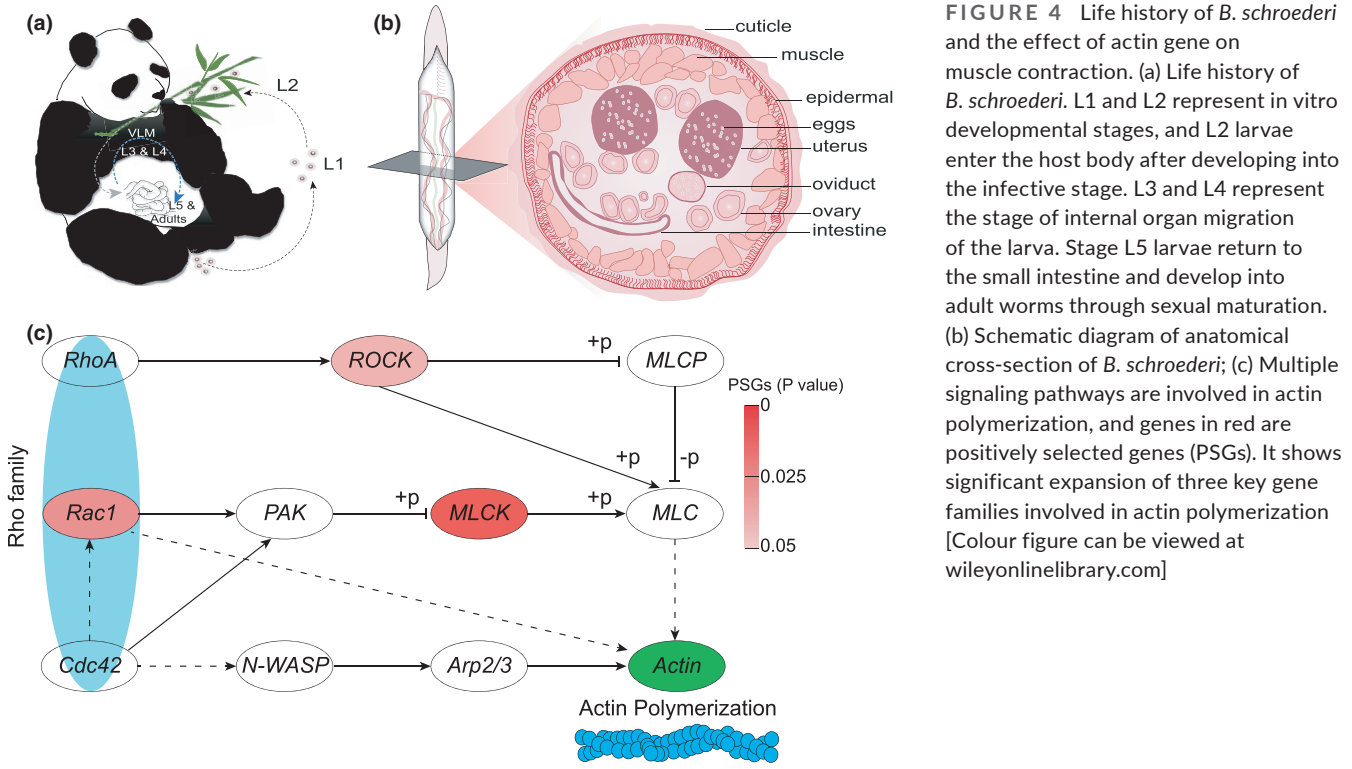

FIGURE 4 Life history of *B. schroederi* and the effect of actin gene on muscle contraction. (a) Life history of *B. schroederi*. L1 and L2 represent in vitro developmental stages, and L2 larvae enter the host body after developing into the infective stage. L3 and L4 represent the stage of internal organ migration of the larva. Stage L5 larvae return to the small intestine and develop into adult worms through sexual maturation. (b) Schematic diagram of anatomical cross-section of *B. schroederi*; (c) Multiple signaling pathways are involved in actin polymerization, and genes in red are positively selected genes (PSGs). It shows significant expansion of three key gene families involved in actin polymerization [Colour figure can be viewed at wileyonlinelibrary.com]

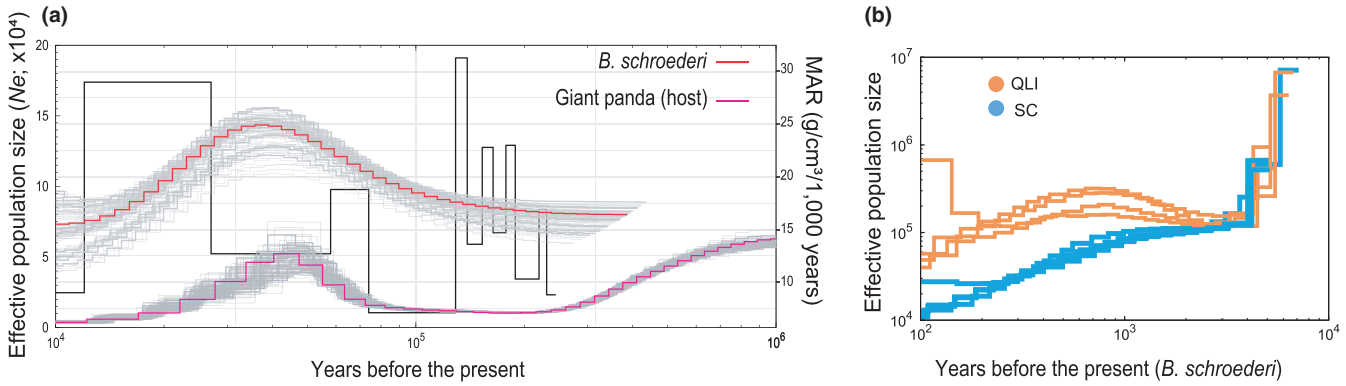

FIGURE 5 Demographic history of the *B. schroederi* reconstructed from the reference and population resequencing genomes. (a) The red and purple line represent the estimated effective population size of *B. schroederi* and host, respectively. The 100 grey curves of *B. schroederi* and host represent the PSMC estimates for 100 sequences randomly resampled from the original sequence. Generation time (g) of *e* and giant panda were 0.17 and 12 years, respectively. The neutral mutation rate per generation (μ) of *B. schroederi* and giant panda were 0.9×10^{-8} and 1.3×10^{-8}, respectively. The black line shows the MAR of Chinese loess. (b) Longitudinal change of the effective population size of the *B. schroederi* populations. The effective population sizes (N_e) were estimated using the MSMC2 method. QLI, Qinling population; SC, Sichuan population [Colour figure can be viewed at wileyonlinelibrary.com]

resistance protein *pgp-3* (Xu et al., 1998) which is related to ivermectin resistance, *glc-1*, *nrf-6*, cytochrome p450 (*CYP*) family members and other drug resistance related genes, were also observed to be under strong positive selection in the captive population (Figure 7c).

3.10 | Identifying anthelmintic resistance-related gene family and drug targets

We use the HMMER3 software to scan several detoxification-related gene families at the whole genome scale, including ATP-binding cassette (ABC) transporters, cytochrome P450 (CYP), glutathione S-transferase (GST), glycoside hydrolase family 18 (CHIA), patched family (PTCHD) and protein tyrosine phosphatase family (PTP) (Figure S12d). A total of 97 ABC transporters, multipass membrane proteins, were identified in *B. schroederi*, and the average number of ABC transporter genes in roundworms was greater than that in *C. elegans* (60) (Schumacher & Benndorf, 2017). The high-quality genome data of *B. schroederi* provides an opportunity to identify biologically active anthelmintic compounds. On the one hand, it enables identification of the targets of existing anthelmintics, on the other hand, it also enables identification of new potential targets for compounds

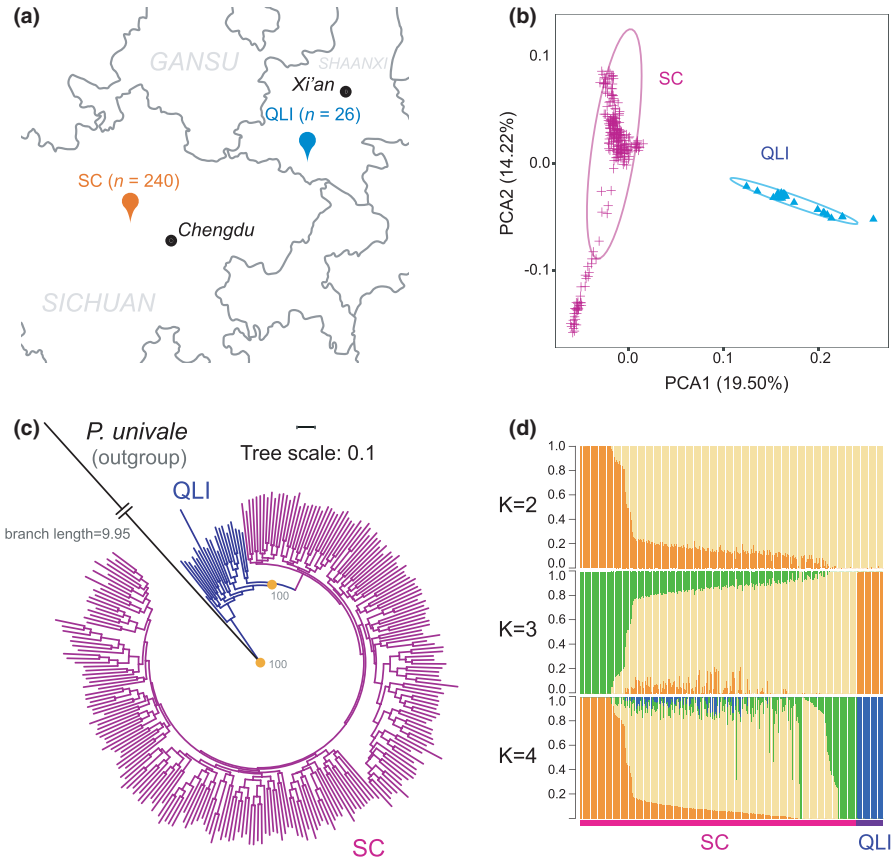

FIGURE 6 Population structure and relationships of Sichuan (SC) in comparison to Qinling (QLI) population. (a) The geographic distribution of the sampling locations for QLI and SC populations. (b) Principal component analysis (PCA) analysis of two populations. (c) A maximum likelihood (ML) phylogenetic tree with 100 bootstrap tests constructed using whole-genome SNPs information. We used *P. univale* as the outgroup. (d) Population structure of SC and QLI populations (K from 2 to 5). The *y*-axis quantifies the proportion of the individual's genome from inferred ancestral population, and *x*-axis shows the different individuals [Colour figure can be viewed at wileyonlinelibrary.com]

from other areas of drug discovery. All compound-related proteins were searched against target proteins from CHEMBL v26 using BLASTP ($E \leq 1 \times 10^{-10}$), and a total of 4554 small molecules with recorded biological activities were identified. By blasting against the ChEMBL databases (Gaulton et al., 2012), a total of 90 known genes, which encode specific drug targets were identified. The corresponding drugs (13 drugs used to treat humans with World Health Organization (WHO) ATC code P02 "WHO anthelmintics" and 10 drugs from DrugBank Wishart et al., 2017) were further collated by searching DrugBank databases and the literature (Data S4a). Some of these drugs have been proven to be effective against *B. schroederi*, such as albendazole (Fu et al., 2011), mebendazole (Bourne et al., 2010), pyrantel (Xie et al., 2011) and ivermectin (Wang et al., 2015). Many existing anthelmintics are compromised by the increase of resistance in roundworm populations (Li et al., 2015). In addition to known drugs, we were committed to identifying new potential drug targets. We focused on single protein ChEMBL targets that may be easier to develop drugs against than protein complexes (Coghlan, Tyagi, et al., 2018). By blasting against target proteins (similarities >80%) in the single protein database ChEMBL, we identified 95 genes encoding single proteins. Then we sat a score of "0/1" considering six main factors to evaluate the potential of the protein as a drug target (see Section 2; Figure 8). Finally, we located the position of all the drug target encoding genes on superscaffolds (Figure 8). Since the existing phase III and above drugs have greater potential for being developed into new anthelmintics, we searched for commercially available compounds against each target protein although these compounds were not originally designed as anthelmintics. Among all the proteins, we found that three target genes (*cmd-1*, *Ap2s1*, *HRAS*) have available compounds with phase III/IV approvals (Data S4b). These potential drug targets and compounds will provide a possible starting point for the development of new anthelmintics.

4 | DISCUSSION

Baylisascaris schroederi exhibits strong environmental adaptability and wide distribution, and is a threat to the health of giant pandas (Zou et al., 1998). In-depth studies of *B. schroederi* have been hampered by the lack of a high-quality genome sequence. The scaffold N50s of published *A. suum*, *P. univalens* and *T. canis* genomes are 290, 1825 and 375 kb, respectively (Table 1). In this study, we present the first chromosome-scale genome assembly of the *B. schroederi* with the scaffold N50 of 12.32 Mb, representing a genome assembly with the best contiguity in Ascarididae. We envisage that this genome will provide a valuable and useful genetic resource for future research on roundworms, as well as drug development for expulsion.

Roundworms have special characteristics that are different from free-living nematodes reflecting the adaptation to the parasitic life. Eggs of roundworms have a tough and elastic polysaccharide chitin shell, which enables eggs to persist in the soil for up to 10 years (Fairbairn, 1970). We have observed a significant expansion of the chitin-binding protein *CPG-2* family in roundworm

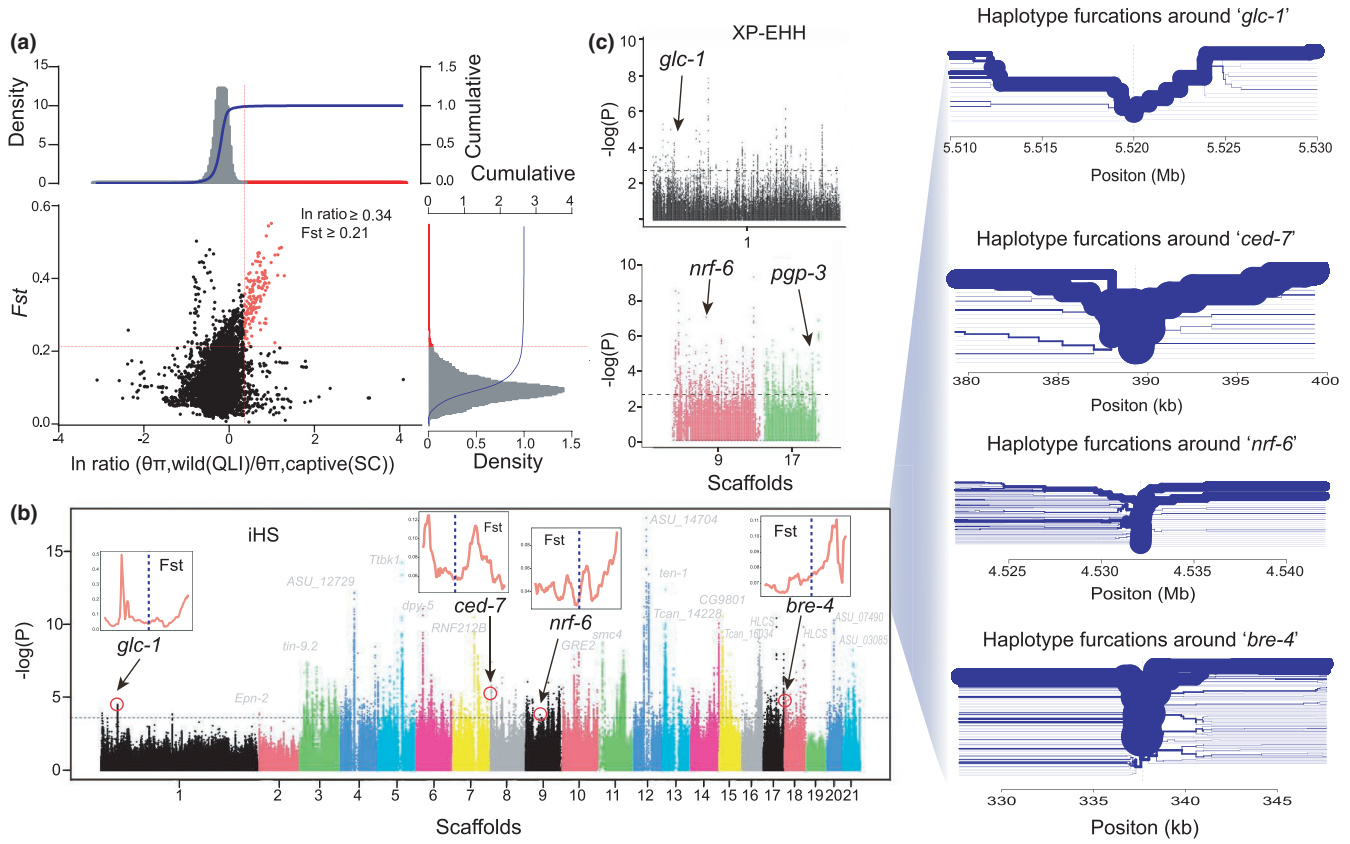

FIGURE 7 Analysis of natural selection in captive populations. (a) Genomic regions with selection sweep signals in captive (SC) and wild (QLI) *B. schroederi* population. Distribution of ln ratio ($\theta_{\pi,\text{wild(QLI)}}/\theta_{\pi,\text{captive(SC)}}$) and F_{ST} of 50 kb windows with 10 kb steps. Red dots represent windows fulfilling the selected regions requirement (corresponding to *Z* test $p < .005$, where $F_{ST} > 0.21$ and ln ratio >0.34). (b) Plot of iHS showing loci under positive selection of captive (SC) population. SNPs with |iHS|≥iHS$_m$ (3.89, top 1%) were shown above the dashed horizontal line. Nucleotide diversity around *glc-1*, *nrf-6*, ABC transporter *ced7* and *bre-4* loci using 10-kb sliding windows were displayed above the genes. The decay of haplotype homozygosity around a focal marker were displayed on the right side of the figure. The furcation structures represent the complete information contained in the concept of extended shared haplotypes EHH (Sabeti et al., 2002). The root (focal marker) is indicated by a vertical dashed line. The thickness of the lines corresponds to the number of scaffolds sharing a haplotype. (c) XP-EHH from each SNP core showing the same nucleotide between the subject and the comparison target, also transformed to *p*-values and plotted in logarithmic scale [Colour figure can be viewed at wileyonlinelibrary.com]

branches, which may be related to the formation of the roundworm eggshell, thereby prolonging the survival of roundworms even in a harsh environment. In addition, in the parasitic stage, larvae enter the intestine, penetrate the intestinal wall, migrate among tissues and organs (Kazacos & Boyce, 1989), molt and develop, finally return to the small intestine to develop into adults, mate, and lay eggs. Some genes potentially involved in tissue invasion and immune evasion have been significantly expanded in roundworms, including genes homologous to metallopeptidase and serine/threonine-protein kinase, respectively. Previous studies have shown that metallopeptidase(s) in the secretory products of astacins in the nematode epidermis can digest collagen in host tissues, and thus be involved in the migration of larvae in viscera (Soblik et al., 2011; Williamson et al., 2006).

Although the morphological characteristics of Ascarididae species are similar, the *B. schroederi* still shows unique molecular evolutionary traits. The giant panda has gradually evolved in response to the bamboo diet during millions of years of evolution (Zhou et al., 1997). In the *B. schroederi* genome, several unique gene families of *B. schroederi* were found to be involved in the metabolism of essential amino acids, especially the degradation of valine, leucine and isoleucine (KO00280; $p < .01$), which is likely to enhance the ability of *B. schroederi* to absorb nutrients. Actin promotes muscle contraction and plays a very important role in the movement and migration of *B. schroederi* in the host. The significant expansion and positive selection of the actin family may have provided the driving force for muscle contraction and cell movement (Hall, 1998). Studies have shown that actin is involved in the repair of nematode epidermis damage (Suhong & Chisholm, 2012), which is of great significance to the migration of *B. schroederi* in the giant panda. The expansion of the actin gene family may, at least to some extent, explain the genetic basis of stronger locomotion ability of *B. schroederi* than other roundworms.

According to a previous investigation, the cause of death of giant pandas in recent decades has shifted from starvation and poaching to VLM-related deaths (Zhang et al., 2008). Frequent use of drugs

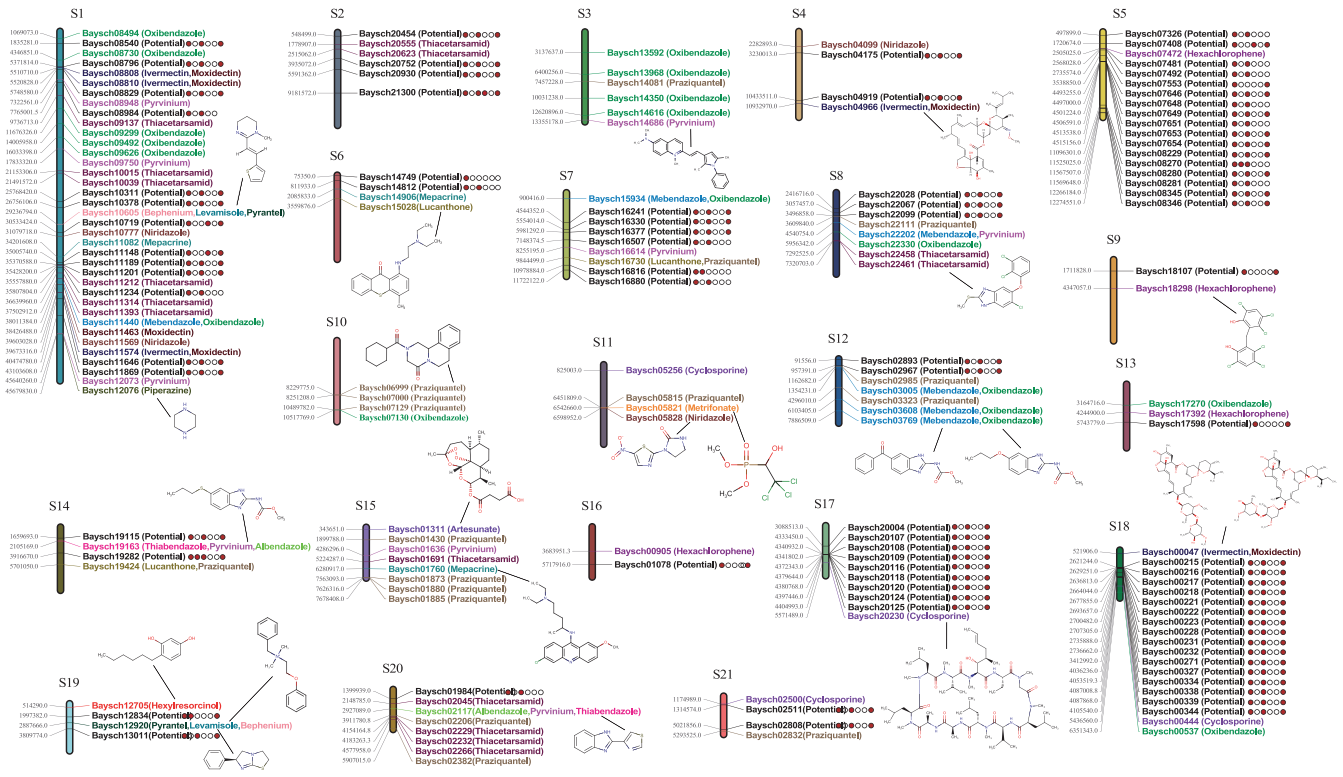

FIGURE 8 The position of known and potential drug target genes on superscaffolds. Different colours indicate different known drugs, and black indicates potential drug targets. The chemical structural formulas of 23 known drugs are drawn. The circles following the potential drug targets represent the six criteria, with a red solid circle indicating match condition and a hollow circle indicating mismatch condition. The six criteria were: (1) Similarity with ChEMBL known drug targets having a highly conserved alignment (>80%). (2) Lack of human homologues. (3) Related to lethal, L3 arrest, flaccid, molt defect or sterile phenotype. (4) A predicted metabolic chokepoint. (5) A predicted excretory/secretory protein (EP). (6) The protein has a structure in the PDBe. Potential drug target proteins encoding genes on each superscaffold and corresponding scores are marked (black) [Colour figure can be viewed at wileyonlinelibrary.com]

may drive the increasing frequency of genes related to drug resistance in the population, leading to widespread drug resistance in the *B. schroederi* population. Furthermore, there have been reports of side effects in giant pandas after the administration of existing anthelmintic drugs (Wang et al., 2015). We observed a recent significant positive selection of *ABC* and *CYP* family members and other resistance-related genes (*glc-1*, *nrf-6*, *pgp-3* and *bre-4*) in captive (SC) populations. Although wild and captive populations were obtained from two different regions (Qinling and Sichuan), natural selection analysis mainly considers recent changes in gene frequency. The two populations are facing completely different selection pressures for deworming, and thus, offer an option for evaluating natural selection trends of a few resistance-related genes. The results indicated an increased frequency of drug resistance-related genes in captive populations. This may be related to the frequent use of drugs in recent decades. Although the degree of natural selection in the current resistance areas cannot be quantified, it is possible that the gene frequency of these genes is still increasing, and it may cause the emergence and increase of resistant individuals. Studies have shown that some new sources of infection may even evolve into potential antibiotic-resistant pathogens (Zumla & Hui, 2019). Therefore, the identification of drug-resistance genes and the detection of drug-resistant individuals are still essential in future works.

There is an urgent need for new anthelmintic drugs for intestinal expulsion of roundworms (James et al.,2009; Jia et al.,2012). Specifically, there is a pressing need for new anthelmintic drugs to protect the giant panda, since existing drugs suffer from low efficacy, serious side effects or rising drug resistance in parasite populations due to increased frequency of use (Wang et al., 2015). The chromosome-scale genome of *B. schroederi* provides a reference for the development of species-specific drugs, and drug targets can be screened at the whole genome level. We identified a total of 90 known drug targets and 95 potential drug targets, providing a basis for the development of follow-up drugs. We searched four compounds (lonafarnib, haloperidol, trifluoperazine and chlorpromazine; Data S4b) that have a phase 3/4 approval. These compounds could be considered for repurposing as novel anthelmintics, which would save considerable effort and expenses. Nevertheless, the anthelmintic activity of these compounds and other potential target compounds needs further testing. We envisage that such works will provide new modalities for the prevention and treatment of baylisascariasis and other parasitic diseases.

5 | CONCLUSIONS

Roundworms have undergone a remarkable evolutionary adaptation to specific hosts accompanying host evolution. The roundworm genome provides a possibility to study the details of gene selection or loss in the process of evolution and adaptation to the intestinal environment, including genes involved in epidermal chitin synthesis, environmental information, and essential amino acid metabolism. In addition, population genomic analysis and drug prediction provide insights revealing the impact of deworming history on population genetic structure and future development of novel anthelmintic drugs.

ACKNOWLEDGEMENTS

This work was supported by the National Key R&D Programme (no. 2017YFD0501702), Open Project of Key Laboratory of SFGA on Conservation Biology of Rare Animals in The Giant Panda National Park (CCRCGP, no. 2020004), Forestry science and technology research project (no. 20180302), the Pearl River Talent Recruitment Programme in Guangdong Province (2019CX01N111), Fundamental Research Funds for the Central Universities (no. 2572020AA30) of China, the Foundation of Key Laboratory of State Forestry and Grassland Administration (State Park Administration) on Conservation Biology of Rare Animals in the Giant Panda National Park (no. KLSFGAGP2020.002), and the Guangdong Provincial Key Laboratory of Genome Read and Write (grant no. 2017B030301011).

CONFLICT OF INTEREST

The authors declare no conflict financial interests.

AUTHOR CONTRIBUTIONS

Z.H., and H.L. designed and initiated the project. Z.P., D.L., S.H., Y.W., H.S., H.C., and L.D. collected the samples. X.C., L.H., L.L., Y.L., Y.Z., J.Y. and H.R.L. performed the DNA extraction, library construction and sequencing. L.H., T.L., H.M.L., R.H., Q.W., Y.Z., and S.Y. performed the data analysis. L.H., T.L. and wrote the manuscript. Z.H., K.K., M.L., D.L., H.L., Q.L., H.Y. and X.X. supervised the manuscript. All authors read and approved the final manuscript. L.H. and T.L. wrote the manuscript with input from Z.H., K.K., M.L., D.L., H.L., Q.L., H.Y. and X.X. Z.H., Z.H., K.K. and Q.L. supervised the project.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study have been deposited into CNGB Sequence Archive (CNSA) (Guo et al., 2020) of China National GeneBank DataBase (CNGBdb) (Feng et al., 2020) with CNSA project ID CNP0001242. All in-house scripts used in this study were available in the github database (https://github.com/HanLei12321/B.schroederi_genome).

ORCID

Zhijun Hou https://orcid.org/0000-0002-8704-1651

REFERENCES

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

Ammassari-Teule, M., & Segal, M. (2017). Dendritic spine plasticity and memory formation. In D. Robin (ed.), *Learning and memory: A comprehensive reference* 2nd ed., (pp. 199–215). Elsevier. https://doi.org/10.1016/B978-0-12-809324-5.21113-9

Amos, B., & Rolf, A. (2000). The SWISS-PROT protein sequence database and its supplement TrEMBL in 2000. *Nucleic Acids Research*, *28*(1), 45–48.

Archibald, A. L. (2017). *Exploiting long read sequencing technologies to establish high quality highly contiguous pig reference genome assemblies*. Plant and Animal Genome Conference.

Benson, G. (1999). Tandem repeats finder: A program to analyze DNA sequences. *Nucleic Acids Research*, *27*(2), 573–580. https://doi.org/10.1093/nar/27.2.573

Bethony, J., Brooker, S., Albonico, M., Geiger, S. M., Loukas, A., Diemert, D., & Hotez, P. J. (2006). Soil-transmitted helminth infections: Ascariasis, trichuriasis, and hookworm. *The Lancet*, *367*(9521), 1521–1532. https://doi.org/10.1016/S0140-6736(06)68653-4

Bie, T. D., Cristianini, N., Demuth, J. P., & Hahn, A. 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

Bo, X., & Yang, Z. (2013). pamlX: A graphical user interface for PAML. *Molecular Biology and Evolution*, *30*(12), 2723–2724.

Bourne, D., Cracknell, J., & Bacon, H. (2010). Veterinary issues related to bears (Ursidae). *International Zoo Yearbook*, *44*(1), 16–32. https://doi.org/10.1111/j.1748-1090.2009.00097.x

Browning, S. R., & Browning, B. L. (2007). Rapid and accurate haplotype phasing and missing-data inference for whole-genome association studies by use of localized haplotype clustering. *The American Journal of Human Genetics*, *81*(5), 1084–1097. https://doi.org/10.1086/521987

Campbell, M. S., Law, M. Y., Holt, C., Stein, J. C., Moghe, G. D., Hufnagel, D. E., Lei, J., Achawanantakun, R., Jiao, D., Lawrence, C. J., Ware, D., Shiu, S.-H., Childs, K. L., Sun, Y., Jiang, N., & Yandell, M. (2013). MAKER-P: A tool kit for the rapid creation, management, and quality control of plant genome annotations. *Plant Physiology*, *164*(2), 513. https://doi.org/10.1104/pp.113.230144

Choy, R. K. M., Kemner, J. M., & Thomas, J. H. (2006). Fluoxetine-resistance genes in *Caenorhabditis elegans* function in the intestine and may act in drug transport. *Genetics*, *172*(2), 885–892.

Coghlan, A., Mutowo, P., O'Boyle, N., Lomax, J., & Berriman, M. (2018). *Creating a screening set of potential anthelmintic compounds using ChEMBL*. PROTOCOL (Version 1).

Coghlan, A., Tyagi, R., Cotton, J. A., Holroyd, N., & Berriman, M. (2018). Comparative genomics of the major parasitic worms. *Nature Genetics*, *51*(1), 163–174.

Conesa, A., Götz, S., García-Gómez, J. M., Terol, J., Talón, M., & Robles, M. (2005). Blast2GO: A universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics*, *21*(18), 3674–3676. https://doi.org/10.1093/bioinformatics/bti610

Cook, A., Aptel, N., Portillo, V., Siney, E., Sihota, R., Holden-Dye, L., & Wolstenholme, A. (2006). *Caenorhabditis elegans* ivermectin receptors regulate locomotor behaviour and are functional orthologues of *Haemonchus contortus* receptors. *Molecular and Biochemical Parasitology*, *147*(1), 118–125. https://doi.org/10.1016/j.molbiopara.2006.02.003

Cutter, A. D. (2008). Divergence times in *Caenorhabditis* and *Drosophila* inferred from direct estimates of the neutral mutation rate. *Molecular Biology and Evolution*, *25*(4), 778–786. https://doi.org/10.1093/molbev/msn024

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. (2011). The variant call format and VCFtools. *Bioinformatics*, *27*(15), 2156–2158. https://doi.org/10.1093/bioinformatics/btr330

De Silva, N. R., Brooker, S., Hotez, P. J., Montresor, A., Engels, D., & Savioli, L. (2003). Soil-transmitted helminth infections: Updating the global picture. *Trends in Parasitology*, *19*(12), 547–551. https://doi.org/10.1016/j.pt.2003.10.002

DePristo, M. A., Banks, E., Poplin, R., Garimella, K. V., Maguire, J. R., Hartl, C., Philippakis, A. A., del Angel, G., Rivas, M. A., Hanna, M., McKenna, A., Fennell, T. J., Kernytsky, A. M., Sivachenko, A. Y., Cibulskis, K., Gabriel, S. B., Altshuler, D., & Daly, M. J. (2011). A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nature Genetics*, *43*(5), 491–498. https://doi.org/10.1038/ng.806

Dudchenko, O., Batra, S. S., Omer, A. D., Nyquist, S. K., Hoeger, M., Durand, N. C., Shamim, M. S., Machol, I., Lander, E. S., Aiden, A. P., & Aiden, E. L. (2017). De novo assembly of the *Aedes aegypti* genome using Hi-C yields chromosome-length scaffolds. *Science*, *356*(6333), 92–95.

Edgar, R. C. (2004). MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research*, *32*(5), 1792–1797. https://doi.org/10.1093/nar/gkh340

Fairbairn, D. (1970). Biochemical adaptation and loss of genetic capacity in helminth parasites. *Biological Reviews*, *45*(1), 29–72. https://doi.org/10.1111/j.1469-185X.1970.tb01075.x.

Fares, H., & Grant, B. (2002). Deciphering endocytosis in *Caenorhabditis elegans*. *Traffic*, *3*(1), 11–19. https://doi.org/10.1034/j.1600-0854.2002.30103.x

Feng, Z. C., Li, J. Y., Fan, Y., Li, N. W., & Xiao, F. W. (2020). CNGBdb: China National GeneBank DataBase. *Hereditas*, *42*(8), 799–809.

Fu, Y., Nie, H. M., Niu, L. L., Xie, Y., Deng, J. B., Wang, Q., Yang, G. Y., Gu, X. B., & Wang, S. X. (2011). Comparative efficacy of ivermectin and levamisole for reduction of migrating and encapsulated larvae of *Baylisascaris transfuga* in mice. *The Korean Journal of Parasitology*, *49*(2), 145.

Gaulton, A., Bellis, L. J., Bento, A. P., Chambers, J., Davies, M., Hersey, A., Light, Y., McGlinchey, S., Michalovich, D., Al-Lazikani, B., & Overington, J. P. (2012). ChEMBL: A large-scale bioactivity database for drug discovery. *Nucleic Acids Research*, *40*(Database issue), D1100–D1107.

Gerard, T., & Jose, C. (2007). Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. *Systematic Biology*, *56*(4), 564–577. https://doi.org/10.1080/10635150701472164

Ghedin, E., Wang, S., Spiro, D., Caler, E., Zhao, Q., Crabtree, J., Allen, J. E., Delcher, A. L., Guiliano, D. B., Miranda-Saavedra, D., Angiuoli, S. V., Creasy, T., Amedeo, P., Haas, B., El-Sayed, N. M., Wortman, J. R., Feldblyum, T., Tallon, L., Schatz, M., … Miranda-Saavedra, D. (2007). Draft genome of the filarial nematode parasite *Brugia malayi*. *Science*, *317*(5845), 1756–1760.

Griffitts, J. S., Huffman, D. L., Whitacre, J. L., Barrows, B. D., Marroquin, L. D., Müller, R., Brown, J. R., Hennet, T., Esko, J. D., & Aroian, R. V. (2003). Resistance to a bacterial toxin is mediated by removal of a conserved glycosylation pathway required for toxin-host interactions. *Journal of Biological Chemistry*, *278*(46), 45594–45602. https://doi.org/10.1074/jbc.M308142200

Guan, D., McCarthy, S. A., Wood, J., Howe, K., Wang, Y., & Durbin, R. (2020). Identifying and removing haplotypic duplication in primary genome assemblies. *Bioinformatics*, *36*(9), 2896–2898. https://doi.org/10.1093/bioinformatics/btaa025

Guo, X., Chen, F., Gao, F., Li, L., Liu, K. E., You, L., Hua, C., Yang, F., Liu, W., Peng, C., Wang, L., Yang, X., Zhou, F., Tong, J., Cai, J., Li, Z., Wan, B. O., Zhang, L., Yang, T., … Xu, X. (2020). CNSA: A data repository for archiving omics data. *Database*, *2020*(2020), baaa055. https://doi.org/10.1093/database/baaa055

Haas, B. J., Papanicolaou, A., Yassour, M., Grabherr, M., Blood, P. D., Bowden, J., Couger, M. B., Eccles, D., Li, B. O., Lieber, M., MacManes, M. D., Ott, M., Orvis, J., Pochet, N., Strozzi, F., Weeks, N., Westerman, R., William, T., Dewey, C. N., … Regev, A. (2013). De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nature Protocols*, *8*(8), 1494–1512. https://doi.org/10.1038/nprot.2013.084

Haas, B. J., Salzberg, S. L., Zhu, W., Pertea, M., Allen, J. E., Orvis, J., White, O., Buell, C. R., & Wortman, J. R. (2008). Automated eukaryotic gene structure annotation using EVidenceModeler and the Program to Assemble Spliced Alignments. *Genome Biology*, *9*(1), R7. https://doi.org/10.1186/gb-2008-9-1-r7

Hall, A. (1998). Rho GTPases and the actin cytoskeleton. *Science*, *279*(5350), 509–514. https://doi.org/10.1126/science.279.5350.509

Hedges, S. B. (2011). TimeTree2: Species divergence times on the iPhone. *Bioinformatics*, *27*(14), 2023–2024.

Hotez, P. P. J., Fenwick, A., Savioli, L., & Molyneux, D. H. (2009). Rescuing the bottom billion through control of neglected tropical diseases. *Lancet*, *373*(9674), 1570–1575.

Hu, H., Zhang, X., Pei, J., Su, L., Zhang, H., Liu, Y., & Wu, X. (2018). Investigation on the morphology and infection situation of intestinal parasites in the wild giant pandas. *Journal of Economic Animal*, *22*(2), 106–111+124.

Hu, J., Fan, J., Sun, Z., & Liu, S. (2019). NextPolish: A fast and efficient genome polishing tool for long-read assembly. *Bioinformatics*, *36*(7), 2253–2255. https://doi.org/10.1093/bioinformatics/btz891

Hu, Y., Yu, L., Fan, H., Huang, G., Wu, Q., Nie, Y., Liu, S., Yan, L., & Wei, F. (2020). Genomic signatures of coevolution between non-model mammals and parasitic roundworms. *Molecular Biology and Evolution*, *38*(4), 531–544.

Jaina, M., Finn, R. D., Eddy, S. R., Alex, B., & Marco, P. (2013). Challenges in homology search: HMMER3 and convergent evolution of coiled-coil regions. *Nucleic Acids Research*, *41*(12), e121.

James, C. E., Hudson, A. L., & Davey, M. W. (2009). Drug resistance mechanisms in helminths: Is it survival of the fittest? *Trends in Parasitology*, *25*(7), 328–335. https://doi.org/10.1016/j.pt.2009.04.004

Jex, A. R., Liu, S., Li, B., Young, N. D., Hall, R. S., Li, Y., Yang, L., Zeng, N., Xu, X., Xiong, Z., Chen, F., Wu, X., Zhang, G., Fang, X., Kang, Y., Anderson, G. A., Harris, T. W., Campbell, B. E., Vlaminck, J., … Gasser, R. B. (2011). Ascaris suum draft genome. *Nature*, *479*(7374), 529–533.

Jia, T.-W., Melville, S., Utzinger, J., King, C. H., & Zhou, X.-N. (2012). Soil-transmitted helminth reinfection after drug treatment: A systematic review and meta-analysis. *PLoS Neglected Tropical Diseases*, *6*(5), e1621. https://doi.org/10.1371/journal.pntd.0001621

Jinchu, H., & Wei, F. (2004). Comparative ecology of giant pandas in the five mountain ranges of their distribution in China. In L. Donald & B. Karen (eds.), *Giant Pandas: Biology and Conservation* (pp. 137–148). University of California Press.

John, B., & Mark, B. (2005). GeneMark: Web software for gene finding in prokaryotes, eukaryotes and viruses. *Nucleic Acids Research*, *33*(suppl_2), W451-W454.

Jones, A. K., Davis, P., Hodgkin, J., & Sattelle, D. B. (2007). The nicotinic acetylcholine receptor gene family of the nematode *Caenorhabditis elegans*: An update on nomenclature. *Invertebrate Neuroscience*, *7*(2), 129–131. https://doi.org/10.1007/s10158-007-0049-z

Jurka, J., Kapitonov, V. V., Pavlicek, A., Klonowski, P., Kohany, O., & Walichiewicz, J. (2005). Repbase update, a database of eukaryotic repetitive elements. *Cytogenetic & Genome Research*, *110*(1–4), 462–467. https://doi.org/10.1159/000084979

Kalvari, I., Nawrocki, E. P., Argasinska, J., Quinones-Olvera, N., Finn, R. D., Bateman, A., & Petrov, A. I. (2018). Non-coding RNA analysis using the Rfam database. *Current Protocols in Bioinformatics*, *62*(1), e51.

Kazacos, K., & Boyce, W. M. (1989). Baylisascaris larva migrans. *Journal of the American Veterinary Medical Association*, *195*(7), 894–903.

Lam-Tung, N., Schmidt, H. A., Arndt, V. H., & Quang, M. B. (2015). IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Molecular Biology and Evolution*, *32*(1), 268–274.

Li, D., He, Y., & Deng, L. (2015). Deworming experiments of ivermectin and pyrantel pamoate on *Baylisascaris schroedari* of captive giant panda. *Animal Husbandry & Veterinary Medicine*, *47*(06), 87–90. (In Chinese)

Li, D., He, Y., Wu, H., Wang, C., Li, C., Lan, J., Cheng, Z., Xie, Y., Han, H., & Yang, G. (2014). Prevalence of helminths in captive giant pandas. *Journal of Economic Animal*, *18*(4), 214–220.

Li, H., Coghlan, A., Ruan, J., Coin, L. J., Hériché, J. K., Osmotherly, L., Li, R., Liu, T., Zhang, Z., Bolund, L., Wong, G. K., Zheng, W., Dehal, P., Wang, J., & Durbin, R. (2006). TreeFam: A curated database of phylogenetic trees of animal gene families. *Nuclc Acids Research*, *34*(Database issue), D572–D580. https://doi.org/10.1093/nar/gkj118

Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*, *25*(14), 1754–1760. https://doi.org/10.1093/bioinformatics/btp324

Li, H., & Durbin, R. (2011). Inference of human population history from individual whole-genome sequences. *Nature*, *475*(7357), 493–496. https://doi.org/10.1038/nature10231

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., & Durbin, R. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics*, *25*(16), 2078–2079. https://doi.org/10.1093/bioinformatics/btp352

Lieberman-Aiden, E., van Berkum, N. L., Williams, L., Imakaev, M., Ragoczy, T., Telling, A., Amit, I., Lajoie, B. R., Sabo, P. J., Dorschner, M. O., Sandstrom, R., Bernstein, B., Bender, M. A., Groudine, M., Gnirke, A., Stamatoyannopoulos, J., Mirny, L. A., Lander, E. S., & Dekker, J. (2009). Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science*, *326*(5950), 289–293.

Lowe, T. M., & Eddy, S. R. (1997). tRNAscan-SE: A program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Research*, *25*(5), 955–964. https://doi.org/10.1093/nar/25.5.955

Maekawa, M., Ishizaki, T., Boku, S., Watanabe, N., Fujita, A., Iwamatsu, A., Obinata, T., Ohashi, K., Mizuno, K., & Narumiya, S. (1999). Signaling from rho to the actin cytoskeleton through protein kinases ROCK and LIM-kinase. *Science*, *285*(5429), 895–898.

Mario, S., Oliver, K., Irfan, G., Alec, H., Stephan, W., & Burkhard, M. (2006). AUGUSTUS: Ab initio prediction of alternative transcripts. *Nucleic Acids Research*, *34*(suppl_2), W435–W439.

Mathieu, G., & Renaud, V. (2012). Rehh: An R package to detect footprints of selection in genome-wide SNP data from haplotype structure. *Bioinformatics*, *28*(8), 1176–1177.

Mcgill, L. M., Fitzpatrick, D. A., Pisani, D., & Burnell, A. M. (2017). Estimation of phylogenetic divergence times in Panagrolaimidae and other nematodes using relaxed molecular clocks calibrated with insect and crustacean fossils. *Nematology*, *19*(8), 899–913.

Murray, W. J. (2002). Human infections caused by the raccoon roundworm, *Baylisascaris procyonis*. *Clinical Microbiology Newsletter*, *24*(1), 1–7. https://doi.org/10.1016/S0196-4399(02)80001-0

Papini, R., Renzoni, G., Malloggi, M., & Casarosa, L. (1995). Visceral larva migrans in mice experimentally infected with *Baylisascaris transfuga* (Ascarididae: Nematoda). *Parassitologia*, *36*(3), 321–329.

Park, D., O'Doherty, I., Somvanshi, R. K., Bethke, A., Schroeder, F. C., Kumar, U., & Riddle, D. L. (2012). Interaction of structure-specific and promiscuous G-protein–coupled receptors mediates small-molecule signaling in *Caenorhabditis elegans*. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(25), 9917–9922. https://doi.org/10.1073/pnas.1202216109

Patterson, N., Price, A. L., & Reich, D. (2006). Population structure and eigenanalysis. *PLoS Genetics*, *2*, e190. https://doi.org/10.1371/journal.pgen.0020190

Pitk, E. (2006). KEGG database. *Novartis Foundation Symposium*, *247*, 91–103.

Price, M. D., & Hurd, D. D. (2019). WormBase: A model organism database. *Medical Reference Services Quarterly*, *38*(1), 70–80. https://doi.org/10.1080/02763869.2019.1548896

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A. R., Bender, D., Maller, J., Sklar, P., de Bakker, P. I. W., 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*(3), 559–575. https://doi.org/10.1086/519795

Sabeti, P. C., Reich, D. E., Higgins, J. M., Levine, H. Z., Richter, D. J., Schaffner, S. F., Gabriel, S. B., Platko, J. V., Patterson, N. J., McDonald, G. J., Ackerman, H. C., Campbell, S. J., Altshuler, D., Cooper, R., Kwiatkowski, D., Ward, R., & Lander, E. S. (2002). Detecting recent positive selection in the human genome from haplotype structure. *Nature*, *419*(6909), 832–837.

Sabeti, P. C., Varilly, P., Fry, B., Lohmueller, J., Hostetter, E., Cotsapas, C., Xie, X., Byrne, E. H., McCarroll, S. A., Gaudet, R., Schaffner, S. F., Lander, E. S., International HapMap Consortium, Frazer, K. A., Ballinger, D. G., Cox, D. R., Hinds, D. A., Stuve, L. L., Gibbs, R. A., … Stewart, J. (2007). Genome-wide detection and characterization of positive selection in human populations. *Nature*, *449*(7164), 913–918.

Velankar, S., van Ginkel, G., Alhroub, Y., Battle, G. M., Berrisford, J. M., Conroy, M. J., Dana, J. M., Gore, S. P., Gutmanas, A., Haslam, P., Hendrickx, P. M., Lagerstedt, I., Mir, S., Fernandez Montecelo, M. A., Mukhopadhyay, A., Oldfield, T. J., Patwardhan, A., Sanz-García, E., Sen, S., … Kleywegt, G. J. (2016). PDBe: Improved accessibility of macromolecular structure data from PDB and EMDB. *Nucleic Acids Research*, *44*(D1), D385–D395.

Schaul, J. (2006). *Baylisascaris transfuga in captive and free-ranging populations of bears (Family: Ursidae)*. The Ohio State University.

Schiffels, S., & Durbin, R. (2014). Inferring human population size and separation history from multiple genome sequences. *Nature Genetics*, *46*(8), 919–925. https://doi.org/10.1038/ng.3015

Schiffer, P. H., Kroiher, M., Kraus, C., Koutsovoulos, G. D., Kumar, S., R Camps, J. I., Nsah, N. A., Stappert, D., Morris, K., Heger, P., Altmüller, J., Frommolt, P., Nürnberg, P., Thomas, W., Blaxter, M. L., & Schierenberg, E. (2013). The genome of *Romanomermis culicivorax*: Revealing fundamental changes in the core developmental genetic toolkit in Nematoda. *Bmc Genomics*, *14*(1), 923. https://doi.org/10.1186/1471-2164-14-923

Schumacher, T., & Benndorf, R. A. (2017). ABC transport proteins in cardiovascular disease—A brief summary. *Molecules*, *22*(4), 589. https://doi.org/10.3390/molecules22040589

Simão, F. A., Waterhouse, R. M., Panagiotis, I., Kriventseva, E. V., & Zdobnov, E. M. (2015). BUSCO: Assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics*, *31*(19), 3210–3212. https://doi.org/10.1093/bioinformatics/btv351

Smit, A., Hubley, R., & Green, P. (2015). *RepeatModeler Open-1.0. 2008–2015*. Institute for Systems Biology. www.repeatmasker.org

Soblik, H., Younis, A. E., Mitreva, M., Renard, B. Y., Kirchner, M., Geisinger, F., Steen, H., & Brattig, N. W. (2011). Life cycle stage-resolved proteomic analysis of the excretome/secretome from *Strongyloides ratti*—Identification of stage-specific proteases. *Molecular & Cellular Proteomics*, *10*(12), M111.010157. https://doi.org/10.1074/mcp.M111.010157

Sterling, T., & Irwin, J. J. (2015). ZINC 15 – Ligand discovery for everyone. *Journal of Chemical Information & Modeling*, *55*, 2324–2337. https://doi.org/10.1021/acs.jcim.5b00559

Suhong, X., & Chisholm, A. D. (2012). A Gαq-Ca signaling pathway promotes actin-mediated epidermal wound closure in *C. elegans*. *Current Biology*, *21*(2), 1960–1967.

Tarailo-Graovac, M., & Chen, N. (2009). Using RepeatMasker to identify repetitive elements in genomic sequences. *Current Protocols in Bioinformatics*, *25*(1), 4.10.1–4.10.14.

Tatusov, R. L., Fedorova, N. D., Jackson, J. D., Jacobs, A. R., Kiryutin, B., Koonin, E. V., Krylov, D. M., Mazumder, R., Mekhedov, S. L., Nikolskaya, A. N., Rao, B. S., Smirnov, S., Sverdlov, A. V., Vasudevan, S., Wolf, Y. I., Yin, J. J., & Natale, D. A. (2003). The COG database: An updated version includes eukaryotes. *BMC Bioinformatics*, *4*(1), 41.

Tempel, S. (2012). Using and understanding RepeatMasker. *Mobile genetic elements*, *859*, 29–51.

Trapnell, C., Pachter, L., & Salzberg, S. L. (2009). TopHat: Discovering splice junctions with RNA-Seq. *Bioinformatics*, *25*(9), 1105–1111. https://doi.org/10.1093/bioinformatics/btp120

Tyagi, R., Tyagi, R., Seshadri, S., Parkinson, J., & Mitreva, M. (2018). Comparative analysis of metabolism in parasitic worms. *Protocol Exchange*, 10. https://doi.org/10.1038/protex.2018.048

Wang, C., Lan, J., Shen, F., Li, L., Huang, W., Zhi, Y., Li, L., Huang, X., Wu, K., & Li, M. (2015). Toxic shock of giant pandas caused by ivermectin. *Chinese Journal of Wildlife*.

Wang, T., Xie, Y., Zheng, Y., Wang, C., Li, D., Koehler, A. V., & Gasser, R. B. (2018). Parasites of the giant panda: A risk factor in the conservation of a species. *Advances in Parasitology*, *99*, 1–33.

Wang, Y., Tang, H., DeBarry, J. D., Tan, X., Li, J., Wang, X., Lee, T.-H., Jin, H., Marler, B., Guo, H., Kissinger, J. C., & Paterson, A. H. (2012). MCScanX: A toolkit for detection and evolutionary analysis of gene synteny and collinearity. *Nucleic Acids Research*, *40*(7), e49. https://doi.org/10.1093/nar/gkr1293

Wildt, D. E., Zhang, A., Zhang, H., Janssen, D. L., & Ellis, S. (2006). *Giant pandas: Biology, veterinary medicine and management*. Cambridge University Press.

Williamson, A. L., Lustigman, S., Oksov, Y., Deumic, V., Plieskatt, J., Mendez, S., Zhan, B., Bottazzi, M. E., Hotez, P. J., & Loukas, A. (2006). Ancylostoma caninum MTP-1, an astacin-like metalloprotease secreted by infective hookworm larvae, is involved in tissue migration. *Infection and Immunity*, *74*(2), 961–967.

Wise, M. E., Sorvillo, F. J., Shafir, S. C., Ash, L. R., & Berlin, O. G. (2005). Severe and fatal central nervous system disease in humans caused by Baylisascaris procyonis, the common roundworm of raccoons: A review of current literature. *Microbes and Infection*, *7*(2), 317–323. https://doi.org/10.1016/j.micinf.2004.12.005

Wishart, D. S., Feunang, Y. D., An, C. G., Lo, E. J., & Wilson, M. (2017). DrugBank 5.0: A major update to the DrugBank database for 2018. *Nucleic Acids Research*, *46*(Database issue), D1074–D1082.

Wolf, K. N., Lock, B., Carpenter, J. W., & Garner, M. M. (2007). Baylisascaris procyonis Infection in a Moluccan Cockatoo (*Cacatua moluccensis*). *Journal of Avian Medicine & Surgery*, *21*(3), 220–225. https://doi.org/10.1647/1082-6742(2007)21[220:BPIIAM]2.0.CO;2

Wu, Y. C., & Horvitz, H. R. (1998). The *C. elegans* cell corpse engulfment gene ced-7 encodes a protein similar to ABC transporters. *Cell*, *93*(6), 951–960.

Xie, Y., Zhang, Z., Wang, C., Lan, J., Li, Y., Chen, Z., Fu, Y., Nie, H., Yan, N., Gu, X., Wang, S., Peng, X., & Yang, G. (2011). Complete mitochondrial genomes *of Baylisascaris schroederi*, *Baylisascaris ailuri* and *Baylisascaris transfuga* from giant panda, red panda and polar bear. *Gene*, *482*(1-2), 59–67. https://doi.org/10.1016/j.gene.2011.05.004

Xu, M., Molento, M., Blackhall, W., Ribeiro, P., Beech, R., & Prichard, R. (1998). Ivermectin resistance in nematodes may be caused by alteration of P-glycoprotein homolog. *Molecular and Biochemical Parasitology*, *91*(2), 327–335.

Yang, G. (1998). Advances on parasites and parasitology of *Ailuropoda melanoleuca*. *Chinese Journal of Veterinary Science*, *18*, 206–208.

Zhang, J.-S., Daszak, P., Huang, H.-L., Yang, G.-Y., Kilpatrick, A. M., & Zhang, S. (2008). Parasite threat to panda conservation. *EcoHealth*, *5*(1), 6–9. https://doi.org/10.1007/s10393-007-0139-8

Zhang, L., Yang, X., Wu, H., Gu, X., Hu, Y., & Wei, F. (2011). The parasites of giant pandas: Individual-based measurement in wild animals. *Journal of Wildlife Diseases*, *47*(1), 164–171. https://doi.org/10.7589/0090-3558-47.1.164

Zhao, S., Zheng, P., Dong, S., Zhan, X., Wu, Q. I., Guo, X., Hu, Y., He, W., Zhang, S., Fan, W., Zhu, L., Li, D., Zhang, X., Chen, Q., Zhang, H., Zhang, Z., Jin, X., Zhang, J., Yang, H., … Wei, F. (2013). Whole-genome sequencing of giant pandas provides insights into demographic history and local adaptation. *Nature Genetics*, *45*(1), 67–71. https://doi.org/10.1038/ng.2494

Zhao, X., & Hao, W. (2007). LTR_FINDER: An efficient tool for the prediction of full-length LTR retrotransposons. *Nucleic Acids Research*, *35*(suppl_2), W265–W268.

Zhou, C., Hu, J., Yuan, C., & Wei, F. (1997). Giant pandas food habits and feeding behaviour in Mabian Dafengding Natural Reserve. *Journal of Sichuan Teachers College*, *18*(4), 273–277.

Zhu, X.-Q., Korhonen, P. K., Cai, H., Young, N. D., Nejsum, P., von Samson-Himmelstjerna, G., Boag, P. R., Tan, P., Li, Q., Min, J., Yang, Y., Wang, X., Fang, X., Hall, R. S., Hofmann, A., Sternberg, P. W., Jex, A. R., & Gasser, R. B. (2015). Genetic blueprint of the zoonotic pathogen *Toxocara canis*. *Nature Communications*, *6*(2), 6145. https://doi.org/10.1038/ncomms7145

Zou, X., Wang, A., Zeng, L., He, G., Wu, K., Chen, Y., & Weng, N. (1998). Lethal factors of diseases and protective countermeasures of wild and penned Giant pandas. *Journal of Forestry Research*, *2*(1998), 77–80.

Zumla, A., & Hui, D. S. (2019). Emerging and reemerging infectious diseases: Global overview. *Infectious Disease Clinics*, *33*(4), xiii–xix. https://doi.org/10.1016/j.idc.2019.09.001

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: Han, L., Lan, T., Li, D., Li, H., Deng, L., Peng, Z., He, S., Zhou, Y., Han, R., Li, L., Lu, Y., Lu, H., Wang, Q., Yang, S., Zhu, Y., Huang, Y., Cheng, X., Yu, J., Wang, Y., … Hou, Z. (2022). Chromosome-scale assembly and whole-genome sequencing of 266 giant panda roundworms provide insights into their evolution, adaptation and potential drug targets. *Molecular Ecology Resources*, 22, 768–785. https://doi.org/10.1111/1755-0998.13504