Check for updates

Received: 31 March 2021 | Revised: 11 September 2021 | Accepted: 22 September 2021

DOI: 10.1111/ddi.13423

RESEARCH ARTICLE

Diversity and Distributions WILEY

Phylogenomics of Northeast Asian *Pungitius* sticklebacks

Yingnan Wang[1,2] | Yu Wang[1,2] | Yahui Zhao[1] | Alexandra Yu Kravchenko[3] | Juha Merilä[4,5] | Baocheng Guo[1,2,6]

[1]Key Laboratory of Zoological Systematics and Evolution, Institute of Zoology, Chinese Academy of Sciences, Beijing, China

[2]University of Chinese Academy of Sciences, Beijing, China

[3]A. V. Zhirmunsky National Scientific Center of Marine Biology, Far Eastern Branch, Russian Academy of Sciences, Vladivostok, Russia

[4]Ecological Genetics Research Unit, Organismal and Evolutionary Biology Research Programme, Faculty of Biological and Environmental Sciences, University of Helsinki, Helsinki, Finland

[5]Research Division for Ecology and Biodiversity, School of Biological Sciences, Faculty of Science, The University of Hong Kong, Hong Kong SAR, China

[6]Center for Excellence in Animal Evolution and Genetics, Chinese Academy of Sciences, Kunming, China

Correspondence
Baocheng Guo, Key Laboratory of Zoological Systematics and Evolution, Institute of Zoology, Chinese Academy of Sciences, Beijing, China.
Email: guobaocheng@ioz.ac.cn

Funding information
National Natural Science Foundation of China, CAS Pioneer Hundred Talents Program and Academy of Finland

Editor: Sophie von der Heyden

Abstract

Aim: Understanding the phylogeography of a species complex can provide important insights into its evolutionary history. However, phylogeographic inference often faces the dilemma of regionally inadequate sampling. *Pungitius* sticklebacks are a case in point: although the highest species diversity is found in Northeast Asia, their phylogeography in this region is still poorly understood.

Location: Northeast Asia.

Methods: With the aid of whole-genome resequencing data, we investigated the phylogeography of Northeast Asian *Pungitius* sticklebacks, with newly sampled 83 worldwide *Pungitius* individuals from 11 locations including eight Chinese locations reported to host only *P. sinensis*.

Results: We discovered that three of the eight Chinese locations hosted populations of *P. kaibarae* and *P. bussei*, species new to the fauna of China. Phylogeographic analyses further clarified the sequence and timing of colonization of Northeast Asia by different *Pungitius* species, shedding new light on their origins and current distribution ranges. Colonization of inland Northeast Asia by *Pungitius* sticklebacks occurred in multiple waves, and the widespread *P. sinensis* expanded its range relatively late in the Pleistocene.

Main conclusions: This study complements our understanding of the phylogeography of *Pungitius* sticklebacks by extending sampling to cover an area that comprises nearly half of the known distribution area of this genus in Northeast Asia. The discovery of three *Pungitus* species from China is of particular interest, as translocations to support locally declining populations have occurred under the assumption that all sticklebacks in China—except the endangered *P. stenurus*—are *P. sinensis*, raising conservation concerns associated with unintentional translocations and possible admixture.

KEYWORDS

mtDNA, phylogenomics, phylogeography, *Pungitius*, SNP

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.
© 2021 The Authors. *Diversity and Distributions* published by John Wiley & Sons Ltd.

1 | INTRODUCTION

Many fundamental questions in evolutionary biology require knowledge about the phylogenetic relationships and geographical distribution patterns in a given species complex (Avise, 2000). Such phylogeographic information provides the key to understand the evolutionary history of organisms, as well as responses of populations and species to geological events (Guisan & Thuiller, 2005; Schluter, 2000). As such, phylogeography has provided valuable contributions to several areas of biology and earth sciences, including the discovery of cryptic biodiversity in many organisms (Beheregaray, 2008; Bickford et al., 2007). However, phylogeographic investigations can be misled by the presence of cryptic diversity if distinct evolutionary lineages are not identified as such, either because of incomplete geographical coverage of sampling, or because taxa within a species complex are phenotypically indistinguishable and therefore left unsampled (Beheregaray, 2008; Bickford et al., 2007).

Pungitius is the most species-rich genus in the stickleback family Gasterosteidae, a group of popular model organisms in evolutionary biology (Bell & Foster, 1995; Wootton, 1976, 1984). The genus is comprised of 11 taxonomically valid species widely distributed across the Northern Hemisphere (Bogutskaya et al., 2008; Denys et al., 2018; Eschmeyer et al., 2016; Guo et al., 2019; Keivany & Nelson, 2000, 2004; Shedko et al., 2005). The phylogeography of *Pungitius* sticklebacks has been extensively investigated in Northeast Asia (Takahashi, & Goto, 2001; Takahashi et al., 2016), Europe (Denys et al., 2018; Shikano et al., 2010; Wang et al., 2015, 2017), North America (Aldenhoven et al., 2010) and globally (Guo et al., 2019). These studies have discovered the existence of two major phylogenetic clades: the *pungitius* clade, which includes the circumpolarly distributed *P. pungitius* and three species (*P. platygaster*, *P. hellenicus* and *P. laevis*) endemic to Europe, and the *sinensis* clade, which includes *P. sinensis*, *P. polyakovi*, *P. tymensis*, *P. kaibarae* and *P. bussei* endemic to Northeast Asia (Guo et al., 2019; Takahashi et al., 2016). In spite of these and related studies (Fang et al., 2021; Feng et al., 2021; Wang et al., 2020; Yamasaki et al., 2020), the regional phylogeography of this genus in Northeast Asia—the diversity cradle of the genus (Guo et al., 2019)—is still incomplete. The same applies to existence different *Pungitius* taxa in local faunas. It is well known that diagnostic characters (e.g., pelvis structure, dorsal spine numbers, lateral plate numbers and presence or absence of keel) used in *Pungitius* taxonomy (Keivany & Nelson, 2000, 2004; Wang et al., 2015) are evolutionarily highly labile and subject to extensive homoplasy, which has frequently led to incorrect species assignments in this genus (Guo et al., 2019).

Earlier phylogeographic studies of *Pungitius* sticklebacks in Northeast Asia have focused on the coastal areas and the Japanese Archipelago (Guo et al., 2019; Takahashi, & Goto, 2001; Takahashi et al., 2016). However, *Pungitius* sticklebacks are known to be widely distributed also in inland Northeast Asia. For example, *Pungitius* sticklebacks are widely distributed in Northern China from the Heilongjiang [Amur] river in the northeast (Berg, 1949; Cheng & Zheng, 1987; Xie, 2007; Zhang, 1995; Zhao, 2018; Zhu, 1995) to the Yangtze river in Central China (Guichenot, 1868), and from the Tumen river in Northeast China (Xie, 2007; Zheng et al., 1980) to the Yellow river in the Hohhot region, Western China (Bou, 2011). Two taxonomically valid species, *P. stenurus* (Bogutskaya et al., 2008) and *P. sinensis* (Zhang & Zhao, 2016), have been recorded from China. *P. stenurus* is known only from a single locality, Hu-lun [Dalai-Nor] Lake in Nei Mongol in Northeast China (Bogutskaya et al., 2008), and all other records of *Pungitius* sticklebacks in China are of *P. sinensis* according to taxonomy (Zhang & Zhao, 2016). However, earlier phylogeographic studies of *Pungitius* sticklebacks have included only one *P. sinensis* population from China (Guo et al., 2019; Takahashi et al., 2016), leaving clear knowledge gaps in (1) our understanding of *Pungitius* species diversity and distribution, and (2) their evolutionary history over an area that comprises nearly half of the known distribution area of this genus in Northeast Asia.

Considering the wide geographic distribution and sparse geographic sampling of *P. sinensis* in earlier phylogeographic studies, there is a high probability for discovering new diversity, as the areas adjacent to Northeast China are known to host several narrowly distributed endemic *Pungitius* species (*viz. P. bussei*, *P. kaibarea*, *P. polyakovi* and *P. stenurus*) in addition to the more widely distributed *P. sinensis* and *P. pungitius* (Guo et al., 2019; Takahashi et al., 2016; see also Liu & Wang, 1974). Discovery of additional diversity and/or species of *Pungitius* sticklebacks from this area could have important implications for their conservation and management, as *Pungitius* sticklebacks from Northeast China have been used to stock water courses in the Beijing area for supplementary releases. Importantly, this has been done under the assumption that they are *P. sinensis*, although it is possible that they are in fact other species. As for the lack of resolution over the evolutionary history of many *Pungitius* sticklebacks in their Asian distribution range, this area can hold the key to understanding their diversification. For example, earlier studies have uncovered frequent introgression among *Pungitius* species in Asia, and all *P. sinensis* lineages studied thus far carry the mitochondrial genome of *P. pungitius* (Guo et al., 2019; Natri et al., 2019; Takahashi et al., 2016; Wang et al., 2015). Whether this is the case for *P. sinensis* from the vast inland area in Northern China is unknown.

The aims of this study were twofold: first, to conduct a species diversity survey of *Pungitius* sticklebacks in China using genomic tools, and second, to investigate the phylogeography of *Pungitius* species in Northeast Asia using the most comprehensive sampling of their distribution area to date. To this end, we generated new genomic data on 83 worldwide *Pungitius* individuals, mostly from northeast Asia, and combined this with publicly available data. As the results show, the findings will also be highly relevant for biodiversity conservation and management of *Pungitius* sticklebacks.

2 | MATERIALS AND METHODS

2.1 | Sample and data collection

We collected 83 worldwide *Pungitius* individuals (two individuals of *P. pungitius* from North America, two individuals of *P. kaibarae*

from South Korea, two individuals of *P. polyakovi* from Russia and 77 *Pungitius* individuals from China). Chinese samples were collected from eight locations covering most of occurrence records of *P. sinensis* in China (Figure 1; Table S1; Zhang & Zhao, 2016). Specifically, five sampling sites were from Northeast China, three of which (YC, YH and MDJ) were from Heilongjiang Province, one (HC) from Jilin Province and one (FS) from Liaoning Province. Two other sites were from Northern China, one of which (CF) was from Nei Mongol Province and the other (XL) from Hebei Province. An additional site (LZ) was from Henan Province in Central China. The XL samples were obtained from the China National Animal Collection Resource Center (http://museum.ioz.ac.cn/), collected in May 2005, whereas all other samples were collected with hand seines and/or minnow traps (mesh size 6 mm) in Spring 2019. All samples from China were initially considered as *P. sinensis*, as only two species of *Pungitius* are known from China (Zhang & Zhao, 2016), and the endangered *P. stenurus* was not included in this study—we failed to capture any individuals from only known occurrence locality of the species in our fieldwork. In addition to newly sampled *Pungitius* individuals, *Pungitius* data available from earlier studies (Figure 1; Table S1; Bae & Suk, 2015; Guo et al., 2019; Hwang et al., 2012; Kawahara et al., 2009; Miya et al., 2001; Nelson & Cresko, 2018) were included for global phylogeographic analysis.

2.2 | DNA extraction and sequencing

Genomic DNA was extracted from ethanol-preserved dorsal muscle tissue from nine to 10 individuals (the same individuals used in morphological analyses) per sampling location, using a QIAGEN DNeasy Kit following the manufacturer instructions. DNA was visualized on 1% agarose gels to assess the quality, and thereafter quantified with a NanoDrop® ND-1000 spectrophotometer and Qubit® fluorometer to ensure success of downstream sequencing steps. DNA library construction and sequencing were done by Novogene CO., LIMITED. The libraries with insertion size of 300–500 bp were sequenced on an Illumina NovaSeq platform with a 150 bp paired-end strategy. Two individuals from each of the eight populations in China and two individuals of *P. pungitius* from North America were whole-genome re-sequenced for ~10× coverage of the *P. pungitius* genome (Varadharajan et al., 2019) to obtain single nucleotide polymorphisms (SNPs) for the nuclear phylogenetic interference. The remaining 61 individuals from the eight populations in China were whole-genome re-sequenced for ~2× of the *P. pungitius* genome to obtain full mitochondrial genomes for phylogenetic interference. Two individuals of *P. polyakovi* from Russia and two individuals of *P. kaibarae* from South Korea included in the nuclear phylogenetic interference were sequenced with the restriction site-associated DNA sequencing (RAD-seq) strategy by following the same method as in

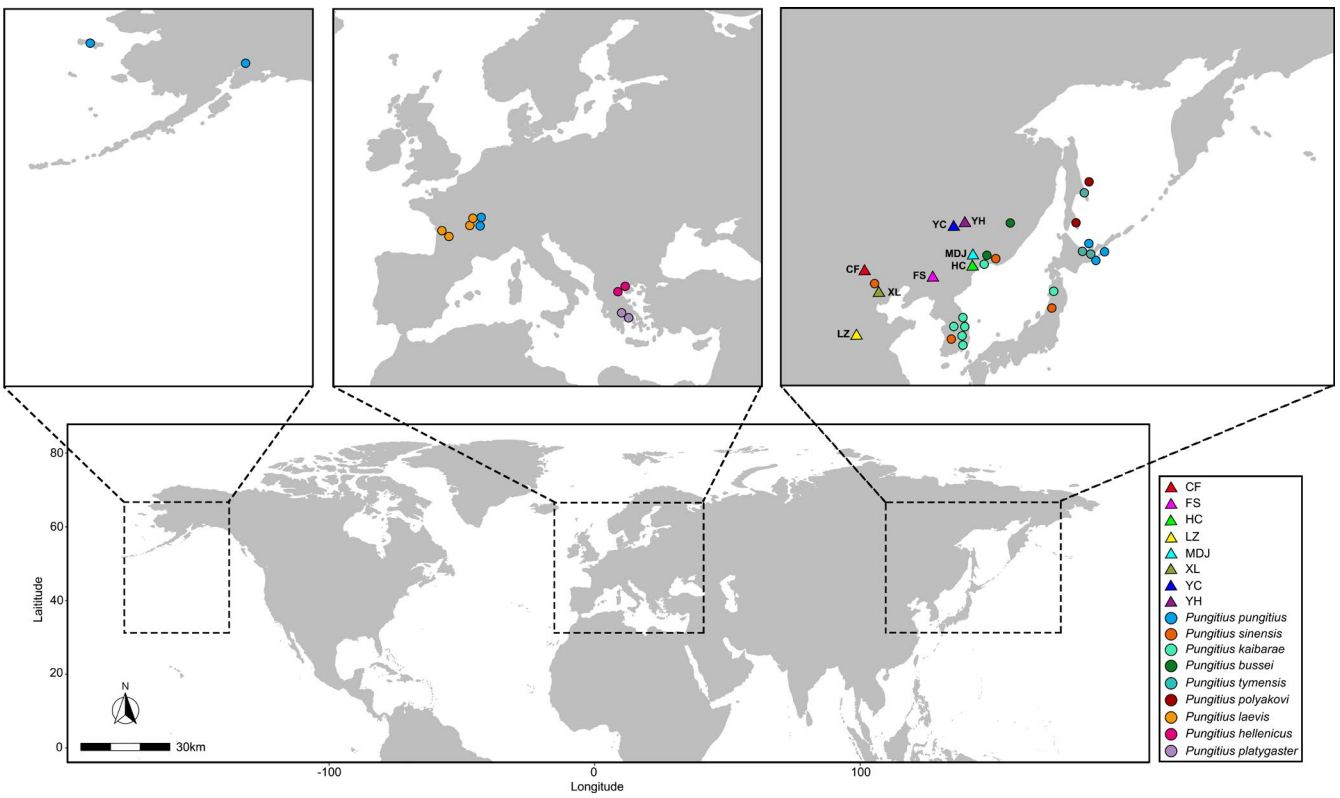

FIGURE 1 Map showing the locations of the *Pungitius* populations used in this study. Triangles represent the eight focal *Pungitius* populations from China collected in this study, and circles represent those for which data were obtained from earlier studies

Guo et al. (2019). In brief, DNA was digested with restriction enzyme *Pst*I, and fragments of 300–500 bp in length after primer ligation were sequenced using the Illumina HiSeq2000 platform with 100-bp paired-end strategy. RAD library preparation and sequencing were done by BGI Hongkong Co., Ltd.

2.3 | Read mapping and SNPs calling

The *P. pungitius* genome (Varadharajan et al., 2019) was used as a reference. Quality filtered reads from each individual were aligned to the reference genome using BWA version 0.7.17 (Li et al., 2009) with default settings. The mapping results in SAM format were converted into BAM format using SAMtools version 1.4 (Li et al., 2009). Sorted and duplicate-removed BAM format mapping results were used for variant calling using BCFtools 1.8 (Li et al., 2009) and SAMtools version 1.4 (Li et al., 2009). Highly reliable SNPs were selected using the VCFtools version 0.1.15 (Danecek et al., 2011) with the following criteria: only biallelic SNPs were kept with genotype quality (GQ) or mapping quality (MQ) no less than 20; biallelic SNPs subjected to phylogenetic inference had a minor allele frequency >0.05 and genotype calls in no less than 90% of the 52 individuals included.

2.4 | Mitochondrial genome assembly and annotation

Reads mapping to the mitochondrial genome were extracted using SAMtools version 1.4. The assembly and annotation of mitochondrial genomes were conducted using MitoZ version 1.04 (Meng et al., 2019) and MITOS (Bernt et al., 2013), respectively. The newly assembled and annotated mitochondrial genomes were further validated with BLAST searches in GenBank and found to be similar to the available *Pungitius* mitochondrial genomes.

2.5 | Phylogenetic analysis

Nuclear phylogeny was inferred with 1,079,030 high-quality SNPs from two individuals from each of the eight populations sampled this study and 38 individuals of eight *Pungitius* species worldwide (Table S1). *Gasterosteus wheatlandi* and *Culaea inconstans* were used as outgroups in the nuclear phylogenetic interference. A maximum likelihood tree was constructed using IQtree version 2.0 (Nguyen et al., 2015) with a TVM + ASC + R7 model to account for the ascertainment bias of SNPs following the suggestion of ModelFinder (Kalyaanamoorthy et al., 2017). Node support was assessed using 100 rapid bootstrap resampling replicates. For the mitochondrial phylogenetic interference, we used 13 mitochondrial coding gene sequences of the 77 individuals from the eight populations sampled in this study, along with 14 individuals from seven *Pungitius* species with publicly available mitochondrial genomes (Table S1). The other three non-*Pungitius* stickleback species—*G. wheatlandi*, *G. aculeatus* and *C. inconstans*—were used as outgroups in the mitochondrial phylogenetic interference. Sequences were aligned for each of the 13 mitochondrial coding genes with MAFFT version 7.450 (Katoh & Standley, 2013). ModelFinder was used to infer the best partitioning scheme and nucleotide substitution models (Kalyaanamoorthy et al., 2017). Maximum likelihood (ML) analyses were conducted using IQtree version 2.0 (Nguyen et al., 2015). Branch support was estimated using the Ultrafast option for bootstrap analysis with 5,000 replicates. As only the *Cytb* gene sequence was available for *P. bussei* and *P. polyakovi* (Takahashi et al., 2016), a phylogeny based on *Cytb* sequences was additionally constructed by including *P. bussei*, *P. polyakovi* and all samples included in the phylogenetic analysis with 13 mitochondrial coding genes, using the same methods as above. The resulting phylogeny was visualized in FigTree version 1.7.1 (http://tree.bio.ed.ac.uk/software/figtree).

2.6 | Divergence time estimation

Divergence time estimation was performed using the 13 mitochondrial coding gene sequences with BEAST2 version 2.6.2 (Bouckaert et al., 2014) using a relaxed clock model. Following earlier studies (Guo et al., 2019; Varadharajan et al., 2019), three calibration nodes were used: a root node of 26.6 Mya (*SD* of 3) with a normal distribution calibration density, the divergence time of 10.97 Mya (*SD* of 1) between *G. aculeatus* and *G. wheatlandi* with a normal distribution calibration density, and the minimum time of the most recent common ancestry (TMRCA) of 7.0 Mya for the genus *Pungitius*, with a uniform distribution from 7 to 10 Mya. Two independent runs of 10,000,000 generations were conducted with a Markov Chain Monte Carlo (MCMC) approach, sampling every 1,000 iterations and discarding the first 50% of sampled iterations as burn-in. Effective sample size stability of the posterior distribution was evaluated by Tracer version 1.7.1 (Rambaut et al., 2018). The maximum clade-credibility tree containing 95% highest posterior density (95% HPD) was obtained using TreeAnnotator 2.5.1 (Bouckaert et al., 2014) and visualized with FigTree version 1.7.1.

2.7 | Morphological analyses

To quantitatively test whether sticklebacks from the eight localities in China are morphologically distinguishable, morphological variation was quantified from 15 individuals from each of the eight populations sampled in this study. A digital photograph was taken from the left side of each fish under standardized lighting conditions with a ruler placed next to the fish to provide a scale. Two series of morphometric analyses were conducted. The first was solely based on landmarks, and the second was based on linear morphological measures obtained from different combinations of the landmarks.

Twenty-nine landmarks (Figure S1) were recorded for each individual using tpsDig version 2.17 (Rohlf, 2013) following previous studies of three-spined (*Gasterosteus aculeatus*; Rogers et al., 2012;

Schluter et al., 2004) and nine-spined sticklebacks (Shimada et al., 2011; Yang et al., 2016). Geometric morphometric analyses of body shape variation were conducted using MorphoJ version 1.06d (Klingenberg, 2011). All landmarks were rescaled and aligned through Generalized Procrustes Analysis, which aims to retain only shape-related information in the dataset by removing variation in position, orientation and size (Rohlf & Slice, 1990; Zelditch et al., 2004). As such, a covariance matrix of the body shape data was extracted for each individual and subjected to a Principal Component Analysis (PCA) to quantify shape variation among the eight populations.

In addition to landmark-based variation in body shape, variation in 14 linear measures of morphology was quantified based on the 29 landmarks as defined in Figure S1. The linear measures included head length, upper jaw length, lower jaw length, orbit diameter, first dorsal fin base length, second dorsal fin base length, anal fin base length, caudal peduncle length, caudal peduncle depth, body depth, snout length, head depth, pectoral fin base length and standard body length. Each trait was measured twice by the same person, and mean values of these repeated measures were used to minimize the effect of measurement error. The linear measures were adjusted to variation in body size by regressing the trait values against standard body length and calculating the residual variation (Yang et al., 2016). PCAs were performed on residuals using the built-in R functions prcomp() and princomp() in the R package FactoMineR (Le et al., 2008) to quantify variation in the 13 linear traits. The results were visualized using the R package factoextra (http://www.sthda.com/english/rpkgs/factoextra). A hierarchical clustering analysis was performed based on squared Euclidean distances using the between-groups linkage method. One-way ANOVAs were used to test whether linear measures differed significantly among samples. In *post hoc* multiple comparisons, the variables that possessed equal variances were analysed with the Least Significant Difference (LSD) method. Otherwise, T2 test of Tamhane was adopted. Linear discriminant function analysis (DFA) of the 13 linear traits was performed to test whether the populations can be distinguished based on variation in these traits. One-way ANOVAs, DFA and clustering analyses were conducted using SPSS version 24.0.

3 | RESULTS

3.1 | Phylogenetic relationships

The well-supported phylogeny based on 1,079,030 nuclear genomic SNPs (left panel, Figure 2) shows that the genus *Pungitius* consists of two major clades, the *pungitius* clade and the *sinensis* clade, consistent with earlier findings (Guo et al., 2019; Takahashi et al., 2016). The *pungitius* clade includes the circumpolarly distributed *P. pungitius* and the species occurring in Eurasia (*P. laevis*, *P. hellenicus* and *P. platygaster*; Figure 2). The *sinensis* clade includes the species endemic to northeast Asia—*P. sinensis*, *P. kaibarae*, *P. tymensis* and *P. polyakovi* (left panel, Figure 2). All eight populations from China sampled in this study were placed in the *sinensis* clade (Figure 2). The Chinese populations CF, FS, HC, LZ and XL were placed in a subclade with *P. sinensis* and *P. polyakovi* (left panel, Figure 2). Notably, the populations CF, FS, XL and LZ were grouped with a previously studied *P. sinensis* population in the Hebei Province, whereas the HC population was grouped with a previously studied *P. sinensis* population from Primorsky Krai in the Russian Far East (left panel, Figure 2). Populations MDJ, YC and YH formed a monophyletic group with *P. kaibarae*, whereas populations YC and YH formed a monophyletic clade in between *P. kaibarea* and *P. tymensis* (left panel, Figure 2).

The mitochondrial phylogeny (right panel, Figure 2) based on 13 mitochondrial coding gene sequences was incongruent with the nuclear phylogeny (left panel, Figure 2)—an observation also made in earlier studies (Guo et al., 2019; Takahashi et al., 2016; Wang et al., 2017). The populations CF, FS, HC, LZ and XL, as well as a *P. sinensis* population from South Korea, were placed in the *pungitius* clade; the HC population and one *P. sinensis* individual from South Korea form a group within this clade (right panel, Figure 2). The populations MDJ, YC and YH form a clade within the *sinensis* lineage and group with *P. kaibarae* from South Korea (right panel, Figure 2). The mitochondrial phylogeny based on the *Cytb* gene is consistent with that based on the 13 mitochondrial coding gene sequences and shows that the populations YH and YC form a monophyletic clade with *P. bussei* from Russia (Figure S2).

3.2 | Divergence time

The divergence time based on 13 mitochondrial coding gene sequences among *Pungitius* species/populations is shown in Figure 3. The TMRCA for the genus *Pungitius* is inferred to be ~7.65 Mya with 95% highest probability density (HPD) intervals of 6.79–8.56 Mya. The TMRCA of populations CF, FS, HC, LZ and XL is inferred to be ~1.32 Mya (95% HPD intervals: 0.89–1.79 Mya) and that of populations CF, FS, HC, LZ and XL ~0.59 Mya (95% HPD intervals: 0.37–0.85 Mya). The populations CF and FS were inferred to have diverged ~0.51 Mya (95% HPD intervals: 0.29–0.75 Mya), whereas the populations XL and LZ were inferred to have diverged ~0.15 Mya (95% HPD intervals: 0.06–0.27 Mya). The TMRCA of the populations MDJ, YC, YH and *P. kaibarae* from South Korea was ~3.05 Mya (95% HPD intervals: 2.08–4 Mya), and the divergence between MDJ (*i.e.* likely *P. kaibarea*) and the two likely *P. bussei* populations (YC and YH) is ~0.53 Mya (95% HPD intervals: 0.26–0.83 Mya).

3.3 | Morphological variation

PCAs of body shape variation based on 29 landmarks showed the first three principal components (PCs), each accounting for >10% of the variation, cumulatively captured 67.7% of the total body shape variation in the eight populations (Figure S3). PC1 captured 35.2% of the total body shape variation, loading mainly on body depth, second dorsal fin base length and caudal peduncle length. PC2 explained 18.8% of the total variation, loading mainly on second dorsal fin

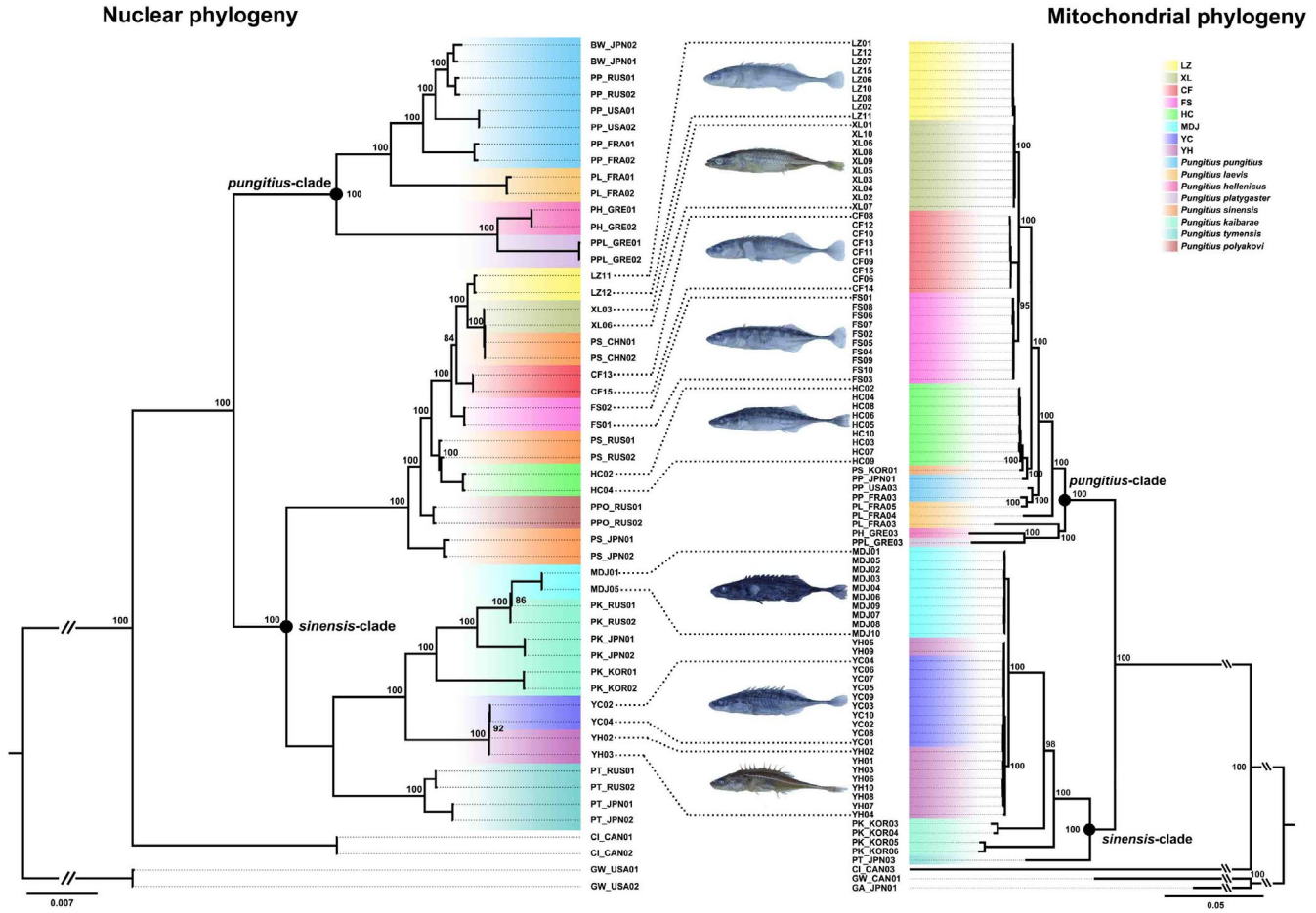

FIGURE 2 The nuclear (left panel) and mitochondrial (right panel) phylogenies of *Pungitius* species. The nuclear phylogeny is based on genome-wide SNP markers, and the mitochondrial phylogeny is based on 13 mitochondrial protein coding genes. GW: *G. aculeatus*; GW: *G. wheatlandi*; CI: *C. inconstans*; PH: *P. hellenicus*; PK: *P. kaibarae*; PL: *P. laevis*; PP: *P. pungitius*; PPL: *P. platygaster*; PPO: *P. polyakovi*; PS: *P. sinensis*; PT: *P. tymensis*; FW: freshwater type of *P. pungitius* now recognized as *P. sinensis* (Guo et al. (2019); OM: morphologically Omono type of *P. pungitius* that has been recognized as *P. kaibarae* in Guo et al. (2019); BW: brackish water type of *P. pungitius*. YC, YH, MDJ, HC, FS, CF, XL and LZ are stickleback populations from China. JPN: Japan; RUS: Russia; USA: United States of America; FRA: France; GRE: Greece; CHN: China; KOR: South Korea; CAN: Canada. Numbers at nodes are bootstrap values

base length and caudal peduncle length. PC3 accounted for 13.7% of the total body shape variation, loading on second dorsal fin base length, body depth and head depth. Considering PC1 and PC2 together, the populations group into three largely overlapping clusters (Figure S3a). The first cluster is comprised of populations YC, YH and MDJ; the second includes populations CF, FS, LZ and XL; the third includes population HC.

The discriminant analysis of the 13 linear traits revealed that individuals can be assigned to the correct sampling locality with 93.3% to 100% accuracy (Table S2). PCAs based on the 13 linear measures identified two PCs that together accounted for 56.1% of the total variation (PC1 = 34.6%, PC2 = 21.5%; Figure S3b). These two PCs primarily captured the variation in head length, upper jaw length, lower jaw length, second dorsal fin base length, caudal peduncle length, caudal peduncle depth, head depth and pectoral fin base length. These two PCs clustered populations to three groups: one consisted of YC and YH, another consisted of CF, FS, HC, LZ, MDJ and XL, while LZ formed its own independent group (Figure S3b). Hierarchical clustering also identified three clusters, one of which also consisted of populations YC and YH. Another cluster contained CF, FS, LZ and XL, and the third contained HC and MDJ (Figure S3c). One-way ANOVA indicated that the LZ population was most frequently and significantly different from other populations, and that caudal peduncle length and second dorsal fin base length were most frequently and significantly different between populations in pairwise comparisons (Figure S3d).

4 | DISCUSSION

One of the most salient findings of our study is that not one but at least three different *Pungitius* species in addition to *P. stenurus* exist in China. This finding is interesting not only from biodiversity and biogeographical perspectives, but also because *Pungitius* sticklebacks

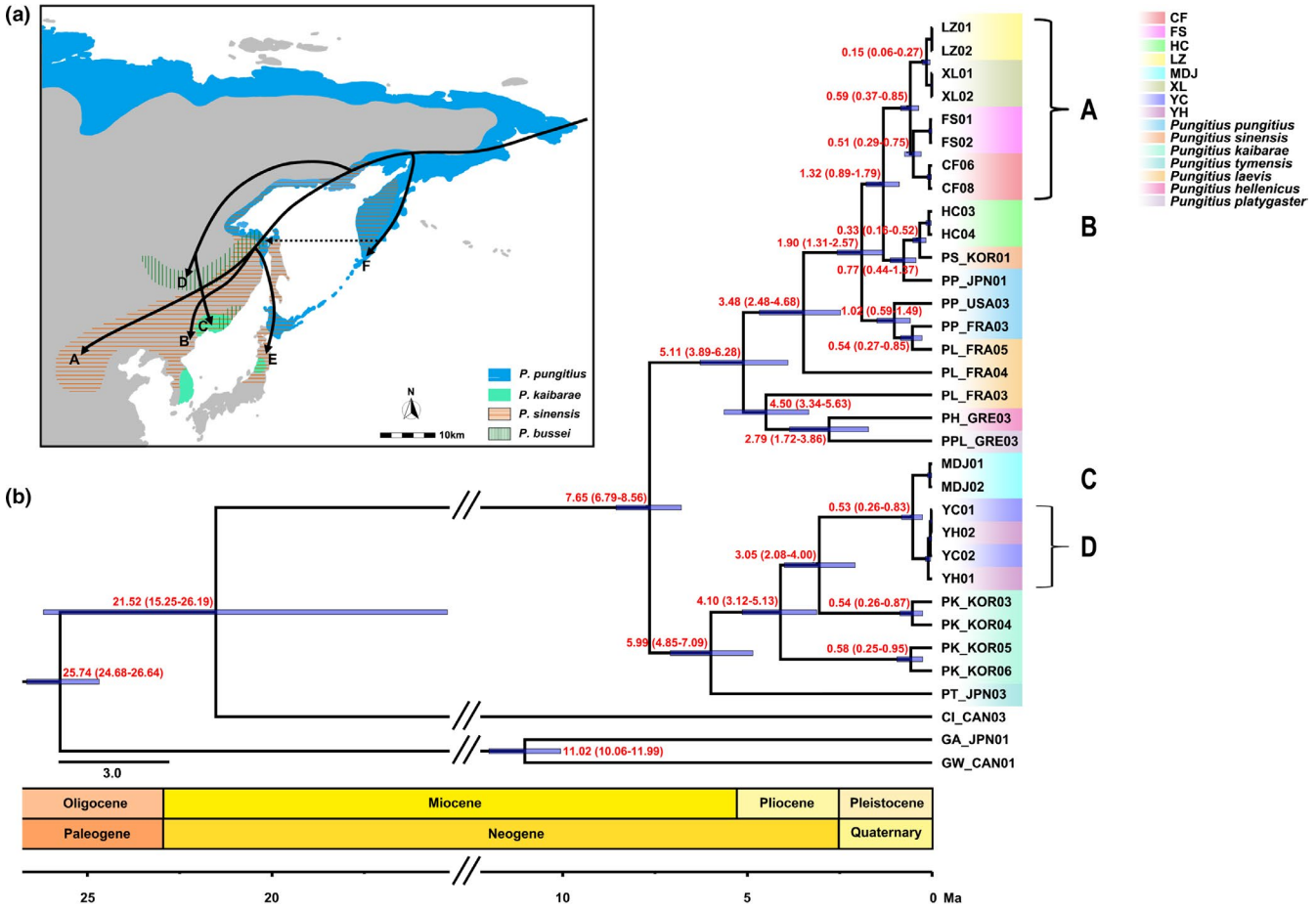

FIGURE 3 (a) Inferred dispersal routes of *Pungtius* to China. A–D correspond to populations in Figure 3b, respectively. "E" represents *P. sinensis* from Japan, and "F" represents *P. pungitius*. The dotted line with arrow indicates the mitochondrial capture event that occurred between *P. pungitius* and *P. sinensis*. (b) Divergence times in *Pungitius* sticklebacks based on 13 mitochondrial protein coding genes. Numbers at nodes are divergence times (Million years ago, Mya), and numbers in brackets are the 95% highest posterior density (HPD) intervals of the divergence time estimates. The bars on nodes indicate the 95% HPD interval of the divergence times

in China have been subject to translocations aiming to support locally dwindling populations under the assumption that all populations included belong to the species *P. sinensis*. This oversight is understandable in the light of our morphological findings: although the species do differ in their characteristics, formal identification is challenging based on simple linear measurement or character counts due to the large overlap in phenotypic variation. Moreover, our comprehensive phylogenomic analysis of *Pungitius* species of Northeast Asia further clarifies the complex sequence and timing of divergence among different taxa in this area and reveals their colonization routes. In the following, we discuss each of these points in more detail.

4.1 | The phylogeography and divergence of *Pungitius* sticklebacks in China

Our phylogeographic analyses based on genome-wide SNPs uncovered previously unrecognized diversity in the genus *Pungitius* in Northern China. Three distinct genetic lineages were identified in China, all of which are members of the *sinensis* clade (Guo et al., 2019). One of these lineages is closely related to *P. sinensis* from China, Russia and Japan, as well as to *P. polyakovi* from its type locality from Sakhalin Island (Shedko et al., 2005). This lineage is comprised of the populations CF, FS, HC, LZ and XL, and we consider that they belong to *P. sinensis*. Within this lineage, the HC population is clustered with *P. sinensis* from Russia, while the remaining populations cluster with *P. sinensis* from the Hebei Province of China (Guo et al., 2019). The second lineage is represented by samples from the locality MDJ. As it clusters with *P. kaibarae* from Russia, Japan and South Korea, we consider it to be *P. kaibarea*. The third lineage is comprised of samples from YC and YH, and clusters with *P. bussei* samples from Russia (Figure S2). Therefore, we consider it to be *P. bussei*. The mitochondrial phylogeny also suggests three genetic *Pungitius* lineages in China, but there was discrepancy with the phylogeny described above. Specifically, the mitochondrial phylogeny shows that the lineage comprising CF, FS, HC, LZ and XL populations clusters with *P. sinensis* from South Korea and belongs to the *pungitius* (rather than *sinensis*) clade (Figure 2). This is due to the well-known mitochondrial

capture of the *P. pungitius* mitogenome to *P. sinensis* (Guo et al., 2019; Takahashi et al., 2016; Takahashi & Takata, 2000; Wang et al., 2015). It is possible that there is yet another distinct evolutionary lineage of sticklebacks in China, but without samples of *P. stenurus* from its type locality we could not explore the possibility of its existence or its affinities to other species in this study.

The existence of three species of *Pungitius* in China is morphologically and biogeographically supported. Morphologically, individuals of the YC population have a long caudal peduncle, well-developed pelvic apparatus, and relatively long dorsal and pelvic spines. Although most of the YH samples were juveniles and hence morphological characteristics could not be fully observed, individuals of this population also have long and thin caudal peduncles. Our quantitative morphological analyses show that YC and YH form a cluster that is different from the other six populations based on the measurements of 13 linear traits (Figure S3). Biogeographically, YC and YH populations belong to the Songhua river basin, which is in the known distribution area of *P. bussei* (Bogutskaya et al., 2008; Keivany, 1996). Morphologically, individuals of the MDJ population have dark bodies, black and shorter dorsal and pelvic spines and smaller lateral scutes as compared to individuals from the YC population. Quantitative morphological analyses show that the MDJ population differs from the lineage comprised of CF, FS, HC, LZ and XL populations in body shape (Figure S3a), and differs from the lineage of YC and YH populations in the measurements of 13 linear traits (Figure S3b). Individuals of the MDJ population seem to be morphologically similar to *P. kaibarae*, as they too have black or blue dorsal colours and black pelvic spines (Figure 2; Bae & Suk, 2015; Takahashi et al., 2016). This is also supported by the fact that the MDJ samples are from the Suifen (Razdolnaya) River that flows into the Sea of Japan, and the rivers draining into the Sea of Japan from Zerkal'naya southward to Tumen River are likely the distribution area of *P. kaibarae* (Bogutskaya et al., 2008). Individuals of the CF, FS, HC, LZ and XL populations are morphologically consistent with the description of *P. sinensis* (Keivany, 1996), with two pelvic soft rays on each side, but lacking large lateral armour plates. However, individuals of the HC population are different from those of the CF, FS, LZ or XL populations, as they have longer and thinner caudal peduncles. Quantitative analyses also show that the HC population is different from the CF, FS, LZ or XL populations in body shape (Figure S3a). These qualitative and quantitative considerations together with biogeographic evidence support our designation of different samples to genetically defined species. Nevertheless, the fact remains that the analysis of shape and linear measurements does not allow unequivocal species demarcation.

The finding of unrecognized diversity in *Pungitius* sticklebacks in China sheds new light on the biogeography of *Pungitius* sticklebacks not only in China, but also in Northeast Asia, which is the diversity hotspot of *Pungitius* sticklebacks, especially considering frequent gene flow during diversification of this genus (Guo et al., 2019; Wang et al., 2017; Yamasaki et al., 2020). According to our molecular clock dating based on mitochondrial sequences (Figure 3b), along with nuclear genetic data from an earlier study (Guo et al., 2019), divergence of the three genetic *Pungitius* lineages in China took place at the Pliocene–Miocene border—long before they colonized China. Considering that there is fossil evidence for the presence of *Pungitius* species in China prior to the late Pliocene (Liu & Wang, 1974), this suggests that *Pungitius* sticklebacks colonized China in multiple waves (Figure 3a). The late Pliocene *Pungitius* fossil was found in the lower part of the Nihowan Formation in Hebei Province where *P. sinensis* is distributed, but the fossil specimen is not thought to be *P. sinensis*, as the shape of lateral bony scutes and the structure of the pelvic arch do not match those of *P. sinensis* (Liu & Wang, 1974). Nevertheless, its existence suggests that *Pungitius* sticklebacks were present in Northern China prior to the late Pliocene, which is long before the currently existing lineages arrived to China. The divergence of the *P. sinensis* lineage containing the Chinese populations CF, FS, HC, LZ and XL diverged from HC in the early Pleistocene (1.32 Mya). Similarly, the divergence between the Chinese *P. kaibarae* lineage (MDJ population) and the *P. bussei* lineage (YC and YH populations) occurred 0.53 Mya. Thus, our results confirm that despite the old age and wide geographical distribution of *P. sinensis*, the species has extended its distribution area to China relatively recently in the Pleistocene (Guo et al., 2019; Takahashi et al., 2016). This is further supported by the incongruence between *P. sinensis* mitochondrial and nuclear phylogenies. Our results, together with those of an earlier study (Guo et al., 2019) suggest that there are at least four divergent lineages of *P. sinensis*: the Northern lineage around the Sea of Okhotsk (PS-RU-TAN and PS-RU-BOL populations in Guo et al., 2019), the Japanese lineage (the freshwater type of *P. pungitius*), the lineage at the west coast of the Sea of Japan (HC population in this study, PS-RU-PRE and PS-RU-KIE populations in Guo et al., 2019, and the South Korean population), and the Chinese inland lineage (CF, FS, LZ or XL populations in this study). According to the nuclear data, *P. sinensis* diverged from *P. tymensis* and *P. kaibarae* ~4.26 Mya (95% HPD interval: 5.33–3.22; Guo et al., 2019), but the mitochondrial data suggest that divergence within *P. sinensis* took place ~1.32 Mya (95% HPD intervals: 0.89–1.79 Mya; Figure 3b) after mitogenome capture from *P. pungitius*. Interestingly, the mitogenome capture from *P. pungitius* to *P. sinensis* might have occurred multiple times according to our mitochondrial phylogeny. This possibility is suggested by the finding that *P. pungitius* from Japan clusters with the lineage of *P. sinensis* from the west coast of the Sea of Japan (Figure 3). If this scenario is true, it suggests that the HC population colonized China independently of the CF, FS, LZ and XL populations.

Diversification of *Pungitius* sticklebacks in Northeast Asia has proceeded mainly in freshwater habitats (Guo et al., 2019) and their current distributions are largely associated with past geomorphological changes in this area (Takahashi, & Goto, 2001), except that the *P. pungitius* is found in both freshwater and marine habitats. While Northeast Asia was not covered by large ice sheets during the Quaternary, except for small glaciers in mountainous regions, it experienced repeated Pleistocene climate oscillations with cold and dry climate during glacial periods and warm and humid

climate during interglacial periods (Barr & Clark, 2012). *Pungitius* sticklebacks that colonized the inland of Northeast Asia in earlier waves existed by late Pliocene but were possibly extinct later in Pleistocene (Liu & Wang, 1974). The Pleistocene glacial and interglacial climate oscillations that forced sea level rises and falls as well as river capture events, have played an important role in shaping geographic distribution of many fish taxa both globally (Bernatchez & Wilson, 1998) and in Northeast Asia (Xu et al., 2014). Colonization of the Japanese archipelago from the Eurasian continent has been suggested for many freshwater fishes (Yuma et al., 1998). The currently wide distribution of *P. sinensis* lineages in Northeast Asia likely came to existence as a result of the Pleistocene glacial and interglacial climate oscillations which re- and disconnected the Japanese archipelago from the Eurasian continent multiple times. The development of low-salinity surface water during Pleistocene glacial periods might have allowed *Pungitius*, especially the highly euryhaline species *P. sinensis*, to disperse through coastal waters in the Sea of Japan and the adjacent Sea of Okhotsk (Takahashi et al., 2016). For example, our results suggest that HC population of *P. sinensis* colonized China independently of the CF, FS, LZ, and XL populations and from the Japanese archipelago. This suggests that migration of freshwater fishes occurred not only from the Eurasian continent to the Japanese archipelago, but also vice versa. Therefore, the results of this study, together with earlier studies (Aldenhoven et al., 2010; Feng et al., 2021; Guo et al., 2019; Shikano et al., 2010; Takahashi et al., 2016; Wang et al., 2015) show that *Pungitius* sticklebacks could serve as a good model system for biogeographical studies aiming to understand the response of populations and species to geological events and changing climatic conditions on both regional and global scales.

4.2 | Taxonomy and its implications for conservation of *Pungitius* sticklebacks in China

Our results show that there are at least four *Pungitius* species in China, namely *P. sinensis*, *P. kaibarae*, *P. bussei* and *P. sternus*. Until now, both *P. kaibarae* and *P. bussei* in China were assumed to be *P. sinensis* and hence represent new records of the fauna in China. This unrecognized diversity can be explained by the following facts. First, the taxonomy and species identification of *Pungitius* relies heavily on osteological and armour traits (e.g., pelvis structures, spine numbers, lateral plate numbers and presence or absence of a keel), as well as coloration as diagnostic characters (Keivany & Nelson, 2000, 2004). However, these traits are labile; correct species identification based on these traits is difficult and sometimes impossible (Guo et al., 2019; Takahashi et al., 2016; Wang et al., 2015). For example, recent studies (Guo et al., 2019; Takahashi et al., 2016) show that the freshwater type of Japanese *P. pungitius* is actually *P. sinensis*, even if it does not have large lateral armour plates typical of *P. pungitius*. With the exception of fish from the MDJ population, all other *Pungitius* populations studied here are consistent with the morphological description of *P. sinensis* (Keivany, 1996), with light body coloration, long and thin caudal peduncles, well-developed pelvic apparatus and reduced lateral armour plates. MDJ individuals have dark bodies, black dorsal spines and short pelvic spines typical of *P. kaibarae* (Bae & Suk, 2015; Takahashi et al., 2016). Therefore, it is difficult to distinguish *P. bussei* (populations of YC and YH) from *P. sinensis* (populations of HC, CF, FS, XL, LZ) on the basis of osteological, armour or colour traits. Our quantitative analyses of morphology support the difficulty of demarcating the three species. A second explanation for the unrecognized diversity relates to the fact that earlier molecular studies of Asian *Pungitius* sticklebacks (Guo et al., 2019; Takahashi et al., 2016) have included only a single Chinese *Pungitius* population in their analyses. As this population happened to be *P. sinensis*, the existence of other species has gone undetected, illustrating the importance of adequate geographical sampling in biogeographic studies.

Further clarification of the distribution of different *Pungitius* stickleback species in China is important for their conservation management. Although *P. stenurus* is the only species currently classified as endangered by the International Union for Conservation of Nature (IUCN), other *Pungitius* populations and/or species might deserve the status of "endangered" or even "critically endangered" according to the demographic criteria used by the IUCN (von Hippel, 2008; Ishikawa et al., 2013; Kitano & Mori, 2016; Merilä, 2013). For example, nearly 50% of native fish species may have been extirpated in Beijing and adjacent areas based on field surveys between 2002 and 2010, including *P. sinensis* that has been enlisted as a class II protected species in Beijing (Zhang et al., 2011). Our results show that *P. sinensis* populations near Beijing (CF and XL) do have low genetic diversity compared to other populations (FS, HC and LZ; Table S3), which might suggest lower capability of those populations to adapt to changing environmental conditions. In order to recover the regional *P. sinensis* population in Beijing, an artificial re-stocking programme was carried out, and 100,000 wild collected *Pungitius* from Northeast China were released to the Beijing area in 2017 and 2018 according to the Beijing Fishery Science Research Institute. However, although the introduced fish were presumed to be *P. sinensis*, genealogical information about the released fish is not available. Given the findings of our study, these releases raise a concern that the introduced species might have been *P. kaibarae*, *P. bussei* or the west coast of the Sea of Japan lineage of *P. sinensis* to where the Chinese inland lineage of *P. sinensis* is distributed. In this case, these releases might stimulate the extirpation of the local *P. sinensis* population(s) that might still exist in the wild in Beijing. Even if they do not go extinct, the local *P. sinensis* populations in the Beijing area might be subject to introgression with unknown consequences to their fitness and evolutionary integrity (Araki et al., 2007; Eldridge & Naish, 2007). In this context, it is worth noting that the population LZ reported in 2018 from Henan Province (Zhou et al., 2018) has been recorded as the southernmost population of *Pungitius* in China, and thought to be artificially established by translocation from elsewhere (Prof. Chunguang

Zhang, personal communication, 16 January 2020). However, our phylogenetic analyses and divergence time estimates suggest that the LZ population has probably a natural, rather than artificial, origin according to its divergence from the XL population (0.06–0.27 Mya; Figure 3). Finally, considering that the subtle phenotypic differences between species pose challenges to species identification in this genus, conservation and management authorities dealing with sticklebacks should resort to molecular species identification when planning translocations.

5 | CONCLUSIONS

Taken together, the results of this study complement our understanding of the evolutionary history and biogeography of the genus *Pungitius* by extending sampling, for the first time, to cover an area that comprises about half of the known distribution area of this genus in Northeast Asia. A genetic survey of this area led to the discovery of two new recorded species (*P. kaibarae* and *P. bussei*) of the fauna in China, as well the discovery of a previously unknown divergent *P. sinensis* lineage in China. The results further suggest that colonization of *Pungitius* sticklebacks in China has occurred in multiple waves, and *P. sinensis* has extended its distribution area relatively recently in the Pleistocene. Finally, the results pinpoint two lines of future research: first, the assessment of evolutionary affiliations of *P. sternus*—a species endemic to China but not included in this study due to lack of samples; second, investigation of whether historical stickleback stocking to the Beijing area has led to establishment of *P. kaibarae* and *P. bussei* populations in this area, and possibly their hybridization with *P. sinensis*.

ACKNOWLEDGEMENTS

We thank Haibo Liu, Chengyi Niu, Dandan Qi and Frank A. von Hippel for help in sampling, Xinxin Li for help in phylogenetic analyses, and Shan Huang for helpful comments. We thank the China National Animal Collection Resource Center (http://museum.ioz.ac.cn/) for providing specimens to this study. The study is supported by the National Natural Science Foundation of China (grant nos. 31672273 and 32022009), the CAS Pioneer Hundred Talents Program and the Academy of Finland (grant nos. 134728 and 218343).

CONFLICT OF INTEREST

The authors claim no conflict of interests.

PEER REVIEW

The peer review history for this article is available at https://publons.com/publon/10.1111/ddi.13423.

DATA AVAILABILITY STATEMENT

Raw sequence reads and assembled mitochondrial genomes underlying this study have been deposited in NCBI, and their accession numbers are listed in Table S1. The SNP data have been deposited in Dryad (https://doi.org/10.5061/dryad.2bvq83bqw).

ORCID

Juha Merilä https://orcid.org/0000-0001-9614-0072
Baocheng Guo https://orcid.org/0000-0002-9796-5700

REFERENCES

Aldenhoven, J. T., Miller, M. A., Corneli, P. S., & Shapiro, M. D. (2010). Phylogeography of ninespine sticklebacks (*Pungitius pungitius*) in North America: Glacial refugia and the origins of adaptive traits. *Molecular Ecology*, *19*(18), 4061–4076. https://doi.org/10.1111/j.1365-294X.2010.04801.x

Araki, H., Cooper, B., & Blouin, M. S. (2007). Genetic effects of captive breeding cause a rapid, cumulative fitness decline in the wild. *Science*, *318*(5847), 100–103. https://doi.org/10.1126/science.1145621

Avise, J. C. (2000). *Phylogeography: The history and formation of species*. Harvard University Press.

Bae, H. G., & Suk, H. Y. (2015). Population genetic structure and colonization history of short ninespine sticklebacks (*Pungitius kaibarae*). *Ecology and Evolution*, *5*(15), 3075–3089. https://doi.org/10.1002/ece3.1594

Barr, I. D., & Clark, C. D. (2012). Late Quaternary glaciations in Far NE Russia; combining moraines, topography and chronology to assess regional and global glaciation synchrony. *Quaternary Science Reviews*, *53*, 72–87. https://doi.org/10.1016/j.quascirev.2012.08.004

Beheregaray, L. B. (2008). Twenty years of phylogeography: The state of the field and the challenges for the Southern Hemisphere. *Molecular Ecology*, *17*(17), 3754–3774. https://doi.org/10.1111/j.1365-294X.2008.03857.x

Bell, M. A., & Foster, S. A. (1995). *The evolutionary biology of the threespine stickleback*. Oxford University Press.

Berg, L. S. (1949). *Freshwater fishes of the USSR and adjacent countries*. Izdatel'stvo Akademii Nauk SSSR.

Bernatchez, L., & Wilson, C. C. (1998). Comparative phylogeography of nearctic and palearctic fishes. *Molecular Ecology*, *7*(4), 431–452. https://doi.org/10.1046/j.1365-294x.1998.00319.x

Bernt, M., Donath, A., Jühling, F., Externbrink, F., Florentz, C., Fritzsch, G., Pütz, J., Middendorf, M., & Stadler, P. F. (2013). MITOS: Improved de novo metazoan mitochondrial genome annotation. *Molecular Phylogenetics and Evolution*, *69*(2), 313–319. https://doi.org/10.1016/j.ympev.2012.08.023

Bickford, D., Lohman, D. J., Sodhi, N. S., Ng, P. K. L., Meier, R., Winker, K., Ingram, K. K., & Das, I. (2007). Cryptic species as a window on diversity and conservation. *Trends in Ecology & Evolution*, *22*(3), 148–155. https://doi.org/10.1016/j.tree.2006.11.004

Bogutskaya, N. G., Naseka, A. M., Shedko, S. V., Vasileva, E. D., & Chereschnev, I. A. (2008). The fishes of the Amur River: Updated check-list and zoogeography. *Ichtyological Exploration of Freshwaters*, *19*(4), 301–366.

Bou, S. (2011). *Fauna inner mongolia (Volume One)*. Inner Mongolia University Press.

Bouckaert, R., Heled, J., Kühnert, D., Vaughan, T., Wu, C.-H., Xie, D., Suchard, M. A., Rambaut, A., & Drummond, A. J. (2014). BEAST 2: A software platform for Bayesian evolutionary analysis. *PLoS Computational Biology*, *10*(4), e1003537. https://doi.org/10.1371/journal.pcbi.1003537

Cheng, T., & Zheng, B. (1987). *The systematic synopsis of Chinese fishes (the first volume)*. Science Press.

Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., Handsaker, R. E., Lunter, G., Marth, G. T., Sherry, S. T., McVean, G., & Durbin, R., & 1000 Genomes Project Analysis Group (2011). The variant call format and VCFtools. *Bioinformatics*, *27*(15), 2156–2158. https://doi.org/10.1093/bioinformatics/btr330

Denys, G. P. J., Persat, H., Dettai, A., Geiger, M. F., Freyhof, J., Fesquet, J., & Keith, P. (2018). Genetic and morphological discrimination of

three species of ninespined stickleback *Pungitius* spp. (Teleostei, Gasterosteidae) in France with the revalidation of *Pungitius vulgaris* (Mauduyt, 1848). *Journal of Zoological Systematics and Evolutionary Research*, *56*(1), 77–101. https://doi.org/10.1111/jzs.12178

Eldridge, W. H., & Naish, K. A. (2007). Long-term effects of translocation and release numbers on fine-scale population structure among coho salmon (*Oncorhynchus kisutch*). *Molecular Ecology*, *16*(12), 2407–2421. https://doi.org/10.1111/j.1365-294X.2007.03271.x

Eschmeyer, W. N., Fricke, R., & Van der Laan, R. (2016). Catalog of fishes: Genera, species, references. In: Electronic version accessed 19.

Fang, B., Kemppainen, P., Momigliano, P., & Merilä, J. (2021). Population structure limits parallel evolution in sticklebacks. *Molecular Biology and Evolution*. https://doi.org/10.1093/molbev/msab144

Feng, X., Merilä, J., & Löytynoja, A. (2021). Detailed phylogeographic analysis of sticklebacks exposes challenges in genetic studies of admixed populations. *bioRxiv*. https://doi.org/10.1101/2021.07.16.452636

Guichenot, M. (1868). Notice sur quelques poissons inédits de Madagascar et da la Chine. *Nouvelles Archives Du Muséum D'histoire Naturelle, Paris*, *5*, 193–206.

Guisan, A., & Thuiller, W. (2005). Predicting species distribution: Offering more than simple habitat models. *Ecology Letters*, *8*, 993–1009. https://doi.org/10.1111/j.1461-0248.2005.00792.x

Guo, B., Fang, B., Shikano, T., Momigliano, P., Wang, C., Kravchenko, A., & Merilä, J. (2019). A phylogenomic perspective on diversity, hybridization and evolutionary affinities in the stickleback genus *Pungitius*. *Molecular Ecology*, *28*(17), 4046–4064. https://doi.org/10.1111/mec.15204

Hwang, D. S., Song, H. B., & Lee, J. S. (2012). Complete mitochondrial genome of the Amur stickleback *Pungitius kaibarae* (Gasterosteiformes, Gasterosteidae). *Mitochondrial DNA*, *23*(4), 313–314. https://doi.org/10.3109/19401736.2012.683183

Ishikawa, A., Takeuchi, N., Kusakabe, M., Kume, M., Mori, S., Takahashi, H., & Kitano, J. (2013). Speciation in ninespine stickleback: Reproductive isolation and phenotypic divergence among cryptic species of Japanese ninespine stickleback. *Journal of Evolutionary Biology*, *26*(7), 1417–1430. https://doi.org/10.1111/jeb.12146

Kalyaanamoorthy, S., Minh, B. Q., Wong, T. K. F., von Haeseler, A., & Jermiin, L. S. (2017). ModelFinder: Fast model selection for accurate phylogenetic estimates. *Nature Methods*, *14*(6), 587–589. https://doi.org/10.1038/nmeth.4285

Katoh, K., & Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Molecular Biology and Evolution*, *30*(4), 772–780. https://doi.org/10.1093/molbev/mst010

Kawahara, R., Miya, M., Mabuchi, K., Near, T. J., & Nishida, M. (2009). Stickleback phylogenies resolved: Evidence from mitochondrial genomes and 11 nuclear genes. *Molecular Phylogenetics and Evolution*, *50*(2), 401–404. https://doi.org/10.1016/j.ympev.2008.10.014

Keivany, Y. (1996). *Taxonomic revision of the genus Pungitius with emphasis on* P. hellenicus (M.Sc.). University of Alberta.

Keivany, Y., & Nelson, J. S. (2000). Taxonomic review of the genus *Pungitius*, ninespine sticklebacks (Gasterosteidae). *Cybium: International Journal of Ichthyology*, *24*, 107–122.

Keivany, Y., & Nelson, J. S. (2004). Phylogenetic relationships of sticklebacks (Gasterosteidae), with emphasis on ninespine sticklebacks (*Pungitius* spp.). *Behaviour*, *141*, 1485–1497. https://doi.org/10.1163/1568539042948187

Kitano, J., & Mori, S. (2016). Toward conservation of genetic and phenotypic diversity in Japanese sticklebacks. *Genes & Genetic Systems*, *91*(2), 77–84. https://doi.org/10.1266/ggs.15-00082

Klingenberg, C. P. (2011). MorphoJ: An integrated software package for geometric morphometrics. *Molecular Ecology Resources*, *11*(2), 353–357. https://doi.org/10.1111/j.1755-0998.2010.02924.x

Le, S., Josse, J., & Husson, F. (2008). FactoMineR: An R package for multivariate analysis. *Journal of Statistical Software*, *25*(1), 1–18. https://doi.org/10.18637/jss.v025.i01

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., & Durbin, R., & 1000 Genome Project Data Processing Subgroup. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics*, *25*(16), 2078–2079. https://doi.org/10.1093/bioinformatics/btp352

Liu, X. T., & Wang, N. C. (1974). A new *Pungitius* from Nihewan formation of North China. *Vertebrata Palasiatica*, *12*(2), 89–96. https://doi.org/10.19615/j.cnki.1000-3118.1974.02.02

Meng, G., Li, Y., Yang, C., & Liu, S. (2019). MitoZ: A toolkit for animal mitochondrial genome assembly, annotation and visualization. *Nucleic Acids Research*, *47*(11), e63. https://doi.org/10.1093/nar/gkz173

Merilä, J. (2013). Nine-spined stickleback (*Pungitius pungitius*): An emerging model for evolutionary biology research. *Annals of the New York Academy of Sciences*, *1289*(1), 18–35. https://doi.org/10.1111/nyas.12089

Miya, M., Kawaguchi, A., & Nishida, M. (2001). Mitogenomic exploration of higher teleostean phylogenies: A case study for moderate-scale evolutionary genomics with 38 newly determined complete mitochondrial DNA sequences. *Molecular Biology and Evolution*, *18*(11), 1993–2009. https://doi.org/10.1093/oxfordjournals.molbev.a003741

Natri, H. M., Merilä, J., & Shikano, T. (2019). The evolution of sex determination associated with a chromosomal inversion. *Nature Communications*, *10*(1), 145. https://doi.org/10.1038/s41467-018-08014-y

Nelson, T. C., & Cresko, W. A. (2018). Ancient genomic variation underlies repeated ecological adaptation in young stickleback populations. *Evolution Letters*, *2*(1), 9–21. https://doi.org/10.1002/evl3.37

Nguyen, L. T., Schmidt, H. A., von Haeseler, A., & Minh, B. Q. (2015). IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Molecular Biology and Evolution*, *32*(1), 268–274. https://doi.org/10.1093/molbev/msu300

Rambaut, A., Drummond, A. J., Xie, D., Baele, G., & Suchard, M. A. (2018). Posterior summarization in Bayesian phylogenetics using Tracer 1.7. *Systematic Biology*, *67*(5), 901–904. https://doi.org/10.1093/sysbio/syy032

Rogers, S. M., Tamkee, P., Summers, B., Balabahadra, S., Marks, M., Kingsley, D. M., & Schluter, D. (2012). Genetic signature of adaptive peak shift in threespine stickleback. *Evolution*, *66*(8), 2439–2450. https://doi.org/10.1111/j.1558-5646.2012.01622.x

Rohlf, F. J. (2013). TpsDig v. 2.17. Retrieved from http://life.bio.sunysb.edu/morph/

Rohlf, F. J., & Slice, D. (1990). Extensions of the procrustes method for the optimal superimposition of landmarks. *Systematic Zoology*, *39*(1), 40–59. https://doi.org/10.2307/2992207

Schluter, D. (2000). *The ecology of adaptive radiation*. Oxford University Press.

Schluter, D., Clifford, E. A., Nemethy, M., & McKinnon, J. S. (2004). Parallel evolution and inheritance of quantitative traits. *The American Naturalist*, *163*(6), 809–822. https://doi.org/10.1086/383621

Shedko, S. V., Shedko, M. B., & Pietsch, T. (2005). *Pungitius polyakovi* sp. n., a new species of ninespine stickleback (Gasterosteiformes, Gasterosteidae) from southeastern Sakhalin Island. Dalnauka, Vladivostok.

Shikano, T., Shimada, Y., Herczeg, G., & Merilä, J. (2010). History vs. habitat type: Explaining the genetic structure of European nine-spined stickleback (*Pungitius pungitius*) populations. *Molecular Ecology*, *19*(6), 1147–1161. https://doi.org/10.1111/j.1365-294X.2010.04553.x

Shimada, Y., Shikano, T., Kuparinen, A., Gonda, A., Leinonen, T., & Merilä, J. (2011). Quantitative genetics of body size and timing of maturation in two nine-spined stickleback (*Pungitius pungitius*)

populations. *PLoS One*, *6*(12), e28859. https://doi.org/10.1371/journal.pone.0028859

Takahashi, H., & Goto, A. (2001). Evolution of East Asian ninespine sticklebacks as shown by mitochondrial DNA control region sequences. *Molecular Phylogenetics and Evolution*, *21*(1), 135–155. https://doi.org/10.1006/mpev.2001.1001

Takahashi, H., Møller, P. R., Shedko, S. V., Ramatulla, T., Joen, S.-R., Zhang, C.-G., Sideleva, V. G., Takata, K., Sakai, H., Goto, A., & Nishida, M. (2016). Species phylogeny and diversification process of Northeast Asian *Pungitius* revealed by AFLP and mtDNA markers. *Molecular Phylogenetics and Evolution*, *99*, 44–52. https://doi.org/10.1016/j.ympev.2016.03.022

Takahashi, H., & Takata, K. (2000). Multiple lineages of the mitochondrial DNA introgression from Pungitius pungitius (L.) to *Pungitius tymensis* (Nikolsky). *Canadian Journal of Fisheries and Aquatic Sciences*, *57*(9), 1814–1833. https://doi.org/10.1139/f00-133

Varadharajan, S., Rastas, P., Löytynoja, A., Matschiner, M., Calboli, F. C. F., Guo, B., Nederbragt, A. J., Jakobsen, K. S., & Merilä, J. (2019). A high-quality assembly of the nine-spined stickleback (*Pungitius pungitius*) genome. *Genome Biology and Evolution*, *11*(11), 3291–3308. https://doi.org/10.1093/gbe/evz240

von Hippel, F. A. (2008). Conservation of threespine and ninespine stickleback radiations in the Cook Inlet basin. *Alaska. Behaviour*, *145*(4–5), 693–724. https://doi.org/10.1163/156853908792451467

Wang, C., Shikano, T., Persat, H., & Merilä, J. (2015). Mitochondrial phylogeography and cryptic divergence in the stickleback genus *Pungitius*. *Journal of Biogeography*, *42*(12), 2334–2348. https://doi.org/10.1111/jbi.12591

Wang, C., Shikano, T., Persat, H., & Merilä, J. (2017). Phylogeography and historical introgression in smoothtail nine-spined sticklebacks, *Pungitius laevis* (Gasterosteiformes: Gasterosteidae). *Biological Journal of the Linnean Society*, *121*(2), 340–354. https://doi.org/10.1093/biolinnean/blw036

Wang, Y., Zhao, Y., Wang, Y., Li, Z., Guo, B., & Merilä, J. (2020). Population transcriptomics reveals weak parallel genetic basis in repeated marine and freshwater divergence in nine-spined sticklebacks. *Molecular Ecology*, *29*(9), 1642–1656. https://doi.org/10.1111/mec.15435

Wootton, R. J. (1976). *The biology of the sticklebacks*. Academic Press.

Wootton, R. J. (1984). *A functional biology of sticklebacks*. Croom Helm.

Xie, Y. (2007). *The freshwater fishes in Northeast China*. Liaoning Science and Technology Press.

Xu, W., Yin, W., Chen, A., Li, J., Lei, G., & Fu, C. (2014). Phylogeographical analysis of a cold-temperate freshwater fish, the Amur sleeper (*Perccottus glenii*) in the Amur and Liaohe River basins of Northeast Asia. *Zoological Science*, *31*(10), 671–679. https://doi.org/10.2108/zs130046

Yamasaki, Y. Y., Kakioka, R., Takahashi, H., Toyoda, A., Nagano, A. J., Machida, Y., & Kitano, J. (2020). Genome-wide patterns of divergence and introgression after secondary contact between *Pungitius* sticklebacks. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *375*(1806), 20190548. https://doi.org/10.1098/rstb.2019.0548

Yang, J., Guo, B., Shikano, T., Liu, X., & Merilä, J. (2016). Quantitative trait locus analysis of body shape divergence in nine-spined sticklebacks based on high-density SNP-panel. *Scientific Reports*, *6*, 26632. https://doi.org/10.1038/srep26632

Yuma, M., Hosoya, K., & Nagata, Y. (1998). Distribution of the freshwater fishes of Japan: An historical overview. *Environmental Biology of Fishes*, *52*, 97–124. https://doi.org/10.1023/A:1007451417556

Zelditch, M., Swiderski, D., Sheets, H., & Fink, W. (2004). *Geometric morphometrics for biologists: A primer.* , 1st ed. Elsevier.

Zhang, C., & Zhao, Y. (2016). *Species diversity and distribution of inland fishes in China*. Science Press.

Zhang, C., Zhao, Y., Xing, Y., Guo, R., Zhang, Q., Feng, Y., & Fan, E. (2011). Fish species diversity and conservation in Beijing and adjacent areas. *Biodiversity Science*, *19*(5), 597–604. https://doi.org/10.3724/sp.J.1003.2011.08072

Zhang, J. (1995). *The ichthyography of Heilongjiang Province*. Heilongjiang Science and Technology Press.

Zhao, W. (2018). *The illustrated of fishes of Heilongjiang province*. Science Press.

Zheng, B., Huang, H., Zhang, Y., & Dai, D. (1980). *The fishes of Tumen River*. Jilin People's Publishing House.

Zhou, C., Gu, Q., Meng, X., Li, B., Zhao, W., Pang, X., Zhang, J., Song, D., & Nie, G. (2018). Analysis of new fish record of *Pungitius sinensis* in Henan Province based on multiple data resource. *Sichuan Journal of Zoology*, *37*(1), 67–73. https://doi.org/10.11984/j.issn.1000-7083.20170037

Zhu, S. (1995). *The synopsis of freshwater fishes of China*. Jiangsu Science and Technology Press.

BIOSKETCH

This team of researchers works collaboratively to investigate genetic basis of ecological adaptation in fishes globally, by combining expertise in population genetics, genomics, biogeography and taxonomy.

Author contributions: B.G. conceived the project. Y.W., Yu.W., Y.Z. and A.K. analysed the data. B.G., Y.W. and J.M. wrote the paper. All authors read and approved the final manuscript.

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: Wang, Y., Wang, Y., Zhao, Y., Kravchenko, A. Y., Merilä, J., & Guo, B. (2021). Phylogenomics of Northeast Asian *Pungitius* sticklebacks. *Diversity and Distributions*, 00, 1–12. https://doi.org/10.1111/ddi.13423