

## ARTICLE OPEN



## Ecological genetics of isolated loach populations indicate compromised adaptive potential

Xi Wang <sup>1</sup>✉, Kerry Reid <sup>1</sup>, Ying Chen <sup>1</sup>, David Dudgeon <sup>1</sup> and Juha Merilä <sup>1,2</sup>✉

© The Author(s) 2024

Many endangered species live in fragmented and isolated populations with low genetic variability, signs of inbreeding, and small effective population sizes - all features elevating their extinction risk. The flat-headed loach (*Oreonectes platycephalus*), a small noemacheilid fish, is widely across southern China, but only in the headwaters of hillstreams; as a result, they are spatially isolated from conspecific populations. We surveyed single nucleotide polymorphisms in 16 Hong Kong populations of *O. platycephalus* to determine whether loach populations from different streams were genetically isolated from each other, showed low levels of genetic diversity, signs of inbreeding, and had small contemporary effective population sizes. Estimates of average observed heterozygosity ( $H_O = 0.0473$ ), average weighted nucleotide diversity ( $\pi_w = 0.0546$ ) and contemporary effective population sizes ( $N_e = 10.2 \sim 129.8$ ) were very low, and several populations showed clear signs of inbreeding as judged from relatedness estimates. The degree of genetic differentiation among populations was very high (average  $F_{ST} = 0.668$ ), even over short geographic distances ( $<1.5$  km), with clear patterns of isolation by distance. These results suggest that Hong Kong populations of *O. platycephalus* have experienced strong genetic drift and loss of genetic variability because sea-level rise after the last glaciation reduced connectedness among paleodrainages, isolating populations in headwaters. All this, together with the fact that the levels of genetic diversity and contemporary effective population sizes within *O. platycephalus* populations are lower than most other freshwater fishes, suggests that they face high local extinction risk and have limited capacity for future adaptation.

Heredity (2024) 133:88–98; <https://doi.org/10.1038/s41437-024-00695-0>

## INTRODUCTION

Changing climatic conditions pose severe challenges for local populations of many organisms and threaten their existence (Albano et al. 2021). Species with high dispersal ability can shift their ranges to include more thermally suitable habitats, leading to the redistribution of global biodiversity (Pech et al. 2017). For instance, it has been predicted that, by 2070, the climatic niches of around 30% of birds and mammals will include countries where they have never lived previously (Tittley et al. 2021). However, migration to new areas is not an option for many non-vagile species in isolated habitat patches, such as lakes and mountaintops. Phenotypic plasticity may provide means for them to acclimate to warmer temperatures (Gienapp et al. 2008) but it is unlikely that such responses will allow populations to keep pace with rising temperatures indefinitely (Cerini et al. 2023). Furthermore, evolutionary adaptation to new environmental conditions is possible only if there is sufficient genetic variation in the population to allow response to natural selection (Lynch and Walsh 1998).

Genetic diversity is the fuel that natural selection needs to allow organismal adaptation to changing environmental conditions. Whether adaptation to changing environmental conditions will occur depends critically on populations' access to genetic variation in the traits influencing fitness. The likelihood of adaptation is the function of effective population size: larger

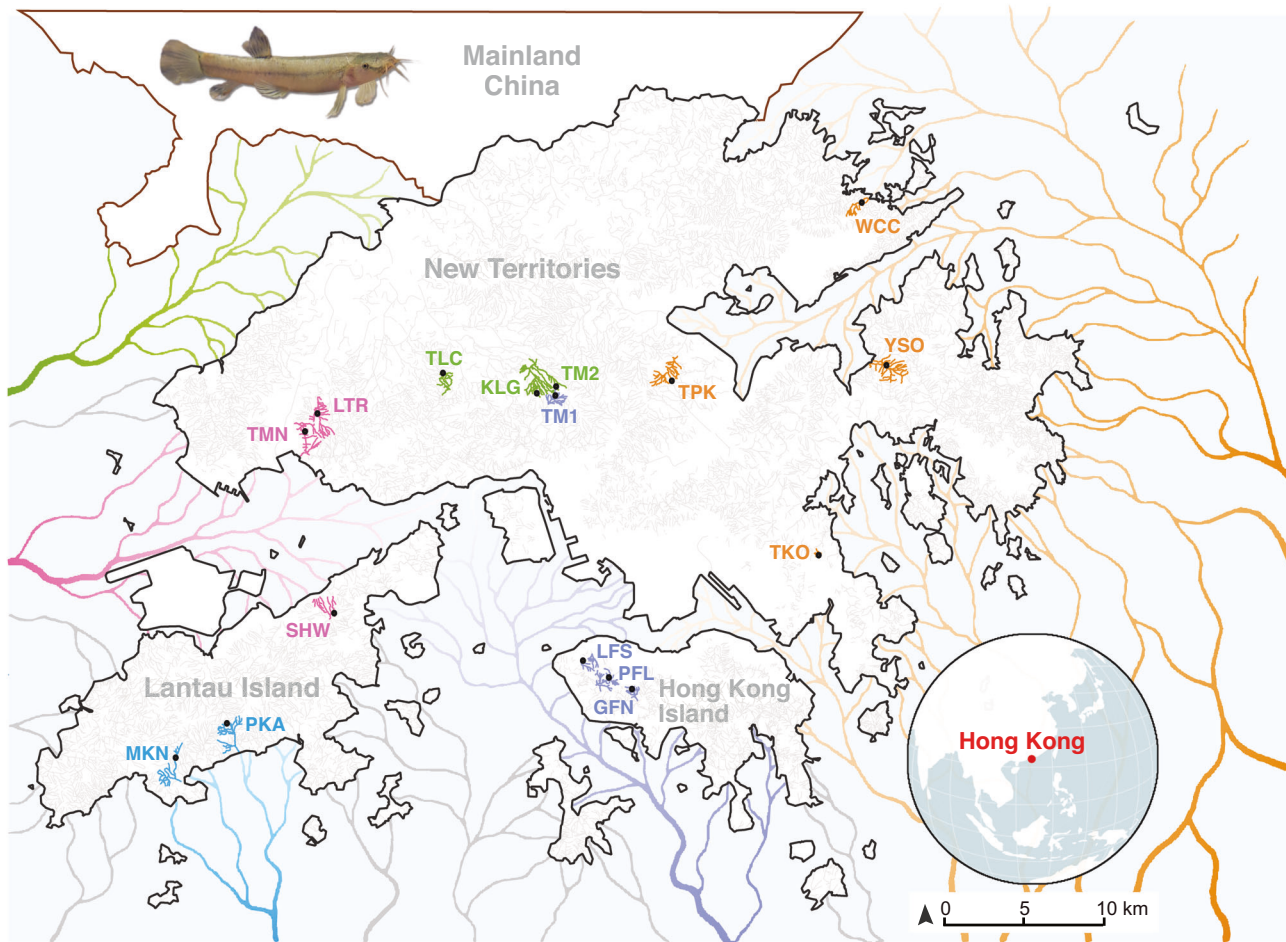
populations harbor more genetic variation than smaller ones, and natural selection is more efficient in the former (Lanfear et al. 2014; Saccheri and Hanski 2006). However, large populations are also larger targets for new mutations and are therefore expected to be burdened with relatively high loads of deleterious recessive mutations (Grossen et al. 2020; Saccheri and Hanski 2006). Furthermore, once the size of formerly large populations becomes reduced – for instance due to climate change-imposed fitness loss – there is a risk that segregating deleterious mutations will become expressed through inbreeding causing significant maladaptation (Framout et al. 2023; Grossen et al. 2020; Robinson et al. 2019).

The flat-headed loach (*Oreonectes platycephalus*: Noemacheilidae; Fig. 1) is a small freshwater fish that lives in the headwaters of hillstreams in Hong Kong and southern China (Du et al. 2008; Dudgeon 2003). The species' wide distribution and confinement to the top parts of hillstreams suggest that it colonized its current habitats post-glacially. Since then, many previously connected streams have become isolated from each other due to ca. 150 m sea level rise after the final glacial maximum at 12,000 years BP (Fyfe et al. 2000). As a result, most populations of this loach in Hong Kong are isolated from each other and hence can be expected to have been subject to attrition of genetic diversity due to genetic drift and inbreeding. Furthermore, given that post-

<sup>1</sup>Area of Ecology and Biodiversity, School of Biological Sciences, The University of Hong Kong, Hong Kong SAR, China. <sup>2</sup>Ecological Genetics Research Unit, Organismal and Evolutionary Biology Programme, University of Helsinki, FI-00014 University of Helsinki, Helsinki, Finland. Associate editor: Bastiaan Star. ✉email: [u3009279@connect.hku.hk](mailto:u3009279@connect.hku.hk); [juha.merila@helsinki.fi](mailto:juha.merila@helsinki.fi)

Received: 10 January 2024 Revised: 18 May 2024 Accepted: 20 May 2024

Published online: 3 July 2024



**Fig. 1** Study populations of flat-headed loaches. Map of the study area showing the sampled hillstreams in the territory of Hong Kong with different palaeo-drainages indicated in color (green: Tai Mo Shan drainage - TMS; pink: West New Territories drainage - NTW; blue: South Lantau drainage - LTS; purple: Hong Kong Island drainage - HKI; orange: East New Territories drainage - NTE; gray: Other paleodrainages not included in this study) redrawn from Fyfe et al. (2000). Light blue depicts the current marine area. Fish insert at the top (left) depicts the study species, the flat-headed loach. For locality abbreviations, see Supplementary Table S1 (Photo courtesy: Chi Kit Yeung).

glacially formed habitats of coastal southeast China are young (Fyfe et al. 1999), there has been little time to restore any lost genetic variation through mutations, because mutation rates in fish are low (Bergeron et al. 2023; Zhang et al. 2023).

Many freshwater systems can be considered to be metapopulations whose evolutionary and genetic dynamics are strongly influenced by river or lake network complexity, as well as by steep physical gradients over short geographic distances (Labonne et al. 2008; Paz-Vinas and Blanchet 2015). The dendritic connectivity of Hong Kong hillstreams is low, and many hillstreams are short, discharging directly to the sea, meaning that the loaches and other freshwater species in different hillstreams are effectively isolated from each other (Tsang et al. 2016; Wong et al. 2017). This isolation traces back to the Pleistocene when the sea levels were 150 m lower and the exposed landmass that extended over 120 km south from the modern coastline (Fyfe et al. 2000). During this time, the hillstreams did not discharge directly to the sea, but joined to several paleodrainages (Fig. 1) allowing connectivity between now isolated loach populations. After these connections were lost, loach populations in different streams would have evolved independently without opportunities for evolutionary or genetic rescue through gene flow (Bell 2017; Whiteley et al. 2015).

This study represents the first step towards understanding the history, population structure, and evolutionary potential of *O. platycephalus* by conducting a population genomic survey of

genetic diversity and differentiation across the territory of Hong Kong. Given what little is known about the biology of this species (mostly its feeding behavior; Dudgeon 1991; 1993), we expected to find a strong population structure attributable to strong genetic drift and lack of gene flow among drainages. We further expected to find low levels of genetic diversity, signs of inbreeding, and hence, reduced adaptive potential in the face of changing environmental conditions. We also assessed the phylogeographic history of the study populations with the aid of the reconstructed geological history of drainage systems in Hong Kong, as well as their contemporary effective population sizes.

## MATERIALS AND METHODS

### Sample collection and DNA extractions

A total of 282 flat-headed loaches were collected from 16 sites in Hong Kong: four sites from East New Territories, six sites from West New Territories, three sites on Hong Kong Island and three sites on Lantau Island, with an average sample size of 18 individuals per population ( $n = 5\text{--}20$ ; Fig. 1, Tables 1 and Supplementary Table S1). The sampling was conducted in January–March 2022 with dip-nets at night when flat-headed loaches are most active. All specimens were euthanized with MS-222 in the field and stored in 95% ethanol maintained at room temperature until genomic DNA was extracted from the pectoral fins using the *DNeasy Blood and Tissue Kit* (Qiagen, Germany) following the manufacturer's recommendations. The DNA concentration for each sample was assessed by

**Table 1.** Estimates of genetic diversity of the 16 flat-headed loach populations.

Population	Paleodrainage system	$N_{IND}$	$H_O$	$H_E$	$Ar$	$\pi_w$	$F_{IS}$	$F_{IS}$ 95%CI Low	$F_{IS}$ 95%CI High
YSO	NTE	5	0.0344	0.0425	1.0398	0.0368	0.1906 <sup>a</sup>	0.1170	0.2057
TKO	NTE	18	0.0558	0.0628	1.0618	0.0595	0.1115 <sup>a</sup>	0.0785	0.1294
TPK	NTE	19	0.0381	0.0446	1.0436	0.0416	0.1457 <sup>a</sup>	0.1058	0.1654
WCC	NTE	19	0.0392	0.0430	1.0423	0.0396	0.0884 <sup>a</sup>	0.0542	0.1037
PKA	LTS	19	0.0345	0.0380	1.0375	0.0361	0.0921 <sup>a</sup>	0.0524	0.1094
MKN	LTS	19	0.0334	0.0349	1.0345	0.0331	0.0430 <sup>a</sup>	0.0061	0.0633
SHW	NTW	15	0.0019	0.0022	1.0020	0.0018	0.1364	−0.1572	0.3088
LTR	NTW	19	0.0627	0.0763	1.0752	0.0745	0.1782 <sup>a</sup>	0.1536	0.1917
TMN	NTW	20	0.0804	0.1072	1.1060	0.1061	0.2500 <sup>a</sup>	0.2314	0.2658
TLC	TMS	19	0.0656	0.0859	1.0849	0.0846	0.2363 <sup>a</sup>	0.2156	0.2551
KLG	TMS	19	0.0521	0.0618	1.0609	0.0600	0.1570 <sup>a</sup>	0.1293	0.1753
TM2	TMS	18	0.0626	0.0782	1.0770	0.0767	0.1995 <sup>a</sup>	0.1743	0.2169
TM1	HKI	19	0.0469	0.0602	1.0584	0.0579	0.2209 <sup>a</sup>	0.1858	0.2386
LFS	HKI	19	0.0536	0.0613	1.0604	0.0591	0.1256 <sup>a</sup>	0.0975	0.1418
GFN	HKI	18	0.0435	0.0483	1.0472	0.0464	0.0994 <sup>a</sup>	0.0641	0.1234
PFL	HKI	17	0.0519	0.0616	1.0608	0.0603	0.1575 <sup>a</sup>	0.1318	0.1764

$N_{IND}$  Number of individuals,  $H_O$  Observed Heterozygosity,  $H_E$  Expected Heterozygosity,  $Ar$  Allelic richness,  $\pi_w$  Average weighted nucleotide diversity,  $F_{IS}$  Inbreeding Coefficient, estimated as  $F_{IS} = 1 - (H_O/H_E)$ .

<sup>a</sup>Indicates statistically significant inbreeding.

NanoDrop One spectrophotometer (Thermo Fisher, US) and diluted to 50 ng/μL for reduced representation library preparations and sequencing.

### DarT sequencing

The 282 loach DNA samples were sequenced at Diversity Arrays Technology Pty Ltd, Australia with DArTseq™ (Diversity Arrays Technology Sequencing Technology) for medium-density sequencing which combines genome complexity reduction methods with next-generation sequencing (Kilian et al. 2012). Libraries were constructed using the DArTseq™ complexity reduction method described here in brief: (i) genomic DNA was digested with two restriction enzymes (*Pst*I and *Mse*I); (ii) DNA fragments were ligated to the barcoded adapters; after which (iii) adapter-ligated fragments were amplified with PCR. DNA libraries were sequenced using Single-Read sequencing runs for 77 cycles (Egea et al. 2017). High-throughput sequencing was implemented with the Illumina HiSeq2500 machine (Illumina, USA) and DArTseq™ marker scoring was performed using DArTsoft (Kilian et al. 2012). Single-nucleotide polymorphisms (SNPs) were scored as “0” = reference allele homozygote, “1” = alternative allele homozygote and “2” = heterozygote (Uba et al. 2021).

### SNP filtering

SNPs called from DArTseq™ were analyzed in R v4.2.1 using package *dartR* v2.7.2 (Gruber et al. 2018; Mijangos et al. 2022) which was developed to access and explore SNP data obtained with DArTseq™. Approximately 20 Gb of DArTseq data (466.78 million raw reads) were generated and calling of sequence variants generated a total of 56,416 SNPs.

The following filtering criteria were applied to exclude SNPs in the given order (steps “i” to “vi” are default settings for SNPs quality controls; details of filtering setting can be found in Supplemental Material): (i) reproducibility < 0.99; (ii) coverage < 5X or > 50X (coverage automatically calculated by *dartR*); (iii) SNPs with sequence tag length < 20 or > 69; (iv) SNP position outside the trimmed sequence tag; (v) secondary SNPs; (vi) pairwise Hamming distance between sequence tags < 0.2; (vii) call rate < 0.7; (viii) minor allele count (MAC) < 3. After these filtering steps, 7045 SNPs were kept for downstream analyses.

### Genetic diversity and population dynamics

The observed heterozygosity ( $H_O$ ), expected heterozygosity ( $H_E$ ), allelic richness ( $Ar$ ), inbreeding coefficient ( $F_{IS}$ ) and 95% CI for  $F_{IS}$  (1000 bootstrap replicates) for each population were calculated with R package *hierfstat* v0.5-11 (Goudet 2005). The nucleotide diversity ( $\pi$ ) for each locus in each population was estimated by *vcftools* v0.1.17 (−site-pi) (Danecek et al.

2011). The average weighted nucleotide diversity ( $\pi_w$ ) of each population was calculated following Konopiński (2023) as:

$$\pi_w = \frac{\sum N_{loci} \pi}{N} \quad (1)$$

where  $N_{loci}$  is the number of loci that do not have all missing data in a population;  $N$  is the number of SNPs in this study ( $N = 7045$ ).

A relatedness statistic of individuals within each population was also calculated by *vcftools* v0.1.17 based on the KING (Kinship-based Inference for Genome-wide association studies) inference (Manichaikul et al. 2010) and the results were plotted by R package *phewas* v1.0.12 (Kolde 2019). According to the KING tutorial, an estimated kinship coefficient range > 0.354, (0.177, 0.354], (0.088, 0.177] and (0.044, 0.088] corresponds to duplicate/twin, first-degree, second-degree, and third-degree relationships, respectively (Manichaikul et al. 2010). Analysis of variance (ANOVA) of relatedness among the 16 populations was conducted in R with  $p$ -value adjusted using the FDR method with the package *agricolae* v1.3-7 (Mendiburu 2019). The contemporary effective population size ( $N_e$ ) of each population was estimated by *NeEstimator* v2.1 based on the Linkage Disequilibrium method with singletons removed (Do et al. 2014). Linear mix-models assessing the effects of geographical factors (altitude, summed length of streams) and relatedness of individuals within each population on  $\pi$  and  $N_e$  were estimated with R package *MCMCglmm* v2.35 (Hadfield 2010) with phylogenetic correlation matrix (converted from neighbor-joining tree generated by *dartR* with default setting; NJ) as a random effect (model: list(G = list(G1 = list(V = 1, nu = 1, alpha.mu = 0, alpha.V = 1000)), R = list(V = 1, nu = 0.002)); nitt = 1001000, thin = 1000, burnin = 10000). The demographic history of each population was estimated with Stairway Plot v2.1.1 (Liu and Fu 2020), based on the folded site-frequency spectrum (SFS) calculated using *dartR*, with a mutation rate of 0.562e-8 mutations/site/generation (estimated for common carp reported by Bergeron et al. 2023) and a generation time of two years. In the existing fish mutation rate studies (Zhang et al. 2023), common carp is the species with the closest phylogenetic relationship to *O. platycephalus*. Since mutation rates in fish are not highly variable (Zhang et al. 2023) it is reasonable to assume that the common carp estimate is a good proxy for *O. platycephalus* mutation rate. As the generation time of *O. platycephalus* is unknown, inferred generation time of *Nemacheilus triangularis* (2.2 years; <https://fishbase.mnhn.fr/summary/Nemacheilus-triangularis>), a closely related species from the same subfamily (Nemacheilidae), was used.

### Population structure

Pairwise  $F_{ST}$  values between populations were estimated with *dartR* based on the implementation in the *StAMPP* package (Pembleton et al. 2013) with



1000 bootstrap replicates and a heatmap showing pairwise  $F_{ST}$  results was plotted with an R package *heatmap*. To see whether levels of differentiation as measured by  $F_{ST}$  differed within and among each paleodrainage, we compared mean  $F_{ST}$  values using *t*-tests. The neighbor-joining (NJ) tree was constructed based on  $F_{ST}$  values obtained with *dartR* using default settings. Population structure was also assessed using a maximum likelihood estimation of individual ancestries by *ADMIXTURE* v1.3.0 (Alexander et al. 2009). Assuming *K* ranging from 1 to 20, cross-validation information of each *K* was collected and the *K* with the lowest cross-validation error was chosen to plot population ancestries. Principal Component Analysis (PCA) was conducted by *dartR* (nfactors = 2) to assess population structure. A map of loach sampling localities with PCA results was drawn by *Adobe Illustrator 2022* (Adobe Inc., US). Isolation by distance (IBD) analyses (geographic distance calculated by coordinates and by shortest possible waterway) based on the Mantel test were performed with *dartR* using pairwise  $F_{ST}$  values with Rousset's correction (Rousset 1997) as genetic distances. The shortest possible waterway between each population pair was estimated based on the map of paleodrainage systems (Fig. 1). If the populations were within the same drainage, the length of the stream connecting the two populations was measured. If the populations were in different drainage systems, the shortest waterway through the sea was measured as a proxy of "paleodistance". The latter measure was adopted as paleodrainage reconstructions were not possible for all drainages (and between drainages) because reconstructions did not extend to cover all relevant areas (cf. Fig. 1).

To compare the level of genetic differentiation among loach populations to that of other freshwater fishes, we compiled data from recent (2013–2023; see Supplementary Table S6 and Supplementary Material) studies which have estimated  $F_{ST}$  among freshwater fish populations using reduced-representation genome sequencing approaches. We found a total of 27 studies (See Supplementary Table S7 and Supplementary Material for details) and extracted the mean pairwise  $F_{ST}$  estimate from each of them. We further estimated the geographic coverage in each of these studies by measuring the area (in km<sup>2</sup>) their sampling covered (i.e. the area contained within the perimeter drawn around furthest sampling localities). Plotting the mean  $F_{ST}$  from each study as a function of the geographic extent of sampling allowed us to explore how the loach  $F_{ST}$  value compared to other fishes. Linear models assessing  $F_{ST}$  as a function of (log) geographic extent of sampling were fitted with R v4.2.1.

## RESULTS

### Genetic diversity and effective population size

Expected heterozygosity ( $H_E$ ) ranged from 0.0022 (SHW) to 0.1072 (TMN) with an average of 0.0568 (Table 1), whereas observed heterozygosity ( $H_O$ ) ranged from low values in SHW (0.0019), MKN (0.0334), and YSO (0.0344) to somewhat higher (but still low) values in LTR (0.0627), TLC (0.0656), and TMN (0.0804; Table 1). The highest average weighted nucleotide diversity ( $\pi_w$ ) was found in population TMN (0.1061) while the lowest nucleotide diversity was found in SHW (0.0018) with an average of 0.0546 (Table 1). Similarly to average weighted nucleotide diversity, allelic richness ( $A_r$ ) values were very low (Table 1). The highest allelic richness ( $A_r$ ) was found in TMN (1.1060) while the lowest was found in SHW (1.0020) with an average of 1.0558 (Table 1). The inbreeding coefficient ( $F_{IS}$ ) was significantly positive in every population, except in SHW where the lower 95% CI was negative, indicating that nearly all of the populations were inbred (Table 1).

Kinship analyses further corroborated the results of genetic diversity analyses: there were significant differences in average relatedness among populations (ANOVA,  $F_{15,2423} = 136.7$ ,  $p < 2e-16$ ; Fig. 2; Supplementary Table S2 and Supplementary Fig. S1) and mean relatedness of individuals within each population correlated negatively with nucleotide diversity (Posterior mean = -0.293,  $pMCMC < 0.001$ ) according to the mixed linear model (Supplementary Table S3). The SHW population had the highest mean relatedness with every individual being related to each other (and mostly sharing ~50% of their genetic information; Fig. 2), whereas all individuals in TM1 and YSO populations were unrelated (Fig. 2).

The contemporary effective population size ( $N_e$ ) estimates were obtained for 15 populations ranging from 10.2 (GFN) to 129.8 (TM1) while estimation for YSO failed likely due to limited

sampling size ( $N_{IND} = 5$ ; Fig. 3). In addition, for SHW, *NeEstimator* failed to obtain the upper limit of 95% Parametric CI possibly because of the limited number of non-singleton loci ( $N_{ns} = 84$ ) used in the estimation (Fig. 3 and Supplementary Table S4). A linear mixed-model was fitted to assess the effects of several variables on  $N_e$ , which revealed that the altitude of the study site was positively associated with  $N_e$  (Posterior mean = 0.101,  $pMCMC = 0.042$ , see in Supplementary Table S3).

### Demographic history

We used a folded SNP frequency spectrum (SFS) model in the *Stairway Plot* to infer the demographic histories of the 16 populations. Almost every population (except SHW) appeared to have experienced at least one population size contraction before the first glaciation event inferred from the local sedimentary record (Fig. 4). During and after glaciation, only populations (PKA and MKN) at south Lantau drainage expanded after being bottlenecked while most (10 out of 16) of the populations continued to decline or appeared to have remained stable (Fig. 4).

### Genetic population structure

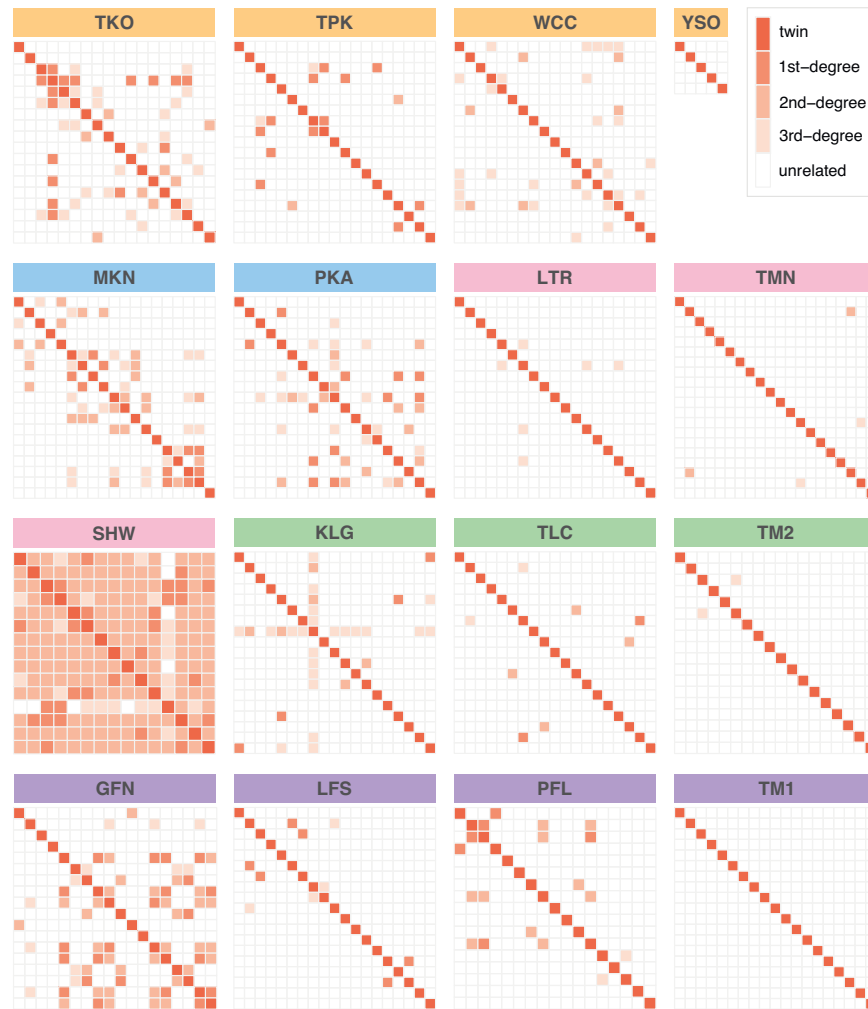
Pairwise  $F_{ST}$  among 16 populations ranged from 0.110 (MKN - PKA) to 0.942 (YSO - SHW; Fig. 5A, Supplementary Table S5) with an average of 0.668. All pairwise estimates were highly significant ( $p < 0.001$ ; Supplementary Table S5). Population SHW showed the highest genetic differentiation from other populations, with an average  $F_{ST}$  of 0.825, followed by WCC with an average  $F_{ST}$  of 0.743. Population TMN exhibited the least genetic differentiation with an average  $F_{ST}$  of 0.372 to the other populations (Fig. 5A). There was nonetheless a clear, albeit relatively weak, pattern of isolation by distance ( $p < 0.05$ ; Fig. 6). *t*-tests of  $F_{ST}$  within and among each paleodrainage system showed that populations originating from New Territories (NTW and NTE drainages) show equal level of differentiation within and among drainage systems while those from other drainages were more similar within the same drainage system (Supplementary Fig. S2). Admixture analyses corroborated the high  $F_{ST}$  values at *K* = 16 (Fig. 5B and Supplementary Fig. S3), none of the clusters exhibited evidence of among population admixture, except that PKA and MKN clustered together (Fig. 5B). Moreover, TKO consisted of two components, indicating that it could be a mixture of two ancestral populations (Fig. 5B).

Plotting the  $F_{ST}$  across the 27 studies (excluding the current study) against geographic extent of sampling shows that the degree of population differentiation increases as a function of geographic area sampled (Fig. 8; linear model:  $F_{1,25} = 4.538$ ,  $p = 0.043$ ). However, when the current study was included to the linear model, the relationship was no longer significant ( $F_{1,26} = 1.054$ ,  $p = 0.314$ ; See Supplementary Table S7 and Supplemental Material for details), indicating that the mean  $F_{ST}$  value of *O. platycephalus* was an outlier (Fig. 8).

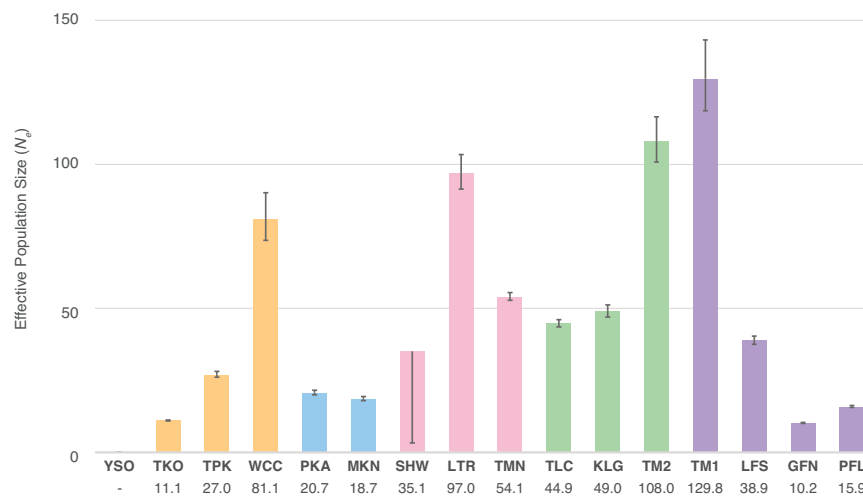
Principal component analysis revealed the clearest distinction at Axis 1 between the two southern Lantau populations (PKA and MKN) and the rest of the populations (Fig. 7). Axis 2 further separated the East New Territories populations (TPK, TKO, WCC and YSO) from the rest (Fig. 7). There was also a clear pattern for populations from the same paleodrainage systems (Fig. 1) to cluster together: Tai Mo Shan system (TMS) - TM2, KLG and TLC; West New Territories system (NTW) - SHW, LTR and TMN; Hong Kong Island system (HKI) - PFL, GFN and LFS; South Lantau system (LTS) - PKA and MKN; East New Territories system (NTE) - TKO, TPK, WCC and YSO (Fig. 7; see Supplementary Fig. S4 for PCA plot excluding PKA and MKN populations). However, TM1 clustered with NTW populations although it was indicated to be part of the HKI drainage system (Fig. 1).

## DISCUSSION

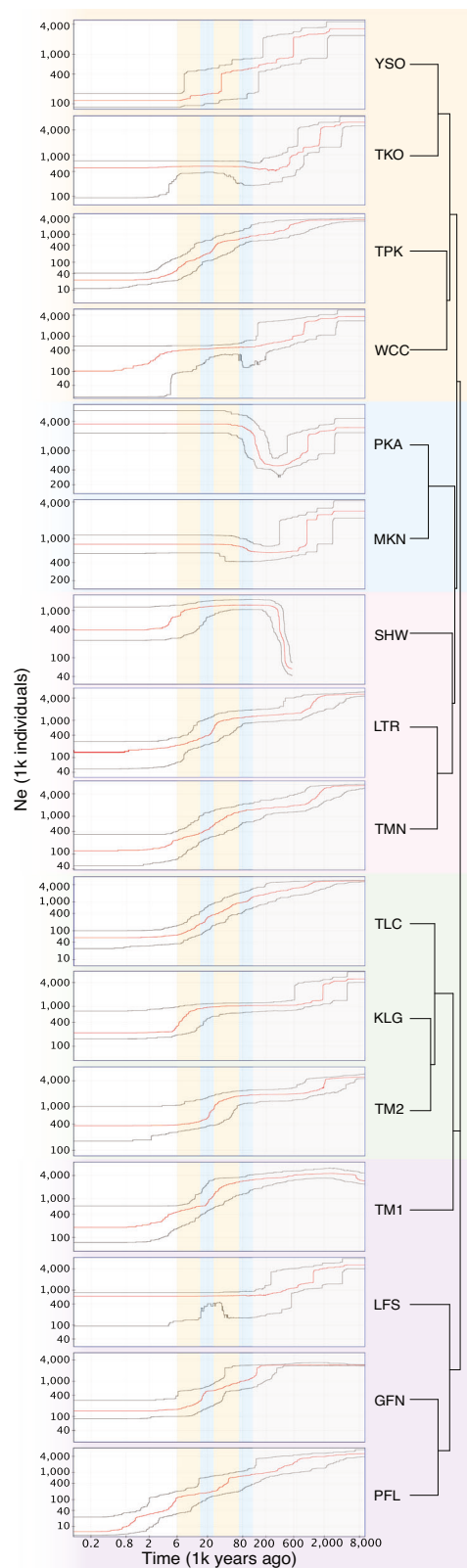
As we expected, *O. platycephalus* populations were genetically highly structured, low in genetic diversity and exhibited very low



**Fig. 2** Relatedness between individuals within each of the 16 flat-headed loach populations. The different color boxes represent populations in different paleodrainage systems (see Fig. 1). Each row and column corresponds to a particular individual in a given population.



**Fig. 3** Estimates of contemporary effective population size ( $N_e$ ) in 16 loach populations as obtained with the program *NeEstimator* (Do et al. 2014). Vertical lines depict 95% Parametric CIs. CIs and  $N_e$  for SHW and YSO, respectively, were not estimable. The different colors represent populations in different paleodrainage systems (see Fig. 1).



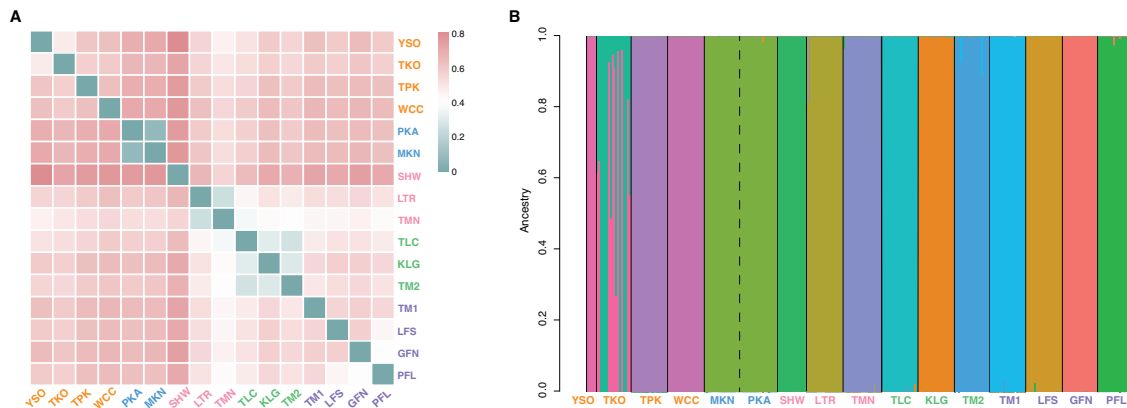
**Fig. 4** The demographic history of 16 flat-headed loach populations inferred by a folded SFS model projected to NJ tree. Different colors represent different paleodrainage systems (see Fig. 1). Red lines: the median estimated historical effective population sizes. Gray lines: 95% confidence interval of the inference. The periods of glaciation and interglaciation are highlighted with blue and orange vertical bars, respectively. The gray vertical area depicts a gap in the sedimentary record of Hong Kong.

contemporary effective population sizes. Not surprisingly, all populations displayed positive inbreeding coefficients, and many showed signs of inbreeding as indicated by high relatedness estimators. Hence, all populations may be at risk of losing their adaptive potential, since  $N_e$  for populations to retain evolutionary potential is estimated to range from 500 to 1000 (Franklin and Frankham 1998; Frankham et al. 2014b), and estimates of less than 50 indicate that populations are in immediate danger (Hoban et al. 2020).

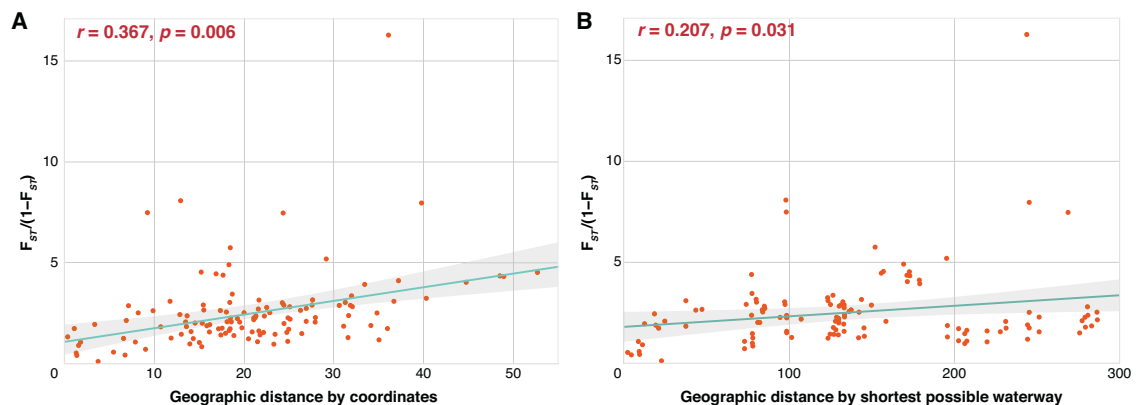
Low level of within-population genetic diversity is typical of freshwater fishes (DeWoody and Avise 2000; Martinez et al. 2018; Ward et al. 1994), and different measures of genetic diversity indicated low, but not exceptionally low genetic diversity within *O. platycephalus* populations. For instance, nucleotide diversities in isolated populations of three-spined stickleback (*Gasterosteus aculeatus*; Coll-Costa et al. 2024) and nine-spined sticklebacks (*Pungitius pungitius*; Kivikoski et al. 2023) are typically much lower than recorded for *O. platycephalus*. However, some caution is needed as different methods and filtering criteria have been used in different studies so genetic diversity estimates may not be strictly comparable (Korunes and Samuk 2021; Konopiński 2023). Nevertheless, loach average weighted nucleotide diversity (mean  $\pi_w = 0.0546$ ) was slightly lower than that of an isolated and near-threatened wels catfish (*Silurus glanis*;  $\pi = 0.0690$ ) population (Palm et al. 2019; Littmann 2022) and a highly genetically structured salmonid fish (*Salvelinus fontinalis*;  $\pi = 0.065$ ; Ferchaud et al. 2020). Likewise, nucleotide diversity in *Achondrostoma salmantinum*, an endangered freshwater cyprinid endemic to Spain, was much higher ( $\pi$  range = 0.133–0.250; Corral-Lou et al. 2021) than in *O. platycephalus*. In the same vein, *Teleogramma*, a small clade of rheophilic cichlids from the Congo River showed similarly high levels of genetic diversity (mean  $\pi$  for each species ranging from 0.067 to 0.155; Alter et al. 2017). All this suggests comparably low levels of genetic diversity in *O. platycephalus* and this inference is backed up by the extremely low estimates of loach effective population sizes. Direct comparisons to earlier genetic studies of Hong Kong freshwater fishes are difficult because they were based on mitochondrial DNA (Wong et al. 2017; Wong et al. 2019; Wu et al. 2019) which has  $\frac{1}{4}$  of that of nuclear markers (Birky et al. 1983), or a handful of microsatellite markers (Wu et al. 2016) which have higher mutation rates than nuclear single nucleotide polymorphisms.

Effective population size is a fundamental parameter in the conservation context as its magnitude indicates the amount of genetic drift and inbreeding taking place (Charlesworth 2009). *O. platycephalus* populations with extremely low contemporary effective population sizes (cf. estimates in Palstra and Fraser 2012; Palstra and Ruzzante 2008) are thereby exposed to further erosion of genetic variability and greater risk of fixation of deleterious mutations. With contemporary effective population sizes below 50, it is likely that selection against deleterious mutations will be ineffective (Brandvain and Wright 2016; Lynch et al. 1995; Robinson et al. 2023) allowing even highly detrimental mutations to become fixed in loach populations. Unfortunately, there is no reference genome for *O. platycephalus* and due to the nature of the reduced representation approach used in this study, we cannot estimate the loads of detrimental mutations in the loach populations. Hence, an obvious and interesting question for follow-up studies would be to characterize the putative deleterious variation in these populations to assess the genomic consequences of prolonged population declines once genomic resources for this become available.

Another important consequence of low contemporary effective population size of *O. platycephalus* populations is their compromised ability to adapt to changing environmental conditions. Although our understanding of likelihood of adaptation to climate change is still limited (Merilä and Hendry 2014; Merilä and Lv 2024), low levels of genetic variability and low effective sizes of *O. platycephalus* populations suggest that they would have limited



**Fig. 5 Population structure among sixteen flat-headed loach populations.** The different font colors represent populations in different paleodrainage systems (see Fig. 1). **A** Pairwise  $F_{ST}$  estimates among populations. **B** Population structure inferred from admixture analysis assigning individuals to the optimal  $K = 16$  (See Supplementary Fig. S3). Each colored bar represents one individual, and colored segments correspond to different ancestral components.



**Fig. 6 Isolation by distance across flat-headed loach populations.** Isolation by distance with **A** geographic distance (km) calculated from site coordinates ( $r = 0.367$ ,  $p = 0.006$ ); and **B** from the shortest possible waterway based on the paleodrainage system map ( $r = 0.207$ ,  $p = 0.031$ ). Pairwise  $F_{ST}$  in both plots obtained with Rousset's (1997) correction.

potential to adapt to sudden changes in conditions within their current range-limited habitats. In the context of a warming climate, these fishes are confined to their current habitats with no refuges to occupy further upstream. Interestingly, we found that populations at higher altitudes had larger contemporary effective population sizes than those from lower altitudes. This may be a reflection of the fact that the species is better adapted to lower-order upstream habitats of hillstreams and has been able to sustain larger effective sizes in these locations than in higher-order reaches further downstream. Such adaptation might be related to both abiotic (e.g. lower temperature) and biotic factors (e.g. a lack of competition with other fish species), but further studies would be needed to resolve this matter.

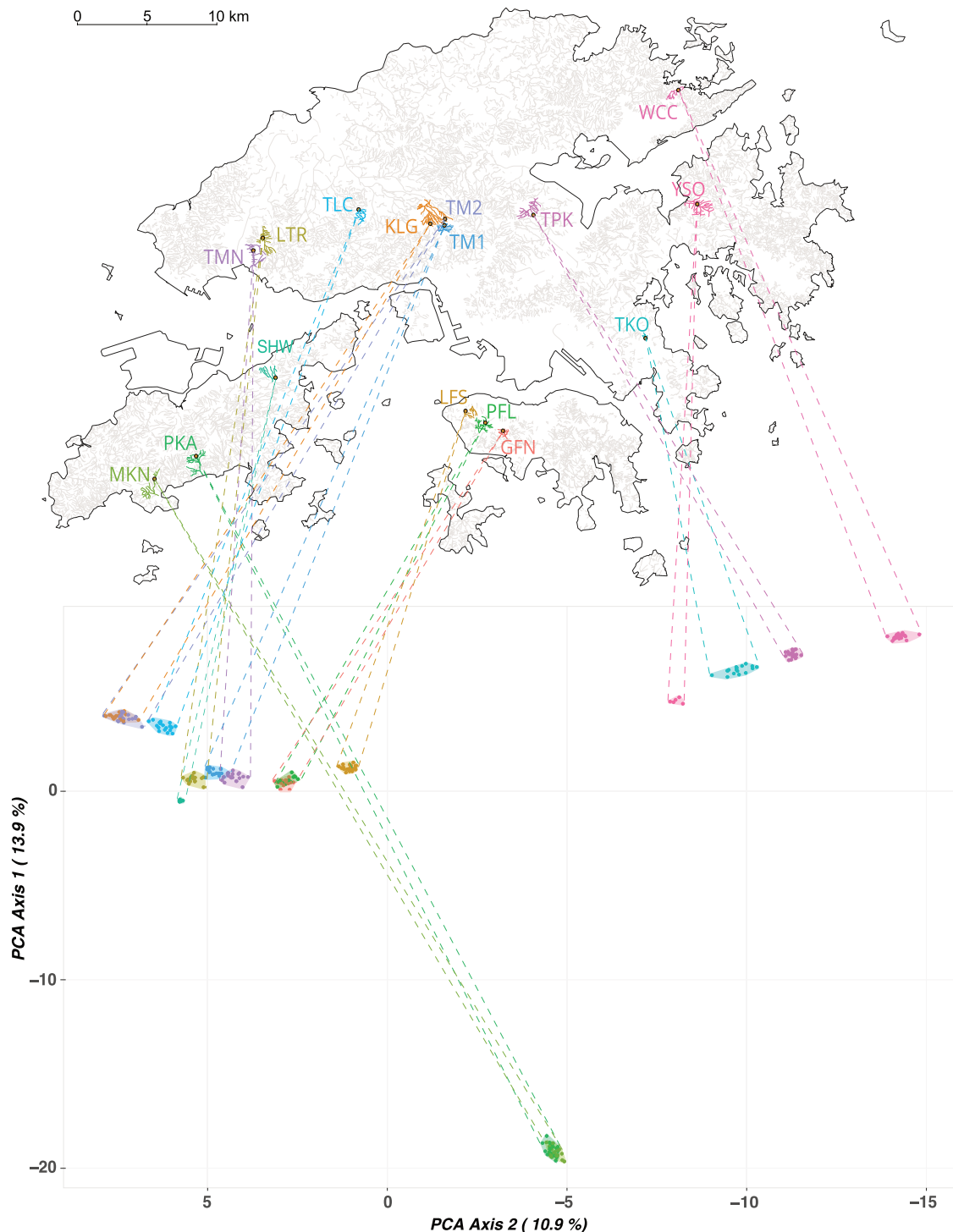
Given the clear footprint of strong genetic drift in *O. platycephalus* populations, it was not surprising that the degree of genetic differentiation among them was high with an average  $F_{ST}$  of 0.668. Although freshwater fish populations are typically more genetically structured than their marine counterparts (Ward et al. 1994), to the best of our knowledge, strong population structuring comparable to that observed in *O. platycephalus* has not been reported for any freshwater fish species from such a small geographic area (Fig. 8 and Supplementary Table S6). However, high levels of population differentiation have been reported among pond populations of nine-spined sticklebacks (*Pungitius pungitius*) which are characterized by strict isolation, with average  $F_{ST} = 0.49$  (Shikano et al. 2010) comparable to - but slightly lower than - *O. platycephalus*. Similarly, two populations of

a landlocked goby *Rhinogobius* sp. in Japan had an  $F_{ST}$  of 0.53 (Ohara et al. 2005). However, both these studies were based on microsatellite markers. We compiled data from published studies that used genome sequencing approaches and found that the degree of genetic differentiation among *O. platycephalus* populations was indeed exceptionally high given the limited geographic extent of sampling within Hong Kong.

Relatively high  $F_{ST}$  values have also been reported for two other Hong Kong loaches that inhabit larger streams than *O. platycephalus*, although these estimates are not directly comparable as they were based on variation in mitochondrial control region (average  $F_{ST} = 0.66$  in *Schistura fasciolata* and  $F_{ST} = 0.88$  in *Pseudogastromyzon myersi*; Wong et al. 2017). Similarly, Wu et al. (2016) reported  $F_{ST} = 0.57$  estimated using five microsatellite markers in the Hong Kong goby *Rhinogobius duospilus* indicating that, as has been noted for shrimps in the same habitats (Tsang et al. 2016; Yam and Dudgeon 2005), the limited connectivity of Hong Kong streams has contributed to high genetic divergence in their inhabitants.

Our reconstructions of historical demography revealed that nearly all *O. platycephalus* populations experienced population size declines before and during the last glaciation. Most but not all populations had also experienced continued population declines since then. These patterns accord with the likelihood that the local populations became isolated following the post-glacial sea level rise (Wong et al. 2017; Wu et al. 2016). The fact that some populations did not exhibit clear downward trends might be





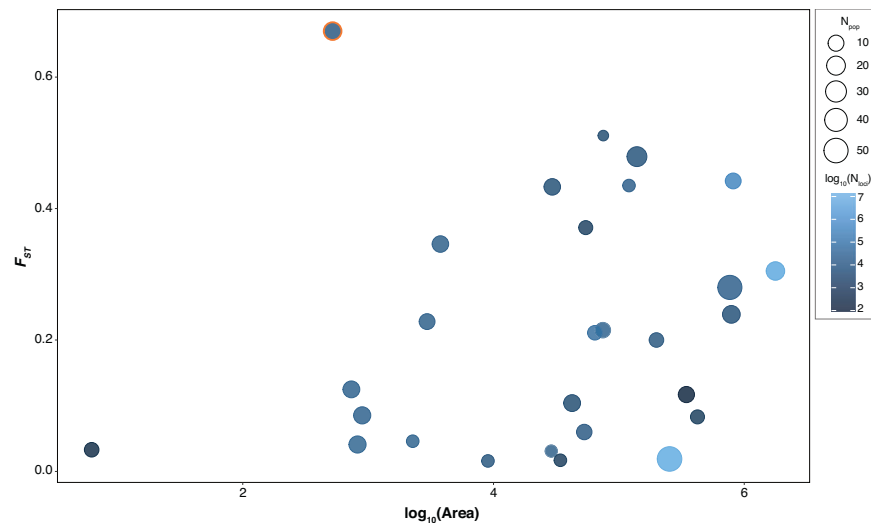
**Fig. 7** Results of principal component analysis (PCA) of allele frequencies projected onto the map of study area. Note that PCA Axis 2 is depicted on the x-axis.

ascribable to methodological problems: the approaches used in historical demography reconstructions tend to lose resolution in modern times especially if sample sizes are low as in our study (Li and Durbin 2011; Liu and Fu 2015; Nadachowska-Brzyska et al. 2022). Some support for this conjecture is provided by the observation that the three populations that did not show downward trends were among those with the smallest contemporary effective population sizes. Assuming an  $N_e/N_c$  ratio of 10% (Frankham 1995; Hoban et al. 2020), our data suggests that

contemporary population sizes ( $N_c$ ) of *O. platycephalus* range from ca. 100 to 1300 individuals. These are very small numbers when compared to many other species (cf. Palstra and Fraser 2012).

Given the findings of this study, what is the outlook for loach populations in Hong Kong hillstreams? Given that most populations have been subject to strong genetic drift and loss of genetic diversity, as well as the predicted further loss of diversity and increased risk of inbreeding depression in populations comprising fewer than 50 individuals (Franklin and Frankham 1998; Hoban





**Fig. 8** Bubble plot illustrating the relationship between  $F_{ST}$  and extent of geographic sampling area for studies of freshwater fishes published between 2013 and 2023 using single nucleotide polymorphism loci (SNPs). Bubble size indicates the number of populations and color indicates the number of SNP loci after  $\log_{10}$  transformation. The  $F_{ST}$  estimate for loaches from the current study is highlighted with an orange circle. For details of the underlying data, see Supplementary Table S6 and Supplementary Material. For linear models fitted into this data, see Supplementary Table S7.

et al. 2020), most (10/16) of the study populations can be expected undergo further loss of variability. Perhaps more seriously, none of the study populations appear to have an effective size large enough to permit adaptation to climate warming or other environmental changes: even the two populations with the highest contemporary  $N_e$  are far from the  $N_e = 500$  mark considered necessary to preserve evolutionary potential, and much smaller than the more conservative  $N_e = 1000$  benchmark (Frankham et al. 2014a). From this perspective, *O. platycephalus* populations in Hong Kong are at high risk of local extinctions, and may even represent an existing extinction debt.

It is interesting to speculate that extinctions have already taken place in hillstreams that we visited for this study where we were unable to find any loaches. Future studies might be able to address this possibility by analyzing environmental DNA from stream sediment cores (e.g. Nelson-Chorney et al. 2019). Possible conservation measures for declining populations might also consider the reciprocal translocation of loaches between streams to enhance genetic diversity, reduce the negative consequences of inbreeding and enhance their adaptive potential in the face of future environmental changes. The two populations (MKN and PKA) in southern Lantau, which are the sole representatives of populations from the South Lantau drainage, should be prioritized for conservation due to their very small contemporary effective population sizes ( $\leq 20$ ) and genetic distinctiveness from all other studied populations. Secondary priorities would be populations on Hong Kong Island (e.g. GFN and PFL) followed by those in the East New Territories (e.g. TKO and TPK).

In conclusion, the picture emerging from our analyses is that local flat-headed loach populations of Hong Kong became isolated after post-glacial sea level rise and have been subject to strong genetic drift in the subsequent absence of connectivity at the metapopulation level, i.e. a natural consequence of epoch rollover. As a consequence, the local populations have experienced loss of genetic diversity and increased genetic differentiation as manifested in extremely high  $F_{ST}$  estimates even over very short geographical distances. The low estimates of contemporary  $N_e$  suggest that local populations have severely compromised ability to remain viable with lowered evolutionary potential to meet the demands of a changing environment. Future studies based on whole genome resequencing data can refine this picture and give deeper insights into the levels (e.g. Robinson et al. 2019)

and consequences (e.g. Keller and Waller 2002) of inbreeding in these populations. We further suggest that the *O. platycephalus* populations in Hong Kong can provide an excellent study system to document whether lowered genetic diversity is associated with increased extinction rate or not - a topic still controversial in conservation genetics (García-Dorado and Caballero 2021; DeWoody et al. 2021; Teixeira and Huber 2021; Willi et al. 2022).

#### DATA AVAILABILITY

DART-seq SNP genotyping datasets (SNPs report and vcf files) are openly available from Figshare (<https://doi.org/10.6084/m9.figshare.24792567>).

#### REFERENCES

- Albano PG, Steger J, Bošnjak M, Dunne B, Guifarro Z, Turapova E et al. (2021) Native biodiversity collapse in the eastern Mediterranean. *Proc R Soc B* 288(1942):20202469
- Alexander DH, Novembre J, Lange K (2009) Fast model-based estimation of ancestry in unrelated individuals. *Genome Res* 19(9):1655–1664
- Alter SE, Munshi-South J, Stiasny MLJ (2017) Genomewide SNP data reveal cryptic phylogeographic structure and microallopatric divergence in a rapids-adapted clade of cichlids from the Congo River. *Mol Ecol* 26(5):1401–1419
- Bell G (2017) Evolutionary rescue. *Annu Rev Ecol Syst* 48:605–627
- Bergeron LA, Besenbacher S, Zheng J, Li P, Bertelsen MF, Quintard B et al. (2023) Evolution of the germline mutation rate across vertebrates. *Nature* 615(7951):285–291
- Birky Jr CW, Maruyama T, Fuerst P (1983) An approach to population and evolutionary genetic theory for genes in mitochondria and chloroplasts, and some results. *Genetics* 103(3):513–527
- Brandvain Y, Wright SI (2016) The limits of natural selection in a nonequilibrium world. *Trends Genet* 32(4):201–210
- Cerini F, Childs DZ, Clements CF (2023) A predictive timeline of wildlife population collapse. *Nat Ecol Evol* 7(3):320–331
- Charlesworth B (2009) Effective population size and patterns of molecular evolution and variation. *Nat Rev Genet* 10(3):195–205
- Coll-Costa C, Dahms C, Kemppainen P, Alexandre CM, Ribeiro F, Zanella D et al. (2024) Parallel evolution despite low genetic diversity in three-spined sticklebacks. *Proc R Soc B* 291:20232617
- Corral-Lou A, Perea S, Doadrio I (2021) High genetic differentiation in the endemic and endangered freshwater fish *Achondrostoma salmantinum* Doadrio and Elvira 2007 from Spain, as revealed by mitochondrial and SNP markers. *Conserv Genet* 22(4):585–600
- Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA et al. (2011) The variant call format and VCFtools. *Bioinformatics* 27(15):2156–2158

- DeWoody J, Avise J (2000) Microsatellite variation in marine, freshwater and anadromous fishes compared with other animals. *J Fish Biol* 56(3):461–473
- DeWoody JA, Harder AM, Mathur S, Willoughby JR (2021) The long-standing significance of genetic diversity in conservation. *Mol Ecol* 30(17):4147–4154
- Do C, Waples RS, Peel D, Macbeth GM, Tillett BJ, Ovenden JR (2014) NeEstimator v2: re-implementation of software for the estimation of contemporary effective population size ( $N_e$ ) from genetic data. *Mol Ecol Resour* 14(1):209–214
- Du LN, Chen XY, Yang JX (2008) A review of the Nemacheilinae genus *Oreonectes* Günther with descriptions of two new species (Teleostei: Balitoridae). *Zootaxa* 1729(1):23–36
- Dudgeon D (1991) An experimental study of abiotic disturbance effects on community structure and function in a tropical stream. *Archiv für Hydrobiologie* 122(4):403–420
- Dudgeon D (1993) The effects of spate-induced disturbance, predation and environmental complexity on macroinvertebrates in a tropical stream. *Freshw Biol* 30(2):189–197
- Dudgeon D (2003) Hillstreams - Hong Kong Field Guides. Wan Li Book, Co. Ltd, Hong Kong
- Egea LA, Mérida-García R, Kilian A, Hernandez P, Dorado G (2017) Assessment of genetic diversity and structure of large garlic (*Allium sativum*) germplasm bank, by diversity arrays technology “genotyping-by-sequencing” platform (DARtseq). *Front Genet* 8:98
- Ferchaud AL, Leitwein M, Laporte M, Boivin-Delisle D, Bougas B, Hernandez C et al. (2020) Adaptive and maladaptive genetic diversity in small populations: Insights from the brook charr (*Salvelinus fontinalis*) case study. *Mol Ecol* 29(18):3429–3445
- Fraimout A, Rastas P, Lv L, Merilä J (2023) Inbreeding depression in an outbred stickleback population. *Mol Ecol* 32(13):3440–3449
- Frankham R, Bradshaw CJ, Brook BW (2014b) Genetics in conservation management: revised recommendations for the 50/500 rules, Red List criteria and population viability analyses. *Biol Conserv* 170:56–63
- Frankham R, Bradshaw CJ, Brook BW (2014a) 50/500 rules need upward revision to 100/1000 – Response to Franklin et al. *Biol Conserv* 176:286–286
- Frankham R (1995) Effective population size/adult population size ratios in wildlife: a review. *Genet Res* 66(2):95–107
- Franklin I, Frankham R (1998) How large must populations be to retain evolutionary potential? *Anim Conserv* 1(1):69–70
- Fyfe JA, Selby IC, Plater AJ, Wright MR (1999) Erosion and sedimentation associated with the last sea level rise offshore Hong Kong, South China Sea. *Quat Int* 55(1):93–100
- Fyfe JA, Shaw R, Campbell SDG, Lai KW, Kirk PA (2000) The Quaternary Geology of Hong Kong. Hong Kong Geological Survey, Geotechnical Engineering Office, Civil Engineering Department, Government of the Hong Kong SAR
- García-Dorado A, Caballero A (2021) Neutral genetic diversity as a useful tool for conservation biology. *Cons Genet* 22(4):541–545
- Gienapp P, Teplitsky C, Alho JS, Mills JA, Merilä J (2008) Climate change and evolution: disentangling environmental and genetic responses. *Mol Ecol* 17(1):167–178
- Goudet J (2005) Hierfstat, a package for R to compute and test hierarchical F-statistics. *Mol Ecol Notes* 5(1):184–186
- Grossen C, Guillaume F, Keller LF, Croll D (2020) Purging of highly deleterious mutations through severe bottlenecks in *Alpine ibex*. *Nat Commun* 11(1):1001
- Gruber B, Unmack PJ, Berry OF, Georges A (2018) darr: An R package to facilitate analysis of SNP data generated from reduced representation genome sequencing. *Mol Ecol Resour* 18(3):691–699
- Hadfield JD (2010) MCMC methods for multi-response generalized linear mixed models: the MCMCglmm R package. *J Stat Softw* 33:1–22
- Hoban S, Bruford M, Jackson JDU, Lopes-Fernandes M, Heuertz M, Hohenlohe PA et al. (2020) Genetic diversity targets and indicators in the CBD post-2020 Global Biodiversity Framework must be improved. *Biol Conserv* 248:108654
- Keller LF, Waller DM (2002) Inbreeding effects in wild populations. *Trends Ecol Evol* 17(5):230–241
- Kilian A, Wenzl P, Huttner E, Carling J, Xia L, Blois H et al. (2012) Diversity arrays technology: a generic genome profiling technology on open platforms. *Methods Mol Biol* 888:67–89
- Kivikoski M, Feng X, Löytynoja A, Momigliano P, Merilä J (2023) Determinants of genetic diversity in sticklebacks. *bioRxiv*, <https://www.biorxiv.org/content/10.1101/2023.03.17.533073v1>
- Kolde R (2019) pheatmap: Pretty Heatmaps. R package version 1(2): 726. <https://cran.r-project.org/package=pheatmap>
- Konopiński MK (2023) Average weighted nucleotide diversity is more precise than pi in estimating the true value of  $\pi$  from sequence sets containing missing data. *Mol Ecol Resour* 23(2):348–354
- Korunes KL, Samuk K (2021) pixy: Unbiased estimation of nucleotide diversity and divergence in the presence of missing data. *Mol Ecol Resour* 21(4):1359–1368
- Labonne J, Ravigné V, Parisi B, Gaucherel C (2008) Linking dendritic network structures to population demogenetics: the downside of connectivity. *Oikos* 117(10):1479–1490
- Lanfear R, Kokko H, Eyre-Walker A (2014) Population size and the rate of evolution. *Trends Ecol Evol* 29(1):33–41
- Li H, Durbin R (2011) Inference of human population history from individual whole-genome sequences. *Nature* 475(7357):493–496
- Littmann L (2022) A recipe for fish and SNPs: Filling the blanks for conservation genomics of Swedish wels catfish (*Silurus glanis*) populations. [Student thesis, DiVA. <http://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-488414>]
- Liu X, Fu YX (2015) Exploring population size changes using SNP frequency spectra. *Nat Genet* 47(5):555–559
- Liu X, Fu YX (2020) Stairway Plot 2: demographic history inference with folded SNP frequency spectra. *Genome Biol* 21(1):1–9
- Lynch M, Conery J, Burger R (1995) Mutation accumulation and the extinction of small populations. *Am Nat* 146(4):489–518
- Lynch M, Walsh B (1998) Genetics and analysis of quantitative traits (Vol. 1). Sinauer, Sunderland, MA
- Manichaikul A, Mychaleckyj JC, Rich SS, Daly K, Sale M, Chen WM (2010) Robust relationship inference in genome-wide association studies. *Bioinformatics* 26(22):2867–2873
- Martinez AS, Willoughby JR, Christie MR (2018) Genetic diversity in fishes is influenced by habitat type and life-history variation. *Ecol Evol* 8(23):12022–12031
- Mendiburu F (2019) Agricolae: statistical procedures for agricultural research. R Package version 1(3):1. <https://cran.r-project.org/package=agricolae>
- Merilä J, Hendry AP (2014) Climate change, adaptation, and phenotypic plasticity: the problem and the evidence. *Evol Appl* 7(1):1–14
- Merilä J, Lv L (2024) Life history traits and climate change. In: Segoli M, Wanberg E (eds.), *Life History Evolution: Traits, Interactions, and Applications*. Wiley, Hoboken, New Jersey
- Mijangos JL, Gruber B, Berry O, Pacioni C, Georges A (2022) dartR v2: An accessible genetic analysis platform for conservation, ecology and agriculture. *Methods Ecol Evol* 13(10):2150–2158
- Nadachowska-Brzyska K, Konczal M, Babik W (2022) Navigating the temporal continuum of effective population size. *Methods Ecol Evol* 13(1):22–41
- Nelson-Chorney HT, Davis CS, Poesch MS, Vinebrooke RD, Carli CM, Taylor MK (2019) Environmental DNA in lake sediment reveals biogeography of native genetic diversity. *Front Ecol Environ* 17(6):313–318
- Ohara K, Takagi M, Hirashima K (2005) Genetic diversity and divergence of the endangered freshwater goby *Rhinogobius* sp. BB in Okinawa Island. *Ichthyol Res* 52:306–310
- Palm S, Vinterstare J, Nathanson JE, Triantafyllidis A, Petersson E (2019) Reduced genetic diversity and low effective size in peripheral northern European catfish *Silurus glanis* populations. *J Fish Biol* 95(6):1407–1421
- Palstra FP, Ruzzante DE (2008) Genetic estimates of contemporary effective population size: what can they tell us about the importance of genetic stochasticity for wild population persistence? *Mol Ecol* 17(15):3428–3447
- Palstra FP, Fraser DJ (2012) Effective/census population size ratio estimation: a compendium and appraisal. *Ecol Evol* 2(9):2357–2365
- Paz-Vinas I, Blanchet S (2015) Dendritic connectivity shapes spatial patterns of genetic diversity: A simulation-based study. *J Evol Biol* 28(4):986–994
- PecI GT, Araújo MB, Bell JD, Blanchard J, Bonebrake TC, Chen IC et al. (2017) Biodiversity redistribution under climate change: Impacts on ecosystems and human well-being. *Science* 355(6332):eaai9214
- Pembleton LW, Cogan NO, Forster JW (2013) StAMPP: an R package for calculation of genetic differentiation and structure of mixed-ploidy level populations. *Mol Ecol Resour* 13(5):946–952
- Robinson J, Kyriazis CC, Yuan SC, Lohmueller KE (2023) Deleterious variation in natural populations and implications for conservation genetics. *Annu Rev Anim Biosci* 11:93–114
- Robinson JA, Räikkönen J, Vucetich LM, Vucetich JA, Peterson RO, Lohmueller KE et al. (2019) Genomic signatures of extensive inbreeding in Isle Royale wolves, a population on the threshold of extinction. *Sci Adv* 5(5):eaau0757
- Rousset F (1997) Genetic differentiation and estimation of gene flow from F-statistics under isolation by distance. *Genetics* 145(4):1219–1228
- Saccheri I, Hanski I (2006) Natural selection and population dynamics. *Trends Ecol Evol* 21(6):341–347
- 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. *Mol Ecol* 19(6):1147–1161
- Teixeira JC, Huber CD (2021) The inflated significance of neutral genetic diversity in conservation genetics. *Proc Natl Acad Sci USA* 118(10):e2015096118

- Titley MA, Butchart SHM, Jones VR, Whittingham MJ, Willis SG (2021) Global inequities and political borders challenge nature conservation under climate change. *Proc Natl Acad Sci USA* 118(7):e2011204118
- Tsang LM, Tsoi KH, Chan SKF, Chan TKT, Chu KH (2016) Strong genetic differentiation among populations of the freshwater shrimp *Caridina cantonensis* in Hong Kong: implications for conservation of freshwater fauna in urban areas. *Mar Freshw Res* 68(1):187–194
- Uba CU, Oselebe HO, Tesfaye AA, Abtey WG (2021) Genetic diversity and population structure analysis of bambara groundnut (*Vigna subteranea* L) landraces using DArT SNP markers. *PLoS One* 16(7):e0253600
- Ward R, Woodward M, Skibinski D (1994) A comparison of genetic diversity levels in marine, freshwater, and anadromous fishes. *J Fish Biol* 44(2):213–232
- Whiteley AR, Fitzpatrick SW, Funk WC, Tallmon DA (2015) Genetic rescue to the rescue. *Trends Ecol Evol* 30(1):42–49
- Willi Y, Kristensen TN, Sgrò CM, Weeks AR, Ørsted M, Hoffmann AA (2022) Conservation genetics as a management tool: The five best-supported paradigms to assist the management of threatened species. *Proc Natl Acad Sci USA* 119(1):e2105076119
- Wong W, Ma K, Tsang LM, Chu KH (2017) Genetic legacy of tertiary climatic change: a case study of two freshwater loaches, *Schistura fasciolata* and *Pseudogastromyzon myersi*, in Hong Kong. *Heredity* 119(5):360–370
- Wong WY, Chow LH, Wu TH, Ma KY, Tsang LM, Chu KH (2019) Conservation of freshwater wildlife in Hong Kong: a genetic perspective. *Aquat Conserv* 29(12):2204–2218
- Wu TH, Tsang LM, Chen IS, Chu KH (2016) Multilocus approach reveals cryptic lineages in the goby *Rhinogobius duospilus* in Hong Kong streams: role of paleodrainage systems in shaping marked population differentiation in a city. *Mol Phylogenet Evol* 104:112–122
- Wu TH, Tsang LM, Chow LH, Chen IS, Chu KH (2019) Cryptic lineages and hybridization of the predaceous chub *Parazacco spilurus* (Actinopterygii, Cypriniformes, Xenocypridae) in Hong Kong. *Hydrobiologia* 826:99–111
- Yam RS, Dudgeon D (2005) Genetic differentiation of *Caridina cantonensis* (Decapoda: Atyidae) in Hong Kong streams. *J North Am Benthol Soc* 24(4):845–857
- Zhang C, Reid K, Sands AF, Fraimout A, Schierup MH, Merilä J (2023) De novo mutation rates in sticklebacks. *Mol Biol Evol* 40(9):msad192

## ACKNOWLEDGEMENTS

Our fieldwork was conducted with *Permission to Make Field Collections for Research Purpose (168 AF GR CON 11/17 Pt.6)* from the Agriculture, Fisheries and Conservation Department of Hong Kong. We thank Jeffery C.F. Chan, Sandy Yau and Alex Lau for help with the fieldwork and Xueling Yi, Chaowei Zhang and Dandan Wang for their helpful suggestions regarding data analyses. Our research was supported by Seed Funding from the University Grants Committee (# 202111159018 to JM, KR and DD).

## AUTHOR CONTRIBUTIONS

Conceived the study: DD, KR and JM; Data collection: XW, YC, DD, KR, JM; Analyzed the data: XW; Lead the writing: XW; Commented and edited the manuscript: DD, JM, KR, YC.

## FUNDING

Open Access funding provided by University of Helsinki (including Helsinki University Central Hospital).

## COMPETING INTERESTS

The authors declare no competing interests.

## ETHICS APPROVAL

All procedures involving animals were in compliance with Cap. 340 Animals (Control of Experiments) Ordinance of 24 May 1963, and ethical approval was granted by the Committee on the Use of Live Animals in Teaching and Research (No 23-390, the University of Hong Kong, Hong Kong SAR).

## ADDITIONAL INFORMATION

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41437-024-00695-0>.

**Correspondence** and requests for materials should be addressed to Xi Wang or Juha Merilä.

**Reprints and permission information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024