Invasive Cyprinid Fish in Europe Originate from the Single Introduction of an Admixed Source Population Followed by a Complex Pattern of Spread

Andrea Simon[1]*, Robert Britton[2], Rodolphe Gozlan[2], Cock van Oosterhout[4], Filip A. M. Volckaert[3], Bernd Hänfling[1]

1 Evolutionary Biology Group, Department of Biological Sciences, University of Hull, Hull, United Kingdom, 2 Centre for Conservation Ecology & Environmental Change, School of Conservation Sciences, Bournemouth University, Poole, Dorset, United Kingdom, 3 Central Laboratory of Animal Diversity and Systematics, Katholieke Universiteit Leuven, Leuven, Belgium, 4 School of Environmental Sciences, University of East Anglia, Norwich, United Kingdom

Abstract

The Asian cyprinid fish, the topmouth gudgeon (*Pseudorasbora parva*), was introduced into Europe in the 1960s. A highly invasive freshwater fish, it is currently found in at least 32 countries outside its native range. Here we analyse a 700 base pair fragment of the mitochondrial *cytochrome b* gene to examine different models of colonisation and spread within the invasive range, and to investigate the factors that may have contributed to their invasion success. Haplotype and nucleotide diversity of the introduced populations from continental Europe was higher than that of the native populations, although two recently introduced populations from the British Isles showed low levels of variability. Based on coalescent theory, all introduced and some native populations showed a relative excess of nucleotide diversity compared to haplotype diversity. This suggests that these populations are not in mutation-drift equilibrium, but rather that the relative inflated level of nucleotide diversity is consistent with recent admixture. This study elucidates the colonisation patterns of *P. parva* in Europe and provides an evolutionary framework of their invasion. It supports the hypothesis that their European colonisation was initiated by their introduction to a single location or small geographic area with subsequent complex pattern of spread including both long distance and stepping-stone dispersal. Furthermore, it was preceded by, or associated with, the admixture of genetically diverse source populations that may have augmented its invasive-potential.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: a.simon@2007.hull.ac.uk

Introduction

Population genetic studies of invasive species have become an instrumental component in the study of biological invasions [1], [2], [3]. The application of neutral molecular markers can elucidate demographic processes during the invasion process and identify colonization pathways and source populations [4], [5]. Such information not only facilitates management and prevention of further invasions but also provides a framework for studies on adaptive evolution during the invasion process [6]. An issue which has recently received much attention but remains poorly understood is the role of genetic diversity in determining the outcome of introductions of non-native species. Introductions of non-native species are often based on the release of a low number of founding propagules containing only a fraction of the genetic variation of the source populations [7]. Such reduced genetic diversity theoretically limits a species' ability to establish invasive populations invoking a genetic paradox [8], [9], [10], [11], [12]. Although many successful invasive species show reduced genetic diversity, recent research suggests that the effects of such bottlenecks are often counteracted by admixture among genetically divergent source populations [3], [13]. For example, multiple introductions have resulted in high genetic diversity of invasive crustaceans [14], fish [3], [15], [16], lizards [17] and plants [18]. Nevertheless, it is currently unknown whether such admixture is merely a side-effect of the invasion process or is actually facilitating the establishment process. Additional population genetic case studies, in combination with studies on ecologically significant traits and genome wide associations are crucial in providing answers to this question.

One of the most compelling fish invasions in the world today is arguably the topmouth gudgeon *Pseudorasbora parva* (Temminck and Schlegel, 1846). This small cyprinid species originating from East Asia was accidentally introduced into Europe in the 1960s in several countries around the Black Sea as part of contingents of Chinese carps for aquaculture [19], [20]. Since then, they have proved highly invasive through a combination of combination of sociological, economical and ecological factors that enabled their rapid human-assisted and natural dispersal throughout the continent. On introduction into a new water body, colonisation is facilitated by their tolerance of degraded aquatic ecosystems and their reproductive traits of early sexual maturity, batch spawning, high reproductive effort and paternal nest guarding that provide a high degree of invasive vigour [20], [21], [22]. Their capacity for

subsequently forming high density populations can then result in sharing of common food resources with native fishes resulting in overlaps in trophic niche [23], with additional concerns over egg predation, disease transmission and facultative parasitism [22].

Whilst this *P. parva* invasion has been traced from the initial point of introduction towards the northern and western parts of Europe, as well as the south towards Turkey and Iran [22], its exact demographic scenario is currently unclear. They are now found in at least 32 countries with contrasting climates (e.g. Algeria, Austria, Poland, Spain), have invaded habitats with a wide range of ecological conditions and their life history traits differ considerably among invasive populations [22]. Possible (non-mutually exclusive) explanations of such variability are: (1) the existence of considerable phenotypic plasticity in life history traits and tolerance to environmental conditions, (2) a rapid evolutionary response, or (3) multiple independent introductions from divergent source populations [19], [22], [24]. Molecular markers have previously been employed to study such questions in other freshwater fish invasions in Europe [25], [26] and North America [7]. For example, using mitochondrial DNA, Vidal *et al.* (2010) [26] showed that the mosquitofish (*Gambusia holbrooki*) was introduced into Europe multiple times from USA. Some *P. parva* populations have also been identified as healthy carriers of pathogens, such as *Anguillicola crassus* [27] and the rosette agent *Sphaerothecum destruens* [28], [29]. It is currently unknown whether other invasive populations or native populations show a similarly low susceptibility to the rosette agent.

Consequently, *P. parva* appear to be a model fish well suited to studying the evolution of ecologically significant traits, disease resistance and the role of genetic diversity in establishment success. Thus, we perform a population genetic analysis of *P. parva* across their native and introduced ranges in order to test different models of colonisation and to determine levels of genetic variation across the invasive range of the species (see Material and Methods for specific hypothesis). This will provide a first population genetic framework for further evolutionary studies on the species.

Materials and Methods

Sampling scheme and hypothesis testing

Samples were collected at a total of 22 sites, 14 in Europe and 8 in Asia (Table 1; Figure 1). Sample size was 15 for the majority of sites with the exception of three sites where 6–10 individuals where sampled. There was also a single sample (Japan) that comprised three individuals; it was excluded from all population-based analyses. The native range of the species is the East Asian sub-region, including the basins of the Huang He, Yangtze, Hai He and Amur Rivers, as well as some Japanese islands, Taiwan and the southern part of Korea [30], [31] and the sampling scheme covers most of the latitudinal space in this range, as well as spanning across the largest part of the European invasive range. The density of the coverage in the native range was appropriate to test some general demographic processes but not the identification of the exact location of potential source populations.

Thus, the aim was to test three non-mutually exclusive models that were proposed to explain the spread of *P. parva* in Europe: i) 'multiple source-sink' model where several independent introduction events from genetically differentiated native source populations to separate European locations would have occurred without involving admixture; ii) 'stepping-stone' model [19] where introduction into a single geographical area would have been followed by gradual expansion from the original introduction; and iii) 'long-distance' model [22] where introduction into a geographical area would have been followed by long-distance translocation within Europe. Furthermore, it was tested whether iv) the invasive populations show signs of a genetic bottleneck or v) might have resulted from an admixture between divergent source populations.

Population genetic theory predicts that these demographic processes will result in different patterns of genetic population structure and therefore molecular approaches can be used to test the likelihood of alternative models. Therefore a number of phylogenetic and population genetic analyses were carried out in order to test the results against the theoretical expectations for the scenarios outlined above. Note that some of these tests assume that a relatively clear phylogeographic subdivision exists in the native range. Therefore the first step was to carry out a network analysis in order test this assumption. Genetic distances and *F*-statistics were used to quantify the degree of differentiation between populations and nucleotide diversity, and haplotype diversity at a standard sample size was used to estimate within population variability. These analyses were complemented by coalescent simulations and a Bayesian estimation of effective population size. The results were then compared with theoretical expectations from the various models and scenarios:

i) 'multiple-source-sink' model: genetic differentiation among invasive populations is high and similar to that found in the native range;

ii) 'stepping stone' model: genetic differentiation in the invasive range is lower than that in the native range, and there is a significant pattern of isolation-by-distance;

iii) 'long-distance' model: genetic differentiation in the invasive range is lower than that in the native range, and there is no pattern of isolation-by-distance;

iv) 'genetic bottleneck' scenario: genetic diversity of invasive populations, in particular haplotype diversity, is lower than that of the source populations; and

v) 'genetic admixture': genetic scenario: genetic variation expressed in nucleotide diversity is higher than that of the source population. Furthermore, recent admixture increases the nucleotide diversity above that expected under equilibrium conditions.

Molecular procedures

The fish were collected and stored in 98% ethanol. Genomic DNA was extracted from the caudal fin tissue using the HotShot method [32]. An approximately 700 bp long section of the mtDNA genome, containing the partial *cytochrome b* gene was amplified applying standard PCR techniques using Verity Thermal Cycler. Primers L15267 and_H15891Ph, previously described by Briolay *et al.* (1998) [33], were used. Thermal cycle amplifications were performed in 15 µL reactions, containing 1.5 µL 160 mM NH_4, 1.5 µL 100 mM dNTPs, 0.4 µL 50 mM $MgCl_2$, 0.075 µL *Taq* polymerase, 0.3 µL each of primers L15267 and H15891Ph, 9.425 µL PCR water and 1.5 µL of template DNA. Cycle parameters were as follows: 2 min at 95°C; 45 s at 94°C, 45 s at 48°C, 1 min at 72°C; 10 min at 72°C. PCR products were directly sequenced in both directions using the PCR primers by Macrogen Inc. Forward and reverse sequences were aligned and edited using CodonCode Aligner [34], (GenBank accession numbers: JF489575-JF489887). Consensus sequences were imported into MEGA v. 4.1 [35] and aligned with ClustalW [36].

Table 1. Sample locations and sample sizes and geographical coordinates for native and invasive population.

Code	Population	N	Geographical co-ordinates	
CG	Guangdong, River, Zhuijang River basin, China	6	23°07′53″N	113°15′59″E
CH	Huairou Reservior, Hai He River basin, China	15	40°18′46″N	116°36′36″E
CK	Kinmen Island population, China	6	24°26′11″N	118°21′27″E
CRH	River Hai He, Hai He River basin, China	15	39°07′15″N	117°12′54″E
CY	Wuhan, Yangtze River Basin, China	10	29°58′20″N	113°53′29″E
JB	Lake Biwa, Yodo River basin, Japan	3	32°20′44″N	136°10′15″E
TI	I-lan county, I-lan River, Lanyang River Basin. Taiwan	15	24°45′00″N	121°45′00″E
TT	Dajia River, Taichung county, Dajia River basin, Taiwan	15	23°09′00″N	120°38′34″E
BS	Slangebeek nean Hasselt, Belgium	15	50°55′48″N	05°15′00″E
EB	Byland Abbey, Yorkshire, UK	15	54°12′10″N	01°09′35″W
FG	Grand Lieu, France	15	47°05′45″N	01°43′46″W
G	River Ammer, Wielenbach, Germany	15	47°52′11″N	11°09′00″E
HA	Hortobagy, Hungary	15	47°36′00″N	21°06′00″E
HE	Ederecsi-patak, Hungary	15	46° 48′04″ N	17° 23′16″E
HG	Gic, Hungary	15	47° 25′32″N	17°44′44″E
HS	Salyi-patak, Hungary	15	47°56′06″N	20°39′58″E
IN	Nestore, Italy	15	43°21′14″N	12°14′10″E
PU	Utrata River, Poland	15	50°35′50″N	18°09′32″E
SC	Vrakuna, Slovakia	15	47°49′24″N	18°49′16″E
SE	Ebro Basin, Spain	15	40°43′12″N	00°51′47″E
SWS	Sylen Lake, Llanelli, South Wales, UK	15	51°40′42″N	04°09′47″W
T	Blanice River, Vodnany, Czech Republic	15	49°08′52″N	14°10′32″E

doi:10.1371/journal.pone.0018560.t001

Phylogenetic analyses and haplotype network

Phylogenetic relationships of haplotypes were reconstructed using the maximum composite likelihood method [37] in combination with Neighbour-Joining as implemented in MEGA v. 4.1 [35]. Furthermore we created a Maximum Likelihood tree, using the RaxML programme [38] using the GTR model optimised for each codon position. Branch support of both was obtained using non-parametric bootstrapping as percent of 1000 repeats and ML support values over 80% were added to tree nodes. Our aim using the phylogenetic tree approach was to show
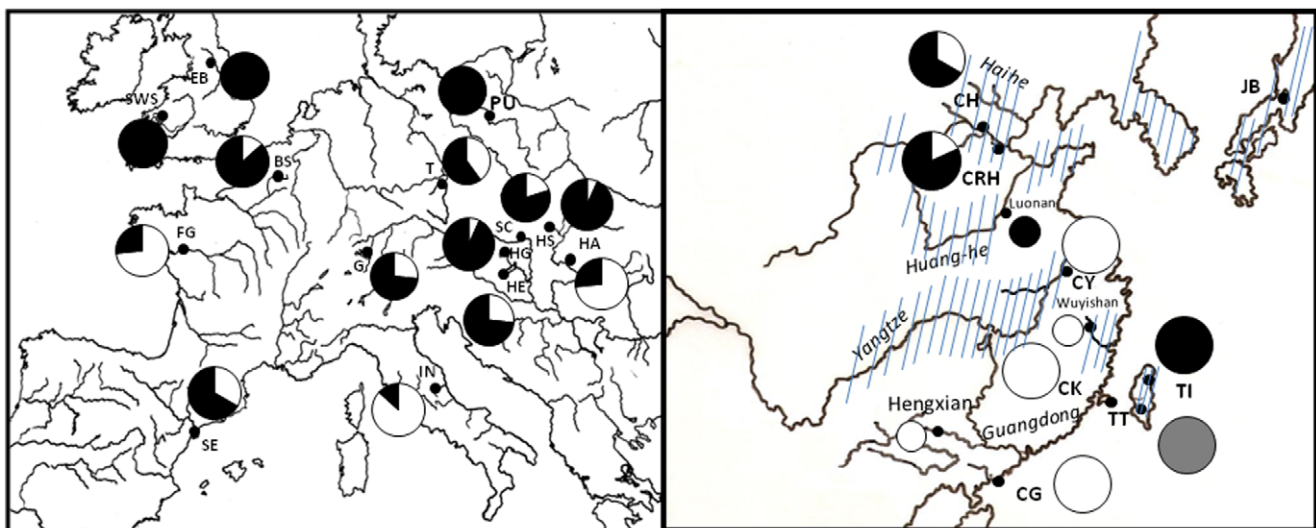

Figure 1. Distribution of *Pseudorasbora parva* samples sites in Europe (left) and in Asia (right), showing the species' native range. Pie charts represent the geographical distribution of major mtDNA lineages (see Figure 4). Lineage 1 = white, Lineage 2 = black, lineage 3 = grey. See Table 1 for population codes. Large pie charts represent samples collected in this study, small pie charts samples from Liu *et al.* 2010.
doi:10.1371/journal.pone.0018560.g001

how distant haplotypes relate to major clades, rather than to provide definite resolution within clades.

In order to increase the geographic coverage, GenBank sequences from five *P. parva* individuals [39] sampled in the Minjiang River at Wuyishan (EU934500), the Pearl River at Hengxian (EU934501 and EU934502) and the Yellow River at Luonan (EU934503 and EU934504) were included in the phylogenetic analysis. Representatives of the main lineages of the cyprinid subfamily Gobioninae according to Tang *et al.* (2010) [40] were included as an outgroup using the same GenBank sequences as Tang *et al* (2010) [40].

A haplotype network was constructed using a median-joining algorithm in Network v. 4.5.10 [41]. Possible homoplastic sites (153, 195, 300, 462, and 585) were weighted down to 1 and all other nucleotide positions were weighted at 50 and we used an ε value of 0. Furthermore, transversions were weighted three times higher than transitions to decrease the likelihood of homoplastic substitutions [42]. A BLAST search of nucleotide sequences [43] was performed in order to confirm that all sequences belonged to *P. parva*.

Population genetic data analysis

DNaSP v. 4.5 [44] was used to estimate within population diversity (nucleotide diversity, π; haplotype diversity, Hs). Standardised measures of genetic diversity were calculated by resampling data sets 1000 times using a bootstrapping procedure [45], [46] based on the size of the smallest sample (6 individuals). Differences in genetic diversity between native and invasive populations were tested using a Mann-Whitney test. The invasive population PU was excluded from the comparison of π because it contained one highly divergent haplotype which is suspected to be derived from hybridisation with *Gobio gobio*. Coalescent based simulations as implemented in DNaSP were used to predict the expected relationship between haplotype diversity (H) and nucleotide diversity (π) under drift-mutation equilibrium and constant population size [47]. Effective population size of native populations assuming mutation-drift equilibrium and absence of migration among watersheds was estimated using MIGRATE-n v. 2.5 (Figure 2) [48]. The option Bayesian inference was used with the default search strategy settings. The rationale of this analysis was to estimate the populations size required to maintain the amount of genetic diversity found in the each population assuming mutation-drift-equilibrium.

Pairwise genetic differentiation among samples was computed as F_{ST} and D_{XY} (using Kimura two-parameter method, 1980 [49]) (Table 1, 2) using DNaSP v. 4.5 [44]. A multi-dimensional scaling (MDS) analysis based on F_{ST} was carried out in order to visualise the genetic relationship between samples. The average pairwise differentiation between native populations was compared to the average pairwise differentiation of invasive populations using a Mann-Whitney test. Isolation by distance (IBD) (Appendix, Table S3.) analysis was then used to test whether the 'stepping-stone' model could explain the spread of *P. parva* within Europe. Pairwise geographic distances among European sites were calculated as Euclidean distances. The theoretical expectation is that a significant correlation should only occur under the 'stepping-stone' model [50], [51]. Three different approaches were used. First, a 'classical' IBD analysis [52] was carried out to test the relationship between matrices of geographical distance and genetic differentiation (F_{ST}) using a Mantel test (1000 permutations) as implemented in the software IBDWS v. 3.16 [53]. The genetic F_{ST} values were log-transformed to achieve a normal distribution. Second, a general linear model (GLM) was used to test the relationship between the geographic distance and genetic

differentiation from the putative site of introduction. Third, a GLM was used to test the relationship the geographic distance from the putative site of introduction and genetic diversity of populations. Under a 'stepping stone' model, genetic diversity is expected to decrease with geographic distance to the original site of introduction, and hence, the genetic distance is expected to increase. The putative site of introduction was Nucet-Dombovita, Romania in the early 1960s [54], however around this time several other introductions took place into Hungary [22], so this population (HA) was used as reference population.

Approximate Bayesian Computation (DIY ABC)

Approximate Bayesian Computation (DIY ABC) [55] was used to estimate the relative likelihood of alternative scenarios of the initial introduction of the species into Europe. In the programme, reference tables (containing parameters based on known values) were used to compare the scenarios and the simulated datasets were then compared to the true values (Cornuet *et al.* 2008). DIY ABC is a computationally intensive approach and therefore only three simplified scenarios where chosen, which appeared most feasible after the initial population genetic analysis. An explicit rationale for choosing specific models will therefore be given in the Results section. The prior distribution of the coalescence time in the evolutionary scenario was partially informed by historical data, such as the date of the first introduction (Appendix, Table S3). The effective population size was set as uniform, 10 and 5×10^4 individuals, covering the full range of biologically feasible values and the Kimura 2 parameters (1980) [49] mutation model was used. For each scenario 10^6 datasets were simulated with the parameter values drawn from the prior distribution (Appendix, Table S3.). The relative likelihoods of the three scenarios were compared by using logistic regression on 1% of the closet simulated data sets.

Results

Phylogenetic and network analysis and distribution of haplotypes

A total of 30 haplotypes were identified using 310 sequences from 8 native and 14 introduced populations (Table 1). The phylogenetic relationship among haplotypes is shown in Figure 3. Both NJ and ML methods yielded the same topology, hence only the NJ tree is displayed but ML support values were added to tree nodes. The two Japanese haplotypes, H23 and H24 were closely related to each other and the phylogenetic analysis (Figure 3) showed that they formed a highly divergent sister group to the remaining *P. parva* haplotypes (sequences divergence ~5–6%). One highly divergent haplotype found in the invasive Polish populations clustered closely to a sequence of *G. gobio*. This haplotype and the Japanese haplotypes were therefore not included in the network analysis. Thirteen haplotypes were found in the invasive populations, five of which were found in more than one invasive population and will be subsequently referred to as common haplotypes. Three of the common haplotypes and two of the rare haplotypes were also found in at least one native population. Three main lineages of *P. parva* haplotypes can be recognised outside of Japan (Figures 3 and 4); a highly diverse central lineage (lineage 2) and two peripheral lineages (lineages 1 and 3) that are separated from the central lineage by 6 and 7 mutations, respectively. Lineage 3 consists of a single haplotype which is fixed in one of the native Taiwanese populations. One native population (TI) sampled in this study and the yellow river sample from Liu *et al* (2010) [39] are restricted to lineage 2 but do not share haplotypes with invasive populations. Three native

Figure 2. Estimates of effective population size (theta) of native populations based equilibrium assumptions.
doi:10.1371/journal.pone.0018560.g002

Table 2. Genetic diversity of *Pseudorasbora parva* populations.

Population	Group	Mh	H	H₆	π	code
CH	native	3	0.44	0.34	0.0056	green
CRH	native	11	0.96	0.75	0.0103	red
CK	native	1	0.00	0.00	0.0000	blue
CG	native	1	0.00	0.00	0.0002	purple
CY	native	2	0.46	0.35	0.0023	orange
TI	native	3	0.59	0.46	0.0011	brown
TT	native	1	0.00	0.00	0.0002	grey
BS	invasive	5	0.68	0.54	0.0067	white
EB	invasive	1	0.00	0.00	0.0000	white
FG	invasive	5	0.78	0.61	0.0109	white
G	invasive	6	0.84	0.66	0.0098	white
HA	invasive	5	0.62	0.48	0.0078	white
HE	invasive	3	0.59	0.46	0.0077	white
HG	invasive	5	0.62	0.49	0.0049	white
HS	invasive	3	0.25	0.20	0.0028	white
IN	invasive	3	0.34	0.27	0.0049	white
SC	invasive	3	0.67	0.53	0.0073	white
SE	invasive	3	0.67	0.52	0.0080	white
SWS	invasive	2	0.24	0.20	0.0012	white
T	invasive	3	0.68	0.55	0.0076	white
PU	invasive	4	0.60	0.47	0.0257	white

Columns represent populations, origin (native or invasive) number of haplotypes found in each population, observed haplotype diversity (H), mean haplotype diversity after bootstrapping based on sample size of 6 and nucleotide diversity (π) and colour code used in Figure 4.
doi:10.1371/journal.pone.0018560.t002

populations (CG, CK, CY) sampled in this study and the Minjiang sample from Liu *et al* (2010) are restricted to lineage 1 and these populations also share a common haplotype with most introduced populations. Furthermore the Pearl River samples from Liu *et al* (2010) [39] fall into lineage 1 but do not share haplotypes with native populations. Two native populations (CRH, CH, Figure 1), however, contained haplotypes from both lineage 1 and 2 but share few haplotypes with the invasive populations. These two populations are from the Hai He River basin at the northern margin of the species distribution. The invasive populations are widely scattered across the network and most populations contain highly divergent haplotypes from both lineage 1 and 2. Among the native populations, the Taiwanese and Japanese (TI, TT, JB) populations do not share haplotypes with any native or invasive populations.

Diversity within populations

After bootstrapping to account for differences in sample size, the genetic diversity of the native populations varied widely among geographical regions. Whereas the two populations of the Hai He drainage showed relatively high diversity (H = 0.34, 0.76; π = 0.006, 0.010), the populations from other drainages of mainland China and Taiwan where much less variable (H = 0.00–0.46; π = 0.000–0.002). (Table 2, Figure 5). Genetic variation in introduced populations also varied considerably. The two recently established British populations showed low levels of variability (H = 0.00, 0.20; π = 0.000, 0.001) whereas the popula-

tions from continental Europe showed relatively high levels of variation (H = 0.20–0.66); π = 0.003–0.026). The highest nucleotide diversity was found in the Polish (PU) population (π = 0.026); this population contained one extremely divergent haplotype that clustered with a *G. gobio* haplotype, suggesting hybridisation and so was excluded from further comparisons. Overall genetic diversity in native populations (mean±SD; Hs = 0.27±0.29; π = 0.003± 0.004) and invasive populations (mean±SD; Hs = 0.43±0.19; π = 0.008±0.006) was not significantly different (H, P = 0.108; π, P = 0.068). However, a more detailed analysis revealed that there were significant differences among certain groups of native and invasive populations. When the recently introduced UK populations were excluded from the analysis, both haplotype diversity and nucleotide diversity and were significantly higher in the invasive populations than native (P = 0.043 and, P = 0.014, respectively).

Next we simulated the nucleotide (π) and haplotype (H) diversity expected in a population that is in mutation-drift equilibrium with constant effective population size (N_e), and we compared this to the empirical data (Figure 5). The simulations show that with increased N_e, both H and π increase, which is predicted from theory, given that larger populations can harbour more nucleotide and haplotype diversity (Figure 5). However, the observed values of π for the introduced populations fall consistently above the theoretically predicted relationship between π and H. Thus, the introduced populations showed a relative excess of nucleotide diversity, given the observed haplotype diversity and assuming mutation-drift equilibrium. Similarly, some native populations also showed a relative excess in π compared to H (Figure 5). This pattern was inconsistent with a mutation-drift equilibrium and can be explained by admixture of populations with diverged nucleotide variation.

Maximum likelihood estimates of theta (N_eμ), using Migrate-n, differ by several orders of magnitude among native populations, ranging from 0.00006 for population TT to 0.01847 for population CRH (Figure 2). Using an average mutation rate for mtDNA of 1% per MY [56], this translates into effective population size estimates between approximately 10^3 and 4.10^5 individuals. This analysis is consistent with the previous simulation study as it shows that the standing nucleotide variation in some populations can only be explained by an exceedingly large effective population size, or more plausibly, by population admixture.

Genetic differentiation and population structure

Pairwise genetic distance (D_{XY}) ranged from 0 to 0.02715 and pairwise genetic differentiation (F_{ST}) ranged from 0 to 1 (Appendix, Table S1), not including the Japanese (JB) population. The pairwise genetic distance among invasive populations (median D_{XY} = 0.009) was only marginally lower than that among native populations (median D_{XY} = 0.012), (P>0.3). Similarly, the native-invasive pairwise comparison expressed in D_{XY} (median D_{XY} = 0.012) (Appendix, Table S2) was neither significantly different from the genetic distance among native populations (P>0.3) nor from that among invasive populations (P>0.3).

In contrast, genetic differentiation (F_{ST}) was considerably lower among the invasive populations (median F_{ST} = 0.21) than among the native populations (median F_{ST} = 0.58) (P<0.001). Furthermore, the F_{ST} between the native-invasive pairwise comparison (median F_{ST} = 0.53) was not significantly different from the genetic distance among native populations (P = 0.27), but it was significantly higher than that among invasive populations (P<0.001). This result is inconsistent with the 'multiple-source-sink' model, and supports both the 'long-distance' and 'stepping stone' models.
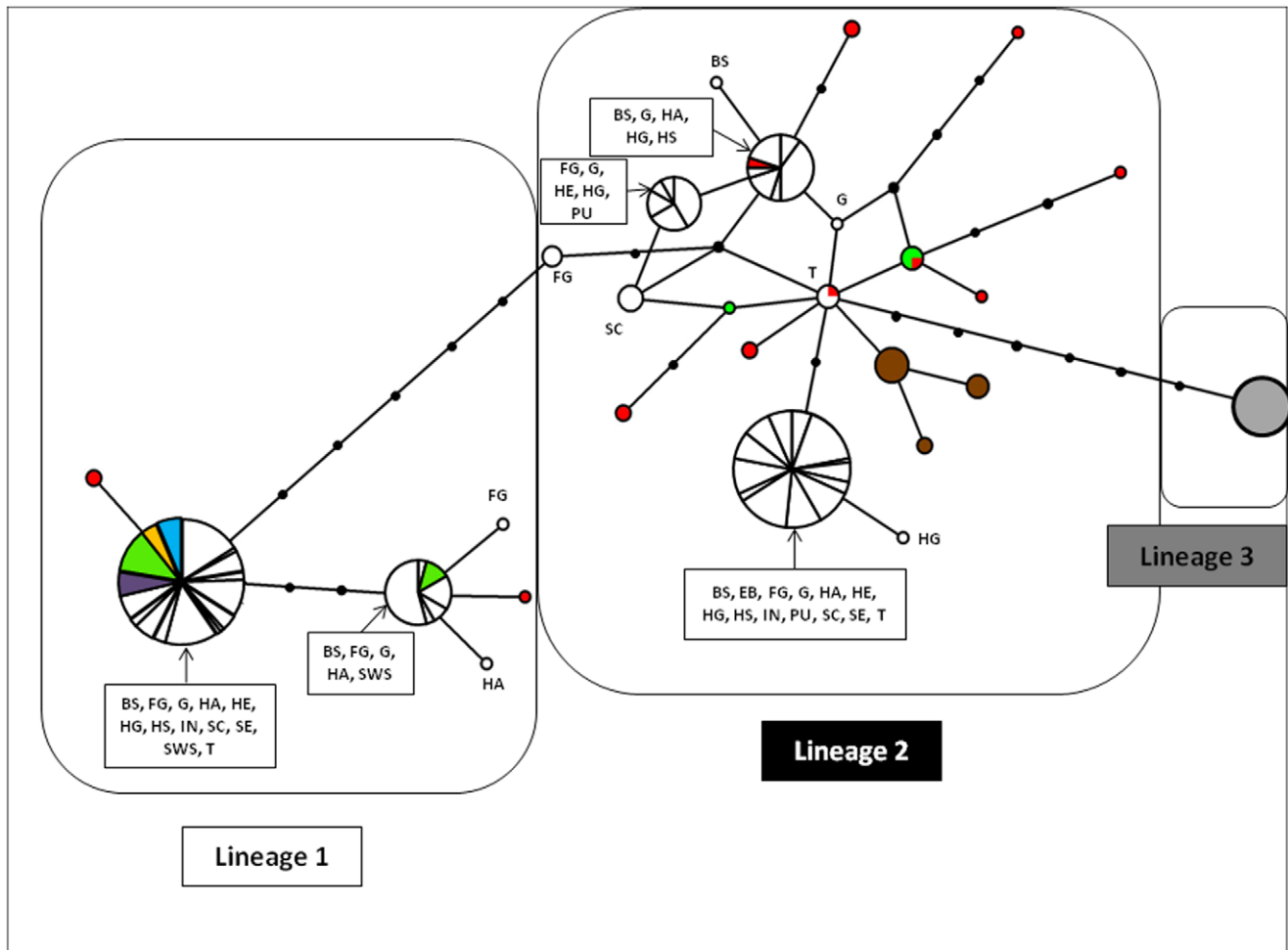
Figure 3. Phylogenetic relationship of haplotypes based on NJ analysis. First value on branches indicate ML support value, second value indicate non-parametric bootstrapping of the NJ-tree. Values are only given for support values >70%.
doi:10.1371/journal.pone.0018560.g003

The multidimensional scaling analysis of the F_{ST} matrix (Figure 6) showed that most of the invasive populations cluster together with two native populations (CRH and CH). This cluster is surrounded by the remaining native populations and two introduced populations (EB and SWS). Overall, there appears to be a pattern that nucleotide diversity increases towards the centre of the plot, i.e. intermediate populations have the highest nucleotide diversity, which again indicates that these populations (invasives and the samples from the river Hai He) are genetically admixed.

A Mantel-test showed a significant relationship between genetic and geographic distance among the European populations (Z = −106685; $r = 0.28$, one sided $P<0.05$). However, when the recently introduced English (EB) and Welsh (SWS) populations were removed, there was no significant genetic isolation-by-distance relationship (Z = 9201, r = 0.06; one sided $P>0.30$). Regression analysis revealed no significant relationship between distance from source and genetic differentiation ($R^2 = 0.005$; $P>0.03$) or genetic diversity ($R^2 = 0.018$; $P>0.03$) respectively. This reveals that the English and Welsh populations are bottlenecked, resulting in the spurious isolation-by-distance signal obtained when including these samples in the Mantel test. However, across continental Europe, topmouth gudgeon does not show evidence of isolation-by-distance and so we conclude that the 'long-distance' model is most consistent with these data.

DIY ABC

Based on the geographic distribution of the haplotype lineages, samples were pooled into three native and one invasive population for which we considered three feasible evolutionary scenarios (Figure 7): (i) pop 1 (native populations of haplotype lineage 1; CG, CK, CY, Minjiang), (ii) pop 2 (admixed native populations from the river Hai He; CH, CRH), (iii) pop 3 (all invasive Hungarian populations; HA, HE, HG, HS), pop 4 (native populations of lineage 2; TI, Yellow River). The Hungarian populations were chosen to represent invasive populations because they were located in close proximity to the original site of introduction. In order to account for the unsampled variation in the native range in lineage 2, one or two ghost population (GH1, GH2) were included in the scenarios (represented as branches without terminal ends in Figure 7). All three scenarios assumed that a founder of size NF that lasted DB generations had event had taken place after introduction into Europe:

> Scenario 1: The source of the invasive population (Pop 3) is the admixed Chinese population (Pop 2) which
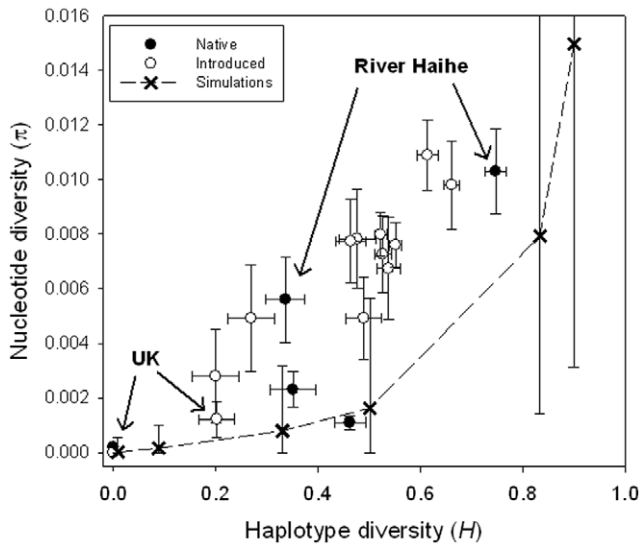
Figure 4. Medium joining network of *cytb* haplotypes from native and introduced populations of *Pseudorasbora parva*, excluding H22, H23 & H24. Adjacent haplotypes are connected through a single point mutation. Each circle represents a single haplotype and its diameter is proportional to the number of individuals with that haplotype. The colour codes represent the locations in which the haplotype is found, filled cricles (●) represents unsampled haplotypes.
doi:10.1371/journal.pone.0018560.g004

originates from an admixture of Pop 1 and a ghost population which split from Pop 4 at time t3.

Scenario 2: The invasive population (Pop 3) is a result of an admixture between pop 1 and an unsampled ghost population which split from pop 4 at time t4. Pop 2 evolved as in scenario 1. Scenario 3: same as Scenario 2 but the admixture of the pop 1 and GH2 populations took place before the admixture of Pop 1 and GH1.

A comparison of posterior probabilities of the three scenarios using local linear regression (Figure 7, Table 3) showed that scenario 1 showed the lowest support with probabilities lower than 0.1. The highest probability was shown for scenario 3. The posterior distribution of model parameters under the most likely scenario was used to make inferences about the timing of events during the colonisation process assuming a generation time of one year. The posterior density of the time of first introduction ($t1$) agrees with historical records (median = 47 generations, 95% credibility interval (CI) = 30–60). Full table of posterior distributions are given in the Appendix, Table S3.

Discussion

The outputs of these analyses revealed that i) there are three evolutionary lineages of the topmouth gudgeon (*P. parva*) in the native range, two of which contributed to the colonisation of Europe; ii) most invasive populations have a higher genetic diversity than their native counterparts and a higher genetic diversity than expected under equilibrium conditions; iii) most native populations have a low genetic diversity typical for riverine fishes, an exception being samples from the Hai He river system which showed very high levels of genetic diversity, which under equilibrium conditions predict extremely high effective population sizes; and iv) the differentiation among invasive populations is much lower than among native populations.

Population genetics of native populations

The existence of four highly divergent haplotype lineages indicates a long isolation among geographic populations of *P. parva*. An approximate estimation of divergence times using a standard molecular clock rate of 1% MY [56] suggests a separation of the Japanese from the Chinese and Taiwanese populations during the Miocene (5–6 MYA) which is consistent with [57]. Accordingly, the remaining lineages will have formed during early Pleistocene (1–1.5 MYA), which implies that multiple glacial refugia must have existed during the ice ages. Although the
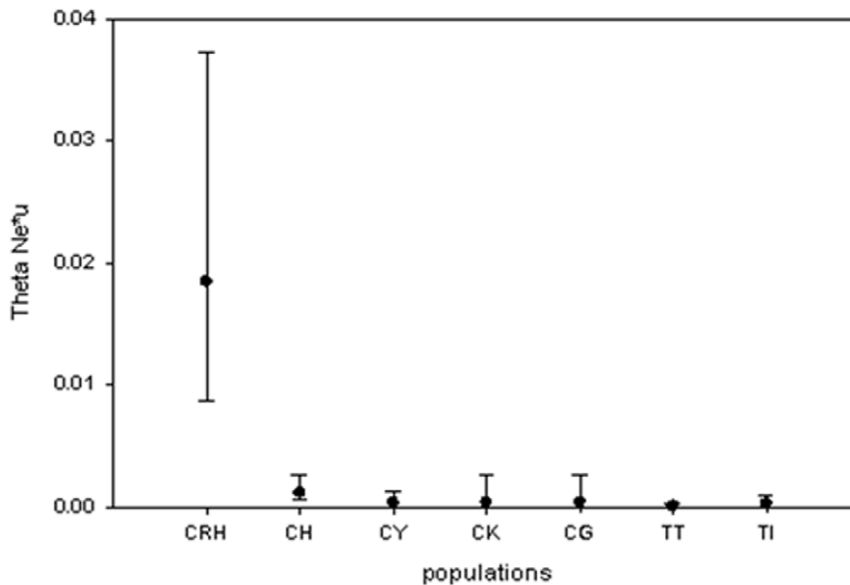

Figure 5. Plot of nucleotide (π) diversity versus haplotype (H) of the introduced (open circles) and native populations (solid circles). Also shown is the expected relationship between nucleotide diversity (±5–95% CI) and haplotype diversity of simulated populations (crosses) under mutation-drift equilibrium for populations. Excluded is the Polish population PU because its high value nucleotide diversity.
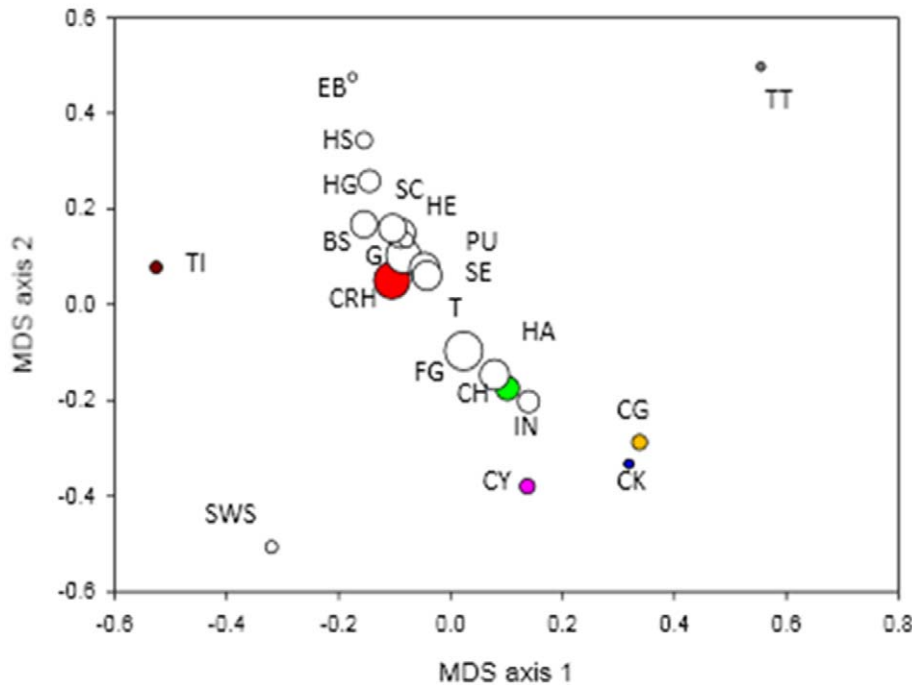doi:10.1371/journal.pone.0018560.g005

Figure 6. Plot of the first and second axis of a multidimensional scaling analysis based on pairwise *F*ST values among populations.
Size of symbols is proportional to the nucleotide diversity of populations. Native populations are colour coded according to Table 1.
doi:10.1371/journal.pone.0018560.g006

sampling scheme limits detailed phylogeograhic inferences, it is apparent that there is a clear geographic association of each lineage across most of the range, but also an area in Northern China where two lineages are found in sympatry. This becomes apparent when analysing genetic diversity within populations. Most *P. parva* populations from their native range showed low haplotype and nucleotide diversity and high levels of differentiation among river systems, which is consistent with the pattern found in many other small freshwater fishes of similar size, such as the European bullhead (*Cottus gobio* L.; [58], [59]) and guppies (*Poecilia reticulata* Peters; [60], [61]). The native populations from the northern range of the distribution were, however, characterised by extremely high genetic diversity, particularly the population CRH. Such high diversity is unusual among freshwater fish populations as they are usually highly structured and show low effective population sizes. Indeed, the effective population size was estimated as requiring approximately 400,000 individuals to maintain the levels of diversity observed in the CRH population and 24,000 individuals in population CH, based on a coalescence approach that assumes mutation-migration-drift equilibrium. Published estimates of effective population size in other freshwater fishes and our own estimates from the remaining native populations ($N_e < 7000$) are several orders of magnitude lower; for example, other cyprinid fishes range around 500 to 1000 individuals [62], guppies range from 100 to 900 [60] and European bullheads between 80 and 500 [63]. This suggests that the populations CRH and possibly CH are not at equilibrium but represent relatively recent secondary contact between divergent populations.

It is possible that the geographic area around the Hai He River basin represents a natural secondary contact zone between divergent phylogeographic lineages. Although the literature on this subject is relatively limited, it seems clear that high tectonic activity and sea level changes during the Pleistocene have created

a complex phylogeographic pattern with little concordance among species [64]. Nevertheless, studies on other freshwater fish, such as *Hemibarbus lameo* [65] and *Salanx ariakensis* [66], found evidence that secondary contact between diverged populations from different major river systems took place during low sea levels at the end of the Pleistocene. Furthermore, the geographic area around the Hai He River basin represents a natural secondary contact zone between divergent phylogeographic lineages of the estuarine, flathead mullet (*Mugil cephalus*) [67], [68].

Alternatively, recent human translocations associated with aquaculture might have caused such an admixture; this may not be considered surprising given that freshwater aquaculture in this area of China is intense [69]. According to Gozlan *et al* (2010b) [22], a high volume of *P. parva* translocations have occurred in China prior to introduction in Europe. These cyprinid translocations coincided with the end of the Chinese civil war and the need for additional sources of animal proteins [22].

Colonisation history

Our data showed that all invasive populations shared at least one of the four common haplotypes and that levels of genetic differentiation were low compared to native populations. Such a pattern would be expected if the invasive populations had spread from a single source. The alternative explanation of high levels of gene flow among initially differentiated invasive populations is extremely unlikely given that this would involve regular gene flow across watersheds. Therefore we reject the possibility that different European populations were independently colonised from divergent source populations ('multiple-source-sink model'). However, a number of results indicated that the introduced populations represented an admixture of divergent source populations. First, the levels of nucleotide diversity of populations in continental Europe were, on average, higher when compared with native populations. Second, the nucleotide diversity of invasive popula-
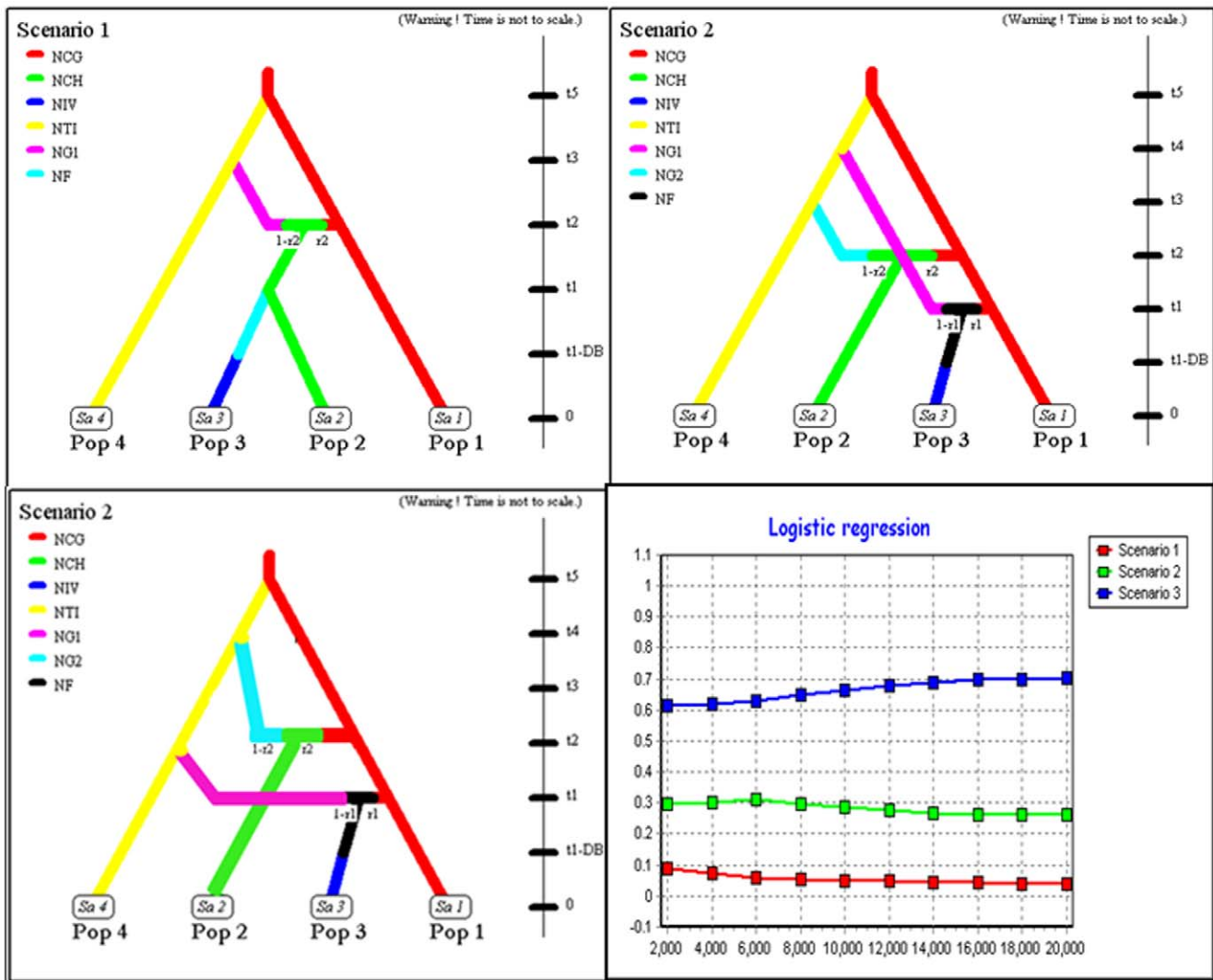
Figure 7. Graphic representation of the three competing invasion scenarios considered in the DIY ABC analysis. (Description of the scenarios are in the Results section.) Graph of linear regression, showing posterior probabilities of the scenarios.
doi:10.1371/journal.pone.0018560.g007

tions was higher than expected from coalescent theory. Finally, the population structure analysis showed that the majority of the invasive populations and two (admixed) native populations occupied central positions in the MDS plot between divergent native populations. The main exceptions to this pattern were the two British populations, which showed a low genetic diversity and high levels of differentiation from other invasive populations, but were fixed for one or two common invasive haplotypes. These populations were founded relatively recently, most likely from sources in Germany [19], [22]. We suggest that this pattern is a result of secondary bottlenecks during spread and translocation within Europe. Our data do not enable us to distinguish whether the admixture event has happened before the introduction into Europe or shortly after the introduction, before the large scale expansion across Europe, but based on the assumption that a single introduction to the same geographical location is more parsimonious than two independent introductions we suggest that it is more likely that the admixture event has happened in the native range.

Although the sampling coverage in the native range was not comprehensive enough to pinpoint the exact location(s) which

acted as a source of invasive European populations, some more general inferences can be drawn. The data outputs suggest that the invasive populations originate from mainland China rather than Taiwan or Japan. The haplotype distribution of invasive populations and populations from northern China raise the possibility that this area is the source of introduction. However, the DIY ABC analysis suggests that this is much less likely than a scenario where the invasive populations in Europe originate from an admixture between populations from lineage 1 (such as the Yangtze) and an unsampled population from lineage 2. Anecdotal reports suggest that *P. parva* were initially translocated to Romania and Hungary from the Yangtze River at Wuhan which is geographically close to our CY sample [22] and most likely originate from an aquaculture pond. Given our genetic results we suggest that these aquaculture populations consisted of a mixture of the local Yangtze population and fish wish were introduced from a different more northern river system possibly a tributary of the Yellow River.

The isolation-by-distance analysis indicated that both 'stepping-stone' and 'long-distance' processes might have contributed to the spread of *P. parva* in Europe. The weak but significant pattern of

Table 3. Output file of the Direct approach, relative proportion of each scenario found in the selected 500 closest dataset; Posterior probabilities of scenarios obtained through a logistic regression computed every 10% of the number of selected datasets.

Direct approach closest

		Scenario 1		Scenario 2		Scenario 3	
50	0.24	[0.0000,0.6144]	0.3	[0.0000,0.7017]	0.46	[0.0231,0.8969]
100	0.27	[0.0000,0.6591]	0.28	[0.0000,0.6736]	0.45	[0.0139,0.8861]
150	0.2667	[0.0000,0.6543]	0.2533	[0.0000,0.6346]	0.48	[0.0421,0.9179]
200	0.26	[0.0000,0.6445]	0.26	[0.0000,0.6445]	0.48	[0.0421,0.9179]
250	0.24	[0.0000,0.6144]	0.268	[0.0000,0.6562]	0.492	[0.0538,0.9302]
300	0.2533	[0.0000,0.6346]	0.2533	[0.0000,0.6346]	0.4933	[0.0551,0.9316]
350	0.2514	[0.0000,0.6317]	0.2771	[0.0000,0.6695]	0.4714	[0.0339,0.9090]
400	0.2625	[0.0000,0.6482]	0.275	[0.0000,0.6664]	0.4625	[0.0255,0.8995]
450	0.2622	[0.0000,0.6478]	0.2733	[0.0000,0.6640]	0.4644	[0.0273,0.9016]
500	0.264	[0.0000,0.6504]	0.264	[0.0000,0.6504]	0.472	[0.0344,0.9096]

Logistic regression closest

		Scenario 1		Scenario 2		Scenario 3	
3000	0.0882	[0.0000,0.1868]	0.2861	[0.0806,0.4916]	0.6257	[0.3878,0.8636]
6000	0.058	[0.0138,0.1022]	0.314	[0.1675,0.4604]	0.6281	[0.4687,0.7874]
9000	0.0519	[0.0199,0.0840]	0.2917	[0.1797,0.4038]	0.6563	[0.5343,0.7784]
12000	0.0471	[0.0219,0.0724]	0.275	[0.1819,0.3680]	0.6779	[0.5767,0.7791]
15000	0.0431	[0.0224,0.0638]	0.2626	[0.1822,0.3430]	0.6943	[0.6070,0.7816]
18000	0.0398	[0.0224,0.0572]	0.2596	[0.1871,0.3322]	0.7005	[0.6222,0.7789]
21000	0.0373	[0.0222,0.0524]	0.2609	[0.1938,0.3280]	0.7018	[0.6296,0.7740]
24000	0.0363	[0.0226,0.0501]	0.2627	[0.1998,0.3255]	0.701	[0.6336,0.7684]
27000	0.0353	[0.0227,0.0478]	0.264	[0.2047,0.3234]	0.7007	[0.6372,0.7641]
30000	0.0341	[0.0226,0.0456]	0.2655	[0.2092,0.3219]	0.7003	[0.6402,0.7605]

doi:10.1371/journal.pone.0018560.t003

IBD across the whole data set was mainly caused by the highly bottlenecked British populations at the margin of the distribution. After excluding these two populations, none of the tests was significant. The 'stepping-stone' colonisation is therefore not likely to be the predominant process for the spread of the species in Europe. We suggest that long-distance dispersal must have played a major role, possibly as a consequence of fish transport associated with aquaculture. This is in agreement with Gozlan *et al.* (2010b) [22], who suggested a *P. parva* dispersal model showing dispersal distances of approximately 250 km from the 1970s to the end of the 1990s, followed by shorter dispersal of 20 km on average since 2000. Additional genetic analyses at the country level with greater resolution of the geographical pattern of haplotypes are likely to confirm this two-stepped invasion process.

Evidence of hybridisation

A single individual from the Polish population contained a highly divergent haplotype. The phylogenetic analysis revealed that the sequence is very closely related to a published GenBank sequence of *G. gobio*. The genus *Gobio* belongs to the same cyprinid subfamily as Pseudorasbora, the Gobioninae and is a close European relative of *P. parva* [40]. Despite the close phylogenetic relationship, the two species show very different phenotypic appearances and misidentification is extremely likely given that only adults were sampled. Although laboratory experiments have not confirmed this, based on these results we therefore conclude

that this indicates mitochondrial introgression and if analysis of nuclear data confirms this, we suggest that the invasive *P. parva* is able to hybridise with at least one native European species. This raises further concerns about the threat which *P. parva* poses to native European fish fauna and corroborates experimental evidence that hybrids between *P. parva* and another European cyprinid *Leucaspius delineatus* are possible [70].

Conclusion

The European introduction of *P. parva* resulted from accidental releases from a human-induced faunal translocation [22]. Their European colonisation was initiated by the introduction to a single location or small geographic area it was preceded by, or associated with, the admixture of genetically diverse source populations. This adds to the existing evidence that many invasive populations show the genetic signature of admixture or of multiple introductions [3], [71]. Although the data available did not fully allow us to disentangle the source populations of the invasive populations, we now have a better perspective of the spread of the species within the native range and the introduction of the species into Europe. It remains to be tested how much of the observed phenotypic variation can be attributed to phenotypic plasticity, but the single origin model supported by our data makes it more likely that the disease resistance reported in some populations of *P. parva*, that potentially will lead to devastating consequences for native fishes [28], [29], is an ubiquitous feature of the invasive populations.

Supporting Information

Table S1 Matrix of F_{ST} values of pairwise genetic comparisons between all populations.
(DOC)

Table S2 Matrix of D_{XY} values of pairwise genetic comparisons between all populations.
(DOC)

Table S3 Set of prior distributions based on historical and demographic data. Posterior probabilities of the selected scenarios in DIY ABC: mean, median and mode values and four quantiles of the posterior distribution. Prior and posterior values of mutation rate used in DIY ABC.
(DOC)

References

1. Geller JB, Darling JA, Carlton JT (2010) Genetic Perspectives on Marine Biological Invasions. Annual Review of Marine Science 2: 367–393. (DOI:10.1146/annurev.marine.010908.163745).
2. Le Roux J, Wieczorek AM (2009) Molecular systematics and population genetics of biological invasions: towards a better understanding of invasive species management. Annals of Applied Biology 154: 1–17. (DOI: 10.1111/j.1744-7348.2008.00280.x).
3. Hänfling B (2007) Understanding the establishment success of non-indigenous fishes: lessons from population genetics. Journal of Fish Biology 71: 115–135. (DOI: 10.1111/j.1095-8649.2007.01685.x).
4. Guillemaud T, Beaumont MA, Ciosi M, Cornuet JM, Estoup A (2010) : Inferring introduction routes of invasive species using approximate Bayesian computation on microsatellite data. Heredity 104: 88–99. (DOI:10.1038/hdy.2009.92).
5. Muirhead JR, Gray DK, Kelly DW, Ellis SM, Heath D, et al. (2008) Identifying the source of species invasions: sampling intensity vs. genetic diversity. Molecular Ecology 17: 1020–1035. (DOI: 10.1111/j.1365-294X.2008.03669.x).
6. Lee CE (2002) Evolutionary genetics of invasive species. Trends in Ecology & Evolution 17: 386–391. (DOI:10.1016/S0169-5348(02)02554-5).
7. Brown JE, Stepien CA (2008) Ancient divisions, recent expansions: Phylogeography and population genetics of the round goby Apollonia melanostoma across Eurasia. Molecular Ecology 11: 2598–615. (DOI: 10.1111/j.1365-294X.2008.03777.x).
8. Nei M, Maruyama T, Chakraborty R (1975) : The bottleneck effect and genetic variability in populations. Evolution 29: 1–10.
9. Williamson M, Fitter A (1996) The characteristics of successful invaders. Biological Conservation 78: 163–170. (DOI:10.1016/0006-3207(96)00025-0).
10. Frankham R (2005) Genetics and extinction. Biological Conservation 126: 131–140. (DOI:10.1016/j.biocon.2005.05.002).
11. Poulin J, Weller SG, Sakai AK (2005) Genetic diversity does not affect the invasiveness of fountain grass (Pennisetum setaceum) in Arizona, California and Hawaii. Diversity and Distributions 11: 241–247. (DOI: 10.1111/j.1366-9516.2005.00136.x).
12. Ficetola GF, Bonin A, Miaud C (2008) Population genetics reveals origin and number of founders in a biological invasion. Molecular Ecology 17: 773–782. (DOI: 10.1111/j.1365-294X.2007.03622.x).
13. Roman J, Darling JA (2007) Paradox lost: genetic diversity and the success of aquatic invasions. Trends in Ecology & Evolution 22: 454–464. (DOI:10.1016/j.tree.2007.07.002).
14. Kelly DW, Muirhead JR, Heath DD, Macisaac HJ (2006) Contrasting patterns in genetic diversity following multiple invasions of fresh and brackish water. Molecular Ecology 15: 3641–3653. (DOI: 10.1111/j.1365-294X.2006.03012.x).
15. Haynes GD, Gilligan DM, Grewe P, Nicholas FW (2009) Population genetics and management units of invasive common carp Cyprinus carpio in the Murray–Darling Basin, Australia. Journal of Fish Biology 75: 295–320. (DOI: 10.1111/j.1095-8649.2009.02276.x).
16. Zidana H, Turner GF, van Oosterhout C, Hänfling B (2009) Elevated mtDNA diversity in introduced populations of Cynotilapia afra (Günther 1894) in Lake Malawi National Park is evidence for multiple source populations and hybridization. Molecular Ecology 18: 4380–4389. (DOI: 10.1111/j.1365-294X.2009.04362.x).
17. Kolbe JJ, Glor RE, Schettino LR, Lara AC, Larson A, et al. (2004) Genetic variation increases during biological invasion by a Cuban lizard. Nature 431: 177–181. (DOI:10.1038/nature02807).
18. Rosenthal DM, Ramakrishnan AP, Cruzan MB (2008) Evidence for multiple sources of invasion and intraspecific hybridization in Brachypodium sylvaticum (Hudson) Beauv. in North America. Molecular Ecology 17: 4657–4669. (DOI: 10.1111/j.1365-294X.2008.03844.x).
19. Gozlan RE, Pinder AC, Shelley J (2002) Occurrence of the Asiatic cyprinid Pseudorasbora parva in England. Journal of Fish Biology 61: 298–300. (DOI: 10.1111/j.1095-8649.2002.tb01755.x).
20. Gozlan RE, Britton JR, Cowx I, Copp GH (2010a) Current knowledge on non-native freshwater fish introductions. (Review) Journal of Fish Biology 76: 751–786. (DOI: 10.1111/j.1095-8649.2010.02566.x).
21. Rosecchi E, Thomas F, Crivelli AJ (2001) : Can life-history traits predict the fate of introduced species? A case study on two cyprinids in southern France. Freshwater Biology 46: 845–853. (DOI: 10.1046/j.1365-2427.2001.00715.x).
22. Gozlan RE, Andreou D, Asaeda T, Beyer K, Bouhadad R, et al. (2010b) : Pan-continental invasion of Pseudorasbora parva: towards a better understanding of freshwater fish invasions. Fish and Fisheries 11: 315–340. (DOI: 10.1111/j.1467-2979.2010.00361.x).
23. Britton JR, Davies GD, Harrod C (2010) : Trophic interactions and consequent impacts of the invasive fish Pseudorasbora parva in a native aquatic foodweb: a field investigation in the UK. Biological Invasions 12: 1533–1542. (DOI 101007/s10530-009-9566-5).
24. Britton JR, Davies GD, Brazier M, Pinder AC (2007) : A case study on the population ecology of a topmouth gudgeon Pseudorasbora parva population in the UK and the implications for native fish communities. Aquatic Conservation: Marine & Freshwater Ecosystems 17: 749–759. (DOI: 101002/aqc809).
25. Benejam L, Carol J, Alcaraz C, Garcia-Berthou E (2005) First record of the common bream (Abramis brama) introduced to the Iberian Peninsula. Limnetica 24: 273–274.
26. Vidal O, Garcia-Berthou E, Tedesco PA, Garcia-Marin J (2010) Origin and genetic diversity of mosquitofish (Gambusia holbrooki) introduced to Europe. Biological Invasions 12: 841–851. (DOI 101007/s10530-009-9505-5).
27. Cesco H, Lambert A, Crivelli AJ (2001) Pseudorasbora parva (téléostéen Cyprinidae) espèce invasive nouvel agent du maintien et de la dissémination de l'anguillicolose en France? Parasite 8: 75–76.
28. Gozlan RE, St-Hilaire S, Feist SW, Martin P, Kent ML (2005) Biodiversity – disease threat to European fish. Nature 435: 1046. (DOI101038/4351046a).
29. Gozlan RE, Whipps CM, Andreou D, Arkush KD (2009) Identification of a rosette-like agent as Sphaerothecum destruens a multi-host fish pathogen. International Journal for Parasitology 39: 1055–1058. (DOI101016/jij-para200904012).
30. Berg LS (1949) Ryby presnych vod SSSR i sopredelnych stran 2 ss 477–925 Izd AN SSSR Moskva-Leningrad.
31. Bănărescu P (1999) Pseudorasbora parva (Temmnick et Schlegel 1846) In P. Bănărescu, ed. The freshwater fishes of Europe 5I Cyprinidae 2/I Aula Verl Wiesbaden.
32. Truett GE, Heeger P, Mynatt RL, Truett AA, Walker JA, et al. (2000) Preparation of PCR-quality mouse genomic DNA with hot sodium hydroxide and tris (hotshot). Biotechniques 29: 52–54.
33. Briolay J, Galtier N, Brito RM, Bouvet Y (1998) Molecular Phylogeny of Cyprinidae Inferred from cytochrome b DNA Sequences. Molecular Phylogenetics and Evolution 9: 100–108. (DOI101006/mpev19970441).
34. Ewing B, Hillier L, Wendl MC, Green P (1998) 'Base calling of automated sequencer traces using PHRED I Accuracy assessment'. Genome Research 8: 175–185: (DOI101101/gr83175).
35. Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4 Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Molecular Biology and Evolution 24: 1596–1599. (DOI101093/molbev/msm092).
36. Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG (1997) The CLUSTAL_X windows interface flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Research 25: 4876–4882. (DOI 101093/nar/25244876).
37. Tamura K, Nei M, Kumar S (2004) Prospects for inferring very large phylogenies by using the neighbor-joining method. Proceedings of the National Academy of Sciences U S A 101: 11030–11035. (DOI101073/pnas0404206101).
38. Stamatakis A, Hoover P, Rougemont J (2008) A Fast Bootstrapping Algorithm for the RAxML Web-Servers. Systematic Biology 57: 758–771. (DOI 101080/10635150802429642).

39. Liu HZ, Yang JQ, Tang QY (2010) Estimated evolutionary tempo of East Asian gobionid fishes (Teleostei Cyprinidae) from mitochondrial DNA sequence data. Chinese Science Bulletin 55: 1501–1510.

40. Tang KL, Agnew MK, Hirt MV, et al. (2010) Systematics of the subfamily Danioninae (Teleostei Cypriniformes Cyprinidae) Molecular Phylogenetics and Evolution 57: 189–214. (DOI101016/jympev201005021).

41. Bandelt HJ, Forster P, Röhl A (1999) Median-joining networks for inferring intraspecific phylogenies. Molecular Biology and Evolution 16: 37–48.

42. Broughton RE, Stanley SE, Durrett RT (2000) Quantification of homoplasy for nucleotide transitions and transversions and a re-examination of assumptions in weighted phylogenetic analysis. Systematic Biology 49: 617–627. (DOI 101080/10635150075004934).

43. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic Local Alignment Search Tool. Journal of Molecular Biology 215: 403–410. (DOI101006/jmbi199 09999).

44. Rozas J, Sánchez-DelBarrio JC, Messeguer X, Rozas R (2003) DnaSP DNA polymorphism analyses by the coalescent and other methods. Bioinformatics 19: 2496–2497. (DOI 101093/bioinformatics/btg359).

45. Nei M, Jin L (1989) Variances of the average numbers of nucleotide substitutions within and between populations. Molecular Biology and Evolution 6: 290–300.

46. Lynch M, Crease TM (1990) The analysis of population survey data on DNA sequence variation. Molecular Biology and Evolution 7: 377–394.

47. Hudson RR (1990) Gene genealogies and the coalescent process. 1–42. in Oxford Surveyin Evolutionary Biology Vol 7 Futuyama D, Antonovics J, eds. Oxford University Press Oxford.

48. Beerli P (2006) Comparison of Bayesian and maximum-likelihood inference of population genetic parameters. Bioinformatics 22: 341–345. (DOI101093/bioinformatics/bti803).

49. Kimura M (1980) A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. Journal of Molecular Evolution 16: 111–120.

50. Ramachandran S, Deshpande O, Roseman CC, Rosenberg NA, Feldman MW, et al. (2005) Support from the relationship of genetic and geographic distance in human populations for a serial founder effect originating in Africa. Proc Nat Acad Sci U S A 102: 15942–15947. (DOI 101073/pnas0507611102).

51. Herborg LM, Weetman D, van Oosterhout C, Hänfling B (2007) Genetic population structure and contemporary dispersal patterns of a recent European invader the Chinese mitten crab Eriocheir sinensis. Molecular Ecology 16: 231–242. (DOI 101111/j1365-294X200603133x).

52. Wright S (1943) Isolation by distance. Genetics 28: 114–138.

53. Jensen JL, Bohonak AJ, Kelley ST (2005) Isolation by distance web service. BMC Genetics 6: 13. v316 http://ibdws.sdsu.edu/.

54. Bănărescu P, Nalbant T (1965) Studies on the systematics of Gobioninae (Pisces Cyprinidae). Rev Roum de Biol Ser de Zoologie 10: 219–229.

55. Cornuet J, Santos F, Beaumont MA, Robert CP, Marin J, et al. (2008) Inferring population history with DIY ABC a user-friendly approach to approximate Bayesian computation. Bioinformatics 24: 2713–2719. (DOI101093/bioinformatics/btn514).

56. Durand JD, Tsigenopoulos CS, Unlu E, Berrebi P (2002) Phylogeny and biogeography of the family Cyprinidae in the Middle East inferred from cytochrome b DNA - Evolutionary significance of this region. Molecular Phylogenetics and Evolution 25: 218-218. (DOI101006/mpev20011040).

57. Watanabe K, Iguchi K, Hosoya K, Nishida M (2000) Phylogenetic relationships of the Japanese minnows Pseudorasbora (Cyprinidae) as inferred from mitochondrial 16S rRNA gene sequences. Ichthyological research 47: 43–50. (DOI 10.1007/BF02674312).

58. Hänfling B, Brandl R (1998) Genetic differentiation of the bullhead Cottus gobio L across watersheds in central Europe evidence for two taxa. Heredity 80: 110–117. (DOI101046/j1365-2540199800279x).

59. Hänfling B, Kollmann J (2002) An evolutionary perspective of biological invasions. Trends in Ecology & Evolution 17: 545–546. (DOI101016/S0169-5347(02)02644-7).

60. Barson NJ, Cable J, van Oosterhout C (2009) Population genetic analysis of microsatellite variation of guppies (Poecilia reticulata) in Trinidad and Tobago evidence for a dynamic source-sink metapopulation structure founder events and population bottlenecks. Journal of Evolutionary Biology 22: 485–497. (DOI101111/j1420-9101200801675x).

61. Willing EM, Bentzen P, van Oosterhout C, Hoffmann M, Cable J, et al. (2010) Genome-wide single nucleotide polymorphisms reveal population history and adaptive divergence in wild guppies. Molecular Ecology 19: 968–984. (DOI 101111/j1365-294X201004528x).

62. Saillant E, Patton JC, Gold JR (2005) Genetic Variation Kinship and Effective Population Size in a Captive Population of the Endangered Cape Fear Shiner Notropis mekistocholas. Copeia 1: 20–28. (DOI 101643/CG-04-200R1).

63. Hänfling B, Weetman D (2006) Concordant genetic estimators of migration reveal anthropogenically enhanced source-sink population structure in the River Sculpin Cottus gobio. Genetics 173: 1487–1501. (DOI101534/genetics105054296).

64. Yang L, He S (2008) Phylogeography of the freshwater catfish Hemibagrus guttatus (Siluriformes Bagridae) Implications for South China biogeography and influence of sea-level changes. Molecular Phylogenetics and Evolution 49: 393–398. (DOI101016/jympev200805032).

65. Lin C, Lin H, Wang J, Chao S, Chiang T (2010) Phylogeography of Hemibarbus labeo (Cyprinidae) secondary contact of ancient lineages of mtDNA. Zoologica Scripta 39: 23–35. (DOI 101111/j1463-6409200900407x).

66. Hua X, Wang W, Yin W, He Q, Jin B, et al. (2009) Phylogeographical analysis of an estuarine fish Salanx ariakensis (Osmeridae Salanginae) in the north-western Pacific. Journal of Fish Biology 75: 354–367. (DOI 101111/j1095-8649200902323x).

67. Jamandre BW, Durand JD, Tzeng WN (2009) Phylogeography of the flathead mullet Mugil cephalus in the north-west Pacific as inferred from the mtDNA control region. Journal of Fish Biology 75: 393–407. (DOI 101111/j1095-8649200902332x).

68. Liu J, Brown CL, Yang T (2009) Population genetic structure and historical demography of grey mullet Mugil cephalus along the coast of China inferred by analysis of the mitochondrial control region. Biochemical Systematics and Ecology 37: 556–566. (DOI101016jbse200909002).

69. Naylor RL, Goldburg RJ, Primavera JH, Kautsky N, Beveridge MCM, et al. (2000) Effect of aquaculture on world fish supplies. Nature 405: 1017–1024. (DOI101038/35016500).

70. Gozlan RE, Beyer K (2006) Hybridisation between Pseudorasbora parva and Leucaspius delineatus. Folia Zoologica 55: 53–60.

71. Roman J (2006) Diluting the founder effect cryptic invasions expand a marine invader's range. Proceedings of the Royal Society B 273: 2453–2459. (DOI101098/rspb20063597).