Mitochondrial DNA Variation and Introgression in Siberian Taimen *Hucho taimen*

Evgeniy S. Balakirev1,2*, Nikolai S. Romanov2, Pavel B. Mikheev3, Francisco J. Ayala1

1 Department of Ecology and Evolutionary Biology, University of California Irvine, Irvine, California, United States of America, 2 A. V. Zhirmunsky Institute of Marine Biology, Far Eastern Branch, Russian Academy of Science, Vladivostok, Russia, 3 Amur River Salmon Sector, Pacific Research Fisheries Centre (TINRO-Centre), Khabarovsk Branch, Khabarovsk, Russia

Abstract

Siberian taimen *Hucho taimen* is the largest representative of the family Salmonidae inhabiting rivers of northern Eurasia. The species is under intensive aquaculture activity. To monitor natural taimen populations we have sequenced a portion (8,141 bp) of the mitochondrial (mt) genome in 28 specimens of *H. taimen* from six localities in the Amur River basin. Nucleotide variability is low (π = 0.0010), but structured in two divergent haplotype groups. A comparison of the data with the GenBank *H. taimen* mt genome (HQ897271) reveals significant differences between them in spite of the fact that the fish specimens come from neighboring geographical areas. The distribution of divergence is non-uniform with two highly pronounced divergent regions centered on two genes, ND3 and ND6. To clarify the pattern of divergence we sequenced the corresponding portion of the mt genome of lenok *Brachymystax tumnensis* and analyzed the GenBank complete mt genomes of related species. We have found that the first and second divergent regions are identical between the GenBank *H. taimen* and two lenok subspecies, *B. lenok* and *B. lenok tsinlingensis*, respectively. Consequently, both divergent regions represent introgressed mtDNA resulting from intergeneric hybridization between the two lenok subspecies and *H. taimen*. Introgression is, however, not detected in our specimens. This plus the precise identity of the introgressed fragments between the donor and the recipient GenBank sequence suggests that the introgression is local and very recent, probably due to artificial manipulations involving taimen – lenok intergeneric hybridization. Human-mediated hybridization may become a major threat to aquatic biodiversity. Consequently we suggest that due attention needs to be given to this threat by means of responsible breeding program management, so as to prevent a potential spread of hybrid fishes that could jeopardize the resilience of locally adapted gene pools of the native *H. taimen* populations.

Introduction

Siberian taimen *Hucho taimen* (Pallas, 1773) is the world’s largest salmonid fish, reaching up to 2 m in length and 105 kg in weight [1–3]. The maximum size and weight of the species are less in the Amur River: 1.5 m and 10 kg (up to 80 kg), respectively [4]. The long life span (20–29 years) and relatively high age at maturity (7–9 years) indicate that the species is vulnerable to overfishing. Indeed, Siberian taimen has significantly declined in abundance throughout their range in the large rivers of the Volga/Caspian, Arctic, and Pacific drainages of northern Eurasia and it is now included as vulnerable species in the Mongolian, Russian, and Chinese red lists [5–7].

The unique biological features and severe decline of taimen populations have stimulated intensive genetic investigations of the species. Intraspecific genetic variation in Siberian taimen has been investigated by means of chromosome analysis [8], allozymes [9], microsatellites [10–13], AFLP markers [14–15], and DNA sequences [16–18]. The level of intraspecific variability detected is generally low and correlated with an increase in habitat degradation and excessive fishing [12], [15]. Significant population differentiation has been found at a large geographic scale, extending through the main Siberian river basins [16–17], as well as at a much smaller geographic scale, between different tributaries within the Amur (Heilongjiang) River basin [12], [13], [15]. Significant genetic differentiation along with low nucleotide sequence diversity have also been found in the European huchen, *H. hucho* (Linnaeus, 1758) [19–20]. This population structure might be explained by the relatively low dispersion capability of the fish and/or the homing like behavior suggested for *H. hucho* [1]. In *H. taimen*, individual home range, obtained with radio and acoustic telemetry, is 0.5–93.2 km (mean ± SD, 23±25.7 km) [21]. Anthropogenic factors may also reduce gene flow between taimen populations [12], [15].

The aquaculture of *H. taimen* is under intensive development; it includes artificial rearing and reproduction [22–24], hybridization [25–28], and propagation [29]. Artificial intergeneric hybrids between *H. taimen* and lenok *Brachymystax lenok* (Pallas, 1773) (which represent the most recent split (~19.9 MY) between Salmonidae genera; [30]) have been reported [25–27]. Heterosis was observed for the hybridization between *H. taimen*♀ and *B. lenok♂ [26]; however the opposite parental combination (*H. taimen♂ and *B. lenok♀) yielded hybrid incompatibility [27].
H. taimen and B. lenok are sympatric in the Amur River drainages [4], [31] but natural hybrids between them are very rare (our own personal observations as well as communications from Dr. Antonov A.L., Institute of Water and Ecological Problems, Far Eastern Branch, Russian Academy of Sciences, Khabarovsk, Russia; and Dr. Ma B., Heilongjiang River Fisheries Research Institute, Chinese Academy of Fishery Sciences, Harbin, China). Natural hybrids between taimen and lenok have, however, been described in Chinese literature [32]. Using mitochondriond DNA (mtDNA) restriction fragment length polymorphism analysis, Shedko [33–34] detected alien haplotypes in lenok, supposedly due to hybridization between lenok and taimen. Indeed, one of the studied lenok specimens had the mtDNA haplotype typical for taimen, but morphologically showed intermediate characteristics between lenok and taimen, suggesting a possible hybridization event between lenok and taimen. Rare hybrid individuals have been detected only in the Khanka Lake basin [33–34]. This area is characterized by relatively high human population density, so that human-mediated hybridization is possible. Hybrid individuals were not detected in any other Primorye Territory sampling points where lenok and taimen occur in sympatry [33–34]. It has been shown that females of taimen H. taimen, different to most other salmonid fishes, do not cover their eggs immediately after having spawned, but rest for a variable number of minutes before covering them [35]. They share this unusual trait with females of lenok B. lenok [36]. Thus, the female reproductive (spawning) behaviors of lenok B. lenok and taimen H. taimen have some similarities which suggests that breeding between these species is theoretically possible.

The H. taimen – B. lenok putative hybrid individuals were characterized by mtDNA restriction fragment length polymorphism analysis [33–34], sequence – related amplified polymorphisms [26], and microsatellite markers [27]. However, direct evidence of hybridization based on the nucleotide sequences was not obtained. Among the different types of nucleotide sequence markers, mtDNA is particularly appropriate to reveal and describe possible introgression in hybrid individuals (e.g., [37–30]). Indeed, several studies of gene flow between closely related species have found that some parts of the genome introgress more easily than others. For example, mtDNA tends to introgress more readily than nuclear DNA (nuDNA) (e.g., [39–42]). It has been suggested that mtDNA may not include loci that contribute to hybrid unfitness or positive assortative mating, and that most mitochondrial genes, unlike nuclear genes, may function fairly well in the genetic background of a related species [43]; see [41–42] for further potential explanations for stronger signatures of introgression of mtDNA compared to nuDNA.

Hybridization and propagation are common in H. taimen aquaculture practice [22–29]. Consequently, negative genetic effects are expected (see Discussion) and should be under control to prevent deterioration of natural taimen populations. In this study, we have sequenced a fragment of the mt genome (8,141 bp) in 28 individuals of H. taimen from six localities of the Amur River basin to test whether B. lenok alleles have introgressed into populations of H. taimen. We found all specimens collected free from introgression. However, the mtDNA of the H. taimen specimen (GenBank HQ97271) from the Hutou range (Ussuri River, China) clearly indicates contemporary introgression due to hybridization with lenok subspecies, B. lenok and B. lenok tsinlingensis Li, 1966. We conclude that genetic monitoring presents an important and reliable approach to identify hybrid individuals and to assess the threat from hybridization to the genetic integrity of H. taimen natural populations.

**Materials and Methods**

**Fish Samples**

The 28 specimens of H. taimen were collected from six rivers: Nora, Bikin, Manoma, Anyuy, Sutara, and Khor, all from the Amur River basin. A single specimen of blunt-nosed lenok Brochymenichthys taimensis Mori, 1930 was collected from the Bikin River. Sampling locations of H. taimen are shown in Fig. S1. The river names, basins, sample sizes, and coordinates are presented in Tables S1 and S2. Additionally we used full mt genomes from GenBank: 1) H. taimen (HQ97271), 2) H. bleekeri Kimura, 1934 (HM804473), 3) B. lenok (JQ686730), and 4) B. lenok tsinlingensis (JQ675732, JQ686731). The nucleotide sequences of Salmo salmo Valenciennes, 1814 (S. salmo Valenciennes, 1814 is a junior synonym of S. salar Linnaeus, 1758) (U12143, AF137010), S. trutta Linnaeus, 1758 (AM910409), S. gairdneri (Mitchell, 1814) (AF154850), and S. alpinus Linnaeus, 1758 (AF154851) were used as outgroups. Full mt genomes of related salmonids, included in the present analysis as outgroup taxa, were selected based on previous molecular evidence of close relationship to H. taimen [30] and screening of nucleotide sequences available in GenBank.

**Ethics Statement**

H. taimen and B. taimensis are listed as “vulnerable” species in the Red Data Book of some constituent territories of the Russian Federation. However, they are considered as “commercial” species in the Amur Region, Primorye Territory, Khabarovsk Territory, and Jewish Autonomous Region, where we collected the specimens for the present study. Limited fishery is officially permitted in these areas. The described field studies were based on the quota limit obtained from the Federal Agency for Fishery of the Russian Federation (order #1172, November 28, 2011; signed by the FAF Deputy Director, V. I. Sokolov; see details, http://www.garant.ru/products/ipo/prime/doc/70002836/). The sampling points are located beyond the protected territories of the Amur River basin (see map and details, http://amur-heilong.net/http04_econet_pas/0405RussiaPas.html). The field studies did not involve endangered or protected species. The locations of the field studies are not privately-owned or protected. The fishes were collected with non-lethal gear and sacrificed after anaesthetization using a solution (75–200 mg/l) of tricaine methanesulfonate (ethyl 3-aminobenzoate methanesulfonate, MS-222; Sigma-Aldrich, E10521) or clove oil (eugenol; Sigma-Aldrich, W246719). The present field study was approved by the Federal Agency for Fishery of the Russian Federation, which has the highest decision authority concerning fish care and use and should be considered as an equivalent to the Institutional Animal Care and Use Committee.

**DNA Amplification, Cloning, and Sequences**

Total genomic DNA was extracted using the DNeasy Blood and Tissue kit (Qiagen, Hilden, Germany) from 96% ethanol-preserved muscle tissue of H. taimen. The procedures for DNA amplification, cloning, and sequencing have been described previously [44–45]. The mtDNA fragment (8.1 kb) was amplified with primers designed with the program mitoPrimer, version 1 [46]. The primers are listed in Table S3. The region of H. taimen mt genome, with the corresponding coordinates of mt genomes of H. taimen [47], H. bleekeri [48] and B. lenok tsinlingensis [49] are given in Table S4. The PCR reactions were carried out in final volumes of 25 µl using Takara Ex Taq” in accordance with the manufacturer’s description (Takara Biotechnology Co., Ltd.). The PCR reaction mixtures were placed in a DNA thermal cycler (Eppendorf,
Mastercycler Gradient), incubated 5 min at 95° and subjected to 32 cycles of denaturation, annealing, and extension: 94° for 30 sec, 54°–60° (depending on the Tm of the particular primer pair, see Table S3) for 30 sec, and 72° for 1.0 min, with a final 7-min extension period at 72°. The PCR products were sequenced by the dideoxy chain-termination technique using Dye Terminator DNA sequencer (Perkin Elmer). The sequences of both strands were determined, using overlapping internal primers spaced 500 nucleotides, on average. At least two independent PCR amplifications were sequenced in both directions to correct for possible errors. The heteroplasmic region was resolved by cloning using the Topo-TA cloning chemistry (Invitrogen). The sequences were annotated with the program DOGMA [50] and deposited in GenBank under accession numbers KC920394–KC920422, KC936238, and KC936240-KC936267.

DNA Sequence Analysis

The sequences were assembled using the program SeqMan (Lasergene, DNASTAR, Inc., version 10.1). Multiple alignments were carried using the program CLUSTAL W [51]; overlapping regions (ND4L - ND4, ND3- ND6; see Table S4) were excluded from the analysis. The computer programs DnaSP, version 5 [52] and PROSEQ, version 2.9 [53] were used for most intraspecific analyses.

Model-based phylogeny reconstructions were performed with concatenated sequence alignments using the neighbor-joining (NJ), maximum-likelihood (ML), and Bayesian algorithms, using respectively the programs MEGA, version 5 [54], GARLI 2.0 [55], and MrBayes 3.2 [56]. For all reconstructions, the best-fit model of nucleotide substitution was chosen with the Akaike Information Criterion (AIC) and the Bayesian information criterion (BIC) in MEGA and jModelTest, version 2 [57] (Table S5). Maximum likelihood bootstrap analyses [58] consisted of 1000 replicates. In ML analysis, only the model specification settings were adjusted according to the respective concatenated dataset, while all other GARLI settings were left at their default value. In Bayesian inference, Markov Chain Monte Carlo (MCMC) was run under a general-time-reversible model plus gamma plus I (GTR+G+I). Analyses were performed as four independent MCMC runs each with three incrementally heated chains and a single cold chain for two million generations. Output trees and data were sampled every 500 generations. Omitting the first 5,000 (25%) burn-in trees, the remaining trees were used to estimate the 50% majority rule consensus trees and the Bayesian posterior probabilities. At the end of the run there were not any trends in a plot of generation versus the log probability of the observed data (the log likelihood value). A convergence diagnostic, the Potential Scale Reduction Factor (PSRF) [59] was between 1.000 and 1.003 indicating a good convergence diagnostic, the Potential Scale Reduction Factor (PSRF) [59] was between 1.000 and 1.003 indicating a good.

Partitioned analyses were performed with GARLI, which allowed the overall rate to be different across each separate mtDNA protein-coding gene included in the present analysis. The partitioned dataset treats each locus separately and each with its own substitution model, while the unpartitioned dataset was regarded as one partition. Unpartitioned data sets included the entire 8,141 bp concatenated dataset (Table S4) and the protein-coding genes only (7,608 bp without stop-codons). Partitioned analyses included (1) each protein-coding gene as a separate partition, (2) the same as (1) but excluding ND3+ ND6, (3) ND3 only, (4) ND6 only, and (5) ND3+ ND6 only (see Results). The topologies obtained with the NJ and ML methods, as well as, by Bayesian inference, were congruent and received high support in most nodes. To be conservative we show the lowest bootstrap values obtained with the ML method.

Recombination Analysis

The alignments were analyzed for evidence of recombination using various recombination detection methods implemented in the program RDP3 [60]. The parental and recombinant sequences were determined using the VisRD method [61], modified version of PHYLPRO [62], and EEEP [63] also implemented in RDP3 (default settings).

Results

Nucleotide Diversity and Divergence

We sequenced a 8,141 bp fragment of mitochondrial genome in 28 H. taimen specimens. The fragment included eight protein-coding genes: COI, COIII, ND3, ND4L, ND4, ND5, ND6, and CYTB, and eight transfer RNAs genes (see Table S4 for coordinates). Fig. 1 shows 32 nucleotide substitution sites (5 singletons and 27 informative sites) detected in the 28 H. taimen specimens (6 sites in COI, 3 sites in COIII, 2 sites in ND3, 7 sites in ND4, 5 sites in ND5, 3 sites in ND6, and 6 sites in CYTB). No polymorphic sites were detected in ND4L or in the transfer RNAs genes. No length polymorphisms were found. Total nucleotide diversity was low (π = 0.0010); the results are in accordance with data on intraspecific variability in H. taimen [16-17] and close species, H. hucho [20] and H. bleekeri [64].

There were 25 synonymous and 7 replacement polymorphic sites in the coding regions of the eight genes studied (7,608 bp; 2,536 codons totally). Replacement substitutions were detected in ND3 (1 site), ND4 (2 sites), ND5 (2 sites), and CYTB (2 sites). All replacements were non-singletons (Fig. 1). Total nucleotide diversity in the coding region was π = 0.0011. Synonymous polymorphism (π = 0.0035) was 17.5 times higher than nonsynonymous variability (π = 0.0002), indicating a high level of negative selection on amino acid substitutions in the mtDNA of Siberian taimen.

Strong haplotype structure was detected with two main haplotype groups. The first and second group include 23 and 5 sequences, respectively (separated by a horizontal line in Fig. 1) with 9 fixed single nucleotide differences between them. The genetic structure is completely congruent for all gene regions studied. The second haplotype group was associated with the replacement polymorphism at positions 2456, 2985, 3408, 5313, 5874, 7373, and 7796 (Fig. 1). The difference between the groups based on the full concatenate (8,141 bp) is highly significant (Fst = 0.7059, P < 0.0001); total sequence divergence (Dxy) between the groups is 0.0021 ± 0.0005. Both haplotype groups, however, can be present within a single locality (Table S1) and have similar values of nucleotide diversity estimates (Table 1).

A comparison of the data obtained in the present study with the publicly available H. taimen mt genome (GenBank HQ897271; the specimen is from Hutou range, Ussuri River, China; thereafter H. taimen (Hutou)) revealed significant divergence between them: the total sequence divergence (Dxy) is 0.0108 ± 0.0013, is five times higher than the divergence between the two groups of sequences found in our specimens, 0.0021 ± 0.0005. The distance between our specimens and H. taimen (Hutou) is below the range of divergence between species, but significantly higher than the divergence between the two groups of sequences in our specimens, and approaches the divergence between two Hucho species, H. taimen.
Fig. 1. Nucleotide polymorphism in the mtDNA of *Hucho taimen*. Numbers above the top sequence give the position of segregating sites. The samples are listed sequentially according to their genetic similarity. Nucleotides are numbered from the beginning of our sequence, position 6,663 in the complete mitochondrial genome of *H. taimen* [47]. Amino acid replacement polymorphisms are marked with asterisks below the reference sequence (Ht5). Dots indicate the same nucleotide as the reference sequence. The coordinates for the samples are listed sequentially according to the same nucleotide as the reference sequence. The dots indicate the same nucleotide as the reference sequence. The coordinates for the fragments used for tree reconstruction (Fig. 4). The trees based on ID1 and ID2 separately showed *H. taimen* (Hutou) identical to *B. lenok* (JQ686730) or to *B. lenok tsinlingensis* (JQ675732, JQ686731), respectively. For the full mtDNA fragment investigated the average divergence between *H. taimen* (Hutou) and the two subspecies of lenok, *B. lenok* (JQ686730) and *B. lenok tsinlingensis* (JQ675732, JQ686731), respectively. For the full mtDNA fragment investigated the average divergence between *H. taimen* (Hutou) and the two subspecies of lenok, *B. lenok* (JQ686730) and *B. lenok tsinlingensis* (JQ675732, JQ686731), respectively. For the full mtDNA fragment investigated the average divergence between *H. taimen* (Hutou) and the two subspecies of lenok, *B. lenok* (JQ686730) and *B. lenok tsinlingensis* (JQ675732, JQ686731), respectively.

Thus, there are two regions (ID1 and ID2) within the mtDNA fragment studied which show sharply discordant phylogenetic signals between *Hucho* and *Brachymystax*. As a consequence, the position of *H. taimen* (Hutou) was sharply different, depending on the fragments used for tree reconstruction (Fig. 4). The trees based on ID1 and ID2 separately showed *H. taimen* (Hutou) identical to *B. lenok* (JQ686730) or to *B. lenok tsinlingensis* (JQ686731), respectively (Fig. 4A and Fig. 4B). The tree based on ID1 plus ID2 displayed *H. taimen* (Hutou) between the two lenok species (Fig. 4C). On the tree excluding the ID1 and ID2 regions, *H. taimen* (Hutou) was within the same cluster as the other *H. taimen* specimens (Fig. 4D); the sequence divergence between them was very low, $D_{xy} = 0.0007 \pm 0.0002$, three times lower than the divergence between the two groups of sequences detected in our specimens ($D_{xy} = 0.0021 \pm 0.0005$). Thus, most of the mtDNA fragment of *H. taimen* (Hutou) has obvious similarity to the *H. taimen* sequences obtained in our study, whereas the ID1 and ID2 regions have unexpected similarity to *Brachymystax* subspecies, and could be explained by introgression of mtDNA resulting from hybridization between lenok and taimen. Other salmonids included in this analysis (*S. salmo*, *S. trutta*, *S. fontinalis*, and *S. alpinus*) did not show any visible discordance in the level of divergence between the ID1 and ID2 regions and the rest of the mtDNA (Fig. 4).

Recombination

We suggest that the phylogenetic inconsistencies might reflect hybridization event(s) between *H. taimen* (Hutou) and the two *Brachymystax* subspecies, which might have resulted in interspecific recombination of mitochondrial DNA. The method of Hudson and Kaplan [68] failed to reveal any signal of recombination when we analyzed our specimens. However, two recombination events were detected when we add *H. taimen* (Hutou) in the recombination analysis. We therefore analyzed the mtDNA alignments for evidence of recombination using various recombination detection methods implemented in the program RDP3 [60] (Table 2). All seven methods detected two recombination events in *H. taimen* (Hutou) within the studied region (Fig. 5), with high statistical support (Table 2). For recombination event 1, the major and minor parents were *H. taimen* (our specimens) and *B. lenok*, respectively.
respective (breakpoint positions 2116–2561; \( P = 2.904 \times 10^{-25} \)); whereas for the recombination event 2 the major and minor parents were \( H.\) taimen (our specimens) and \( B.\) lenok tsinlingensis, respectively (breakpoint positions 6188–6870; \( P = 8.528 \times 10^{-14} \)) (Fig. 5).

Thus, the recombinational analysis confirms that the mtDNA of \( H.\) taimen (Hutou) is a recombinant product between the mtDNA of \( H.\) taimen and each of two lenok subspecies, \( B.\) lenok and \( B.\) lenok tsinlingensis. The location of the recombinant fragments corresponds to the ID1 and ID2 regions described above (see Figs. 3 and 5). Interestingly, recombination was not detected in our \( H.\) taimen specimens, which, together with full identity of the introgressed fragments between \( H.\) taimen (Hutou) and the two \( Brachymystax\) subspecies, suggests contemporary local introgression. These results are relevant concerning the DNA barcoding for fishes (and possibly other organisms where introgression may have happened). Interspecific mtDNA recombination such as detected in the \( ND3\) and \( ND6\) genes of \( H.\) taimen (the present data) and \( ND1\) gene of \( Salmo\) [77] and \( Salvelinus\) [78] (see Discussion), may importantly limit the usefulness of mtDNA barcoding for species identification.

**Discussion**

Introgression is frequent in fish but most observations are based on indirect evidence.

There are multiple examples of introgression in salmonid (as well as in many others) fishes (review in [38], [79–80]). There are also multiple evidence of mtDNA recombination in plants, fungi, and animals, including humans (review in [81–82]). However, instances of interspecific mtDNA recombination have been rarely detected in hybridizing conifers [83], salmonids, \( Salmo\) and \( Salvelinus\) [77–78], and primates [84], which could probably be explained by fact that in most cases the data obtained involve relatively short DNA fragments or indirect genetic markers like allozymes, PCR-RFLP or microsatellites. The indirect approaches, albeit may yield correct conclusions about introgression, do not elucidate the precise architecture of the introgression events. The use of full (or big segments of) mtDNA genome allows to reveal and analyze introgression more precisely.

**Localization and Size of the Introgressed Fragments**

Using the 8,141 bp fragment of mt genome we detected two recombinant fragments, which could be accounted for by introgression of mtDNA due to hybridization between each of two subspecies of lenok, \( B.\) lenok and \( B.\) lenok tsinlingensis, and Siberian taimen \( H.\) taimen. The introgression is limited to two relatively small regions (around 0.9 kb) including two genes, \( ND3\) and \( ND6\). Our results are consistent with recent data from a genome scan of the bivalve mollusk *Mytilus galloprovincialis* [83]. Using AFLP markers Gosset and Bierne [85] detected a small fraction (around 2%) of outlier loci with high \( F_{ST}\) and proved that the high divergence was due to differential introgression of alleles from the sister-hybridizing species *Mytilus edulis*. Consequently, introgression would seem possible and frequent enough to account for “genomic islands of differentiation” [67], which are often interpreted as a result of spatially heterogeneous selection [86].

**The ND6 Gene is Usually Excluded from Phylogenetic Analysis**

The \( ND6\) mtDNA gene is usually excluded from phylogenetic analyses of vertebrates [87–88]. The reasons are diverse: 1) \( ND6\) is the only protein-coding gene encoded on the light strand and consequently has quite different evolutionary properties from those of the other 12 protein genes, making it an inappropriate mix with model-based methods; 2) the presence of many indels; 3) heterogeneous base composition; 4) consistently poor phylogenetic performance; and finally 5) difficulties with \( Kappa\) (transition-transversion ratio) estimation. Consequently, \( ND6\) is usually ignored (references include a large majority of papers in the field of “mitogenomics” published during the last twelve years). The \( ND6\) gene was also excluded by Wang et al. [48] and Si et al. [49] investigating mt genomes of taimen \( H.\) bleekeri and lenok subspecies.
B. lenok and B. lenok tsinlingensis. The mt genome of H. taimen (Hutou) was not analyzed using a phylogenetic approach [47]. Our results show that ND6 (along with ND3) provides clear evidence of gene flow between Brachymystax subspecies and H. taimen and could be used as a valuable marker to detect intergeneric hybridization between them. Similar patterns could occur in other vertebrates, suggesting the ND6 gene should be included in at least preliminary phylogenetic analysis. Otherwise an interesting feature of fish mitochondrial genome evolution with important practical consequences might be missed.

Is Introgression due to Natural or Artificial Hybridization?

We detected full identity in two short introgressed mtDNA fragments between H. taimen (Hutou) and two lenok subspecies, B. lenok and B. lenok tsinlingensis. Full identity suggests contemporary introgression in both cases (B. lenok - H. taimen and B. lenok tsinlingensis - H. taimen). A remaining question concerns the mechanism of hybridization. Natural hybrids between H. taimen and B. lenok have been registered in the areas of sympathy [32–34]. Consequently, the introgressed mtDNA fragments in H. taimen (Hutou) could be explained by natural hybridization events. However, the involvement of B. lenok tsinlingensis can hardly be explained by natural hybridization, because this subspecies is endemic to the Huanghe River basin [3], [49], which is disjunct from the Amur River basin at the present time. Moreover, this subspecies is rare; it has been listed in the China Red Data Book of Endangered Animals [6]. It might be suggested that the rate of mtDNA evolution is extremely low in H. taimen and that...
introgressed fragment is a “frozen” relict from Pleistocene time, when the Huanghe and Amur River basins were connected [89–90]. We consider this alternative hypothesis (low rate of mtDNA evolution in *H. taimen*) highly improbable, especially taking into account the data of Xia et al. [91] on the phylogeographic structure of *B. lenok* (which is close to *B. lenok tsinlingensis*). Using two mtDNA fragments (control region, 835 bp and *CYTB*, 1,069 bp) they detected strong geographic differentiation among *B. lenok* populations. Particularly, they found significant difference between lenok inhabiting the Huanghe and Amur River basins, with the

Figure 3. Sliding-window plot of divergence along the concatenate mitochondrial DNA sequences (8,141 bp) between the GenBank *Hucho taimen* mt genome (HQ897271) and the *H. taimen* sequences obtained in the present study. Window sizes are 250 nucleotides with 25-nucleotide increments. The order of the mitochondrial genes is schematically shown at bottom. Note the two significant peaks of divergence centered on two genes, ND3 and ND6.

doi:10.1371/journal.pone.0071147.g003

Figure 4. Phylogenetic trees based on different fragments of the mitochondrial DNA sequences: (A) first island of divergence only, (B) second island of divergence only, (C) first plus second islands of divergence, (D) without the two islands of divergence. The tree topologies obtained with Maximum likelihood and Bayesian inference are congruent. Two sequences of *H. taimen*, Ht5 and Ht16, representing haplotype group 1 and 2, are included. The *Salvelinus* and *Salmo* sequences (GenBank accession numbers in Materials and Methods) are used as outgroups. Note the changed position of *H. taimen* from the Hutou range (GenBank HQ897271), depending on the region used for the tree reconstruction.

doi:10.1371/journal.pone.0071147.g004
average sequence divergence 1.93%. Consequently, the full identity of the introgressed fragments identified in *H. taimen* (Hutou) could not likely be a product of natural hybridization, but rather it would likely be due to artificial hybridization experiments under laboratory conditions. The absence of introgression in the *H. taimen* specimens collected from close localities in the Ussuri River (but in Russian territory) further supports this hypothesis. Indeed, a technique for artificial reproduction between *H. taimen* and *B. lenok* has been developed [25]. Artificial intergeneric hybrids have been reported between *H. taimen* and *B. lenok* [25–27] but not between *H. taimen* and *B. lenok tsinlingensis*. The fact that the hybrid *H. taimen* individual was captured in a natural environment (Hutou range, Ussuri River, Heilongjiang Province) [47] suggests the possibility of hybrid individuals escaping from fish rearing farms.

The Negative Consequences of Human-mediated Hybridization

Interspecific hybridization has been described for many fish species [review in [79], [92–94]]. Hybrid speciation is widespread in plant and animal evolution (e.g., [95–97] and references therein) and mosaic genomes due to hybrid speciation are common in microorganisms, plants and animals [98]. Natural introgressive hybridization does not require special attention in terms of conservation and can indeed play a positive role as a source of variability [99–100]. However, human-mediated hybridization represents a major threat to aquatic biodiversity worldwide, leading to the erosion of local genetic diversity with loss of adaptive potential and negative fitness consequences [review in [92], [101–103]]. Captive propagation can help reduce the risk of short-term extinction, but ecological and genetic interactions between wild and hatchery fishes may have negative consequences for wild populations, including loss of genetic variability, loss of adaptations, and change of population make-up and structure (review in [104–106]). Thus, artificial breeding, hybridization, and captive propagation cannot substitute for conserving the species in the wild.

There have been some important successes for *H. taimen* in aquaculture practice and propagation [22–29], but these activities should proceed with great caution and in conjunction with a full assessment of risks and benefits. Human-induced genetic erosion of *H. taimen* populations could be provoked by releases of farm-reared *H. taimen* and *B. lenok* — *B. lenok* hybrids; genetic and ecological risks associated with the releases should not be neglected in management and policy. The significant threat of human-mediated hybridization demands a responsible breeding program and management, in order to prevent potential spread of hybrid fishes that can jeopardize the resilience of locally adapted gene pools of native *H. taimen* populations. Genetic monitoring offers an important and reliable approach to discriminate hybrid individuals and to assess the threat from hybridization to the genetic integrity of *H. taimen* natural populations.

The mtDNA data clearly show hybridization and introgression between *H. taimen* and two lenok subspecies, *B. lenok* and *B. lenok tsinlingensis*. Further evidence from biparentally inherited nuclear DNA will be required to critically evaluate the revealed patterns, in particular to determine what fraction of the extant *H. taimen* genome has been impacted by gene flow between *H. taimen* and lenok subspecies.

### Supporting Information

**Figure S1** Sampling locations of *Hucho taimen*. Arabic numerals (1–6) correspond to sample sites. The sample site from China (Hutou range, the Ussuri River) is marked by an asterisk. See Table S2 for river names, basins, sample sizes, and coordinates. (TIF)

**Table S1** *Hucho taimen* (Ht) specimens and collection sites. (DOCX)

**Table S2** River names, basins, sample sizes, and coordinates for the *Hucho taimen* specimens used in this study. (DOCX)

**Table S3** Primers used in the present study. (DOCX)

**Table S4** Region of the *Hucho taimen* mitochondrial genome studied in the present work, with corresponding coordinates in the mitochondrial genomes of *H. taimen*, *H. bleekeri*, and *Brachymystax lenok tsinlingensis*. (DOCX)
Table S5 Best-fit models of mitochondrial protein-coding genes evolution in Hucho taimen.

Acknowledgments

We thank Elena Balakireva, Dr. Zehao Shen, Dr. Mary Ann Bruce, and Dr. Nancy Aguilar-Roca for encouragement and help.

References

1. Holčík J (1995) Threatened fishes of the world - Hucho taimen (Linnaeus, 1758) (Salmonidae). DordrechtThe Netherlands: Dr W Junk Publishers. 254 p.
2. Matveev AN, Pronin NM, Samusenok VP, Bronte CR (1998) Ecology of Siberian taimen Hucho taimen in the Lake Baikal basin. J Great Lakes Res 24: 905–916.
3. Froese R, Pauly D (2012) FishBase. World Wide Web electronic publication. Available at: http://www.fishbase.org, version 04/2012.
4. Nikolsky GV (1956) Fishes of the Amur River basin. Moscow: USSR Academy of Sciences. 551 p.
5. Holčík J (1995) Threatened fishes of the world - Hucho taimen (Linnaeus, 1758) (Salmonidae). Environ Biol Fishes 43: 105–106.
6. Yue PQ, Chen YG (1998) China red data book of endangered animals, Pisces. Beijing: Science Press. 247 p.
7. Ocock J, Baasanjav G, Baillie G, Erdenebat M, Kotellat M, et al. (compilers and editors) (2006) Summary conservation action plans for Mongolian fishes. Regional Red List Series. Vol. 4. London: Zoological Society of London. 37 p.
8. Frolov SV, Fodova VN (2000) Polymorphism and karyotype divergence in taimens of the genus Hucho. Russian J Genet 36: 175–178.
9. Sun Y, Sun ZW, Yin HB (2010) The analysis of isozymes in Hucho taimen. Chinese J Fish 27: 46–50.
10. Liang LQ, Chang YM, Dong CZ, Sun XW (2004) Genetic analysis for Hucho taimen in Wusuli River with microsatellites. J Fish China 28: 241–244.
11. Tong GX, Kuang YY, Yin JS, Liang LQ, Sun XW (2006) Isolation of microsatellite DNA and analysis on genetic diversity of endangered fish, Hucho taimen (Pallas). Mol Ecol Notes 6: 1099–1101.
12. Kuang YY, Tong GX, Xu W, Yin JS, Sun XW (2009) Analysis of genetic diversity in the endangered Hucho taimen from China. Acta Ecol Sinica 29: 92–97.
13. Liu B, Kuang YY, Tong GX, Yin JS (2011) Analysis of genetic diversity on 9 wild stocks of taimen (Hucho taimen) by microsatellite markers. Zool Res 32: 397–404.
14. Tong GX, Bao YL, Kuang YY, Yin JS (2008) Study on technology system built for AFLP of Hucho taimen (Pallas). Acta Agri Univ Jiangxiensis 30: 405–410.
15. Tong GX, Kuang YY, Yin JS (2009) AFLP analysis of genetic diversity of taimen (Hucho taimen) in wild populations. J Fish Sci China 6: 833–841.
16. Knudsen K, Yoneda Y, Nishimura K, Yukimura M, Kurita T (2007) Genetic diversity of wild and captive taimens (Hucho taimen) from Western Mongolia. J Fish Sci China 10: 914–917.
17. Xie ZY, Huang SW, Yan RY (1959) Lenok and taimen in Heilongjiang province. Chinese J Fish 2: 69–75.
18. Xu LX, Kuang YY, Tong GX, Ding L, Liu B, et al. (2011) An analysis on genetic constitution of Hucho taimen, Hucho taimen and hybrids (Hucho taimen [female] x Bichthys taimen [male]) by SNP markers. Acta Agric Univ Jiangxiensis 33: 1187–1194.
19. Wang J, Kuang YY, Tong GX, Yin JH (2011) Genomics analysis on incompatibility of intergeneric hybridizations between Hucho taimen (Guichenot & Poey, 1858) and Bichthys taimen (Kawam. & Steindachner, 1871) by using 30 polymorphic SSR markers. J Fish Sci China 3: 547–555.
20. Lou YD, Li QX (2006) Distinct hybridization of fish and its application in aquaculture in China. J Fish Sci China 1: 159–164.
21. Xu W, Yin JS, Kuang YY, Jiang GG (2012) Artificial propagation of huchen, Hucho taimen Pallae. Shanghai Shuishan Daxue Xueba 17: 452–456.
22. Côté-Lafrenière A, Weir IK, Bernatchez L (2012) Framing the salmonid family phylogenetic portrait: a more complete picture from increased taxon sampling. PLoS ONE 7(10): e46662.
23. Zolotukhin SF, Semenchenko AU, Belyaev VA (2000) Taimen and lenok in the Russian Far East. Khabarovsk: Khabarovskaya Kraevaya Tipografiya. 128 p.
24. Xie ZY, Huang SW, Yan RV (1959) Lenok and taimen in Heilongiang delta and their natural hybrids. Acta Hydrobiol Sinica 3: 215–220.
25. Shelko SV (2003) Phylogenetic relationships of lenok genus Bichthys (Salmonidae, Salmoniformes) and features of their speciation. Ph.D. thesis. Vladivostok. 207 p.
26. Shelko SV (2012) Phylogenetic relationships of lenok genus Bichthys (Salmonidae, Salmoniformes) and features of their speciation. Verlag: LAP LAMBERT Academic Publishing. 216 p.
27. Esteve M, McMinnan DA (2008) Spawning behavior of lenok, Bichthys taimen (Salmoniformes) from the Uur River, Northern Mongolia. J Ichthyol 48: 1031–1036.
28. Esteve M, Gilroy D, McMinnan DA (2009) Spawning behaviour of taimen (Hucho taimen) from the Uur River, Northern Mongolia. Environ Biol Fish 84: 107–109.
29. Avise JC, Saunders NC (1984) Hybridization and introgression among species of sunfish (Lepomis): analysis by mitochondrial DNA and allozyme markers. Genetics 108: 237–255.
30. Smith GR (1992) Introgression in fishes: significance for paleontology, cladistics, and evolutionary radiation. Ph.D. thesis. San Diego State University. 414 p.
31. Bachtrog D, Thornton K, Clark A, Andolfatto P (2006) Extensive introgression of mitochondrial DNA relative to nuclear genes in the Drosophila yakuba species group. Evolution 60: 292–302.
32. Gompert Z, Forister ML, Fordey JA, Nice CC (2008) Widespread mito-nuclear discordance with evidence for introgressive hybridization and selective ascertainment. Mol Biol Evol 17: 3221–3244.
33. Toews DP, Berdys A (2012) The biogeography of mitochondrial and nuclear discordance in animals. Mol Ecol 21: 3907–3910.
34. Coyne JA, Orr HA (2004) Speciation. SunderlandMA: Sinauer Associates. 545 p.
35. Balakirev ES, Chechulin VR, Lohzin VV, Ayala FJ (2003) DNA polymorphism in the F-strerase gene cluster of Drosophila melanogaster. Genetics 164: 533–544.
36. Balakirev ES, Pavlyukhov VA, Ayala FJ (2008) DNA variation and endosymbiotic associations in phenotypically-diverse sea urchin Strongylocentrotus franciscanus. Proc Nat Acad Sci USA 105: 16218–16223.
37. Yang CH, Chang HW, Ho CH, Chou YC, Chang LY (2011) Conserved PCR primer set designing for closely-related species to complete mitochondrial genome sequencing using a sliding window-based PSO algorithm. PLoS ONE 6(5): e17729.
38. Wang Y, Zhang XG, Yang SY, Song ZB (2011a) The complete mitochondrial genome of the taimen, Hucho taimen, and its unusual features in the control region. Mitochon DNA 22: 111–119.
39. Yang Y, Guo R, Li H, Zhang XY, Du J, et al. (2011b) The complete mitochondrial genome of the Sichuan taimen (Hucho bleekeri), repetitive sequences in the control region and phylogenetic implications for Salmonidae. Mar Genom 4: 221–228.
40. Si SJ, Wang Y, Xu GF, Yang SY, Mou ZB, et al. (2012) Complete mitochondrial genomes of two lenoks, Bichthys taimen and Bichthys taimen, using the same 1536 bp of the control region. Mitochon DNA 23: 338–340.
41. Wyman SK, Jensen RK, Boro JE (2004) Automatic annotation of organellar genomes with DOGMA. Bioinformatics 20: 3252–3255.
42. Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence

Author Contributions

Conceived and designed the experiments: ESB. Performed the experiments: ESB. Analyzed the data: ESB. Contributed reagents/materials/analysis tools: ESB NSR PBM FJA. Wrote the paper: ESB FJA.
weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 22: 4673–4680.

52. Liberato P, Rozas J (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics 25: 1411–1412.

53. Filatov DA (2002) PROSEQ: a software for preparation and evolutionary analysis of DNA sequence data sets. Mol Ecol Notes 2: 621–624.

54. Tamura K, Peterson D, Peterson N, Stecher G, Nei M, et al. (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol Biol Evol 28: 2731–2739.

55. Zwickl DJ (2006) Genetic algorithm approaches for the phylogenetic analysis of large biological sequence data sets under the maximum likelihood criterion. PhD Thesis. Austin: University of Texas.

56. Ronquist F, Teslenko M, van der Mark P, Ayres DL, Barrowcliffe S, et al. (2012) MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. Syst Biol 61: 539–542.

57. Darriba D, Taboada GL, Doallo R, Posada D (2012) jModelTest 2: more models, new heuristics and parallel computing. Nat Methods 9: 772.

58. Felsenstein J (1985) Confidence limits on phylogenies: an approach using the bootstrap. Evolution 39: 783–791.

59. Gelman A, Rubin DB (1992) Inference from iterative simulation using multiple sequences. Stat Sci 7: 457–511.

60. Martin DP, Lemez P, Lott M, Moulton V, Posada D, et al. (2010) RDP3: a flexible and fast computer program for analyzing recombination. Bioinformatics 26: 2462–2463.

61. Lemez P, Lott M, Martin DP, Moulton V (2009) Identifying recombinants in human and primate immunodeficiency virus sequence alignments using quartet scanning. BMC Bioinformatics 10: 126.

62. Weiller GF (1996) Phylogenetic profiles: a graphical method for detecting generic recombinations in homologous sequences. Mol Biol Evol 15: 326–335.

63. Beko RG, Hamilton N (2006) Phylogenetic identification of lateral gene transfer events. BMC Evol Biol 6: 15.

64. Wu X-C (2006) The loss of genetic diversity in Sichuan taimen as revealed by DNA fingerprinting. Biochem Genet 44: 177–185.

65. Nei M (1987) Molecular evolutionary genetics. New York: Columbia University Press. 512 p.

66. Martin DP, Rybicki E (2000) RDP: detection of recombination amongst aligned sequences. J Mol Evol 11: 147–164.

67. Luikart G, England PR, Tallmon D, Jordan S, Taberlet P (2003) The power and promise of population genomics: from genotyping to genome typing. Nat Rev Genet 4: 981–994.

68. Hudson RR, Kaplan N (1985) Statistical properties of the number of recombinant events in the history of a sample of DNA sequences. Genetics 111: 147–164.

69. Martin D, Rybicki E (2000) RDP: detection of recombination amongst aligned sequences. Bioinformatics 16: 562–563.

70. Padidam M, Sawyer S, Fauquet CM (1999) Possible emergence of new gemiviruses by frequent recombination. Virology 265: 218–223.

71. Martin DP, Posada D, Crandall KA, Williamson C (2005) A modified bootscan algorithm for automated identification of recombinant sequences and geminiviruses by frequent recombination. Virology 265: 218–225.

72. Bredell H, Martin DP, Van Harmelen J, Varsani A, Sheppard HW, et al. (2010) Evidence for hybridization in Brachymystax lenok Pallais (Salmonidae, Salmonidae) populations in water systems of Eastern China, inferred from mitochondrial DNA sequences. Zool Stud 49: 190–200.

73. Martin DP, Lemez P, Moulton V, Posada D, et al. (2010) RDP3: a flexible and fast computer program for analyzing recombination. Bioinformatics 26: 2462–2463.

74. Posada D, Crandall KA (2001) Evaluation of methods for detecting hybridization in fish: new methods of detection and genetic interpretations. In: Ryman N, Utter F, editors. Population genetics and fishery management. Seattle, WA: University of Washington Press. 161–192.

75. Dawding TE, DeMazur BD (1993) Evolutionary significance of introgressive hybridization in cyprinid fishes. Nature 362: 444–446.

76. Ballard JW, Whithlock MC (2004) The incomplete natural history of mitochondria. Mol Ecol 13: 729–744.

77. Barr CM, Neiman M, Taylor DR (2005) Inheritance and recombination of mitochondrial genomes in plants, fungi and animals. New Phylol 168: 39–50.

78. Pilgrim BL, Perry RC, Barron JL, Marshall HD (2012) Nucleotide variation in mtDNA Variation and Introgression in Hucho taimen special issue. Environ Biol Fish 94: 1–6.