

# First indication of Japanese mitten crabs in Europe and cryptic genetic diversity of invasive Chinese mitten crabs

Sarah Hayer<sup>1</sup>, Dirk Brandis<sup>1</sup>, Günther B. Hartl<sup>2</sup>, Christine Ewers-Saucedo<sup>1</sup>

**1** Zoological Museum of the Christian-Albrechts University, Hegewischstrasse 3, 24105 Kiel, Germany **2** Institute of Zoology, Christian-Albrechts University, Am Botanischen Garten 1-9, 24118 Kiel, Germany

Corresponding author: Christine Ewers-Saucedo ([ewers-saucedo@zoolmuseum.uni-kiel.de](mailto:ewers-saucedo@zoolmuseum.uni-kiel.de))

---

Academic editor: Adam Petrusek | Received 1 April 2019 | Accepted 15 July 2019 | Published 28 August 2019

**Citation:** Hayer S, Brandis D, Hartl GB, Ewers-Saucedo C (2019) First indication of Japanese mitten crabs in Europe and cryptic genetic diversity of invasive Chinese mitten crabs. *NeoBiota* 50: 1–29. <https://doi.org/10.3897/neobiota.50.34881>

---

## Abstract

The Chinese mitten crab (*Eriocheir sinensis*) is a prominent aquatic invader with substantial negative economic and environmental impacts. The aim of the present study was to re-evaluate the genetic diversity of mitten crabs throughout their native and invaded ranges based on publicly available sequence data, and assess if multiple introductions or rapid adaptation could be responsible for biologically divergent mitten crabs in Northern Europe. We assembled available genetic data of a fragment of the mitochondrial cytochrome c oxidase subunit one gene (COI) for all species of the genus *Eriocheir*. We applied phylogenetic and population genetic analyses to compare native and invasive populations, and to identify possible source populations. The phylogenetic reconstruction revealed that five COI sequences from Europe, morphologically identified as Chinese mitten crab, actually belong to the Japanese mitten crab (*Eriocheir japonica*), representing the first indication of its presence in European waters. All other COI sequences from Europe could unambiguously be assigned to the Chinese mitten crab. In some Northern German populations of Chinese mitten crabs, genetic diversity was surprisingly high, due to seven unique haplotypes encoding several amino acid substitutions. This diversity may reflect a cryptic introduction from an unsampled native location, or rapid adaptation in the invaded range. Based on the genetic diversity shared between native and introduced range, Feiyunjiang, a tributary of the Yangtze River, emerges as a plausible source population for the original introduction of Chinese mitten crabs to Europe. This study highlights the complex and dynamic invasion processes of mitten crabs in Europe. We urge to further monitor mitten crab invasions using genetic tools.

## Keywords

Species invasion, amino acid substitution, barcoding, COI, source population, rapid adaptation

## Introduction

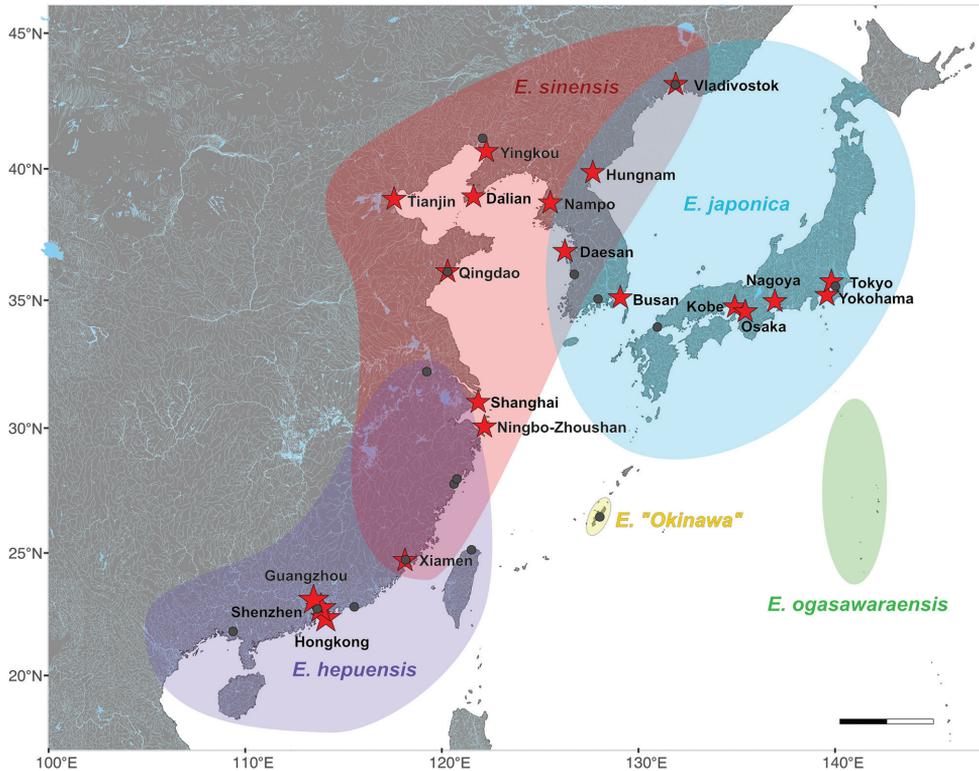
Species invasions have altered the global ecological landscape dramatically in the past centuries. Their impacts are exemplified by the Chinese mitten crab (*Eriocheir sinensis* H. Milne Edwards, 1853), one of the taxa included in the list of the “world’s 100 worst invasive alien species” (Lowe et al. 2000). Native to Russia, China, Korea and Taiwan, it has been introduced to Europe at the beginning of the 20<sup>th</sup> century and subsequently to the United States via the ballast water of large shipping vessels (Panning and Peters 1933; Clark et al. 1998; Rudnick et al. 2000; 2003; Herborg et al. 2003, 2005; Dittel and Epifanio 2009). The economic and ecological effects of its invasion are staggering. While declining in abundance in its native range, where it is considered a delicacy and farmed extensively (Yuan 2005), it has become an unprecedented nuisance in its introduced range. During past mass occurrences, when thousands of crabs migrated from their adult inland freshwater habitats to marine spawning grounds, fishermen lost nets and even abandoned certain fishing grounds (Rudnick et al. 2000). River banks were destabilized due to the extensive burrowing activity of adult crabs (Rudnick et al. 2005). Moreover, ecological communities have the potential to be altered by competition with native crabs and crayfish (Rudnick et al. 2000; Gilbey et al. 2008; Dittel and Epifanio 2009).

Identifying the source or sources of such widespread invasions is an important task for risk assessment and species management. Assigning the geographic sources of invasive populations requires geographic differentiation within the native range. Such differentiation allowed, for example, to pinpoint the source populations of introduced olive populations in Hawaii and Australia as South Africa and western or central Mediterranean, respectively (Besnard et al. 2007). In contrast, if native populations are genetically homogeneous, or the employed genetic markers are not variable enough to detect genetic structure, source populations can only be assigned to broad geographic regions. The European shore crab (*Carcinus maenas*), for example, has invaded the East coast of North America twice (Roman 2006). The source of each invasion could be broadly categorized as Northern and Southern European, respectively, based on slight genetic structure in the native range and ecological differences of the two invading populations. More detailed assignment was hampered by broadly distributed common haplotypes of the studied marker in the native range, possibly caused by anthropogenic reshuffling of native diversity. In the case of another well-dated species invasion, the source of introduction of the North American spiny-cheek crayfish (*Faxonius limosus*) to Poland in 1890 could not be determined because the invaded range is dominated by a haplotype common throughout the native range (Filipová et al. 2011). Similarly, Hänfling et al. (2002) were unable to pinpoint a source population for the mitten crab invasion to Europe due to apparent genetic homogeneity in the native range. This could have been due to small sample sizes (6 to 10 individuals sampled per population) and observed low diversity (5 haplotypes only) of the employed genetic marker, a fragment of the mitochondrial cytochrome oxidase subunit one (COI). Under-sampling the native range is a general problem, not only with regard to the number of populations, but also with number of individu-

als analyzed per population (Muirhead et al. 2008). If only a few individuals are sampled, much of the genetic diversity present at any one site might be missed, obscuring assignment probabilities (Muirhead et al. 2008).

Since the initial assessment by Hänfling et al. (2002), several phylogeographic studies have characterized the genetic population structure of mitten crabs in their large native range (Sui et al. 2009; Wang et al. 2008; Xu et al. 2009). These studies detected five monophyletic lineages (Fig. 1). Some authors refer to these lineages as different species (e.g. Chu et al. 2003; Wang et al. 2008; Naser et al. 2012; Chen et al. 2017), others as subspecies or lineages (e.g. Tang et al. 2003; Xu et al. 2009, Zhang et al. 2012). We do not aim to resolve these taxonomic issues, but use the species names as unambiguous labels throughout the text. These lineages have distinct but partially overlapping ranges (Komai et al. 2006; Xu et al. 2009): the Hepu mitten crab (*Eriocheir hepuensis* Dai, 1991) is present in Southern China from Hepu to Oujiang, the Chinese mitten crab from Tongan in China to Vladivostok in Russia including Korea, the Japanese mitten crab (*Eriocheir japonica* (De Haan, 1835)) is the only lineage present in Japan, but occurs in Russia and Korea as well, *Eriocheir ogasawaraensis* Komai, Yamasaki, Kobayashi, Yamamoto & Watanabe, 2006 is restricted to the Ogasawara Islands and an additional formally undescribed Japanese mitten crab lineage is endemic to Okinawa (Fig. 1) (Wang et al. 2008). Based on combined sequence data for two mitochondrial gene fragments, population structure was significant in Chinese and Japanese mitten crabs, but less so in the Hepu mitten crab (Wang et al. 2008 using cytochrome oxidase subunit two and cytochrome b as genetic markers; Xu et al. 2009 using COI and cytochrome b). The significant population differentiation of both Japanese and Chinese mitten crab provides a working baseline for assigning source populations. To date, Japanese mitten crabs have sporadically occurred outside their native range in the United States only (Benson and Fuller 2019; Jensen and Armstrong 2004), but have not been detected in Europe yet. Chinese mitten crabs, on the other hand, have invaded both Europe and the United States successfully. Population genetic studies of Chinese mitten crabs from the invaded range came to quite discordant conclusions. Some studies found that populations within Europe were admixed (Hänfling et al. 2002; Czerniejewski et al. 2012), while others found significant levels of differentiation (Herborg et al. 2007; Otto 2012). In contrast, only a single COI haplotype of this species has been reported from the United States (Hänfling et al. 2002). While the source populations remained obscure, Hänfling et al. (2002) identified three haplotypes that were shared between the native and invaded ranges as well as a widespread invasive haplotype unknown in the native range, but found in both Europe and the United States. Given the presence of both invasive haplotypes detected in the native range as well as invasive haplotypes not detected in the native range, they concluded that multiple invasions occurred in Europe, and that the USA were likely invaded via Europe.

Invasive populations can become melting pots of novel genetic combinations with unforeseen adaptive potential (Geller et al. 2010). While species invasions are often associated with a loss of genetic diversity in the introduced range, either be-



**Figure 1.** Native distribution of mitten crabs of the genus *Eriocheir*. Map redrawn from Xu et al. (2009), and species ranges interpolated around their sampling sites (marked by grey points) for each lineage to highlight range overlaps. Large shipping ports (marked by red asterisks) are most likely invasion sources.

cause only a few individuals invaded the range, and/or because genetic drift in these small populations subsequently reduces diversity (Nei et al. 1975), multiple invasions can alleviate the effects of these founding events (Dlugosch and Parker 2008). The brown anole lizard (*Anolis sagrei*), for example, has invaded Florida multiple times from geographically distinct source populations. Each source population had a unique genetic makeup, and admixture in the invaded range led to genetic diversity higher than in populations of the native range (Kolbe et al. 2004, 2007). Similar observations have been made for the common ragweed (Genton et al. 2005). Each new invasion might bring in new genetic variation, accompanied by novel ecological and physiological strategies that warrant attention (Geller et al. 2010).

This could be the case for Chinese mitten crabs. Otto (2012) reported on a novel reproductive behavior and physiology in invasive mitten crabs. In contrast to earlier reports (Anger 1991), mitten crabs completed their life cycle in the brackish Baltic Sea, and post-spawning females migrated back into the rivers, and did not die as in other populations. Otto (2012) concluded that this novel behavior might have been caused by a cryptic invasion of mitten crabs with different physiological requirements, in line with an earlier conclusion of cryptic invasions based on genetic data (Hänfling et al.

2002). Both lines of evidence, however, allow for an alternative explanation: invasive mitten crabs could have adapted rapidly to the brackish water conditions of the Baltic Sea, leading to novel genetic, physiological and behavioral diversity.

Rapid adaptation is emerging as a common feature of species invasions (Card et al. 2018; Dlugosch and Parker 2008; Prentis et al. 2008). In animals, hybridization between distinct introduced lineages, allele shifts due to bottlenecks and standing genetic variation are likely agents of rapid adaptation. Selection acts fastest on standing variation, and this process is suggested to be the most common driver of rapid adaptation (Prentis et al. 2008). The Burmese Python (*Python molurus bivittatus*), for example, is native to Southeast Asia and was introduced to Southern Florida in the early 1980s (Card et al. 2018). Despite several freezing periods that caused high python mortality, this species has become a successful invader into North America. Investigations showed that these freezing events resulted in a shift of python physiology caused by changes in allele frequencies in functional genes (Card et al. 2018).

The goal of this study was to re-evaluate the genetic diversity of mitten crabs of the genus *Eriocheir* throughout its native and invaded ranges, and to assess if multiple introductions or rapid adaptation might have caused the recent appearance of biologically distinct mitten crabs in Northern Europe. Given a large body of previous work, we utilized publicly available mitochondrial sequence data. We employed different approaches to assign source populations to the invasive populations in Europe and the United States. First, we reconstructed phylogenetic relationships among *Eriocheir* sequences, grouping thereby invasive individuals into the evolutionary lineages known from the native range. In the next step, we assessed the native distribution of haplotypes also present in the introduced range. Then, we calculated genetic distances for all population pairs, which we use on the one hand to evaluate population genetic structure in the native range, and on the other hand to identify which native populations are most similar to introduced populations. We assume thereby that allele frequencies have not shifted significantly since the invasion, and that enough individuals invaded the new range to mirror native allele frequencies. Bayesian assignment tests have been proposed as a suitable alternative to assign invasive individuals to source populations (Geller et al. 2010). They may not assume that populations are in migration-drift-equilibrium (Aktas 2015), but are limited to detect very recent gene flow between populations within the past few generations (Herborg et al. 2007). Given that the initial invasion occurred around 30 generations ago (generation times taken from Dittel and Epifanio 2009), neither the assumptions of genetic distance measures nor assignment tests are likely to be met. Thus, neither approach provides an ideal fit to the pattern of mitten crab invasion, but our inferences are reinforced when several approaches point to the same source populations. Lastly, we assessed the potential for adaptive evolution in the investigated mitochondrial DNA fragment by identifying amino acid substitution patterns in the introduced haplotypes. Mitochondrial genes are not known to be directly involved in osmoregulations or other adaptations to low-salinity conditions, but the amino acid sequence is highly conserved, being under strong purifying se-

lection (Meiklejohn et al. 2007; Pentinsaari et al. 2016). Any changes in the amino acid sequence we detect are at least unusual and warrant further investigation. Based on our findings, we form several hypotheses regarding the mitten crab invasions that should be followed up using expanded geographic sampling, genomic approaches and historical collections.

## Methods

### Data preparation

We downloaded all available sequences for the genus *Eriocheir* and for two outgroup species, *Neoeriocheir leptognathus* and *Platyeriocheir formosa* (acc. nos. AF316537, AF317326; Tang et al. 2003), from NCBI GenBank (<http://www.ncbi.nlm.nih.gov/Genbank>), and sequences of the Chinese mitten crab from the Barcode of Life Data System (BOLD, <http://www.barcodinglife.org>), ignoring sequences already available in GenBank. An initial screening showed that most sequences were mitochondrial. Thus, we mapped all sequences to a complete mitochondrial genome sequence of a Chinese mitten crab (GenBank acc. KY041629) in Geneious v. 9.1.8 (Kearse et al. 2012). We chose the genetic locus for subsequent analyses for which data existed from both the native and introduced range, a fragment of the gene for the cytochrome c oxidase subunit one (COI).

For the phylogenetic reconstruction, we included COI sequences of all *Eriocheir* species and two outgroups. Most GenBank data consisted of haplotypes, not the actual sequences for each sampled individual. For the population genetic analyses, we reconstructed original haplotype frequencies by replicating haplotype sequences according to the data reported in the publications, attaching locality information to these sequences, and excluding sequences without sampling site information.

### Phylogenetic reconstruction

The first step was to assert the species affinities of all sequences within a phylogenetic framework. For this, we used all sequences available in GenBank and BOLD. We built a maximum likelihood tree with the PHYML (Guindon et al. 2010) plugin in Geneious v. 9.1.8 (Kearse et al. 2012) under the General Time Reversible substitution model, estimating the transition/transversion ratio, the proportion of invariable sites and the gamma distribution parameter. The number of substitution rate categories was set to four. Branch support was calculated from 100 bootstrap replicates. The goal of this phylogenetic reconstruction was not to understand interspecific evolutionary relationships, but to ensure the lineage affinities of the sequences given the recent taxonomic confusion (Chu et al. 2003; Tang et al. 2003; Wang et al. 2008; Xu et al. 2009), and similar morphology of the species (Naser et al. 2012).

## Population genetic analyses: The search for invasion sources

All population genetic analyses were conducted in R version 3.3.3 (R Development Core Team 2018). We constructed a parsimony network of all haplotypes for each species that was found in Europe or the United States with the function ‘haplotype’ in the R package ‘haplotypes’ (Aktas 2015), highlighting native and introduced populations. We visualized the geographic haplotype distribution in the native and introduced range by adapting available scripts for haplotype networks. We then identified haplotypes found in both the native and introduced range and evaluated their distribution in the native range to identify possible sources for invasion.

For Chinese mitten crabs, the species with the most complex invasion history, we conducted further population genetic analyses to understand invasion patterns and get additional support for probable source populations. We compared haplotype and nucleotide diversity for all sampling sites, which we also refer to as populations. We wrote our own function to calculate haplotype diversity of each population based on the formula of Nei and Tajima (1981), and used the function ‘nuc.div’ of the ‘pegas’ package (Paradis 2010) to calculate nucleotide diversity (Nei 1987). We assessed if diversity indices were significantly different between native and introduced populations using standard ANOVA. We calculated Tajima’s *D*, which may indicate population size changes or selective sweeps, using the function ‘tajima.test’ of the ‘pegas’ package (Paradis 2010). Significant deviations from zero were estimated based on a beta distribution (Tajima 1989).

We calculated genetic differentiation between all population pairs as  $\Phi_{st}$  and Jost’s *D* with the functions ‘pairwiseTest’ of the package ‘strataG’ (Archer et al. 2016), and the function ‘pairwise\_D’ of the ‘mmod’ package (Winter 2012), respectively.  $\Phi_{st}$  is a derivative of the classical fixation index *F<sub>st</sub>*, developed for mitochondrial haplotype data (Excoffier et al. 1992). We estimated significant deviations from zero (no differentiation between population pairs) by comparison with an empirical distribution of  $\Phi_{st}$  values based on 1000 permutations. Jost’s *D* provides a more accurate measure of population differentiation when genetic diversity is high and the number of unique alleles per population is large (Jost 2008). Significance was assessed by bootstrapping populations across 1000 replicates. Each measure of differentiation resulted in a large number of pairwise comparisons, which are difficult to interpret. Therefore, we visualized overall population similarity with Metric Multidimensional Scaling, using the function ‘cmdscale’, and with hierarchical cluster analysis, using the function ‘hclust’. Both functions are part of the R package ‘stats’. The analyses of population structure served on the one hand to assess how differentiated the native populations were, and therefore to indicate how narrowly we might be able to pinpoint the sources of introduction. On the other hand, we identified the native populations that were most similar to introduced populations as candidate source populations for the invasions.

We tested whether the native populations were sufficiently diverse to confidently assign individuals from the invaded range using the R package ‘assignPOP’ (Chen et al. 2018), which employs supervised machine learning to evaluate the discrimina-

tory power of genetic or non-genetic data by resampling cross-validation. This means that individuals from each population were randomly divided into training and test sets, and assignment tests were repeated through resampling training individuals 100 times. We used the R function ‘assign.MC’ to conduct Monte Carlo cross validation using 80% of individuals from each population as training data. The proportion of correctly assigned individuals provides an estimate of assignment accuracy, which we calculated with the function ‘accuracy.MC’. In case of sufficiently high discriminatory power, we would assign individuals from the invaded range to native populations using the function ‘assign.X’.

### **Amino acid substitutions: Indication of the potential for adaptation**

We extracted the DNA sequence alignment for the haplotypes, which was generated during the construction of the parsimony network (function ‘haplotype’ of package ‘haplotypes’ and function ‘write.dna’ of package ‘APE’) (Aktas 2015; Paradis et al. 2004) and imported it into Geneious v. 9.1.8 (Kearse et al. 2012). We mapped it to the complete mitochondrial sequence of a Chinese mitten crab from China (GenBank acc. KY041629), and translated the DNA sequences to their amino acid sequence. This translation allowed us to identify amino acid substitutions. We assessed the directionality of change from the most parsimonious ancestral haplotype using the parsimony network constructed earlier.

## **Results**

### **Data sources**

On September 25, 2018, we downloaded a total of 1020 sequences for the genus *Eriocheir* from GenBank, including 11 complete mitochondrial genome sequences. From these, we extracted 106 COI sequences after aligning all sequences to the complete mitochondrial genome sequence of a Chinese mitten crab from China (Li et al. 2016). In addition, we included seven COI sequences of Chinese mitten crabs from BOLD that were not available in GenBank. Prior to analyses, we removed sequence AF317334 of a Hepu mitten crab because it had unusually many substitutions towards one end of the sequence (a sign of poor sequence quality), and discarded sequence CBCC039-11 from BOLD because it was shorter than the other sequences, but otherwise of the same haplotype and sampling site as CBCC049-11. None of the remaining sequences translated any stop codons, which would have indicated sequencing errors or the presence of nuclear pseudogenes.

The final alignment for the phylogenetic reconstruction contained 141 COI sequences (553 bp long) deposited in GenBank and BOLD under the names of the

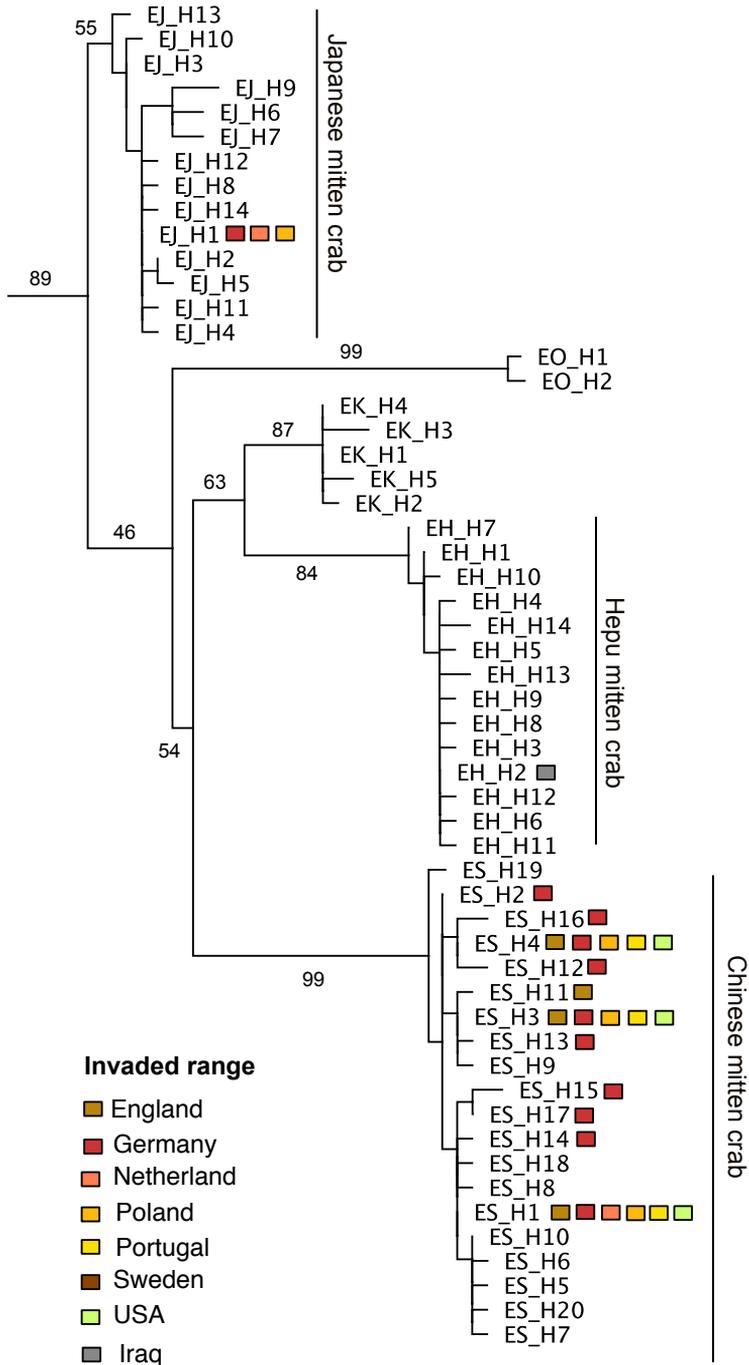
following species: 2 sequences of *E. ogaswaraensis* (Xu et al. 2009), 45 sequences of Japanese mitten crabs (Tang et al. 2003; Xu et al. 2009), 6 sequences of Hepu mitten crabs (Chu et al. 2003; Tang et al. 2003; Naser et al. 2012; Wang et al. 2016), 82 sequences of Chinese mitten crabs (Hänfling et al. 2002; Chu et al. 2003; Tang et al. 2003; Sun et al. 2005; Xu et al. 2009; Czerniejewski et al. 2012; Otto 2012; Zhang et al. 2015; Raupach et al. 2015; Li et al. 2016; Wang et al. 2016) and two sequences of the outgroup species.

For the population genetic analyses, we removed the following sequences without sampling site information: AF105247, FJ455507, NC\_011597, and FJ455505. The final COI dataset for population genetic analyses contained 455 sequences of Chinese mitten crabs from 45 populations belonging to 20 haplotypes and 38 COI sequences of Japanese mitten crabs from 8 populations belonging to 14 haplotypes. We reconstructed the population haplotype frequencies only for these two species because we wanted to infer the invasion sources of European and US invasions. The population-specific sequence information for Japanese and Chinese mitten crabs are summarized in Suppl. material 1: Tables S1 and S2, respectively.

### Phylogenetic reconstruction

Our phylogenetic reconstruction recovered five main lineages of the genus *Eriocheir*, in agreement with previous phylogenetic studies (Xu et al. 2009; Naser et al. 2012) (Suppl. material 2: Fig. S1). For legibility, we provide the phylogenetic reconstruction with haplotypes only, highlighting their occurrence in the invaded range (Fig. 2). Haplotypes from invasive individuals of *Eriocheir* belonged to Japanese mitten crabs (Europe), Chinese mitten crabs (Europe and North America) and Hepu mitten crabs (Western Asia). The invasion of Hepu mitten crabs into Iraq has been discussed in detail elsewhere (Naser et al. 2012), and we focus our analyses on Chinese and Japanese mitten crab lineages.

The occurrence of Japanese mitten crabs outside of their native range has not been reported previously. The phylogenetic reconstruction recovered that five European individuals identified as Chinese mitten crab actually grouped with Japanese mitten crabs (Fig. 2). One crab was collected in Germany in 2009 (Raupach et al. 2015), one in Poland in 2015 by Dagmara Wojcik-Fudalewska (BOLD acc. OZ-IMP066-15), and three individuals caught in Holland and obtained from a seafood retailer were studied by Cristian Bernardi of the Università degli Studi di Milano in 2011 (BOLD acc. CBCC038-11, CBCC039-11, CBCC040-11). Sequences of the Polish and Dutch individuals were only available in BOLD (<http://www.boldsystems.org>), not in NCBI GenBank, and have, to our knowledge, not been part of scientifically published studies. The origin of the latter was indicated in BOLD ambiguously as “Italy, Holland” without geographic coordinates but clarified in direct communication with C. Bernardi.



**Figure 2.** Maximum likelihood tree of mitten crab COI haplotypes. Haplotypes found in the invaded range are color-coded based on the country of collection. Numbers on branches represent bootstrap support. Branches without numbers have less than 50% bootstrap support. We omitted the outgroups from the figure to save space.

## Population genetic analyses

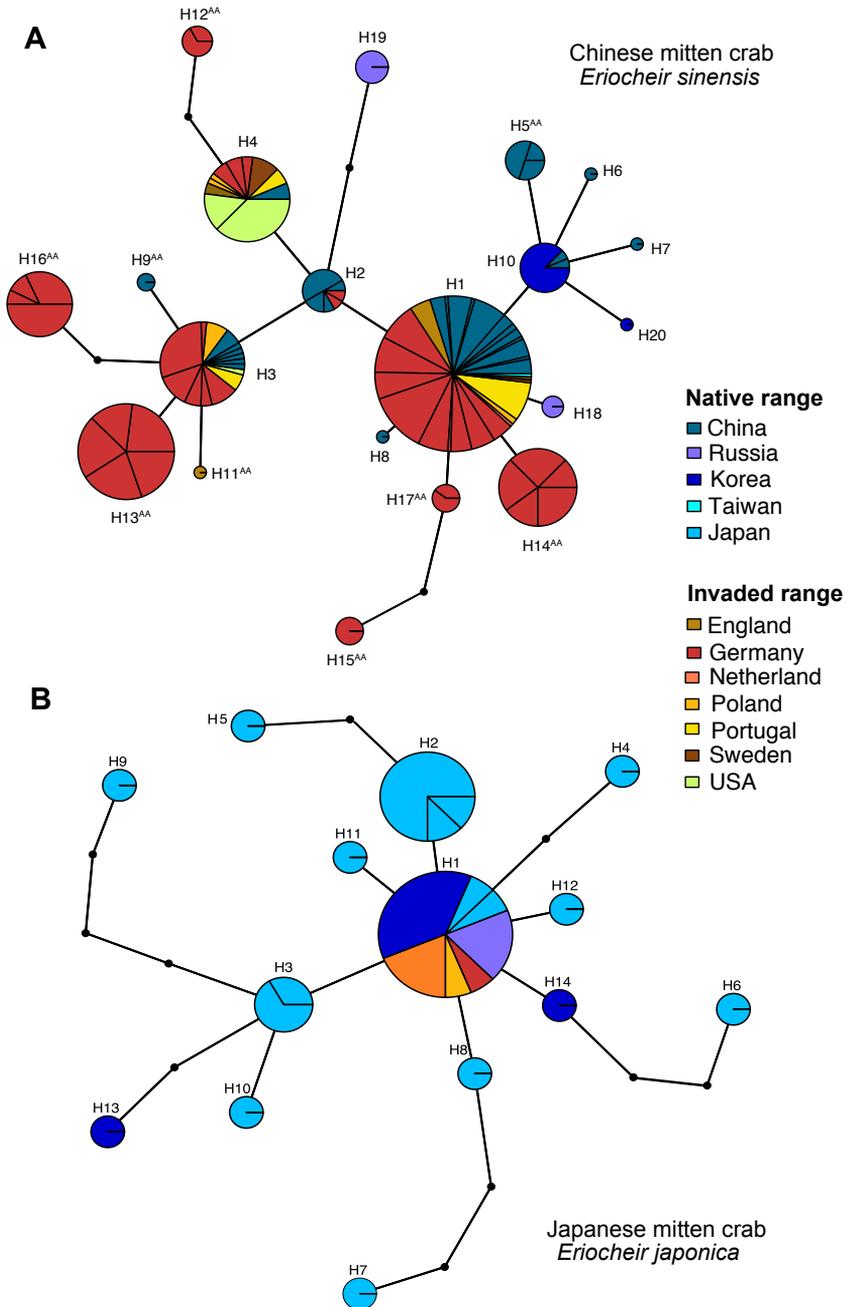
We identified 14 haplotypes for Japanese mitten crabs, labelled H1 to H14 (Fig. 3B). The most common haplotype, H1, occurred in all populations but Shimonoseki, Japan. The invasive individuals found in Germany, Poland and Holland also had this haplotype, thus limiting our ability to assign a more detailed source population to invasive Japanese mitten crabs, and not meriting further population genetic analyses.

We identified a total of 20 haplotypes for Chinese mitten crabs (Fig. 3A). Of these, nine haplotypes were found only in the native range, seven haplotypes only in Europe and four haplotypes in both native and introduced ranges. The geographic distribution maps visualize that haplotypes found only in the native region (blue colors) are common in central and northern Asia (Fig. 4A). Most of Europe is dominated by three of the four haplotypes shared between the native and introduced region (H1, H2, H4), which are more or less common in the native range: H1 was widespread in both the native and introduced range, thus providing little detail about the source of invasion (Fig. 4). H2 has a widespread distribution in its native range, found in two northern locations, Dalian City and Wuhu, and two central locations, Liaohe and an unspecified part of the Yangtze River (Fig. 4A, Suppl. material 2: Fig. S2).

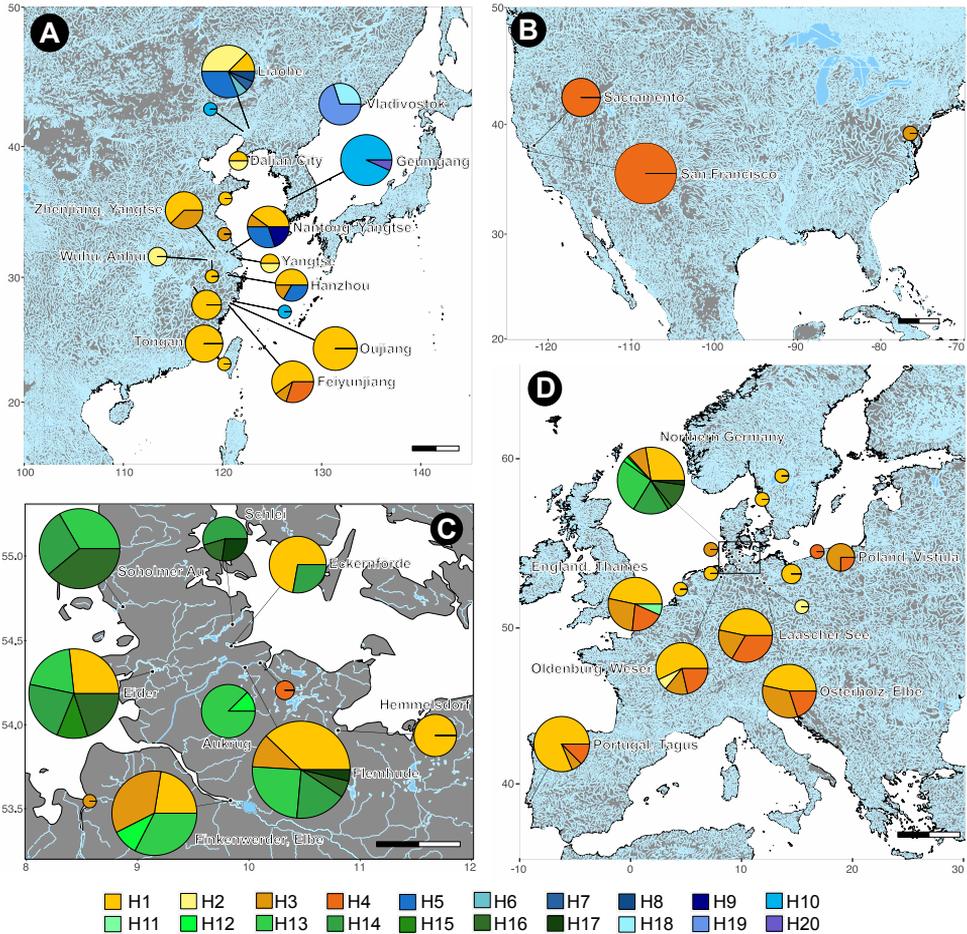
In the introduced range, H2 was found in two individuals only, one sampled in the Weser river near Oldenburg and the other in the Elbe river in Brandenburg, suggestive of its overall low frequency in the introduced range (Fig. 4D, Suppl. material 2: Fig. S2). H3 was reported from several central Chinese locations (Feiyunjiang, Hangzhou, Nantong, Yancheng, and Xhenjiang), and was widespread in Europe (Fig. 4, Suppl. material 2: Fig. S3). H4 was also widespread in Europe but reported in the native range only from Feiyunjiang (Fig. 4, Suppl. material 2: Fig. S4). In summary, three widespread invasive haplotypes were found in Feiyunjiang, making it a plausible main source for the invasion.

Several Northern German populations are genetically distinct: Aukrug, Eckernförde, Eider, Finkenwerder, Flemhude, Schlei, and Soholmer Au (marked with an asterisk in Fig. 4C, D). These populations contained most of the haplotypes not detected in the native range, which we colored in green (compare Fig. 4A and C). In contrast to the European populations, the documented diversity in the US populations is lower. The large, established populations of the West coast seem to consist of a single haplotype, H4 (Fig. 4B), while a single individual sampled from an unestablished population on the East coast of the United States has a different haplotype (H3).

Overall haplotype diversity for Chinese mitten crabs was 0.832, and ranged from 0 to 0.805 per population (Table 1). Overall nucleotide diversity was 0.00384, and ranged from 0 to 0.00475 (Table 1). Introduced populations did not have lower haplotype diversity than native populations ( $Df = 1$ ,  $F$ -value = 0.46,  $p$ -value = 0.505), nor did the nucleotide diversity between native and introduced populations differ ( $Df = 1$ ,  $F$ -value = 0.453,  $p$ -value = 0.508). Tajima's  $D$  ranged from -1.159 to 2.315 per population (Table 1), and was not significantly different from zero in all but one population



**Figure 3.** Parsimony networks for Chinese and Japanese mitten crabs. Each circle represents a haplotype, and the size of the circle is proportional to the abundance of this haplotype **A** Chinese mitten crab: The smallest circle (e.g. H8) contains a single sequence, while the largest circle (H1) contains 161 sequences **B** Japanese mitten crab: The smallest circle (e.g. H4) contains a single sequence, while the largest circle (H1) contains 18 sequences. The colors represent sampling sites: blue colors are native sites, yellow-orange-red colors are European sites and green colors are US sites.



**Figure 4.** Geographic distribution of COI haplotype frequencies of Chinese mitten crabs. The distribution is shown in the native range (A), the United States (B), Northern Germany (C) and Europe in general (D). Haplotypes only found in the native range are colored in blue tones, haplotypes only found in the introduced range are colored in green tones, and haplotypes found in both the native and introduced range are colored in yellow and orange tones. The smallest pie chart in each graphic represents a single individual. All populations with five or more sampled individuals are named. Additionally, in a we also point out the populations of Dalian City, Wuhu and Yangtze, as they contain the otherwise rare but invasive haplotype H2. Scale bars: 250 km (A, B, D). 25 km (C).

(Soholmer Au,  $D = 2.315$ ,  $p$ -value = 0.021). In some cases, this could be the result of low sample size, which reduces the power to detect deviations from the null expectation.

A total of 9 haplotypes were private. They were distributed among four native sites (Liaohu: H6, H7, H8; Nantong: H9; Vladivostok: H18, H19; Geumgang: H20) and two introduced sites (Thames: H11; Eider: H15). Estimates of population differentiation among native populations with five or more sampled individuals revealed significant population structure across the native range (Suppl. material 1: Table S3).

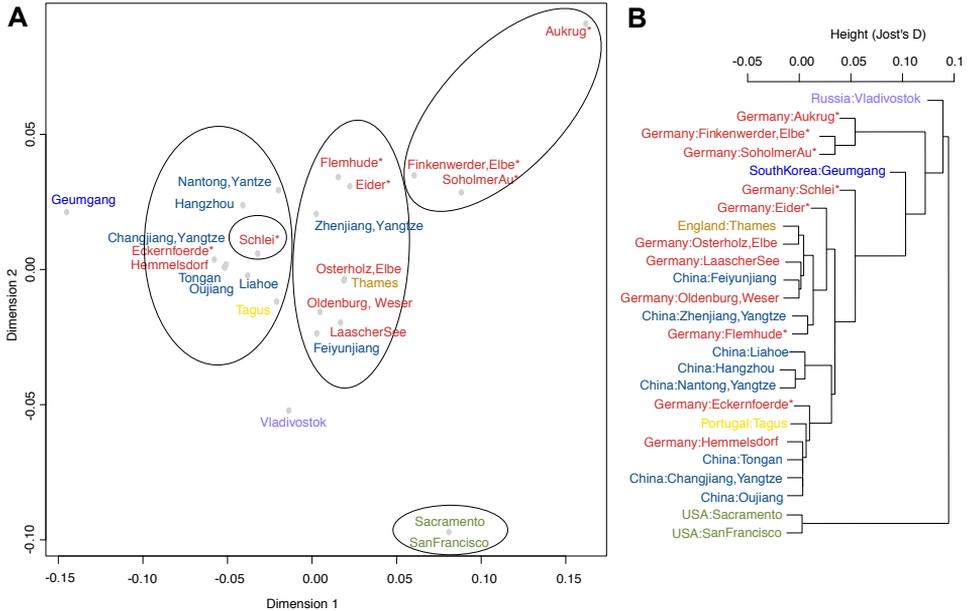
**Table 1.** Results of population genetic analyses of Chinese mitten crab (*Eriocheir sinensis*) populations from sampling sites with more than five sampled individuals.

Sampling site	n	h	Haplotype diversity	Nucleotide diversity	Tajima's D	D p-value
Native range						
Feiyunjiang, China	10	3	0.600	0.002	0.473	0.636
Hangzhou, China	6	3	0.733	0.003	-0.057	0.954
Liaohe, China	16	6	0.783	0.003	0.095	0.924
Nantong, Yangtze, China	10	4	0.778	0.004	1.032	0.302
Oujiang, China	11	1	0.000	0.000	NA	NA
Tongan, China	8	1	0.000	0.000	NA	NA
Zhenjiang, Yangtze, China	8	2	0.536	0.002	1.449	0.147
Vladivostok, Russia	10	2	0.467	0.003	1.229	0.219
Geumgang, China	15	2	0.133	0.000	-1.159	0.246
Invaded range						
Thames, England	15	4	0.714	0.003	0.666	0.505
Aukrug, Germany	16	2	0.233	0.002	-0.744	0.457
Eckernfoerde, Germany	18	2	0.425	0.001	0.870	0.385
Eider, Germany	45	5	0.805	0.005	1.925	0.054
Finkenwerder, Elbe	40	4	0.729	0.003	0.758	0.448
Flemhude, Germany	53	6	0.766	0.004	1.171	0.241
Hemmelsdorf, Germany	10	1	0.000	0.000	NA	NA
Laascher See, Germany	15	3	0.676	0.002	1.386	0.166
Oldenburg, Weser, Germany	14	4	0.648	0.002	0.683	0.495
Osterholz, Elbe, Germany	15	3	0.676	0.002	1.386	0.166
Schlei, Germany	11	3	0.655	0.004	0.821	0.412
Soholmer Au, Germany	36	3	0.679	0.005	2.315	0.021
Tagus, Portugal	16	3	0.342	0.001	-0.708	0.479
Sacramento, USA	7	1	0.000	0.000	NA	NA
San Francisco, USA	18	1	0.000	0.000	NA	NA

Abbreviations: n: number of available sequences/individuals sampled; h: number of haplotypes.

Jost's D ranged from 0 to 0.215 and  $\Phi_{st}$  from 0 to 1 (Suppl. material 1: Table S3). The overall pattern of relative differentiation was very similar between the two measures (Suppl. material 1: Table S3). Vladivostok in Russia and Geumgang in South Korea were significantly differentiated from all other native populations, and from each other. Some populations were undifferentiated with either distance measure, and clustered closely: 1) Oujiang and Tongan, both monotypic for haplotype H1, and 2) Hangzhou, Nantong and Liaohe.

We used these pairwise genetic distances to identify which introduced populations were genetically similar to native populations, representing potential sources of the invasion. In general, populations dominated by the same haplotype cluster together. The two monotypic Chinese populations, Oujiang and Tongan, cluster together with the German populations from Hemmelsdorf, Tagus and Eckernfoerde (Fig. 5). A second large cluster consists of several non-native populations from Germany and England and the Chinese population from Feiyunjiang and Zhenjiang (Fig. 5). These populations are undifferentiated with regard to Jost's D, which ranged from 0 to 0.003, and  $\Phi_{st}$ , which ranged from 0 to 0.035. The Northern German populations Aukrug, Eckernfoerde, Schlei and Soholmer Au are significantly differentiated from



**Figure 5.** Visualization of genetic population similarity based on pairwise genetic differences calculated as Jost's D values. **A** multi-dimensional scaling plot **B** hierarchical cluster analysis dendrogram. Asterisks denote Northern German populations that are dominated by haplotypes not found in the native range.

each other and all other populations. In the MDS plot of the first two coordinate axes (Fig. 5A), the Schlei population appears to lie within the first large cluster, but it is differentiated well by the third axis (Suppl. material 2: Fig. S5). Similarly, the US populations are differentiated from all other populations. In the introduced range at large, Jost's D ranged from 0 to 0.234, and  $\Phi_{st}$  from 0 to 1. Within Europe, Jost's D ranged from 0 to 0.234, and  $\Phi_{st}$  from 0 to 0.691, making European populations much more differentiated than native populations.

The Monte Carlo cross validation procedure revealed little power to discriminate between source populations with assignment tests. The assignment accuracy averaged across replicates was 0.032. Thus, we did not attempt to assign invasive individuals to any particular native population with this method.

### Amino acid substitutions

Amino acid substitutions took place in eight COI haplotypes: H5, H9, and H12 to H17 (Suppl. material 2: Fig. S6). Of these, the haplotypes H5 and H9 were only found in China, while H12 to H17 were the haplotypes only detected in Northern Germany. Based on the parsimony network, most substitutions occurred convergently. Only H15 evolved directly from H17. We can further infer the directionality of these substitutions from the haplotype network. It stands out that both proline and threonine evolved repeatedly in this small fragment of the COI gene.

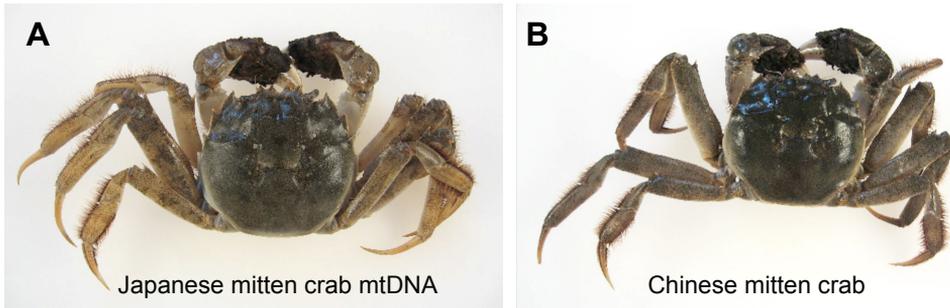
## Discussion

### The Japanese mitten crab entered the European stage more than a decade ago

To our knowledge, we provide the first report of Japanese mitten crabs (*Eriocheir japonica*) outside their native range. Our phylogenetic reconstruction placed five sequences identified as Chinese mitten crabs clearly within the Japanese mitten crab lineage. The sequences were collected in Holland, Germany and Poland between 2009 and 2015. The German individual was collected inland in the Rhine river, and may not have necessarily migrated successfully to the North Sea for reproduction. The Dutch and Polish individuals were collected closer to the North and Baltic Sea, suggestive of an established, reproducing population of Japanese mitten crabs in Europe for the past ten years or more.

At first, it seems surprising that this invasion of Japanese mitten crabs has remained cryptic for at least a decade, but the morphological similarity between Chinese and Japanese mitten crabs did not make it obvious (Fig. 6) (Guo et al. 1997; Jensen and Armstrong 2004; Naser et al. 2012). Moreover, all European sequences of Japanese mitten crabs were generated as part of sequencing efforts of between one and a few *Eriocheir* specimens (see BOLD records), diluting the meaning of the high genetic dissimilarity to other Chinese mitten crab sequences. Raupach et al. (2015), for example, actually discussed the high intraspecific genetic diversity of mitten crab sequences in their large barcoding study from German waters. They noted that the observed high genetic distances in their sample of Chinese mitten crabs were caused by a single specimen, which we assigned to the Japanese mitten crab based on its COI sequence. Morphological species identification placed this individual clearly as a Chinese mitten crab (Raupach et al. 2015). Similarly, according to the shape of the interocular carapace rim, the Dutch individuals would be identified as Chinese mitten crabs (Guo et al. 1997), but their COI sequences belong to Japanese mitten crabs. Interestingly, the Barcoding of Life Data System itself recognized that the five sequences in question clustered with Japanese mitten crabs, not with Chinese mitten crabs ([http://www.boldsystems.org/index.php/Public\\_BarcodeCluster?clusteruri=BOLD:AAA8754](http://www.boldsystems.org/index.php/Public_BarcodeCluster?clusteruri=BOLD:AAA8754)). This discordance between morphology and mitochondrial sequence data may be due to the taxonomic confusion among *Eriocheir* species (Costa et al. 2007).

Cryptic morphology is a general problem in biological invasions that can only be resolved with molecular data. Bastrop and Blank (2006), for example, used mitochondrial sequence data to show that in addition to the invasive polychaete *Marenzelleria neglecta*, two more species of the genus had invaded the Baltic Sea unnoticed. The invasive populations of the virile crayfish consist completely of a lineage not yet identified in the native North American range (Filipová et al. 2010). The cosmopolitan reed *Phragmites australis* represents an unusual case of cryptic invasion, where a non-native haplotype is currently replacing the native genetic diversity in North America (Saltonstall 2002). The hypothesis of cryptic morphology is therefore clearly appealing



**Figure 6.** Mitten crabs collected in Holland in 2011. **A** Individual carrying mitochondrial DNA (mtDNA) of the Japanese mitten crab, morphologically identified as Chinese mitten crab, (BOLD accession: CBCC040-11, museum catalogue no: Università degli Studi di Milano, Ispezione degli Alimenti di Origine Animale MALAC:00040) **B** Chinese mitten crab (BOLD accession: CBCC037-11, museum catalogue no: Università degli Studi di Milano, Ispezione degli Alimenti di Origine Animale MALAC:00037). Photographs by Cristian Bernardi published under CC BY-NC 3.0 license.

and plausible. However, it is puzzling that all individuals identified as Japanese mitten crab at the sequence level were morphologically undoubtedly identified as Chinese mitten crabs. This discordance could hint at hybridization and subsequent introgression between Chinese and Japanese mitten crabs, resulting in morphological Chinese mitten crab hybrids carrying Japanese mitten crab mitochondrial genomes. This hybridization could have taken place either in the native or the invaded range. In such a case, pure Japanese mitten crabs do not necessarily have to form a stable population in Europe. Rather, their mitochondrial genomes would occur in some proportion of individuals with predominantly Chinese mitten crab genomes. Interspecific hybridization of another global invader has been confirmed for the Shore crab genus *Carcinus* using a combination of mitochondrial sequence data and nuclear microsatellite data (Darling 2011). To understand the current distribution of Japanese mitten crabs, possible hybrids between Japanese and Chinese mitten crabs or introgressed individuals in Europe, future systematic sampling, mitochondrial and nuclear sequencing of mitten crabs is highly warranted.

### Significant genetic structure in the native and introduced range of Chinese mitten crabs

Much work has been conducted on the phylogeography of mitten crabs in their native range (Hänfling et al. 2002; Wang et al. 2008; Sui et al. 2009; Xu et al. 2009; Zhang et al. 2012; Zhang et al. 2014). Our re-analysis was therefore only aimed at assessing how useful the COI marker alone is to differentiate between populations, a prerequisite for identifying source populations with certainty (Geller et al. 2010), and to compare native and introduced diversity. We found that native populations are weakly but

significantly differentiated, but this differentiation does not align with geography or river system, as noted previously (Sui et al. 2009). One reason might be the exchange of crabs for commercial farming. Their extended planktonic larval period could also contribute to weak levels of genetic differentiation. This somewhat "chaotic" distribution of haplotypes makes it impossible to extrapolate the geographic distribution of genetic diversity, and precludes the assignment of broader regions as source regions. We can only discuss specific sampling sites as being more or less likely sources of introduction, as the genetic makeup of even the closest neighbor of any one site can be very different, e.g. in the case of Feiyunjiang and Oujiang. A more extensive sampling with regard to number of individuals and populations is certainly desirable to understand the patterns of diversity in the native range better. The fact that most populations were differentiated provides nonetheless a working baseline to assign source populations.

Most of the native populations had positive Tajima's  $D$  values, albeit not significantly different from zero, which is generally interpreted as populations being in mutation-drift equilibrium. It suggests that populations did not expand, shrink, or undergo recent selective sweeps at the mitochondrial genome. This pattern of genetic stability is anticipated for native populations. That we find the same pattern in most introduced populations is unexpected. We would expect to find negative Tajima's  $D$  values, indicative of recent bottlenecks. It seems unlikely that the introduced populations are already at equilibrium. Instead, an invasion of sufficient number of individuals that brought over a substantial amount of the native diversity could explain the observed pattern, either in a single or in multiple introduction events. In concordance with this idea, genetic diversity is not significantly lower in invasive populations, as would be expected when few individuals invade a new range.

The most distinct feature of the introduced range is the presence of seven haplotypes that have not been sampled in the native range. These haplotypes appear restricted to Northern Germany. Their distribution dominates the population structure in Europe, which divides populations with and without those unique alleles. We recovered more population structure than identified by Hänfling et al. (2002), and echo the findings of Otto (2012), who generated and analyzed the Northern German COI data initially. She did not take the other known data into account, however, thus limiting her conclusions.

### **Plausible source populations of the Chinese mitten crab invasion**

Hänfling et al. (2002) conducted the first search for source populations of the European and US invasion. Using COI sequence data, they identified five haplotypes that occurred in both China (three populations sampled) and Europe (five populations sampled). They did not find significant population structure in either the native or introduced range, but observed a significant differentiation between those two. This was due to a haplotype common in all European and US populations, but absent in the native range. They used the presence of this haplotype as evidence for multiple introduc-

tions into Europe, and a secondary introduction of the United States from Europe. We identified this haplotype (H4) in one site in the native range, in Feiyunjiang. The other two haplotypes from Feiyunjiang (H1 and H3) are also common in Europe and the US, making this location the most plausible source of the invasion of those sampled so far. Feiyunjiang is located between the large ports of Shanghai and Xiamen (compare Figs 1–5), which are suitable donor locations, each of the many departing commercial vessels from their ports acting as potential invasion vectors. The results of the analysis of pairwise population differentiation are concordant with these findings. Feiyunjiang clusters with several of the introduced populations, and is not significantly differentiated from them. The last haplotype found in both native and introduced range is H2. It was not detected in Feiyunjiang, but was present in Dalian City, Wuhu, part of the Yangtze River and Liaohe. Any of these locations could therefore be the source population of a second independent invasion into Europe. Unfortunately, the first three sites are only represented by two individuals each, making detailed comparisons of haplotype compositions between these native and invaded sites impossible. Alternatively to a second independent introduction event from a different location, this haplotype might occur in low frequencies in Feiyunjiang, but was not recovered there due to small sample sizes. In this case, the colonization of Europe by all the above-mentioned haplotypes could have been due to a single successful invasion event. Given the low frequency of the haplotype H2 in the invaded range, this is clearly a possibility. In general, the large native range is under-sampled with regard to the number of individuals and number of populations (Muirhead et al. 2008; Geller et al. 2010). This becomes especially important given the weak and chaotic population structure across the native range, which limits our power to predict the region of origin. It is, however, highly unlikely that the northern range of Chinese mitten crabs, e.g. Russia and South Korea, where we recovered only haplotypes absent from Europe and the United States, was the source of the invasion.

The US populations of Chinese mitten crabs have been speculated to be secondarily introduced from Europe (Hänfling et al. 2002). Based on our analyses, this remains a possibility, as the West coast populations are of a single haplotype, which was only found in Feiyunjiang in the native range, but is widespread in Europe. However, whatever led to the successful invasion of Europe from Feiyunjiang (or another native population with similar haplotype composition) might also have led to the successful invasion of the United States. The low genetic diversity of these populations certainly argues for the invasion of few individuals. In contrast, the genetic diversity of European invaders is indicative of the invasion of several individuals. A single sequence available for a Chinese mitten crab from the East coast of the United States from an unestablished population (Benson and Fuller 2019) is genetically distinct from the monotypic West coast population, advocating for an independent invasion of the East coast. The East coast haplotype (H3) is present in Europe, but is also relatively widespread in China, making both a secondary invasion from Europe, or a direct invasion from Asia, equally likely. A secondary invasion from the West coast of the United States is, however, highly unlikely.

The analyses of genetic distances between populations echo on the one hand some of the results obtained by the comparison of haplotype identity between native and invaded ranges, and highlight, on the other hand, some of the difficulties associated with population genetic analyses of non-equilibrium scenarios pervasive during invasions. We found two clusters of mixed native and introduced populations: the first cluster contained populations from across Europe and Feiyunjiang. In line with the results of haplotype identity, Feiyunjiang is therefore a plausible source population. The second cluster contains populations monotypic for the most widespread haplotype. In this case, the native populations of Tongan and Oujiang have the same genetic makeup as the introduced populations of Hemmelsdorf, Eckernförde and Tagus, but this similarity may well be due to small populations and strong drift in the introduced populations, which could have eradicated much of the genetic diversity. Thus the second cluster of native and introduced populations cannot be interpreted as a separate introduction.

### **The distribution and origin of the uniquely Northern German haplotypes**

The restricted distribution of the haplotypes H12 to H17 in northern Germany could either reflect a snapshot taken during an ongoing expansion or ecological restrictions. The most recent samples included in our analyses are those Northern German samples with unique haplotypes collected between 2008 and 2010 (Otto 2012). No sites outside of Northern Germany were sampled, which means these haplotypes may have already spread throughout the remaining European range. One indication of a recent origin or arrival of these haplotypes can be gained via a comparison with other studies that included sites close by. Hänfling et al. (2002) included a site in the Elbe river near Osterholz, from which they collected 15 crabs between 1999 and 2000 (see suppl. material 1: table S1). This site is downstream from the Finkenwerder site sampled by Otto (2012). None of the crabs collected in the Elbe between 1999 and 2000 had any of these uniquely Northern German haplotypes that were common between 2008 and 2010. Herborg et al. (2007) analyzed six microsatellite markers for six European populations, including the same Osterholz site. Otto (2012) also analyzed her samples with microsatellite markers, including four of the markers analyzed by Herborg et al. (2007). The raw data are not available (as is so often the case for microsatellite data), but at the four common microsatellite loci, the Finkenwerder samples from 2008–2010 show higher allelic richness ( $A = 9.2\text{--}10.2$ ) than the Osterholz samples collected ten years earlier ( $A = 4.3\text{--}9.1$ ) across all four loci. While this is suggestive of a recent addition of genetic diversity, it does not preclude a restricted distribution of those haplotypes either, as Chinese mitten crabs commonly show genetic structure within the same river systems (Herborg et al. 2007; Sui et al. 2009). A broad sampling of current mitten crab genetic diversity in the invaded range would clarify how widely distributed those haplotypes really are.

The origin of the haplotypes only found in Northern Germany remains mysterious. If we interpret the absence of those haplotypes in the Osterholz samples, and the presence of two of these haplotypes in the Finkenwerder samples ten years later as the recent and simultaneous addition of these haplotypes to Europe, the most plausible scenario is a cryptic invasion from an unsampled native site. The source of such a cryptic invasion might be located in the northern range of Chinese mitten crabs. Overall, the number of analyzed native populations was rather small given the large range of Chinese mitten crabs (Fig. 4A). The available data did not include some large ports, such as Tianjin, Nampo, Daesan and Hungnam, all located in the northern range of Chinese mitten crabs (Fig. 1), which could be suitable donor areas. Based on the known distribution of Chinese mitten crabs, the very large ports around Hong Kong can be excluded as the invasion source because Hepu mitten crabs, not Chinese mitten crabs, occur in southern China (Wang et al. 2008; Xu et al. 2009) (Fig. 1).

Under a scenario of multiple invasions, the amino acid substitutions we found in all of the uniquely Northern German haplotypes evolved in the native range, and were introduced during the cryptic invasion. Whether these haplotypes confer indeed a selective advantage cannot be answered with certainty. They may carry, in fact, neutral or slightly deleterious mutations but have been swept to high frequencies during a recent strong selection event at a linked region of the genome (Smith and Haigh 1974). The contemporaneous discovery of a novel physiology and behavior in Northern German mitten crabs, which allows them to complete the larval cycle in the brackish Baltic Sea (Otto 2012), may not be a coincidence. The expected range expansion caused by this novel ecology has already been documented by the recent and widespread occurrence of ovigerous females in the Eastern Baltic Sea (Ojaveer et al. 2007). Similarly, the occurrence of mitten crabs throughout much of the freshwater system of Sweden was hard to explain, invoking long-distance migration of crabs from their North Sea spawning grounds (Drotz et al. 2010). We suspect that these crabs belong to the same physiological type as the crabs investigated by Otto (2012), and are able to complete their larval cycle in the Baltic Sea.

Alternative hypotheses to a novel introduction can explain the origin of these uniquely Northern German haplotypes. In our opinion, the second most likely explanation is that the unique haplotypes evolved in the introduced range. Given that all of these haplotypes had one to three AAS, these haplotypes might have evolved rapidly in the introduced range in response to novel ecological conditions. Moreover, all of these uniquely Northern German haplotypes are closest related to a haplotype that was already present in Northern Europe (Fig. 3A). An argument against this hypothesis is that we would have expected to find some temporal sequence of haplotype evolution, with at least a few of the haplotypes occurring in earlier samples. Another hypothesis is that these haplotypes have been in Europe since the initial introduction, but either only recently increased in abundance, or were always restricted to Northern Germany, which had not been sampled before 2008. While genetic structure among sites or sampling years of the same river system exists (Herborg et al. 2007; Sui et al. 2009), crabs

have to migrate along rivers to get to their marine mating and breeding grounds. Thus some mixing of haplotypes should occur along the river. Given the small sample sizes of older sites, a recent expansion of these haplotypes from standing variation is clearly possible. It is not obvious, however, why these haplotypes would have remained at very low frequencies for about 100 years, since their introduction.

Lastly, we cannot ignore the fact that all sequences with uniquely Northern German haplotypes were collected by Otto (2012). If she had fallen prey to sequencing errors, we might expect randomly changed bases throughout each sequence, leading to many haplotypes present only once, and/or to the presence of stop codons. The fact that she sequenced many individuals with the same haplotype, and these translate to functional amino-acid sequences, makes sequencing error an unlikely source of these haplotypes. Furthermore, DNA fragments with uncertain base calls were sequenced in both directions (Otto 2012), which should remove possible sequencing artifacts caused by faulty sequencing chemistry (Wares, pers. comm.).

At this point, we cannot determine if the high and unique haplotype diversity of Northern Germany is due to novel, potentially adaptive mutations that occurred after introduction, or due to multiple invasions. To clarify the origin of the unique haplotypes, we propose three approaches. Firstly, a more extensive sampling of the native range should identify if these haplotypes are present in the native range. Such sampling has already taken place (Tang et al. 2003; Sun et al. 2005), but we were unable to incorporate these data into our study because they used genetic markers not yet applied to the introduced populations. Thus we propose that future genotyping efforts of invasive specimens should include these genetic markers as well. Secondly, population genetic analyses of invasive mitten crabs from museum collections could identify the temporal pattern of haplotype appearance. A sudden appearance of all unique haplotypes during the invasion history would hint at a new invasion event, while a stepwise appearance of novel haplotypes would be consistent with their evolution within the introduced range. Such pattern would, however, also be consistent with multiple additional invasion events, each introducing one or a few novel haplotypes. Lastly, population genomic analyses of native and invasive mitten crabs might reveal if the potentially adaptive haplotypes arose within invasive populations, in which case most of the genome of introduced Chinese mitten crabs should be undifferentiated and only small regions be highly differentiated. In contrast, a second introduction should show a more or less even differentiation across the genome. These efforts are aided by the recent publication of the nuclear genome of Chinese mitten crabs (Song et al. 2016) as well as the complete mitogenomes of several mitten crab species (Liu et al. 2015; Li et al. 2016).

## Conclusions

This study uncovered complex population genetic pattern of invasive mitten crabs. Some of our findings are unambiguous, such as the presence of the mitochondrial genome of a second mitten crab species, the Japanese mitten crab, in Europe, suggest-

ing either a cryptic invasion of this species or previous hybridization between Chinese and Japanese mitten crabs. This new European addition was only revealed by our data synthesis, which included barcoding data collected from various entities of a few individuals. The genetic diversity within European populations of Chinese mitten crabs remains puzzling, including the presence of several amino acid substitutions in haplotypes found only in Northern Germany. Taken together with the contemporaneous occurrence of a novel physiology and behavior in the same populations, it is possible that carriers of this haplotype have an adaptive advantage. Given the negative impacts of mitten crabs as an invasive species, we can only urge to monitor these invasive populations closely, using genetic tools such as the commonly used barcoding locus COI (Darling and Blum 2007). Simultaneously, genomic and historical data could greatly enhance our understanding of the invasion process. We show that mitten crabs in Europe are melting pots of genetic diversity (Geller et al. 2010), making them prime targets to study cryptic invasions and possibly also rapid adaptations.

## Acknowledgements

We would like to thank Dr. John Wares (UGA) for his improvements to the manuscripts, and his ongoing support for CES' work. We would also like to thank the reviewers Charles Fransen and Milan Koch, as well as the Neobiota subject editor Adam Petrušek. This study was supported by a grant from the German Federal Ministry of Education and Research (Bundesministerium für Bildung und Forschung, project number 01UQ1711).

## References

- Aktas C (2015) haplotypes: Haplotype Inference and Statistical analysis of genetic Variation.
- Anger K (1991) Effects of temperature and salinity on the larval development of the Chinese mitten crab *Eriocheir sinensis* (Decapoda: Grapsidae). *Marine Ecology Progress Series* 72: 103–110. <https://doi.org/10.3354/meps072103>
- Archer FI, Adams PE, Schneiders BB (2017) strataG: An R package for manipulating, summarizing and analysing population genetic data. *Molecular Ecology Resources* 17: 5–11. <https://doi.org/10.1111/1755-0998.12559>
- Bastrop R, Blank M (2006) Multiple Invasions – A Polychaete genus enters the Baltic Sea. *Biological Invasions* 8: 1195–1200. <https://doi.org/10.1007/s10530-005-6186-6>
- Benson AJ, Fuller PL (2019) *Eriocheir sinensis*: U.S. Geological Survey, Nonindigenous Aquatic Species Database, Gainesville, FL. <https://nas.er.usgs.gov/queries/FactSheet.aspx?SpeciesID=182>
- Besnard G, Henry P, Wille L, Cooke D, Chapuis E (2007) On the origin of the invasive olives (*Olea europaea* L., Oleaceae). *Heredity* 99: 608–619. <https://doi.org/10.1038/sj.hdy.6801037>

- Bouma S, Soes D (2010) A risk analysis of the Chinese mitten crab in The Netherlands.
- Card DC, Perry BW, Adams RH, Schield DR, Young AS, Andrew AL, Jezkova T, Pasquesi GIM, Hales NR, Walsh MR, Rochford MR, Mazzotti FJ, Hart KM, Hunter ME, Castoe TA (2018) Novel ecological and climatic conditions drive rapid adaptation in invasive Florida Burmese pythons. *Molecular Ecology* 27: 4744–4757. <https://doi.org/10.1111/mec.14885>
- Chen K-Y, Marschall EA, Sovic MG, Fries AC, Gibbs HL, Ludsin SA (2018) assignPOP: An R package for population assignment using genetic, non-genetic, or integrated data in a machine-learning framework. *Methods in Ecology and Evolution* 9: 439–446. <https://doi.org/10.1111/2041-210X.12897>
- Chen X, Wang J, Huang L, Yue W, Zou J, Yuan C, Lu G, Wang C (2017) Evolutionary relationship of three mitten crabs (*Eriocheir* sp) revealed by mitogenome and 5S ribosomal DNA analysis. *Aquaculture and Fisheries* 2: 256–261. <https://doi.org/10.1016/j.aaf.2017.10.004>
- Chu K, Ho H, Li C, Chan T (2003) Molecular phylogenetics of the mitten crab species in *Eriocheir*, sensu lato (Brachyura: Grapsidae). *Journal of Crustacean Biology* 23: 738–746. <https://doi.org/10.1651/C-2347>
- Clark PF, Rainbow PS, Robbins RS, Smith B, Yeomans WE, Thomas M, Dobson G (1998) The alien Chinese mitten crab, *Eriocheir sinensis* (Crustacea: Decapoda: Brachyura), in the Thames catchment. *Journal of the Marine Biological Association of the United Kingdom* 78: 1215–1221. <https://doi.org/10.1017/S002531540004443X>
- Costa, FO, DeWaard JR, Boutillier J, Ratnasingham S, Dooh RT, Hajibabaei M, Hebert PD (2007). Biological identifications through DNA barcodes: the case of the Crustacea. *Canadian Journal of Fisheries and Aquatic Sciences* 64: 272–295. <https://doi.org/10.1139/f07-008>
- Czerniejewski P, Skuza L, Drotz MK, Berggren M (2012) Molecular connectedness between self and non self-sustainable populations of Chinese mitten crab (*Eriocheir sinensis*, H. Milne Edwards, 1853) with focus to the Swedish Lake Vänern and the Oder and Vistula River in Poland. *Hereditas* 149: 55–61. <https://doi.org/10.1111/j.1601-5223.2012.02246.x>
- Darling JA, Blum MJ (2007) DNA-based methods for monitoring invasive species: a review and prospectus. *Biological Invasions* 9: 751–765. <https://doi.org/10.1007/s10530-006-9079-4>
- Darling JA (2011) Interspecific Hybridization and Mitochondrial introgression in invasive *Carcinus* shore crabs. *PLOS ONE* 6: e17828. <https://doi.org/10.1371/journal.pone.0017828>
- Dittel AI, Epifanio CE (2009) Invasion biology of the Chinese mitten crab *Eriocheir sinensis*: a brief review. *Journal of Experimental Marine Biology and Ecology* 374: 79–92. <https://doi.org/10.1016/j.jembe.2009.04.012>
- Glugosch K, Parker I (2008) Founding events in species invasions: genetic variation, adaptive evolution, and the role of multiple introductions. *Molecular Ecology* 17: 431–449. <https://doi.org/10.1111/j.1365-294X.2007.03538.x>
- Drotz MK, Berggren M, Lundberg S, Lundin K, von Proschwitz T (2010) Invasion routes, current and historical distribution of the Chinese mitten crab (*Eriocheir sinensis* H. Milne Edwards, 1853) in Sweden. *Aquatic Invasions* 5: 387–396. <https://doi.org/10.3391/ai.2010.5.4.08>

- Excoffier L, Smouse PE, Quattro JM (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics* 131: 479–91.
- Filipová L, Holdich DM, Lesobre J, Grandjean F, Petrusek A (2010). Cryptic diversity within the invasive virile crayfish *Orconectes virilis* (Hagen, 1870) species complex: new lineages recorded in both native and introduced ranges. *Biological Invasions* 12: 983–989.
- Filipová L, Lieb DA, Grandjean F, Petrusek A (2011) Haplotype variation in the spiny-cheek crayfish *Orconectes limosus*: colonization of Europe and genetic diversity of native stocks. *Journal of the North American Benthological Society* 30: 871–881. <https://doi.org/10.1007/s10530-009-9526-0>
- Geller JB, Darling JA, Carlton JT (2010) Genetic perspectives on marine biological invasions. *Annual Review of Marine Science* 2: 367–393. <https://doi.org/10.1146/annurev.marine.010908.163745>
- Genton B, Shykoff J, Giraud T (2005) High genetic diversity in French invasive populations of common ragweed, *Ambrosia artemisiifolia*, as a result of multiple sources of introduction. *Molecular Ecology* 14: 4275–4285. <https://doi.org/10.1111/j.1365-294X.2005.02750.x>
- Gilbey V, Attrill MJ, Coleman RA (2008) Juvenile Chinese mitten crabs (*Eriocheir sinensis*) in the Thames estuary: distribution, movement and possible interactions with the native crab *Carcinus maenas*. *Biological Invasions* 10: 67–77. <https://doi.org/10.1007/s10530-007-9110-4>
- Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, Gascuel O (2010) New algorithms and methods to estimate Maximum-Likelihood Phylogenies: Assessing the performance of PhyML 3.0. *Systematic Biology* 59: 307–321. <https://doi.org/10.1093/sysbio/syq010>
- Guo J, Ng N, Dai A, Ng PK (1997) The taxonomy of three commercially important species of mitten crabs of the genus *Eriocheir* De Haan, 1835 (Crustacea: Decapoda: Brachyura: Grapsidae). *Raffles Bulletin of Zoology* 45: 445–476.
- Hänfling B, Carvalho GR, Brandl R (2002) mt-DNA sequences and possible invasion pathways of the Chinese mitten crab. *Marine Ecology Progress Series* 238: 307–310. <https://doi.org/10.3354/meps238307>
- Herborg L-M, Rushton S, Clare A, Bentley M (2003) Spread of the Chinese mitten crab (*Eriocheir sinensis* H. Milne Edwards) in continental Europe: analysis of a historical data set. *Migrations and Dispersal of Marine Organisms*. Springer, pp 21–28. [https://doi.org/10.1007/978-94-017-2276-6\\_3](https://doi.org/10.1007/978-94-017-2276-6_3)
- Herborg L-M, Rushton S, Clare A, Bentley M (2005) The invasion of the Chinese mitten crab (*Eriocheir sinensis*) in the United Kingdom and its comparison to continental Europe. *Biological Invasions* 7: 959–968. <https://doi.org/10.1007/s10530-004-2999-y>
- Herborg L-M, 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. <https://doi.org/10.1111/j.1365-294X.2006.03133.x>
- Jensen GC, Armstrong DA (2004) The occurrence of the Japanese mitten crab, *Eriocheir japonica* (De Haan), on the west coast of North America. *California Fish and Game* 90: 94–99.
- Jost L (2008) GST and its relatives do not measure differentiation. *Molecular Ecology* 17: 4015–4026. <https://doi.org/10.1111/j.1365-294X.2008.03887.x>

- Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C, Thierer T, Ashton B, Mentjies P, Drummond A (2012) Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data (version 8.0.3). *Bioinformatics* 28: 1647–1649. <https://doi.org/10.1093/bioinformatics/bts199>
- Kolbe JJ, Glor RE, Schettino LR, Lara AC, Larson A, Losos JB (2007) Multiple sources, admixture, and genetic variation in introduced *Anolis* lizard populations. *Conservation Biology* 21: 1612–1625. <https://doi.org/10.1111/j.1523-1739.2007.00826.x>
- Kolbe JJ, Glor RE, Rodríguez Schettino L, Lara AC, Larson A, Losos JB (2004) Genetic variation increases during biological invasion by a Cuban lizard. *Nature* 431: 177. <https://doi.org/10.1038/nature02807>
- Komai T, Yamasaki I, Kobayashi S, Yamamoto T, Watanabe S (2006) *Eriocheir ogasawaraensis* Komai, a new species of mitten crab (Crustacea: Decapoda: Brachyura: Varunidae) from the Ogasawara Islands, Japan, with notes on the systematics of *Eriocheir* De Haan, 1835. *Zootaxa* 1168: 1–20. <https://doi.org/10.11646/zootaxa.1168.1.1>
- Li G, Xu J, Wang H, Chen X, Xu Y, Zheng Y, Mao X, Yang S (2016) Identification and analysis of the complete mitochondrial genome sequence of *Eriocheir sinensis*. *Mitochondrial DNA Part A* 27: 3923–3924. <https://doi.org/10.3109/19401736.2014.987267>
- Lowe S, Browne M, Boudjelas S, De Poorter M (2000) 100 of the world's worst invasive alien species: a selection from the global invasive species database. Invasive Species Specialist Group Auckland.
- Meiklejohn CD, Montooth KL, Rand DM (2007) Positive and negative selection on the mitochondrial genome. *Trends in Genetics*, 23: 259–263. <https://doi.org/10.1016/j.tig.2007.03.008>
- Muirhead JR, Gray DK, Kelly DW, Ellis SM, Heath DD, MacIsaac HJ (2008) Identifying the source of species invasions: sampling intensity vs. genetic diversity. *Molecular Ecology* 17: 1020–1035. <https://doi.org/10.1111/j.1365-294X.2008.03669.x>
- Naser MD, Page TJ, Ng N, Apel M, Yasser AG, Bishop JM, Ng PK, Clark PF (2012) Invasive records of *Eriocheir hepuensis* Dai, 1991 (Crustacea: Brachyura: Grapsoidea: Varunidae): Implications and taxonomic considerations. *BioInvasions Records* 1: 71–86. <https://doi.org/10.3391/bir.2012.1.1.15>
- Nei M (1987) *Molecular evolutionary genetics*. Columbia University Press, New York. <https://doi.org/10.7312/nei-92038>
- Nei M, Maruyama T, Chakraborty R (1975) The bottleneck effect and genetic variability in populations. *Evolution* 29: 1–10. <https://doi.org/10.1111/j.1558-5646.1975.tb00807.x>
- Nei M, Tajima F (1981) Genetic drift and estimation of effective population size. *Genetics* 98: 625–640.
- Ojaveer H, Gollasch S, Jaanus A, Kotta J, Laine AO, Minde A, Normant M, Panov VE (2007) Chinese mitten crab *Eriocheir sinensis* in the Baltic Sea—a supply-side invader? *Biological Invasions* 9: 409–418. <https://doi.org/10.1007/s10530-006-9047-z>
- Otto T (2012) *Reproduction Biology and Population Genetics of the alien Chinese mitten crab (Eriocheir sinensis) in Schleswig-Holstein*. Dissertation, Christian-Albrechts University.

- Panning A, Peters N (1933) Die chinesische Wollhandkrabbe (*Eriocheir sinensis* H. Milne Edwards) in Deutschland. Zoologischer Anzeiger (Suppl. material) 104: 1–180.
- Paradis E (2010) PEGAS: an R package for population genetics with an integrated-modular approach. *Bioinformatics* 26: 419–420. <https://doi.org/10.1093/bioinformatics/btp696>
- Paradis E, Claude J, Strimmer K (2004) APE: analyses of phylogenetics and evolution in R language. *Bioinformatics* 20: 289–290. <https://doi.org/10.1093/bioinformatics/btg412>
- Pentinsaari M, Salmel, H, Mutanen M, Roslin T (2016) Molecular evolution of a widely-adopted taxonomic marker (COI) across the animal tree of life. *Scientific Reports* 6: 35275. <https://doi.org/10.1038/srep35275>
- Prentis PJ, Wilson JR, Dormontt EE, Richardson DM, Lowe AJ (2008) Adaptive evolution in invasive species. *Trends in Plant Science* 13: 288–294. <https://doi.org/10.1016/j.tplants.2008.03.004>
- R Development Core Team (2018) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna. <https://www.R-project.org/>
- Raupach MJ, Barco A, Steinke D, Beermann J, Laakmann S, Mohrbeck I, Neumann H, Kihara TC, Pointner K, Radulovici A, Segelken-Voigt A, Wesse C, Knebelberger T (2015) The application of DNA barcodes for the identification of marine crustaceans from the North Sea and adjacent regions. *PLOS ONE* 10: e0139421. <https://doi.org/10.1371/journal.pone.0139421>
- Roman J (2006) Diluting the founder effect: cryptic invasions expand a marine invader's range. *Proceedings of the Royal Society of London B: Biological Sciences* 273: 2453–2459. <https://doi.org/10.1098/rspb.2006.3597>
- Rudnick DA, Halat KM, Resh VH (2000) Distribution, ecology and potential impacts of the Chinese mitten crab (*Eriocheir sinensis*) in San Francisco Bay. UC Berkeley: University of California Water Resources Center.
- Rudnick DA, Hieb K, Grimmer KF, Resh VH (2003) Patterns and processes of biological invasion: the Chinese mitten crab in San Francisco Bay. *Basic and Applied Ecology* 4: 249–262. <https://doi.org/10.1078/1439-1791-00152>
- Rudnick DA, Chan V, Resh VH (2005) Morphology and impacts of the burrows of the Chinese mitten crab, *Eriocheir sinensis* H. Milne Edwards (Decapoda, Grapsoidea), in south San Francisco Bay, California, USA. *Crustaceana* 78: 787. <https://doi.org/10.1163/156854005774445500>
- Saltonstall K (2002) Cryptic invasion by a non-native genotype of the common reed, *Phragmites australis*, into North America. *Proceedings of the National Academy of Sciences USA* 99: 2445. <https://doi.org/10.1073/pnas.032477999>
- Song L, Bian C, Luo Y, Wang L, You X, Li J, Qiu Y, Ma X, Zhu Z, Ma L, Wang Z, Lei Y, Qiang J, Li H, Yu J, Wong A, Xu J, Shi Q, Xu P (2016) Draft genome of the Chinese mitten crab, *Eriocheir sinensis*. *GigaScience* 5: 5. <https://doi.org/10.1186/s13742-016-0112-y>
- Smith JM, Haigh J (1974) The hitch-hiking effect of a favourable gene. *Genetical Research*, 23: 23–35. <https://doi.org/10.1017/S0016672300014634>
- Sui L, Zhang F, Wang X, Bossier P, Sorgeloos P, Hänfling B (2009) Genetic diversity and population structure of the Chinese mitten crab *Eriocheir sinensis* in its native range. *Marine Biology* 156: 1573–1583. <https://doi.org/10.1007/s00227-009-1193-2>

- Sun H, Zhou K, Song D (2005) Mitochondrial genome of the Chinese mitten crab *Eriocheir japonica sinensis* (Brachyura: Thoracotremata: Grapsoidea) reveals a novel gene order and two target regions of gene rearrangements. *Gene* 349: 207–217. <https://doi.org/10.1016/j.gene.2004.12.036>
- Tajima F (1989) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* 123: 595–595.
- Tang B, Zhou K, Song D, Yang G, Dai A (2003) Molecular systematics of the Asian mitten crabs, genus *Eriocheir* (Crustacea: Brachyura). *Molecular Phylogenetics and Evolution* 29: 309–316. [https://doi.org/10.1016/S1055-7903\(03\)00112-X](https://doi.org/10.1016/S1055-7903(03)00112-X)
- Wang C, Li C, Li S (2008) Mitochondrial DNA-inferred population structure and demographic history of the mitten crab (*Eriocheir sensu stricto*) found along the coast of mainland China. *Molecular Ecology* 17: 3515–3527. <https://doi.org/10.1111/j.1365-294X.2008.03850.x>
- Wang J, Huang L, Cheng Q, Lu G, Wang C (2016) Complete mitochondrial genomes of three mitten crabs, *Eriocheir sinensis*, *E. hepuensis*, and *E. japonica*. *Mitochondrial DNA Part A* 27: 1175–1176. <https://doi.org/10.3109/19401736.2014.936425>
- Winter DJ (2012) mmod: an R library for the calculation of population differentiation statistics. *Molecular Ecology Resources* 12: 1158–1160. <https://doi.org/10.1111/j.1755-0998.2012.03174.x>
- Xu J, Chan T-Y, Tsang LM, Chu KH (2009) Phylogeography of the mitten crab *Eriocheir sensu stricto* in East Asia: Pleistocene isolation, population expansion and secondary contact. *Molecular Phylogenetics and Evolution* 52: 45–56. <https://doi.org/10.1016/j.ympev.2009.02.007>
- Yuan X (2005) China national fishery statistic and analysis in 2004. *Fishery and Wealth Guide* 13: 13–16.
- Zhang D, Ding G, Ge B, Zhang H, Tang B, Yang G (2014) Comparative phylogeography of two marine species of crustacean: Recent divergence and expansion due to environmental changes. *Gene* 550: 141–147. <https://doi.org/10.1016/j.gene.2014.08.006>
- Zhang D, Ding G, Ge B, Zhang H, Zhou C, Tang B (2012) The redivision of geographic population and genetic structure of *Eriocheir* in the West-Pacific Ocean. *Gene* 503: 126–130. <https://doi.org/10.1016/j.gene.2012.04.070>
- Zhang H-B, Jiang S-H, Xuan F-J, Li C-F, Zhang D-Z, Zhou C-L, Tang B-P (2015) The complete mitochondrial genome of *Eriocheir japonica sinensis* (Decapoda: Varunidae) and its phylogenetic analysis. *Biochemical Systematics and Ecology* 62: 241–248 <https://doi.org/10.1016/j.bse.2015.09.008>

## Supplementary material 1

### Tables S1–S3

Authors: Sarah Hayer, Dirk Brandis, Günther B. Hartl, Christine Ewers-Saucedo

Data type: species data

Explanation note: Table S1. *Eriocheir japonica*. Table S2. *Eriocheir sinensis*. Table S3. Pairwise population.

Copyright notice: This dataset is made available under the Open Database License (<http://opendatacommons.org/licenses/odbl/1.0/>). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.

Link: <https://doi.org/10.3897/neobiota.50.34881.suppl1>

## Supplementary material 2

### Figures S1–S6

Authors: Sarah Hayer, Dirk Brandis, Günther B. Hartl, Christine Ewers-Saucedo

Data type: multimedia

Explanation note: Figure S1. Maximum likelihood tree of mitten crab COI sequences from GenBank and BOLD. Figure S2. Geographic distribution of COI haplotype H2 of the Chinese mitten crab (*Eriocheir sinensis*). Figure S3. Geographic distribution of COI haplotype H3 of the Chinese mitten crab (*Eriocheir sinensis*). Figure S4. Geographic distribution of COI haplotype H4 of the Chinese mitten crab (*Eriocheir sinensis*). Figure S5. Multidimensional scaling plot of the first and third axes based on Jost's D distances between sampling locations. Figure S6. Codons of the COI sequence of Chinese mitten crabs that translate to amino acid substitutions (AAS).

Copyright notice: This dataset is made available under the Open Database License (<http://opendatacommons.org/licenses/odbl/1.0/>). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.

Link: <https://doi.org/10.3897/neobiota.50.34881.suppl2>