Journal of FISH BIOLOGY

Published for the Fisheries Society of the British Isles

*Journal of Fish Biology* (2012) **80**, 181–194
doi:10.1111/j.1095-8649.2011.03165.x, available online at wileyonlinelibrary.com

# Individual genetic tagging for teleosts: an empirical validation and a guideline for ecologists

D. Andreou*, J. Vacquie-Garcia†‡, J. Cucherousset*†‡§, S. Blanchet†‖, R. E. Gozlan* and G. Loot‡‖

**Centre for Conservation Ecology and Environmental Sciences, School of Applied Sciences, Bournemouth University, Poole, Dorset BH12 5BB, U.K., †CNRS, UPS, ENFA, UMR5174 EDB (Laboratoire Évolution et Diversité Biologique), 118 route de Narbonne, F-31062 Toulouse, France, ‡Université de Toulouse, UPS, UMR5174 EDB, F-31062 Toulouse, France and ‖Station d'Ecologie Expérimentale du CNRS à Moulis, USR 2936, 09200 Moulis, France*

*(Received 20 October 2010, Accepted 15 October 2011)*

The efficiency of individual genetic tagging was determined by using passive integrated transponders (PIT) as a comparative conventional tagging method. Fifty-five common dace *Leuciscus leuciscus* were captured in the wild, PIT tagged and fin clipped (for DNA analysis). Thirty fish were recaptured on three occasions and tissue samples were collected. Using 18 microsatellite loci, 79–94% of the recaptures were correctly assigned. Experience with scoring *L. leuciscus* microsatellites led to more individuals correctly assigned. Allowing matches that differed by one or two alleles resulted in 100% of all recaptures successfully assigned irrespective of the observer. Reducing the set of loci to five to six loci appropriately selected did not affect the assignment rate, demonstrating that costs can be subsequently reduced. Despite their potential benefits, the application of genetic tags for teleosts has been limited. Here, it was demonstrated that genetic tagging could be applied, and a clear guideline (flowchart) is provided on how this method can be developed for teleosts and other organisms, with subsequent practical applications to ecology, evolutionary biology and conservation management.

© 2011 The Authors
Journal of Fish Biology © 2011 The Fisheries Society of the British Isles

Key words: *Leuciscus leuciscus*; mark–recapture; microsatellite; passive integrated transponder.

## INTRODUCTION

Identifying and monitoring individuals is a cornerstone in ecological, evolutionary and conservation research (Palsbøll, 1999). Many types of external and internal tagging methods have been used to provide valuable information for many animal species. For example, bird migration has been extensively studied by the use of banding (Buckley *et al*., 1998) while coded wire tags have helped to determine fish population structure and have facilitated fisheries management decisions (Hammer & Blackenship, 2001). More recently, passive integrated transponder (PIT) tags have

§Author to whom correspondence should be addressed. Tel.: +33 5 61 55 84 61; email: julien.cucherousset@cict.fr



© 2011 The Authors
Journal of Fish Biology © 2011 The Fisheries Society of the British Isles

been used in the study of habitat use, growth and population characteristics of fishes, mammals, reptiles and amphibians (Gibbons & Andrews, 2004; Cucherousset *et al*., 2008, 2009).

These conventional methods suffer from several limitations arising, for instance, from potential tag loss, alteration, misreading or misidentification (Gibbons & Andrews, 2004). Tags such as visible implant elastomers can be lost, for example, during the metamorphosis of amphibians from larval to adult stages (Grant, 2008) and tag retention can be highly variable between physical tagging methods (Gibbons & Andrews, 2004; Bolland *et al*., 2009; Campbell *et al*., 2009). Moreover, the use of physical tagging methods is often limited by the minimal size of the organism that can be tagged (Acolas *et al*., 2007), *i.e.* small species and young individuals, potentially leading to a biased view of the population.

Advancement of molecular techniques and their use in ecology has led to a better understanding of population dynamics (DeYoung & Honeycutt, 2005). With the use of DNA fragments such as microsatellites, genetic fingerprinting has been applied to numerous eukaryotes allowing individual-based genetic analyses and individual tagging (Palsbøll, 1999). Individual genetic tagging is based on capture–recapture (Lukacs & Burnham, 2005) and consists of building a genetic profile for each captured and recaptured individual using multiple loci. These individual genetic profiles are subsequently used to reassign individuals and identify recaptures. Genetic tagging offers a number of advantages compared to conventional tagging methods, including 100% tag retention, no minimum individual size and the use of the genetic data to answer further ecological and evolutionary questions (Palsbøll, 1999).

Individual genetic tagging has also been extensively used with mammals and notable examples include North Atlantic humpback whales *Megaptera novaeangliae* (Palsbøll *et al*., 1997), North Atlantic right whales *Eubalaena glacialis* (Frasier *et al*., 2009), black bears *Ursus americanus* and brown bears *Ursus arctos* (Woods *et al*., 1999). In fishes, genetic tagging has been used in identifying released fishes in stock enhancement programmes (Perez-Enriquez & Taniguchi, 1999) and determining the reproductive success of individual fishes (Gross & Kapuscinski, 1997). As far as is known, individual genetic tagging has rarely been used for mark–recapture studies in fishes, and the few exceptions only concern large-bodied non-teleost species; *i.e.* lemon sharks *Negaprion brevirostris* (Poey 1868) (Feldheim *et al*., 2002) and pallid sturgeon *Scaphirhynchus albus* (Forbes & Richardson 1905) (DeHaan *et al*., 2008). Teleosts comprise the largest and most diverse group of vertebrates (Ravi & Venkatesh, 2008) and this group includes species of high economic and ecological value such as salmonids and cyprinids. Individual genetic tagging therefore has the potential to become a new and promising tool that could provide insightful information for the study of teleost ecology, evolution and conservation.

Palsbøll (1999) stressed the need for the technique to be optimized and thoroughly validated prior to its application. The present study describes the validation for the use of individual genetic tagging in teleosts using a cyprinid [common dace *Leuciscus leuciscus* (L. 1758)] as the model species. Specifically, the aims of the present study were (1) to use PIT tagging as a comparative conventional tagging method to determine the efficiency of individual genetic tagging and (2) to develop an optimized protocol for the use of individual genetic tagging in *L. leuciscus*. Finally, a

© 2011 The Authors

guideline for developing individual genetic tagging is provided for ecologists aiming to use this approach with a new model species.

## MATERIALS AND METHODS

### FIELD SURVEY

The study was conducted in the lower reach (520 m long) of a millstream on the River Frome (Dorset, U.K.). First, the study consisted of capturing and PIT-tagging individuals (16 June and 17 July 2008) by dividing the site into 50 m reaches with stop-nets and sampling the fish by electrofishing (50 Hz pulsed D.C.) using two successive passes. Captured *L. leuciscus* were anaesthetized with 2-phenoxyethanol, measured (fork length, $L_F$, to the nearest mm), weighed (nearest 0·1 g) and a PIT tag (11·5 mm × 2·1 mm; ID 100, EID Aalten; www.dorset.nu) was inserted into the peritoneal cavity using a sterile scalpel. A small sample of the pectoral fin was collected using sterile scissors and was stored in 98% molecular grade ethanol for molecular analysis. After recovery in oxygenated water, *L. leuciscus* (mean ± S.E. $L_F = 123{\cdot}3 \pm 3{\cdot}7$ mm, range = 80–207 mm, $n = 55$) were released where they were captured (Cucherousset *et al.*, 2010). Second, three recapture surveys were performed using the same protocol (19 August, 23 September and 10 October 2008) and captured fish were checked for PIT tags using a portable antenna (Cucherousset *et al.*, 2005). Recaptured PIT-tagged individuals ($n = 30$) were sampled for genetic material and subsequently released.

### MICROSATELLITE GENOTYPING

Total DNA was extracted from fin tissue using salt procedures (Aljanabi & Martinez, 1997). Individual genotypes were obtained at 18 microsatellite loci (Table I). Fifteen of these

TABLE I. Description of the 18 microsatellite loci used in the study of *Leuciscus leuciscus*

| Loci | Allele number | $H_e$ | $H_o$ | $F_{is}$ (W & C) | References |
|---|---|---|---|---|---|
| *LleC-184* | 18 | 0·884 | 0·963 | −0·081 | Dubut *et al.* (2009) |
| *Z21908* | 19 | 0·818 | 0·854 | −0·035 | Shimoda *et al.* (1999) |
| *LC27* | 16 | 0·824 | 0·836 | −0·006 | Vyskocilova *et al.* (2007) |
| *LleA-071* | 10 | 0·824 | 0·796 | 0·043 | Dubut *et al.* (2009) |
| *LceC1* | 11 | 0·813 | 0·763 | 0·071 | Larno *et al.* (2005) |
| *Rru4* | 8 | 0·739 | 0·763 | −0·023 | Barinova *et al.* (2004) |
| *CypG24* | 11 | 0·733 | 0·763 | −0·031 | Baerwald & May (2004) |
| *BL1-30* | 15 | 0·686 | 0·763 | −0·103 | Dubut *et al.* (2009) |
| *Lsou08* | 7 | 0·713 | 0·709 | 0·015 | Dubut *et al.* (2009) |
| *MFW1* | 8 | 0·668 | 0·654 | 0·030 | Crooijmans *et al.* (1997) |
| *LleC-049* | 8 | 0·648 | 0·636 | 0·028 | Dubut *et al.* (2009) |
| *CypG03* | 7 | 0·620 | 0·618 | 0·013 | Baerwald & May (2004) |
| *LC290* | 5 | 0·619 | 0·600 | 0·040 | Vyskocilova *et al.* (2007) |
| *CypG30* | 10 | 0·569 | 0·581 | −0·012 | Baerwald & May (2004) |
| *Ca12* | 5 | 0·524 | 0·436 | 0·177 | Dimsoski *et al.* (2000) |
| *Lid8* | 8 | 0·402 | 0·418 | −0·029 | Barinova *et al.* (2004) |
| *Rhca20* | 5 | 0·404 | 0·400 | 0·021 | Girard & Angers (2006) |
| *Lco5* | 7 | 0·746 | 0·127 | 0·832 | Turner *et al.* (2004) |

$H_e$, expected heterozygosity; $H_o$, observed heterozygosity; $F_{is}$, deviation from the Hardy–Weinberg equilibrium; W & C, Weir & Cockerham.

© 2011 The Authors
Journal of Fish Biology © 2011 The Fisheries Society of the British Isles, 

loci were isolated and developed on closely related cyprinid species and have been previously used in a study involving *L. leuciscus* (Blanchet *et al*., 2009). An additional three were recently developed specifically for *L. leuciscus* (Dubut *et al*., 2009). Allele numbers varied from five to 19 per locus (Table I). Loci were co-amplified using the QIAGEN® Multiplex PCR Kit (Qiagen; www.qiagen.com). PCR reactions were carried out in a 10 μl final volume containing 5–20 ng of genomic DNA, 5 μl of 2× QIAGEN Multiplex PCR Master Mix and locus-specific optimized combination of primers. PCR amplifications were performed in a Mastercycler PCR machine (Eppendorf®; www.eppendorf.com) under the following conditions: 15 min at 95° C followed by 30 cycles of a 1 min at 94° C, 1 min at 60° C and 1 min at 72° C and finally followed by a 60 min final elongation step at 72° C. Amplified fragments were then separated on an ABI PRISM 3730 automated capillary sequencer (Applied Biosystems; www.appliedbiosystems.com). Allelic sizes were then scored using GeneMapper v.4.0 (Applied Biosystems).

## DATA ANALYSES

To test whether scoring experience affected error rate in individual genetic reassignment, scoring was carried out independently by four observers. Two observers with previous experience with microsatellite scoring for *L. leuciscus* and two novice observers without previous experience with microsatellite scoring for *L. leuciscus* genotyped the data. The online software GENECAP (Wilberg & Dreher, 2004) was used to blindly assess the recapture rate according to individual genotypes. GENECAP is a Microsoft Excel macro that compares each individual genotype with all other genotypes within the dataset (*i.e.* all possible couples of individuals) to determine matching genotypes. The individual matches obtained from GENECAP can then be compared to the actual individual matches (*i.e.* those found using PIT tags) to determine the percentage of correct and incorrect matches. With reference to the statistical literature, two possible sources of mistakes, *i.e*. type I and type II errors, were identified and quantified. Type I errors were defined as matches identified from genotypes that were not observed using PIT tags. Type I errors can arise either due to tag loss or limitation of the genetic tagging (*i.e.* a recapture was falsely identified). Type II errors were defined as non-identified matches from genotypes that were observed using PIT tags (*i.e.* a match was not found although it existed). Type II errors can arise because of genotyping and scoring errors. GENECAP also provides matches that differ by one and two alleles, which allows matches to be refined. Type I and type II errors were also quantified on these incomplete matches. Match errors were assessed for each observer independently to assess whether scoring experience influenced error rates.

A re-sampling procedure was then used to assess the influence of the number of loci chosen for genotyping individuals and their expected heterozygosity on the percentage of match errors (*i.e*. percentage of errors). To do so, the corrected database was used (*i.e.* the database was corrected for potential typing errors) and 1000 independent combinations of $n$ loci, with $n$ varying from 1 to 18 (*i.e.* 18 000 combinations in total) were haphazardly created. For each combination, the mean expected heterozygosity ($H_e$) was calculated and the number of individual matches (*i.e.* observed matches) was assessed. The proportion of observed matches was calculated by dividing the number of observed matches by the total number of potential matches (*i.e.* all pair-wise comparisons). The proportion of expected matches was set as the number of true (and known) genetic reassignments, divided by the total number of potential matches. The percentage of errors was finally calculated as the proportion of observed matches less the proportion of expected matches.

Finally, a cumulative test was performed to complete the analysis presented above. The probability of two individuals within a population to share the same genotype is referred to as the probability of identity ($P_I$; Woods *et al*., 1999). $P_I$ is highly correlated to $H_e$ ($r > 0{\cdot}90$ over the 18 loci) and provides an estimate of the accuracy of a set of loci to correctly match individuals. In this test, the 18 loci were ranked according to $H_e$ and the minimum number of loci required to correctly assign all recaptured individuals was tested. Starting with a single locus (the one with the highest $H_e$), the number of loci was gradually increased by adding at each step the loci with the next highest $H_e$. The $P_I$ was calculated for all set of loci. Two measurements of $P_I$ values were calculated, $P_{IHW}$ and $P_{Isib}$ (according to Woods *et al*., 1999).

© 2011 The Authors

$P_{\text{Isib}}$ allows for related individuals to be present in the sample and is a more conservative measure of $P_{\text{I}}$ compared to $P_{\text{IHW}}$. All calculations were performed using GENECAP.

## RESULTS

A total of 55 individuals were individually PIT tagged, released and used for genetic analyses. Thirty individuals were recaptured of which nine were captured twice, leading to a total number of 48 true genetic reassignments (*i.e.* $30 + 9 \times 2$).

Individual genetic tagging was successful at identifying recaptured PIT-tagged individuals. The data were checked for PCR errors and these were absent. The experience of the observer performing the allele scoring influenced the number of marked individuals matched correctly and resulted in 79–94% of marked individuals successfully matched (Table II). When allowing for matches that differ by one and two alleles, however, 100% of individuals were successfully matched irrespective of the observer (Table II). In addition, all individuals matched using genotypes were also matched using PIT tags (*i.e.* there were no type I errors), supporting the robustness of this method (Table II).

The re-sampling procedure indicated that all individuals would be correctly assigned with 0% error using a minimum of four to six loci (Fig. 1), greatly reducing the time and cost of this procedure. It also further supported the importance of considering expected heterozygosity when selecting candidate microsatellite loci. Specifically, the chance of having a set of loci that failed to assign all recaptured individuals was greatly reduced with increasing expected heterozygosity (Fig. 2).

Accordingly, no matching errors occurred when the six loci with the highest $H_{\text{e}}$ were selected (Table III). Using this set of loci $P_{\text{Isib}}$ was found to be 0·002, indicating that there were fewer than two individuals in a total of 1000 which shared the same genotypes in this population. This probability was reduced to $10^{-4}$ for a set of nine loci and to $10^{-5}$ for 12 loci (Table III).

Table II. Comparison between observed genotype matches and actual passive integrated transponder (PIT) tag matches with *Leuciscus leuciscus*. Observer's scoring experience indicates whether the observer was trained to score microsatellites or not. Correct matches is the proportion (percentage) of PIT tag matches that were correctly identified based on genotypes (the inverse of this proportion measures type II errors as defined in the main text). Matches with one and two errors are the proportion of PIT tag matches that were correctly identified based on genotypes but that differ by one and two alleles, respectively (*i.e.* incomplete matching). Incorrect matches are the proportion of matches identified based on genotypes but were not observed using PIT tags (*i.e.* type I error according to definition in the main text)

| Observer's scoring experience | Observer's identity | Correct matches | Matches with one error | Matches with two errors | Incorrect matches (type I error) |
|---|---|---|---|---|---|
| Experienced | Observer 1 | 45/48 (94%) | 2/48 | 1/48 | 0/48 |
| | Observer 2 | 43/48 (90%) | 4/48 | 1/48 | 0/48 |
| Novice | Observer 3 | 39/48 (81%) | 5/48 | 4/48 | 0/48 |
| | Observer 4 | 38/48 (79%) | 6/48 | 4/48 | 0/48 |

© 2011 The Authors

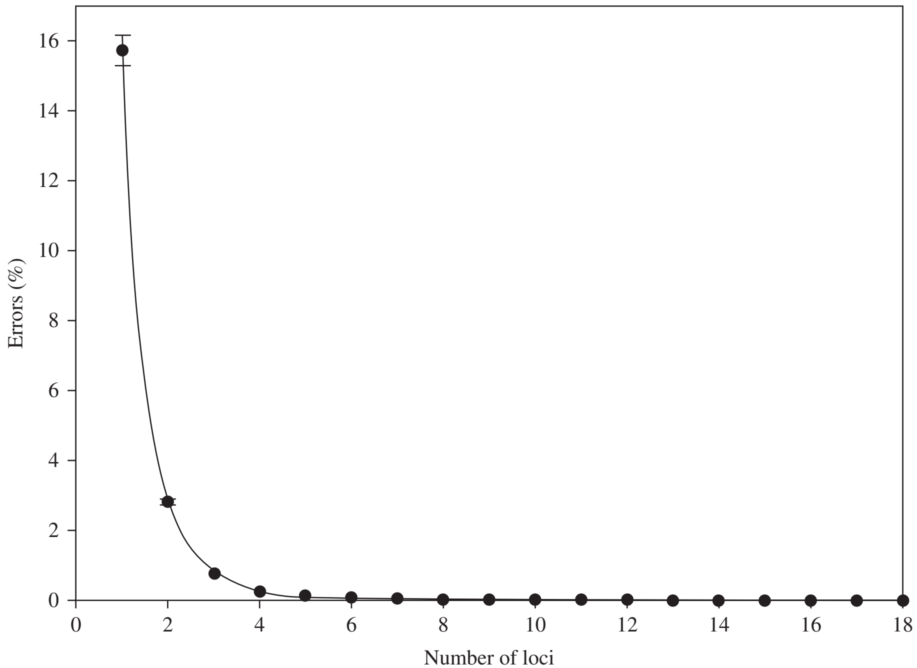


Fig. 1. Results of a re-sampling procedure to establish the relationship between per cent errors of individual assignments and the number of loci chosen for genotyping individual *Leuciscus leuciscus*. Percentage of errors over all possible individual matches was calculated as the percentage of observed matches at each re-sampling iteration less the percentage of expected matches.

## DISCUSSION

Since the establishment of the field of molecular ecology in the 1980s and 1990s, the use of molecular tools to address ecological questions has increased. In the present study, the use of molecular tools for individual assignment in a teleost is described and validated. Individual genetic tagging was successfully used to identify recaptured *L. leuciscus*. Whilst the observers' experience influenced the accuracy of individual genetic profiles and led to <100% of the individuals being correctly reassigned, 100% reassignment was successful when allowing for scored genotypes that differed by one or two alleles. Such discrepancies can be easily avoided through a careful screen of the allele peak profiles at each locus and the establishment of clear scoring rules and criteria. Alleles at specific loci have characteristic patterns which are mainly influenced by the DNA repeat unit. Loci with messy allele patterns, *i.e.* patterns in which the true allele cannot be easily identified and can thus be misidentified by inexperienced observers, should be removed during the loci selection process. This would increase the percentage of individuals reassigned correctly and hence avoid not detecting a recapture.

Prior to commencing a study using individual identification with DNA markers, the significance level of $P_I$, *i.e.* the number of correctly reassigned individuals (Waits *et al*., 2001), must be estimated. This can be achieved by using reported mean expected heterozygosity values for the loci used. The re-sampling analysis performed

© 2011 The Authors

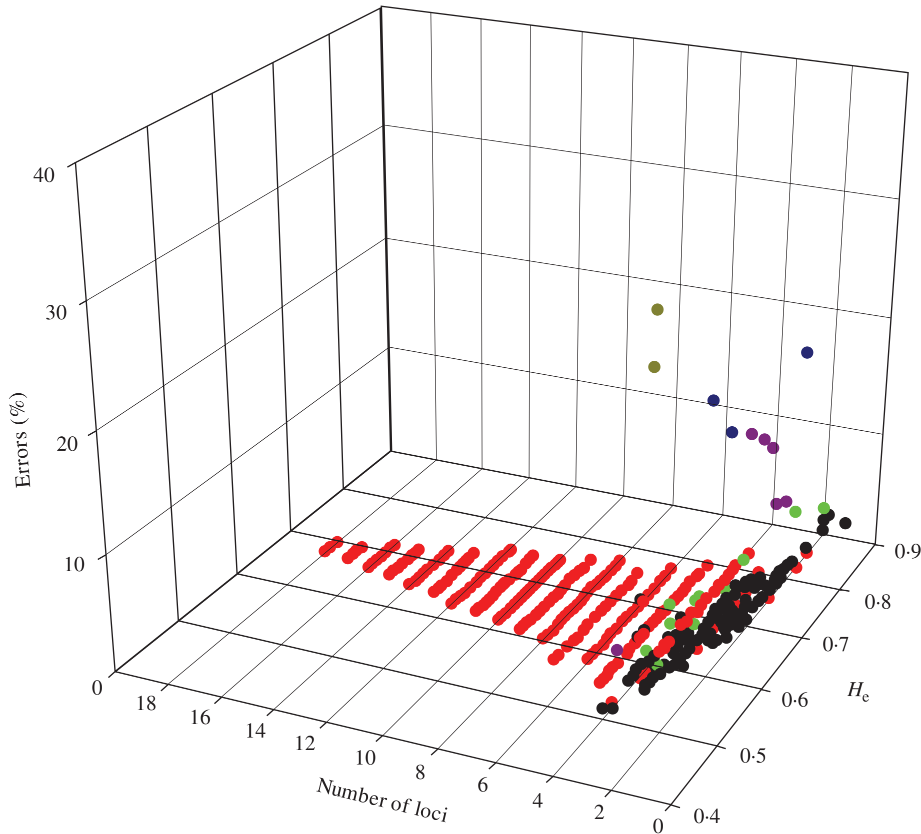


FIG. 2. Results of a re-sampling procedure aiming at visualizing the relationship between the percentages of errors of individual assignments, the numbers of loci chosen for genotyping individuals and the mean expected heterozygosity ($H_e$) at each locus set in samples of *Leuciscus leuciscus*. Percentage of errors over all possible individual matches was calculated as the percentage of observed matches at each re-sampling iteration less the percentage of expected matches. Zero per cent (●) indicates iterations where the combination between the number of loci and their mean heterozygosity provides no assignments errors. (●, 0%; ●, ≤5%; ●, ≤10%; ●, ≤20%; ●, ≤30%; ●, ≤40%).

in this study indicated that a much lower number of loci (compared to the 18 initially used here) could allow a correct identification of individuals. Such analysis should be utilized during the development of a genetic tagging protocol as it can greatly decrease the costs of the procedure. In the case of the *L. leuciscus* population studied here, a minimum of four to six loci provided sufficient discriminatory power to identify all individuals with 0% error. This analysis, however, used the corrected genotypes and the percentage error would be higher if genotype errors were allowed.

Mean allelic richness and $H_e$ can greatly influence the $P_I$ and Waits *et al.* (2001) have illustrated that population structure and mean $H_e$ can lead to lower observed $P_I$ values compared to the theoretically determined ones. As mean $H_e$ values can vary among populations and species, it is important to determine mean allelic richness and $H_e$ at candidate loci *a priori* and incorporate such information when developing individual genetic tagging. Waits *et al.* (2001) proposed to use a more conservative

© 2011 The Authors

TABLE III. The probability of identity ($P_I$) was calculated for different combinations of the 18 loci in *Leuciscus leuciscus*. Loci were combined in order of their expected heterozygosity ($H_e$) values, *i.e.* every additional locus added had the next highest $H_e$ (see Table I for $H_e$ values). Two $P_I$ values were calculated $P_{Isib}$ and $P_{IHW}$ (Woods *et al.*, 1999) using GENECAP. Number of false identification is the number of individuals identified as recaptures using the different microsatellite combinations minus the known recaptured individuals (in this case 30)

| Number of loci | $P_{Isib}$ | $P_{IHW}$ | Number of false identification |
|---|---|---|---|
| 1 | 0·2984 | 0·0148 | >100 |
| 2 | 0·1117 | 0·0013 | >10 |
| 3 | 0·0397 | 7·5E-05 | 2 |
| 4 | 0·0141 | 4·1E-06 | 2 |
| 5 | 0·0049 | 2E-07 | 1 |
| 6 | 0·0020 | 2·4E-08 | 0 |
| 7 | 0·0008 | 2·4E-09 | 0 |
| 8 | 0·0003 | 2·3E-10 | 0 |
| 9 | 0·0001 | 2·8E-11 | 0 |
| 10 | 6·2E-05 | 3·3E-12 | 0 |
| 11 | 2·9E-05 | 5·6E-13 | 0 |
| 12 | 1·3E-05 | 1E-13 | 0 |
| 13 | 6·6E-06 | 2E-14 | 0 |
| 14 | 3·2E-06 | 4·1E-15 | 0 |
| 15 | 1·6E-06 | 7·5E-16 | 0 |
| 16 | 8·8E-07 | 2E-16 | 0 |
| 17 | 5·7E-07 | 7·9E-17 | 0 |
| 18 | 3·7E-07 | 3E-17 | 0 |

$P_I$ calculation ($P_{Isib}$) which allows for siblings to be present within the sample. Using $H_e$ values, $P_{Isib}$ values for different loci combinations were calculated and when all 18 loci (mean observed $H_e$ 0·649) are used the $P_{Isib}$ for the data was $3·7 \times 10^{-7}$ whereas a combination of six loci (mean observed $H_e$ 0·829) was 0·002. Thus, with 18 loci the probability of misidentifying an individual as a recapture is $<10^{-6}$ sampled fish.

Genotyping errors can confound the results of molecular work and especially genetic fingerprinting (Bonin *et al.*, 2004). In addition to carefully selecting marker profiles to avoid mis-scoring, the appropriate positive and negative controls must be included in each PCR reaction. Blind samples represent the ideal controls as they include tissue samples that are taken through the entire genotyping procedure, *i.e.* from DNA extraction to genotyping scoring (Bonin *et al.*, 2004). Positive samples can include previously typed samples that can be used as an amplification reference and control for PCR and scoring errors. Negative controls should be used at each step in order to control for cross sample and reagent contamination. The increased use of automated procedures (from DNA extraction to scoring), however, decreases the rate of such potential errors. Additionally, freely available software such as Micro-Checker (von Oosterhout *et al.*, 2006) can be used to select loci minimizing scoring artefacts.

One of the disadvantages associated with genetic tagging is its general higher cost in comparison to physical tags. Such costs include those associated with developing

© 2011 The Authors

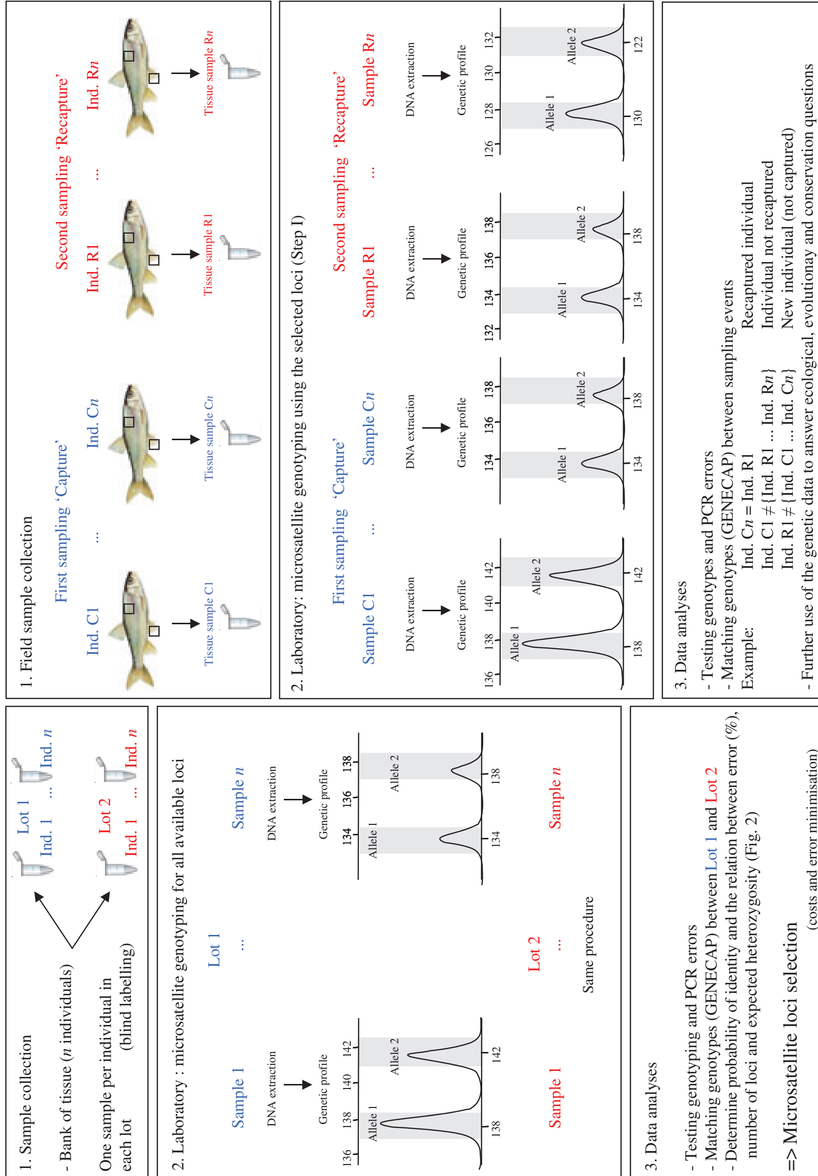


FIG. 3. Legend on next page.

the method for a new species and the cost of genotyping every recaptured individual. These costs can be substantial in large populations with low recapture rates as is the case of marine fish populations. In addition to the higher financial costs, information from physical tags is less time-consuming to obtain and does not require having any additional molecular expertize. Thus, depending on the questions asked, the costs and benefits associated to each method should be assessed to determine which one is the most appropriate for the study.

Although more planning and initial work is required when developing individual genetic tagging protocols compared to physical tagging, the benefits of genetic tagging should outweigh these costs. For example, genetic tags do not suffer from tag loss and their use is not limited by organism size. This allows for the study of individual life-history traits and population dynamics from the earliest life stages, as well as for small-bodied species. In studies where individual physical tagging is required, small-bodied juveniles could also be included by initially using individual genetic tagging until they reach the minimal size at tagging. The individual genetic data obtained can also be used at the population level to address additional ecological and evolutionary questions such as population admixing (gene flow) and effective population size (Castric & Bernatchez, 2004; Waples & Do, 2010). Long-term studies utilizing physical tags can suffer from tag loss as such tags can be lost over time (Oosthuizen *et al*., 2010). Genetic tags, however, can be used to monitor long-lived species over long periods of time with no tag loss, increasing the understanding of their ecology whilst ensuring sufficient sample sizes.

Furthermore, the easy storage and long-term life of isolated DNA at $-70^{\circ}$ C should attract the establishment of sample collections. In addition, the success of DNA isolation from historic samples, *e.g.* fish scales (Yue & Orban, 2001) and formalin-fixed samples (Zardus *et al*., 2006), can be used to increase the available information on species and populations of special interest. The wealth of historic fish scale archives for numerous teleosts provides ample opportunity for incorporating such data into current studies. The importance of molecular data obtained in ecological and evolutionary studies is further supported by the introduction of a new data archiving policy by the leading journals in ecology and evolution (Whitlock *et al*.,

FIG. 3. Guideline for developing individual genetic tagging for a new species. The approach consists of two steps: (a) optimizing the costs by selecting the number of microsatellite loci based on the expected heterozygosity, allelic richness, the estimated probability of identity and the per cent error; (b) using the selected loci to reassign individuals from different sampling events. (a) Step I: Using an existing bank of tissue (*e.g.* preserved specimens, scales and fin-clips), each sample is split into two sub-samples assigned to two lots. Lot 1 will correspond to the capture samples and lot 2 the recapture samples that are blind labeled. DNA is then extracted from each sample and the genetic profile is established for all available loci for the species. Genotypes are matched between lots using GENECAP (or similar software) and the relation between the per cent error, loci number, expected heterozygosity and probability of identity is established to select the microsatellite loci. (b) Step II: Samples are collected in the field (non-lethal sampling) at different sampling events (initial capture and subsequent recaptures). DNA is then extracted from each sample and the genetic profile is established for the selected loci (step I). Genotypes are matched between lots using GENECAP (or similar software) and individuals are attributed to three possible cases: recaptured individuals, individual not recaptured and new individual that was not initially captured. After individual reassignment, genetic data can be used to answer additional questions. Further developments can notably include the non-invasive collection of recapture samples. C, capture sample; Ind., individual; R, recapture sample.

© 2011 The Authors

2010). Molecular data including microsatellite allele frequency data will be archived in respective public databases. This will in turn ensure the availability of such data increasing their use in meta-analysis whilst facilitating the design of new studies. It is thus proposed that where applicable, individual genetic tagging should be employed. To support this, a clear guideline is provided (Fig. 3) which can be used to develop individual genetic tagging for teleosts and other organisms.

Over the last decade, individual genetic tagging has greatly benefited from non-invasive sampling techniques (Taberlet & Luikart, 1999; Lukacs & Burnham, 2005). In non-invasive genetic sampling, samples are collected without catching, disturbing or even observing the animal as DNA is extracted from faeces, hair, feathers and other tissue (Taberlet *et al.*, 1999; Lukacs & Burnham, 2005; Broquet *et al.*, 2007). Although this tissue might contain a low quantity of DNA, its use can be optimized by carefully selecting the genetic markers (Broquet *et al.*, 2007). Non-invasive sampling techniques are not commonly used by fish (and aquatic) ecologists (Beja-Pereira *et al.*, 2009) that usually rely on destructive and non-destructive approaches instead (Taberlet & Luikart, 1999).

By demonstrating that individual genetic tagging can be successfully used with teleosts, it is hoped that this study will provide incentives for the development of non-invasive genetic sampling as a way to identify and monitor individuals. Although non-invasive samples might be more difficult to obtain in aquatic than in terrestrial environments (*e.g.* transport by the flow and environmental degradation), Ficetola *et al.* (2008) recently demonstrated that DNA is preserved and collected in fresh water for subsequent analyses. Consequently, the development of new technologies and tools allowing a hands-off collection of DNA in the aquatic environment is suggested. Indeed, in the case of fishes, mucus (Livia *et al.*, 2006) and scales (Yue & Orban, 2001) have already been reported to be suitable for DNA analyses, and, along with other sources of DNA (*e.g.* faeces and urine), could be used in individual genetic tagging. This approach could be used in locations where fishes and aquatic organisms are present in high density and in a restricted space (*e.g.* nursery habitats, spawning grounds and fish passes). This approach might provide a new insight into the understanding of the ecology, evolution and conservation of many rare, endangered and small-bodied aquatic species.

We are grateful to the numerous persons who participated in field and laboratory work and to three anonymous reviewers who provided insightful comments. The research leading to these results has received funding from the European Community's Seventh Framework Programme FP7-PEOPLE-2007-2-1-IEF under grant agreement number PIEF-GA-2008-219558 (J.C.). This research was also funded by the Agence de l'Eau Adour-Garonne (France) and a 'Small Research Grant' by the Fisheries Society of the British Isles.

## References

Acolas, M. L., Roussel, J. M., Lebel, J. M. & Bagliniere, J. L. (2007). Laboratory experiment on survival, growth and tag retention following PIT injection into the body cavity of juvenile brown trout (*Salmo trutta*). *Fisheries Research* **86,** 280–284.

Aljanabi, S. M. & Martinez, I. (1997). Universal and rapid salt-extraction of high quality genomic DNA for PCR-based techniques. *Nucleic Acids Research* **25,** 4692–4693.

Baerwald, M. R. & May, B. (2004). Characterization of microsatellite loci for five members of the minnow family Cyprinidae found in the Sacramento–San Joaquin Delta and its tributaries. *Molecular Ecology Notes* **4,** 385–390.

© 2011 The Authors

Barinova, A., Yadrenkina, E., Nakajima, M. & Taniguchi, N. (2004). Identification and characterization of microsatellite DNA markers developed in ide *Leuciscus idus* and Siberian roach *Rutilus rutilus*. *Molecular Ecology Notes* **4,** 86–88.

Beja-Pereira, A., Oliveira, R., Alves, P. C., Schwartz, M. K. & Luikart, G. (2009). Advancing ecological understandings through technological transformations in noninvasive genetics. *Molecular Ecology Notes* **9,** 1279–1301.

Blanchet, S., Rey, O., Berthier, P., Lek, S. & Loot, G. (2009). Evidence of parasite-mediated disruptive selection on genetic diversity in a wild fish population. *Molecular Ecology* **18,** 1112–1123.

Bolland, J. D., Cowx, I. G. & Lucas, M. C. (2009). Evaluation of the VIE and PIT tagging methods for juvenile cyprinid fishes. *Journal of Applied Ichthyology* **25,** 381–386.

Bonin, A., Bellemain, E., Bronken Eidesen, P., Pompanon, F., Brochmann, C. & Taberlet, P. (2004). How to track and assess genotyping errors in population genetics studies. *Molecular Ecology* **13,** 3261–3273.

Broquet, T., Menard, N. & Petit, E. (2007). Noninvasive population genetics: a review of sample source, diet, fragment length and microsatellite motif effects on amplification success and genotyping error rates. *Conservation Genetics* **8,** 249–260.

Buckley, P. A., Francis, C. M., Blancher, P., DeSante, D. F., Robbins, C. S., Smith, G. & Cannell, P. (1998). The North American bird banding program: into the 21$^{st}$ century. *Journal of Field Ornithology* **69,** 511–529.

Campbell, T. S., Irvin, P., Campbell, K. R., Hoffmann, K., Dykes, M. E., Harding, A. J. & Johnson, S. A. (2009). Evaluation of a new technique for marking anurans. *Applied Herpetology* **6,** 247–256.

Castric, V. & Bernatchez, L. (2004). Individual assignment test reveals differential restriction to dispersal between two salmonids despite no increase of genetic differences with distance. *Molecular Ecology* **13,** 1299–1312.

Crooijmans, R. P. M. A., Bierbooms, V. A. F., Komen, J., VanderPoel, J. J. & Groenen, M. A. M. (1997). Microsatellite markers in common carp (*Cyprinus carpio* L). *Animal Genetics* **28,** 129–134.

Cucherousset, J., Roussel, J.-M., Keeler, R., Cunjak, R. A. & Stump, R. (2005). The use of two new portable 12-mm PIT Tag detectors to track small fish in shallow streams. *North American Journal of Fisheries Management* **25,** 270–274.

Cucherousset, J., Marty, P., Pelozuelo, L. & Roussel, J. M. (2008). Portable PIT detector as a new tool for non-disruptively locating tagged amphibians in the field: a case study with Pyrenean brook salamanders (*Calotriton asper*). *Wildlife Research* **35,** 780–787.

Cucherousset, J., Pallisson, J. M., Cuzol, A. & Roussel, J. M. (2009). Spatial behaviour of young-of-the-year northern pike (*Esox lucius* L.) in a temporarily flooded nursery area. *Ecology of Freshwater Fish* **18,** 314–322.

Cucherousset, J., Britton, J. R., Beaumont, W. R. C., Nyqvist, M., Sievers, K. & Gozlan, R. E. (2010). Determining the effects of species, environmental conditions and tracking method on the detection efficiency of portable PIT telemetry. *Journal of Fish Biology* **76,** 1039–1045.

DeHaan, P. W., Jordan, G. R. & Arden, W. R. (2008). Use of genetic tags to identify captive-bred pallid sturgeon (*Scaphirhynchus albus*) in the wild: improving abundance estimates for an endangered species. *Conservation Genetics* **9,** 691–697.

DeYoung, R. W. & Honeycutt, R. L. (2005). The molecular toolbox: genetic techniques in wildlife ecology and management. *Journal of Wildlife Management* **69,** 1362–1384.

Dimsoski, P., Toth, G. P. & Bagley, M. J. (2000). Microsatellite characterization in central stoneroller *Campostoma anomalum* (Pisces, Cyprinidae). *Molecular Ecology* **9,** 2187–2189.

Dubut, V., Martin, J. F., Gilles, A., Van Houdt, J., Chappaz, R. & Costedoat, C. (2009). Isolation and characterization of polymorphic microsatellite loci for the dace complex: *Leuciscus leuciscus* (Teleostei: Cyprinidae). *Molecular Ecology Resources* **9,** 1179–1183.

Feldheim, K. A., Gruber, S. H., de Marignac, J. R. C. & Ashley, M. V. (2002). Genetic tagging to determine passive integrated transporter tag loss in lemon sharks. *Journal of Fish Biology* **61,** 1309–1313.

© 2011 The Authors

Ficetola, G. F., Miaud, C., Pompanon, F. & Taberlet, P. (2008). Species detection using environmental DNA from water samples. *Biology Letters* **4,** 423–425.
Frasier, T. R., Hamilton, P. K., Brown, M. W., Kraus, S. D. & White, B. N. (2009). Sources and rates of errors in methods of individual identification for the North Atlantic right whales. *Journal of Mammalogy* **90,** 1246–1255.
Gibbons, J. W. & Andrews, K. M. (2004). PIT tagging: simple technology at its best. *Bioscience* **54,** 447–454.
Girard, P. & Angers, B. (2006). Characterization of microsatellite loci in longnose dace (*Rhinichthys cataractae*) and interspecific amplification in five other Leuciscinae species. *Molecular Ecology Notes* **6,** 69–71.
Grant, E. H. C. (2008). Visual implant elastomer mark retention through metamorphosis in amphibian larvae. *Journal of Wildlife Management* **72,** 1247–1252.
Gross, M. L. & Kapuscinski, A. R. (1997). Reproductive success of smallmouth bass estimated and evaluated from family-specific DNA fingerprints. *Ecology* **78,** 1424–1430.
Hammer, S. A. & Blackenship, H. L. (2001). Cost comparison of marks, tags, and mark-with-tag combinations used in salmonid research. *Aquaculture* **63,** 171–178.
Larno, V., Launey, S., Devaux, A. & Laroche, J. (2005). Isolation and characterization of microsatellite loci from chub *Leuciscus cephalus* (Pisces:Cyprinidae). *Molecular Ecology Notes* **5,** 752–754.
Livia, L., Antonella, P., Hovirag, L., Mauro, N. & Panara, F. (2006). A non-destructive, rapid, reliable and inexpensive method to sample, store and extract high-quality DNA from fish mucus and buccal cells. *Molecular Ecology Notes* **6,** 257–260.
Lukacs, P. M. & Burnham, K. P. (2005). Review of capture–recapture methods applicable to noninvasive genetic sampling. *Molecular Ecology* **14,** 3909–3919.
von Oosterhout, C., Weetman, D. & Hutchinson, W. F. (2006). Estimation and adjustment of microsatellite null alleles in nonequilibrium populations. *Molecular Ecology Notes* **6,** 255–256.
Oosthuizen, W. C., de Bruyn, P. J. N., Bester, M. N. & Girondot, M. (2010). Cohort and tag-site-specific tag-loss rates in mark–recapture studies: a southern elephant seal cautionary cause. *Marine Mammal Science* **26,** 350–369.
Palsbøll, P. J. (1999). Genetic tagging: contemporary molecular ecology. *Biological Journal of the Linnean Society* **68,** 3–22.
Palsbøll, P. J., Allen, J., Berube, M., Clapham, P. J., Feddersen, T. P., Hammond, P. S., Hudson, R. R., Jorgensen, H., Katona, S., Larsen, F., Lien, J., Mattila, D. K., Sigurjohnsson, J., Sears, R., Smith, T., Sponer, R., Stevick, P. & Oien, N. (1997). Genetic tagging of humpback whales. *Nature* **388,** 767–769.
Perez-Enriquez, R. & Taniguchi, N. (1999). Use of microsatellite DNA as genetic tags for the assessment of a stock enhancement program of red sea bream. *Fisheries Science* **65,** 374–379.
Ravi, V. & Venkatesh, B. (2008). Rapidly evolving fish genomes and teleost diversity. *Current Opinion in Genetics and Development* **18,** 544–550.
Shimoda, N., Knapik, E. W., Ziniti, J., Sim, C., Yamada, E., Kaplan, S., Jackson, D., de Sauvage, F., Jacob, H. & Fishman, M. C. (1999). Zebrafish genetic map with 2000 microsatellite markers. *Genomics* **58,** 219–232.
Taberlet, P. & Luikart, G. (1999). Non-invasive genetic sampling and individual identification. *Biological Journal of the Linnean Society* **68,** 41–55.
Taberlet, P., Waits, L. P. & Luikart, G. (1999). Noninvasive genetic sampling: look before you leap. *Trends in Ecology and Evolution* **14,** 323–327.
Turner, T. F., Dowling, T. E., Broughton, R. E. & Gold, J. R. (2004). Variable microsatellite markers amplify across divergent lineages of cyprinid fishes (subfamily Leusicinae). *Conservation Genetics* **5,** 279–281.
Vyskocilova, M., Simkova, A. & Martin, J. F. (2007). Isolation and characterization of microsatellites in *Leusiscus cephalus* (Cypriniformes, Cyprinidae) and cross-species amplification within the family Cyprinidae. *Molecular Ecology Notes* **7,** 1150–1154.
Waples, R. S. & Do, C. (2010). Linkage disequilibrium estimates of contemporary $N_e$ using highly variable genetic markers: a largely untapped resource for applied conservation and evolution. *Evolutionary Applications* **3,** 244–262.

© 2011 The Authors
Journal of Fish Biology © 2011 The Fisheries Society of the British Isles, 

Waits, L. P., Luikart, G. & Taberlet, P. (2001). Estimating the probability of identity among genotypes in natural populations: cautions and guidelines. *Molecular Ecology* **10,** 249–256.

Whitlock, M. C., McPeek, M. A., Rauscher, M. D., Rieseberg, L. & Moore, A. J. (2010). Data archiving. *American Naturalist* **175,** 145–146.

Wilberg, M. J. & Dreher, B. P. (2004). GENECAP: a program for analysis of multilocus genotype data for non-invasive sampling and capture–recapture population estimation. *Molecular Ecology Notes* **4,** 783–785.

Woods, J. G., Paetkau, D., Lewis, D., McLellan, B. N., Proctor, M. & Strobeck, C. (1999). Genetic tagging of free-ranging black and brown bears. *Wildlife Society Bulletin* **27,** 616–627.

Yue, G. H. & Orban, L. (2001). Rapid isolation of DNA from fresh and preserved fish scales for polymerase chain reaction. *Marine Biotechnology* **3,** 199–2004.

Zardus, J. D., Etter, R. J., Chase, M. R., Rex, M. A. & Boyle, E. E. (2006). Bathymetric and geographic population structure in the pan-Atlantic deep-sea bivalve *Deminucula atecellana* (Schenck, 1939). *Molecular Ecology* **15,** 639–651.

© 2011 The Authors