Phylogeographic structure of a protogynous hermaphrodite species, the ballan wrasse *Labrus bergylta*, in Ireland, Scotland, and Norway, using mitochondrial DNA sequence data

J. D’Arcy†, L. Mirimin*†, and R. FitzGerald

Carna Research Station, Ryan Institute, National University of Ireland, Galway, Republic of Ireland

*Corresponding author: tel: +27 (0)21 808 5833; fax: +27 (0)21 808 5837; e-mail: lmirimin@gmail.com


Received 3 December 2012; accepted 24 January 2013; advance access publication 19 February 2013.

The ballan wrasse, *Labrus bergylta*, is a protogynous hermaphrodite marine fish species that inhabits coastal waters of the eastern North Atlantic. Sequential hermaphrodite species tend to be characterized by a skewed sex ratio, which is thought to lead to marked population structuring due to a reduced effective number of breeders. Furthermore, due to its large body size (compared with other wrasse species) and its peculiar feeding behaviour, this species has been identified as a candidate cleaner fish to be used in parasite control of farmed finfish. In the present study, we used mitochondrial DNA (control region) sequence data to investigate the genetic diversity and population structuring of ballan wrasse in waters around the British Isles and southern Norway. Ballan wrasse in southern Norway showed lower levels of genetic diversity than around the British Isles, which appear to be the result of historical demographic events (population bottleneck followed by expansion). Analysis of mismatch distributions and the presence of two highly divergent clades unevenly represented in Atlantic and Scandinavian regions suggest distinct recolonization patterns in these two regions. These results provide a first insight on the status of wild populations of ballan wrasse in the eastern North Atlantic, with implications for conservation and management.

**Keywords:** control region, parasite control, protogyny, wrasse.

**Introduction**

Population genetic structure of marine species with high dispersal potential can be shaped by many and complex factors (Palumbi, 1996). Current levels of population structuring and genetic diversity often carry signatures of past demographic events, which occur in response to both recent (Perry et al., 2005) and historical climatic fluctuations, such as the Ice Ages that dominated the Quaternary Period (Hewitt, 2004). In addition, wild populations have been subject to direct and indirect pressure deriving from human activities affecting marine ecosystems worldwide (Halpern et al., 2008).

Wrasse species (family Labridae) are a large assemblage of marine fish composed of ~504 species in 70 genera (Parenti and Randall, 2011), which are generally characterized by a sequential hermaphrodite developmental system, where gonads of the primary sex develop earlier in time than the secondary sex (Devlin and Nagahama, 2002; Frisch, 2004). Molecular approaches have improved our understanding of the evolution (Cowman et al., 2009), classification (Hanel et al., 2002; Chen et al., 2004), as well as population structure of wrasse species (Chen et al., 2004; Purcell et al., 2006; Froukh and Kochzius, 2007; Haney et al., 2007), where it appears that both life-history traits and environmental conditions may play a significant role in shaping population structuring (Chen et al., 2004; Froukh and Kochzius, 2007). Since the early 1970s, several wrasse species have been used as biological cleaner fish for the control of ectoparasites (Copepoda, Caligidae) in farmed Atlantic salmon (*Salmo salar*) in Ireland, UK, and Norway (Bjordal, 1990, 1991; Tully et al., 2006).
in the aquaculture sector, the harvest and relocation of a large number of fish from natal areas to farming sites poses some questions regarding the potential impact of such practice to wild populations (Sundt and Jorstad, 1998). In fact, up to 10 million wild individuals are stocked annually onto Norwegian salmon farms (Treasurer, 2012, and references therein). Although such practice has affected mainly smaller species, such as goldsinny (Ctenolabrus rupestris Linnaeus, 1758), there has been an increasing interest in using larger wrasse species, such as ballan wrasse (Labrus bergylta Ascanius 1767) to treat larger salmon (3–6 kg; Kvenseth et al., 1996), which has resulted in the initiation of dedicated breeding programmes (Ottesen et al., 2011).

The ballan wrasse is a protogynous (female primary sex) hermaphrodite species found in coastal waters along the Eastern Atlantic from Morocco to Norway, including Madeira, the Azores, and the Canary Islands, and it is the largest of the wrasse species in waters around the British Isles (Quignard and Pras, 1986; Elofsson et al., 1999). Although information about ballan wrasse life history has started to emerge in recent times (D’Arcy et al., 2012; Ottesen et al., 2012), the population structure and demographic history of this species is still unknown in the wild. Marine fish populations inhabiting the eastern North Atlantic have been affected by past climatic fluctuations, with the last glacial maxima (LGM) ending around 25–18 thousand years ago (Bernatchez and Wilson, 1998; Taberlet et al., 1998; Hewitt, 2004). The effect of such environmental changes can be reflected in current patterns of genetic diversity, where signatures of population expansion or reduction can be detected especially in coastal species characterized by temperate-cold distribution ranges (Ware and Cunningham, 2001; Wilson and Veraguth, 2010). Empirical results have shown a great variety of diversity patterns, indicating that such signatures of demographic changes are the result of complex environmental and biological factors (Francisco et al., 2009; Almada et al., 2012; Robalo et al., 2012).

The aim of the present study is to use mitochondrial DNA (mtDNA) control region data to investigate genetic diversity and population structure of ballan wrasse in three main regions (Ireland, Scotland, and Norway), where populations have been recently targeted as source for cleaner fish for the aquaculture industry.

### Material and methods

#### Sample collection and DNA extraction

In all, 279 wild ballan wrasse were sampled from western Ireland (two locations: MWEENISH and BERTRAGHBOY BAY), Northern Ireland (one location: PORTAFERRY), western Scotland (two locations: LOCHALINE and LOCH SUNART), and southern Norway (two locations: Hidra and Søgne) from March to September in 2010 and 2011 (Figure 1). All fish were captured using trammel-nets, pots, and/or fish traps. Fish were anaesthetized using tricaine methanesulphonate (MS-222) before the collection of small amounts of tissue (fin clips). DNA was isolated from each individual sample via the proteinase K digestion and Chelex extraction procedure (2 h at 56°C, total volume of 200 μl, with 0.1 mg proteinase K and 10 min at 99°C in 10% Chelex; modified from Walsh et al., 1991). Extracted DNA was stored at −20°C until further use.

Due to reliability in the amplifying mtDNA control region in various fish species, primer A (Lee et al., 1995) and primer TDKD (Kocher et al., 1993) were chosen to amplify and sequence a variable portion of the mtDNA control region of ballan wrasse. Following a number of optimization steps, polymerase chain reaction conditions were modified from original publications as follows: each amplification reaction was carried out in a total volume of 25 μl, containing 50–100 ng of genomic DNA, 1 x Green GoTaq buffer (Promega), 1.5 mM MgCl2, 1 μM of each primer, 0.25 mM of each dNTP, and 1 U of GoTaq DNA polymerase (Promega). Thermal cycling conditions included an initial denaturation step of 3 min at 95°C, followed by 40 cycles of 60 s at 95°C, 60 s at 52°C, 90 s at 72°C, followed by a final extension step of 10 min at 72°C. Each amplified DNA fragment was sequenced from both 5’- and 3’-ends, using the same forward (A) and reverse (TDKD) primers, by GATC-Biotech (Konstanz, Germany; www.gatc-biotech.com) using an ABI3730xl platform. For each individual, complementary sequence data were assembled into consensus sequences using CODONCODE ALIGNER (v. 3.7.1 CodonCode Corporation, Dedham, MA, USA), and subsequently, homologous sequences from all individuals were aligned using MEGA 5.0 (Tamura et al., 2011). A number of randomly selected individuals (3% of the total sample size) was used in duplicate runs, where DNA was re-extracted and target fragments resequenced ex novo, to ensure the reliability of the sequencing process.

The number and the composition of polymorphic sites, haplotype, and nucleotide diversity were calculated using ARLEQUIN 3.5 (Excoffier and Lischer, 2010). The same program was used to perform Tajima’s D (Tajima, 1989) and Fu’s Fs (Fu, 1997) neutrality tests and mismatch distribution analysis between pairs of haplotypes (Rogers and Harpending, 1992). The observed mismatch distribution was compared with the distribution expected in...
populations affected by sudden expansion (1000 replicates), under the assumption of selective neutrality, in which a unimodal distribution is expected (Rogers and Harpending, 1992). The sum of squared deviation (SSD) was used to detect departure between observed and expected distributions (Schneider and Excoffier, 1999). Past demographic growth patterns at the species level were also tested using the linear Bayesian skyline plot model (Drummond et al., 2005), as implemented in BEAST v.1.7.4 (Drummond and Rambaut, 2007).

Spatial genetic structure among and within sampled areas was carried out by an analysis of molecular variance (AMOVA) using ARLEQUIN, where conventional $F$-statistics were calculated using haplotype frequencies and significance tests were carried out after 1023 permutations. To test the hypothesis of the population subdivision, an AMOVA was performed among the seven sampling locations nested into four groups each representing a sampled region (i.e. western Ireland, Northern Ireland, western Scotland, and southern Norway). In addition, pairwise $F$-statistics ($F_{ST}$) were calculated between sampled areas based on haplotype frequencies (Excoffier and Lischer, 2010) and the Tamura and Nei estimator (Tamura and Nei, 1993). The significance of $F$-statistics was evaluated after 1023 permutations. A Tamura and Nei model was also used to estimate the mean and net divergence between groups of haplotypes using MEGA (Tamura et al., 2011).

An inference of phylogenetic relationships between haplotypes was produced using NETWORK 4.6 (available at www.fluxus-engineering.com). Maximum parsimony trees were reconstructed using the median-joining network algorithm (Bandelt et al., 1999), where weights of transversions and transitions were set to three and one, respectively, and the epsilon value was set to zero.

The times to the most recent common ancestor (tMRCA) of $L. bergylta$ and intraspecific mtDNA lineages were estimated by Bayesian inference, as implemented in the computer program BEAST v.1.7.4. The Bayesian distribution was generated by combining results from three independent runs, characterized by 30 million Markov chain Monte Carlo steps (10% burn-in; chosen to obtain an EES value equal or greater to 200 for each parameter estimated), using a strict clock and the HKY with Gamma site heterogeneity model of substitution; as previously identified as the most suitable model by applying the Akaike Information Criterion using jModelTest2 (Guindon and Gascuel, 2003; Darriba et al., 2012). As no clock calibration was available for $L. bergylta$ control region, we assumed a 5% per million year (My) divergence rate, as previously hypothesized as a potentially realistic rate for the same mtDNA region in another wrasse species, $Symphodus melops$ (Robalo et al., 2012). All BEAST results were displayed and inspected using Tracer v.1.4 (Rambaut and Drummond, 2007).

**Results**

Following the removal of non-overlapping 3- and 5-end sequence data, a 332-bp fragment of the mtDNA control region was obtained from a total of 279 individuals. Comparison of aligned sequences allowed the identification of 72 polymorphic sites, which included 67 transitions, 10 transversions, and 1 insertion–deletion (Supplementary Table S1). In all, 91 different haplotypes were identified among the sampled areas (GeneBank accession numbers: KC477846 –KC477936; Supplementary Table S2). A BLAST search among online DNA databases (e.g. GenBank) confirmed that the closest (95–100%) publicly available homologous sequence originated from $L. bergylta$ used in a phylogenetic study of wrasse species (Hanel et al., 2002). The most common haplotype was strongly represented in the two southern Norway sampling areas, but rare in all other areas. Conversely, western Ireland and western Scotland were represented by several moderately represented haplotypes, which in turn were absent or rare in the Norwegian sites. All sampled areas were characterized by a large proportion of unique haplotypes, which accounted for a total of 64 of 91 (70.3%; Supplementary Table S2). The overall haplotype diversity was $0.914 \pm 0.012$ and the overall nucleotide diversity was $0.032 \pm 0.016$. Haplotype diversity within sampling areas ranged between 0.909 and 0.989 in western Ireland, Northern Ireland, and western Scotland regions, while it was lower in the two southern Norway sites (0.481 and 0.605). Similarly, nucleotide diversity

| Table 1. Descriptive statistics and diversity indices of 332 bp of the mitochondrial control region of ballan wrasse from seven locations. |
|-----------------|-----------------|-----------------|-----------------|-----------------|
|                 | Western Ireland | Northern Ireland | Western Scotland | Southern Norway |
|                 |                 |                 |                 |                 |
|                 | Mweenish | Bertraghboy | Portaferry | Lochaline | Loch Sunart | Segne | Hidra |
| Longitude       | 09° 49'9"W | 09° 51'54"W | 05° 33'00"W | 05° 46'58"W | 05° 56'02"W | 07° 48'53"E | 06° 31'55"E |
| Latitude        | 53° 17'18"N | 53° 19'00"N | 54° 23'00"N | 56° 31'30"N | 56° 40'22"N | 58° 04'01"N | 58° 13'20"N |
| n               | 72       | 44       | 14       | 26       | 24       | 50       | 49       |
| $N_{45}$        | 39       | 32       | 13       | 18       | 15       | 10       | 12       |
| clade I vs. clade II frequencies |
| Polymorphic sites | 52      | 50      | 31      | 39      | 24      | 28      |
| $h$             | 0.064 (+0.010) | 0.070 (+0.016) | 0.098 (+0.031) | 0.942 (+0.034) | 0.909 (+0.048) | 0.481 (+0.087) | 0.605 (+0.075) |
| $\pi$           | 0.026 (+0.013) | 0.029 (+0.015) | 0.027 (+0.015) | 0.025 (+0.013) | 0.028 (+0.015) | 0.008 (+0.005) | 0.029 (+0.011) |
| $D$             | -0.681 | -0.555 | -0.242 | -0.708 | -0.384 | -1.704* | 0.182 |
| $F_{st}$        | -16.145*** | -14.307*** | -4.667* | -14.553 | -1.683 | -0.947 | 1.798 |
| SSD             | 0.167    | 0.017   | 0.038   | 0.037   | 0.337   | 0.016   | 0.452   |
| Model (SSD)     | 0.001    | 0.484   | 0.240   | 0.388   | 0.397   | 0.061   | 0.000   |

$n$, number of individuals; $N_{45}$, number of haplotypes; grey, clade I; black, clade II; $h$, haplotype diversity; $\pi$, nucleotide diversity; $D$, Tajima’s $D$; $F_{st}$, Fu’s $F_{st}$; SSD, sum of squared deviation. Fu’s $F_{st}$ minimum significance level was corrected to 0.02.

* $p < 0.01$.

** *** $p < 0.001$. 

---

**n** Downloaded from https://academic.oup.com/icesjms/article-abstract/70/3/685/918519 by guest on 11 December 2018
within sites was higher in Irish and Scottish regions (0.025–0.029) than in the two Norwegian sites (0.008 and 0.020; Table 1).

Analysis of the distribution of pairwise haplotype differences (mismatch analysis) and tests of selective neutrality (Fu’s $F_s$ and Tajima’s $D$) were initially carried out for each of the seven sampling locations. Most locations displayed a bi- or multimodal pattern (data not shown); nonetheless, the observed mismatch distribution was not significantly different (SSD, $p > 0.05$) from a distribution expected under expansion, except Mweenish and Hidra (Table 1). Departure from selective neutrality (i.e. negative and significant values) was observed in the two western Ireland sites, the Northern Ireland site (Fu’s $Fs$) and the southern Norway site Søgne (Tajima’s $D$), whereas no deviation from selective neutrality was observed in the two western Scotland sites and in the southern Norway site Hidra (Table 1). It is worth noting that Fu’s $F_s$ is regarded as a more sensitive indicator of population expansion than Tajima’s $D$ (Fu, 1997).

Due to the presence of two highly divergent clades (which may affect mismatch distribution and interpretation) and following the identification of two main population clusters (i.e. British Isles and southern Norway, see below), mismatch analysis was also carried out for each population cluster for each clade separately. This analysis revealed evidence of sudden population expansion and a unimodal mismatch distribution always, except clade II in southern Norway, which, however, consisted only of two haplotypes (Figure 2). Mismatch distributions observed in the Scandinavian population cluster tended to be shallower than that in British Isles, reflecting the lower diversity and closer relationship among haplotypes found in southern Norway.

A network of haplotypes was drawn to infer phylogenetic relationships among mtDNA control region sequence data, as well as to deduce possible demographic events (Figure 3). The network depicted a complex branching pattern that was dominated by two highly divergent clusters (clades) of haplotypes. The mean and the net genetic divergence (Tamura and Nei, 1993) between these two clades were 5.5% (s.e. = 0.011) and 3.8% (s.e. = 0.010), respectively. Each clade included representatives from all seven sampling locations, suggesting that these two main lineages were already present before wrasse colonized the studied areas. The proportion of haplotypes from each clade in each sampling location is shown in Table 1. Clade I was characterized by many unique or low-frequency haplotypes and one very common haplotype, which was very frequent in the southern Norway region but not in the Irish and Scottish regions. Conversely, clade II presented several moderately frequent haplotypes, which were mostly present in the Irish and Scottish regions, but rare in the southern Norway region (Figure 3). Furthermore, the most common haplotype in southern Norway was characterized by a star-shaped structure, where many low-frequency haplotypes differ by a single mutational step from a main common haplotype. Similarly, star-like

![Figure 2. Mismatch distribution for clades I and II in the British Isles and southern Norway.](https://academic.oup.com/icesjms/article-abstract/70/3/685/918519)

![Figure 3. Network of haplotypes. Node size reflects the frequency of its occurrence (yellow, Mweenish; green, Bertraghboy Bay; white, Portaferry; black, Lochaline; blue, Loch Sunart; dashed line, Søgne; crossed line, Hidra). Dashes on lines indicate the number (two or more) of mutational steps between nodes.](https://academic.oup.com/icesjms/article-abstract/70/3/685/918519)
patterns were evident in some of the most common haplotypes in the Irish and Scottish sampled areas. Such patterns are often associated with populations that originated from a small number of individuals leading to reduced genetic diversity (founder effect), in which enough time has elapsed for single-step mutations to arise from common (central) haplotypes.

At the species level, assuming a divergence rate of 5% per My, the tMRCA for L. bergylta mtDNA control region data was estimated at 693 ky (95% highest posterior density: 405–1015 ky). The same estimate was obtained for each sampling locality (data not shown), which is due to the presence of highly divergent haplotypes from each of the two divergent clades in each location (as shown in Table 1). When analysing each clade separately, tMRCA for clade I (more frequent in Norway) was 287 ky (165–430 ky) and for clade II (more frequent around the British Isles) was 693 ky (405–1015 ky). Although not significant, these results suggest that clade I could have a more recent origin than clade II. Assuming that the chosen divergence rate (5% per My) is realistic, the Bayesian skyline plot suggests that ballan wrasse (in the sampled areas) experienced a population expansion before the LGM (between 100 and 50 ky) and it has been stable in the last 50 ky (Figure 4).

A hierarchical AMOVA was carried out among sampled locations nested within the western Ireland, Northern Ireland, western Scotland, and southern Norway regions. The largest proportion of variance was observed within sampling locations (83.81%) with associated F-statistic \( \Phi_{ST} \) equal to 0.162 (\( p < 0.001 \)). A substantial proportion of variance was also attributed to differentiation among regions (16.31%) with associated F-statistic \( \Phi_{SC} \) equal to 0.163 (\( p < 0.01 \)). In contrast, no significant variation among sampling locations within regions was detectable (\( \Phi_{ST} \) = 0.00136, \( p = 0.535 \); Table 2). Similarly, pairwise \( F_{ST} \) estimates based on mtDNA control region haplotype frequencies and the Tamura and Nei (1993) estimator showed a stronger differentiation between regions than within regions (Table 3). More specifically, the two southern Norway sites showed strong segregation from all other sites, using both the haplotypic frequency \( F_{ST} \) as well as the Tamura and Nei (1993) estimates. The latter estimator also showed significant population structuring between the two southern Norway sites, Hidra and Søgne (\( F_{ST} = 0.072, p < 0.05 \)). Population structuring was moderate to weak between the Scottish and Irish regions, where the only significantly positive \( F_{ST} \) was detected between Loch Sunart and Mweenish (\( F_{ST} = 0.017, p < 0.05 \); Table 3).

**Discussion**

**Genetic diversity and demographic events**

Wild populations of ballan wrasse showed variable levels of genetic diversity among the studied areas. The high haplotypic diversity encountered in the Irish and Scottish sites is similar to the values reported in other wrasse species (Chen et al., 2004; Froukh and Kochzius, 2007; Haney et al., 2007). In contrast, the low levels of diversity found in Norway have not been previously reported in any wild wrasse population, except corkwing wrasse (\( S. melops; \) Robalo et al., 2012). Similar to the present study, Scandinavian corkwing wrasse populations showed lower

![Bayesian skyline plot](https://academic.oup.com/icesjms/article-abstract/70/3/685/918519)

**Figure 4.** Bayesian skyline plot.

### Table 2. Results from AMOVA.

<table>
<thead>
<tr>
<th>Source of variation</th>
<th>d.f.</th>
<th>Sum of squares</th>
<th>Variance components</th>
<th>% variation</th>
<th>( \Phi )-statistic</th>
<th>( p )-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Among groups</td>
<td>3</td>
<td>15.797</td>
<td>0.07872</td>
<td>16.31</td>
<td>( \Phi_{CT} = 0.16305 )</td>
<td>0.00684 ± 0.00271</td>
</tr>
<tr>
<td>Among populations within groups</td>
<td>3</td>
<td>1.143</td>
<td>-0.00055</td>
<td>-0.11</td>
<td>( \Phi_{SC} = -0.00136 )</td>
<td>0.53470 ± 0.01701</td>
</tr>
<tr>
<td>Within populations</td>
<td>272</td>
<td>110.053</td>
<td>0.40461</td>
<td>83.81</td>
<td>( \Phi_{ST} = 0.16191 )</td>
<td>0.00000 ± 0.00000</td>
</tr>
<tr>
<td>Total</td>
<td>278</td>
<td>126.993</td>
<td>0.48277</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Table 3. Population pairwise \( F_{ST} \).

<table>
<thead>
<tr>
<th></th>
<th>Mweenish</th>
<th>Bertraghboy Bay</th>
<th>Portaferry</th>
<th>Lochaline</th>
<th>Loch Sunart</th>
<th>Søgne</th>
<th>Hidra</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mweenish</td>
<td>-0.0006</td>
<td>-0.0006</td>
<td>-0.0087</td>
<td>0.01678*</td>
<td>-0.25765***</td>
<td>0.19792***</td>
<td></td>
</tr>
<tr>
<td>Bertraghboy Bay</td>
<td>-0.01106</td>
<td>-0.01216</td>
<td>-0.00817</td>
<td>-0.00619</td>
<td>-0.26716***</td>
<td>0.19804***</td>
<td></td>
</tr>
<tr>
<td>Portaferry</td>
<td>-0.0329</td>
<td>-0.03685</td>
<td>-0.00597</td>
<td>-0.00487</td>
<td>0.32153***</td>
<td>0.2314***</td>
<td></td>
</tr>
<tr>
<td>Lochaline</td>
<td>-0.01407</td>
<td>-0.01481</td>
<td>-0.03796</td>
<td>-0.00966</td>
<td>0.27196***</td>
<td>0.1927***</td>
<td></td>
</tr>
<tr>
<td>Loch Sunart</td>
<td>-0.01399</td>
<td>-0.0199</td>
<td>-0.03823</td>
<td>-0.03166</td>
<td>-0.2877***</td>
<td>0.20626***</td>
<td></td>
</tr>
<tr>
<td>Søgne</td>
<td>0.6022***</td>
<td>0.59216***</td>
<td>0.70076***</td>
<td>0.66935***</td>
<td>0.65673***</td>
<td>-</td>
<td>0.00776***</td>
</tr>
<tr>
<td>Hidra</td>
<td>0.42048***</td>
<td>0.38638***</td>
<td>0.43868***</td>
<td>0.43023***</td>
<td>0.41971***</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Values calculated from haplotype frequencies are above the diagonal, and values calculated using the Tamura and Nei estimator are below the diagonal. 

*\( p < 0.01 \).

***\( p < 0.001 \).
haplotype and nucleotide diversity than Atlantic populations, which was suggested to be the result of a very recent (post-LGM) colonization of northern areas (Robalo et al., 2012). Although caution has to be exercised when estimating times of divergence in the absence of an accurate mutation rate, the present study used the same divergence rate (5%) for the same DNA region as the corkwing wrasse study, but obtained older divergence times that pre-date the LGM. Assuming that the ballan and corkwing wrasse mtDNA control region is characterized by similar mutation rates, these results indicate that ballan and corkwing wrasse populations found in Scandinavian waters are characterized by distinct phylogeographical patterns during recolonization of the eastern North Atlantic following glacial times. This could be the result of the presence of distinct refugia around the British Isles, as hypothesized by recent studies on seaweed (Provan et al., 2005; Hoarau et al., 2007), which may have played a significant role in affecting the post-glacial colonization of different coastal species into the North Sea and Scandinavian waters. Nonetheless, divergence times as calculated in the present study should be corroborated by using additional mitochondrial and nuclear markers to obtain more reliable and accurate estimates. Furthermore, additional samples should be collected from areas throughout the species range in the eastern and central North Atlantic to enable a more comprehensive understanding of post-glacial colonization patterns for this species.

For ballan wrasse, the occurrence of two highly divergent clades which both show the evidence of expansion may be the results of the retention of at least two main haplotypes in possibly more than one refugium during the LGM, followed by post-glacial expansion. Differential representation of distinct mtDNA lineages is not uncommon in marine pelagic species (Finnerty and Block, 1992; Graves and McDowell, 1995; Chow et al., 2000; Martinez et al., 2006). For instance, two distinct mtDNA clades were identified in populations of swordfish (Xiphias gladius) in the North Atlantic and the Mediterranean Sea, where the frequency of haplotypes within each clade was used to determine levels of the population subdivision (Alvarado Bremer et al., 1996; Viñas et al., 2010). Such patterns of restricted gene flow have been suggested to be the result of oceanographic conditions and natality phenotypy (Viñas et al., 2010).

Differing intraspecific levels of genetic diversity indicate that ballan wrasse populations found in Scandinavian waters and around the British Isles have been shaped by distinct demographic histories. High haplotype diversity coupled with low nucleotide diversity has been considered as a signature of a genetic bottleneck followed by population expansion and accumulation of mutations (Avise, 2000). This could be the case in some of the populations investigated in the present study, for which the calculation of neutrality tests (Fu’s F, and Tajima’s D) and the investigation of the mismatch distribution indicated some degree of departure from mutation-drift equilibrium. Overall, the evidence of population expansion was detected in both British Isles (Atlantic) and Scandinavian populations, although it appeared to be more marked in the latter region. It is worth noting that the presence of two highly divergent clades may cause misinterpretation of patterns of mismatch distributions when calculated together (Marjoram and Donnelly, 1994) due to potential type II errors (i.e. failure to reject a false null hypothesis of equilibrium; Alvarado Bremer et al., 2005, and references therein). This can be the case for locations like Mweenish and Hidra, which showed a significant departure from a distribution under the model of expansion probably due to the presence of haplotypes from both clades and not because of true deviation from the expansion.

These findings indicate that population expansion may have occurred more recently in Norwegian than in Atlantic populations; hence, the studied Norwegian areas (located close to the northernmost limit of the species’ distribution range) were probably colonized by fewer individuals and more recently than the Irish and Scottish areas. This is in line with predictions that species characterized by a temperate-cold distribution should display lower levels of diversity in areas that were not accessible during the glacial periods (Avise, 2000; Hewitt, 2004), although exceptions to this trend have started to emerge (Almada et al., 2012). Interestingly, evidence of population expansion appears to be a recurrent feature in populations studies of wrasse species (Chen et al., 2004; Froukh and Kochzius, 2007; Haney et al., 2007). Such recurrent pattern could be the result of ongoing demographic fluctuations in which population expansion/contraction is driven by factors such as life history (e.g. egg type and larval dispersal time), reproductive mode (e.g. sequential hermaphroditism), and by variation of available habitats induced by sea-level changes.

Population structure based on mtDNA control region data

The present study provided the first evaluation of levels of population structuring in ballan wrasse in the eastern North Atlantic, where population differentiation was strong between regions but weak within regions. The low population structuring found around the British Isles is comparable with levels of mtDNA differentiation found in sixbar wrasse in the South China Sea (Chen et al., 2004), fourline wrasse in the Red Sea (Froukh and Kochzius, 2007), and bluehead wrasse in the Caribbean (Haney et al., 2007). In contrast, the stronger level of differentiation found between Norway and the British Isles tend to be less common in other wrasse species. This corroborates a previous phylogenetic analysis which indicated that ballan wrasse shows higher intraspecific sequence divergence than other wrasse species (Hanel et al., 2002).

The difference in levels of within-region differentiation (with population structuring being more marked in Norwegian sites) can be explained by local environmental factors (e.g. water currents and deep water) and life-history parameters (e.g. planktonic larval duration and adult dispersal). Ballan wrasse’s reproductive cycle includes benthic eggs, a relatively long larval stage of 37–49 d (Ottesen et al., 2012) and adults that can live for up to 25 years of age and to a maximum recorded length of 60 cm (Quignard and Pras, 1986; Darwall et al., 1992). Sexually mature ballan wrasse males tend to be territorial and form harems in localized areas (Darwall et al., 1992). Although the egg type (i.e. attached to the benthos) and territoriality may suggest limited dispersal from natal areas, the relatively long planktonic larval stage may contribute to lowering genetic differentiation between adjacent areas. Water currents can vary in inshore waters and may be responsible for larval transportation along the shore, which would explain the low levels of differentiation around the Irish coasts. Such variable intraspecific patterns of genetic differentiation indicate that local environmental conditions may play a significant role in population structuring of ballan wrasse. At a larger geographical scale, findings from the present study indicate that the major population break occurs in the North Sea and only...
marginally between Ireland and Scotland, indicating that deep waters are among the main barriers to gene flow for this species in the studied regions. This is not uncommon in demersal fish species that occur within certain depth ranges, where bathymetric barriers and limited egg dispersal are among the major factors shaping the population structure (Knutsen and Sannæs, 2009).

It has been suggested that selection would favour a sequential hermaphroditic life history to increase the reproductive success of each sex as different suitable size is reached within the life time of an individual; also known as “the size advantage hypothesis” (Ghiselin, 1969; Warner, 1975). A review of many sex-changing species has revealed that such life-history traits tend to lead to skewed population sex ratios in protogynous species (Allsop and West, 2004), which in turn may lead to smaller effective population size ($N_e$) and increased population structure (Hartl and Clark, 1997). A review of marine species of varying reproductive mode found that sex-changing strategies were not a good predictor of high levels of population structuring (Chopelet et al., 2009), though urging the need for further empirical data especially from hermaphroditic species. In the present study, levels of intra-specific genetic diversity and population structuring were variable across the sampled areas, further supporting the fact that reproductive mode cannot be used as a general indicator of species-wide levels of genetic structuring. However, results presented here should be corroborated by using additional nuclear markers, such as microsatellites and single nucleotide polymorphisms. At present, no such markers have been developed directly from ballan wrasse, and only very few studies have attempted the cross-species amplification of microsatellite loci (Knutsen and Sannæs, 2009). Preliminary testing of nine microsatellite markers from other wrasse species (Arigoni and Largiadèr, 2000; Galarza et al., 2006; Knutsen and Sannæs, 2009) showed poor cross-species amplification in the populations under study in the present manuscript (data not shown); thus, further novel markers should be developed ex novo from ballan wrasse.

**Implications for conservation and management**

Findings from the present study revealed some important aspects of ballan wrasse biology that will aid future conservation and management of the species. The harvesting of wild wrasse to be supplied to farms has been more intense in Norway, where large numbers of wild individuals are stocked annually onto salmon farms (Treasurer, 2012). Although this has affected mainly small wrasse species such as goldsinn, there is an increasing interest in using ballan wrasse due to its large body size, which makes it more suitable for treatment of bigger salmon (Kvenseth et al., 1996), and consequently, it may lead to a heavier use of this species in the future. Analyses of mtDNA data from the present study showed variable levels of genetic diversity in ballan wrasse populations, which are the result of past natural events (e.g. historical bottleneck followed by population expansion) more so than by contemporary human activities (e.g. recent bottleneck due to overfishing). However, the lower genetic diversity observed in southern Norway indicates that these populations will be more vulnerable to future pressure from human activities than populations around the British Isles. This is because populations with lower genetic diversity are regarded as more susceptible to inbreeding depression, reduced fitness, and lower evolutionary potential (Frankham, 2002).

Wrasse species may also be more susceptible to human-related pressure due to their peculiar reproductive mode (sequential hermaphroditism). In fact, it has been suggested that harvesting larger individuals from wild populations of sequential hermaphro- dite species may result in the removal of the less abundant secondary sex, which may have severe consequences to population structure and viability (Darwall et al., 1992; Sattar et al., 2008). Although this is the case for smaller wrasse species such as goldsinn wrasse, size-related pressure may not be as strong in ballan wrasse, where larger (probably male) individuals tend to exceed the maximum target size required for cleaner fish and therefore are not captured. Nonetheless, as sequential hermaphroditism is thought to be the result of an evolutionary process that maximizes fitness of both primary and secondary sex depending on body size (Kazancıog˘lu and Alonzo, 2010), caution should be exercised when applying size-selective fishing pressure to such species.

In areas where local populations are not abundant or outside the species’ distribution range (e.g. North of Norway), cleaner fish are generally supplied from wild populations occurring in southern areas, which leads to the translocation of fish from natal areas to distant farming sites (Sundt and Jørstad, 1998). A population analysis using tissue enzyme data of goldsinn wrasse found significant population structuring between local wild populations occurring in proximity to farms in northern Norway and captive fish originally transferred from the southern Skagerrak area (Sundt and Jørstad, 1998), indicating that any escapee would result in potential mixing of distinct populations. Based on the levels of population structuring uncovered in the present study, the translocation of wild ballan wrasse could lead to population mixing especially in Norwegian waters, and therefore such practice should be carried out with caution. Moreover, the transferring of fish from the wild into captivity leads to higher stress levels and potential for disease outbreaks (see review in Treasurer, 2012). To satisfy the increasing demand for cleaner fish and to provide disease-free individuals to farms, a number of programmes has been initiated for commercial rearing of ballan wrasse in captivity in countries of the Northeast Atlantic where farmed organic salmon is commonly produced (Norway, Scotland, and Ireland; e.g. Ottesen et al., 2011). Thus, based on the results from the present study, the population structure should be investigated in source populations for broodstock fish using mtDNA and additional markers to aid breeding programmes and to minimize the potential impacts of interactions between wild and farmed fish.

**Supplementary data**

Supplementary material is available at the ICESJMS online version of the manuscript.

**Acknowledgements**

The authors would like to acknowledge all people who helped with sample collection, including A. O’Conghaile, S. Amey, M. Yusof (Faizzj), J. Treasurer, L. Turner, J. Caldwell, R. Scotland, J. O’Donnell, and A. Lindhom.

**References**


Population genetics of ballan wrasse


Handling editor: Mikko Heino