

SHORT COMMUNICATION

Is the ballan wrasse (*Labrus bergylta*) two species? Genetic analysis reveals within-species divergence associated with plain and spotted morphotype frequencies

María QUINTELA,¹ Elin Annie DANIELSEN,¹ Lua LOPEZ,² Rodolfo BARREIRO,² Terje SVÅSAND,¹ Halvor KNUTSEN,^{1,3,4} Anne Berit SKIFTESVIK⁵ and Kevin A. GLOVER¹

¹Institute of Marine Research, Bergen, Norway, ²Grupo de investigación BIOCOST, University of A Coruña, Spain, ³Centre for Ecological and Evolutionary Synthesis, Department of Biology, University of Oslo, Oslo, Norway, ⁴University of Agder, Kristiansand, Norway and ⁵Austevoll Research Station, Storebø, Norway

Abstract

The ballan wrasse (*Labrus bergylta*) is a marine fish belonging to the family Labridae characterized by 2 main morphotypes that occur in sympatry: spotty and plain. Previous studies have revealed differences in their life-history traits, such as growth and maturation; however, the genetic relationship between forms is presently unknown. Using 20 recently developed microsatellite markers, we conducted a genetic analysis of 41 and 48 spotty and plain ballan wrasse collected in Galicia (northwest Spain). The 2 morphotypes displayed highly significant genetic differences to each other ($F_{ST} = 0.018$, $P < 0.0001$). A similar degree of genetic differentiation ($F_{ST} = 0.025$, $P < 0.0001$) was shown using the STRUCTURE clustering approach with no priors at $K = 2$. In this case, the frequency of spotty and plain morphotypes was significantly different ($\chi^2 = 9.46$, $P = 0.002$). It is concluded that there is significant genetic heterogeneity within this species, which appears to be highly associated with the spotty and plain forms, but not completely explained by them. Given the previously demonstrated biological differences between morphotypes, and the present genetic analyses, we speculate about the convenience of a taxonomic re-evaluation of this species.

Keywords: ballan wrasse, color morphotypes, *Labrus bergylta*, microsatellites, speciation

INTRODUCTION

Ballan wrasse (*Labrus bergylta* Ascanius, 1767, Labridae) is a protogynous hermaphrodite marine fish showing highly variable color patterns that are not related with the sex of the individuals, as opposed to other wrasse species (Villegas-Ríos *et al.* 2013a). In Galicia (northwest Spain), 2 distinct morphotypes of this species are known to coexist in sympatry. These include

Correspondence: María Quintela, Institute of Marine Research, PO Box 1870, 5817 Bergen, Norway.
Email: maria.quintela.sanchez@imr.no

spotted colored individuals (Fig. 1a) that display a dark orange or reddish body patterned with white dots, and plain colored individuals (i.e. non-spotted; Fig. 1b) that are characterized by a uniform, although variable, body color (mainly greenish, brownish or reddish), darker in the back and whitish in the abdomen (Villegas-Ríos *et al.* 2013b). In Galicia, where this has been investigated, both types display overlapping but, nevertheless, different life-history strategies, with plain morphs investing more in reproduction at the expense of growing less versus spotted morphs investing less in reproduction but growing bigger (Villegas-Ríos 2013). These differentiated types not only hold different common names in this region, but are commercialized separately (spotted individuals being more expensive) (Villegas-Ríos *et al.* 2013b). Together, these observations indicate that the 2 morphotypes may belong to genetically different groups. Indeed, the possibility of them representing 2 separate species has been put forward based upon their slightly different life-history strategies (Villegas-Ríos 2013). However, this remains to be investigated.

In spite of the growing body of literature about ballan wrasse life history in recent times (e.g. D'Arcy *et al.* 2012; Villegas-Ríos *et al.* 2013a,b; Villegas-Ríos *et al.* 2014), the population structure and demographic history of this species is still poorly known in the wild, as it is the basis of differences between plain and spotted individuals. These divergent patterns of coloration have also been observed in other geographic areas

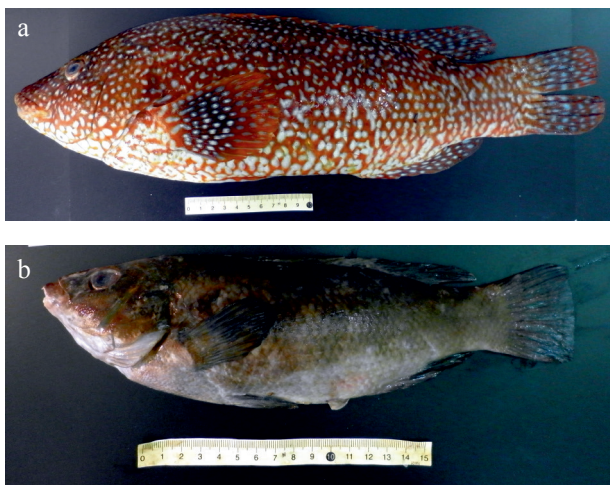


Figure 1 Ballan wrasse morphotypes: Pictures of (a) spotted and (b) non-spotted morphs.

of the distribution range, from Azores Islands to Norway (see Villegas-Ríos *et al.* [2013b] and references therein), and, thus, it has been recommended to consider plain and spotted morphotypes as 2 independent management units throughout its distribution range, at least while the taxonomic status of the species remains unresolved (Villegas-Ríos *et al.* 2013b). The persistence of color morphs within populations is usually attributed to a number of factors, such as non-random mating, fluctuations in selection regimes, and frequency-dependent selection by predation or intrasexual competition (e.g. Hoekstra *et al.* 2004). In addition, within-population sexual selection can generate negative frequency dependence, initiating reproductive isolation and, thus, allowing color morphs to speciate even in sympatry (e.g. Gray & McKinnon 2007).

Recently, we developed a set of 20 microsatellite markers that enables genetic studies of this species (Quintela *et al.* 2014). Here, the main aim was to evaluate whether there were any genetic differences between plain and spotted morphotypes of ballan wrasse in order to investigate the existence and, eventually, the degree of disruptive selection between color patterns as well as to measure the gene flow between them. A second aim was to provide some insight about the taxonomic status of both phenotypes.

MATERIAL AND METHODS

Biological samples and DNA extraction

A total of 89 individuals belonging to the spotted (41) and non-spotted (48) morphotypes were obtained in the Galician coast (northwest Spain) from a local fish market in A Coruña (43°21'33.98"N, 8°24'7.20"W). No specific permission for sampling was required as the individuals sampled for this study were obtained from commercial artisanal inshore fishing. All samples were photographed, weight and lengths were taken, and genomic DNA was extracted from muscle using the Qia-gen DNeasy Blood & Tissue Kit following the manufacturer's instructions.

Genetic analyses with mtDNA

We used the barcoding approach to assess whether there were differences in mtDNA sequence that could support the hypothesis of ballan wrasse morphotypes belonging to 2 different species. Thus, a 648-bp region of mitochondrial COI gene was amplified in all the individuals using the universal primers described by Folmer

et al. (1994). Amplifications were performed in 25 μ L of a solution containing 0.5 mM of each primer, 0.2 mM of each dNTP, 2 mM MgCl₂, 1 \times AmpliTaq buffer, 1.25 U AmpliTaq DNA polymerase (Applied Biosystems, Foster City, CA, USA) and 1.5 μ L DNA. Cycling conditions were 2 min denaturing at 95°C followed by 30 cycles of 30 s at 95°C, 30 s at 45°C and 1 min at 72°C. After removing the excess of primers and nucleotides (with shrimp alkaline phosphatase and exonuclease I enzymes), samples were sequenced by MacroGen (MacroGen, Seoul, Korea) on ABI 3730 instruments (Applied Biosystems). Sequences were checked and edited using the program Genious v. 6.1.7 (Biomatters, Auckland, New Zealand). After alignment and trimming, the final sequence length was 648 bp.

Genetic analyses with microsatellites

Twenty microsatellite loci (WrA103, WrA107, WrA111, WrA112, WrA113, WrA203, WrA223, WrA224, WrA228, WrA236, WrA237, WrA254, WrA255, WrA256, WrA259, WrA261, WrB102, WrB212, WrB213, WrB215) were used as described in Quintela *et al.* (2014). Six multiplexed PCR reactions were performed in final volume of 10 μ L containing 50 ng DNA template, 1 \times buffer, 2 mM MgCl₂, 1.25 mM dNTPs, 0.06–0.12 μ M of each primer and 1U Go-Taq polymerase. PCR profiles included an initial 4-min denaturation at 94°C followed by 24 cycles of 50 s at 94°C, 90 s at an annealing temperature of 56°C, 1 min of extension at 72°C and a final extension of 72°C for 10 min. Forward primers were labeled with fluorescent dyes and PCR products were electrophoresed on an ABI Prism 377 Genetic Analyzer (Applied Biosystems). The 500LIZ size standard (Applied Biosystems) was used to accurately determine the size of the fragments and allelic variation. Fragments were analyzed with the software GeneMapper v5 (Applied Biosystems). All samples were genotyped twice to evaluate genotyping consistency, and repeatability was found to be 99.86%. The laboratory in which the microsatellite analysis was conducted performs analyses for the regulatory authorities, and routinely checks genotyping quality.

Statistical analyses

In order to detect any sign of divergence between plain and spotted morphotypes, we conducted outlier detection tests aiming to identify loci under directional selection. To minimize the risk of detecting false positives, we used 2 different procedures to detect loci with a pattern deviating from neutrality. First, we used the hierarchical Bayesian method implemented in BayeScan

software (Foll & Gaggiotti 2008) and, second, we used the Fdist approach implemented in LOSITAN (Antao *et al.* 2008). However, as none of the outlier detection procedures performed on the 89 fish divided according to morphotypes revealed any locus deviating from neutral expectations, the total set of 20 microsatellite loci was regarded as neutral for all subsequent analyses.

Samples sorted into their respective morphotypes (i.e. spotty and plain) were used to estimate the total number of alleles, the number of private alleles, observed (H_o) and unbiased expected heterozygosity (U_{H_e}), and the inbreeding coefficient F_{IS} per morphotype using the program GenAlEx (Peakall & Smouse 2006). The genotype distribution of each locus per morph and its direction (heterozygote deficit or excess) was compared with the expected Hardy–Weinberg distribution using the program GENEPOP on the web (Rousset 2008) as was linkage disequilibrium. Both were examined using the following Markov chain parameters: 10 000 steps of dememorization, 1000 batches and 10 000 iterations per batch. The same analyses were also conducted on STRUCTURE-produced clusters (see below).

We used BAPS 6.0 (Corander *et al.* 2004) in a blind manner to identify the most likely number of individual clusters (from $K = 1$ to $K = 20$) and then we performed the most likely admixture of genotypes: an approach that is powerful in identifying hidden structure within populations (Corander & Marttinen 2006).

To investigate if the division according phenotypes would match any potential genetic differentiation, we performed a blind clustering of the 89 samples (i.e. without the use of any prior) using the Bayesian model-based clustering algorithms implemented in STRUCTURE v. 2.3.4 (Pritchard *et al.* 2000). We chose a model assuming admixture and correlated allele frequencies without using population information and with all individuals regarded as 1 single population. Ten runs with a burn-in period consisting of 100 000 replications and a run length of 1 000 000 Markov chain Monte Carlo (MCMC) iterations were performed for a number of clusters ranging between $K1$ and $K5$. We then used STRUCTURE Harvester (Earl & von Holdt 2012) to calculate the Evanno *et al.* (2005) *ad hoc* summary statistic ΔK , which is based on the rate of change of the “estimated likelihood” between successive K values. Runs were averaged with CLUMPP version 1.1.1 (Jakobsson & Rosenberg 2007) using the LargeKGreedy algorithm and the G' pairwise matrix similarity statistics. Averaged runs were displayed using barplots. Furthermore, STRUCTURE was also conducted under al-

most identical conditions but with a slightly different approach; that is, assisting the clustering with popinfo *a priori*.

To compare the genetic differentiation between morphotypes on the one hand, and the clusters of individuals on the other, we computed traditional pairwise F_{ST} (Weir & Cockerham 1984) analysis and F_{ST} per locus using ARLEQUIN v.3.5.1.2 (Excoffier *et al.* 2005), and G'_{ST} using GenAEx (Peakall & Smouse, 2006). G'_{ST} is useful in situations like the present one when the number of sampled populations is small, especially for pairwise comparisons. In all cases, signification was calculated using 10 000 permutations. In addition, a factorial correspondence analysis was applied using the software GENETIX (Belkhir *et al.* 2004) to further assess the distribution of samples by putative groups based on individual genotypes.

The pattern and magnitude of gene flow between morphotypes was assessed by using the Bayesian inference with the assistance of BayesAss v. 1.3 (Wilson & Rannala 2003). This approach does not rely on a dataset that is on Hardy–Weinberg or in migration-drift equilibrium, and it uses a MCMC algorithm to estimate the posterior probability distribution of the proportion of migrants from one population to another (M). The analysis was conducted with 3×10^6 iterations, with a burn-in of 10^6 iterations, and a sampling frequency of 2000 to ensure that the model's starting parameters were sufficiently randomized.

RESULTS

mtDNA

Eighty-seven individuals were successfully sequenced for COI (2 of the non-spotted ones failed). In total, the 87 samples revealed 4 mutations in 4 polymorphic sites (one of them parsimoniously informative), which generated 5 different haplotypes. Haplotype 1, present in 79 of the fish, clearly dominated in both morphotypes: 87% in the non-spotted versus 95.1% in the spotted one (Fig. 2). Individuals belonging to the plain morphotype showed 3 unique haplotypes not detected in the spotted morph (Haplotypes 3, 4 and 5 that were present in 1 individual each).

Microsatellite genetic analysis sorting samples according to morphotypes

All 89 individuals were amplified for the suite of microsatellites. Thus, while 2 of the samples failed to am-

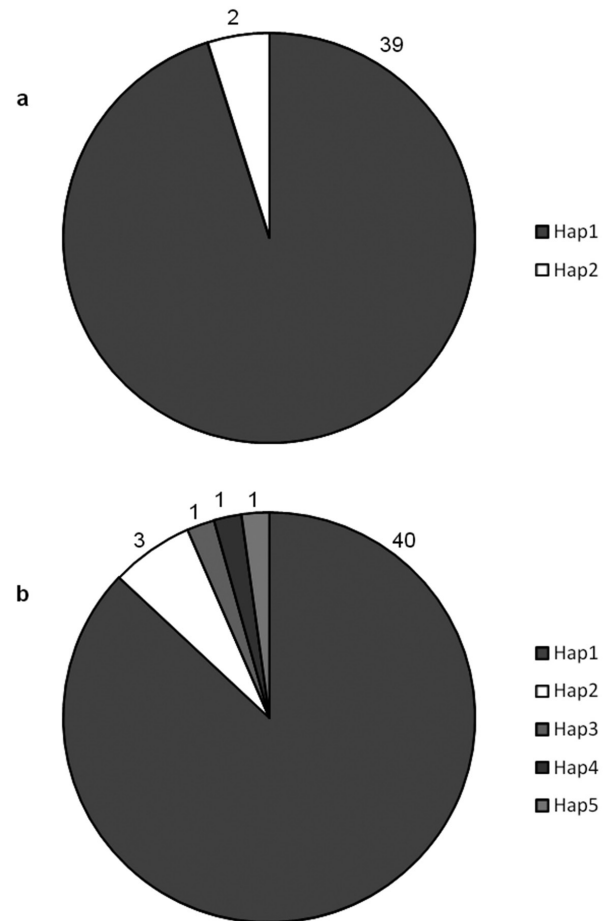


Figure 2 Ballan wrasse mtDNA - COI haplotype distribution according to morphotypes: (a) spotted and (b) non-spotted. The numbers of individuals per haplotype are depicted besides the pie chart.

plify for COI, they provided microsatellite profiles consistent with the species (Quintela *et al.* 2014). A total of 150 microsatellite alleles across the 20 markers were observed in the 89 samples. Individuals belonging to the non-spotted morphotype displayed smaller size than the spotted ones, higher number of total and private alleles and slightly higher allelic richness (Table 1a). Observed and unbiased expected heterozygosity together with the inbreeding coefficient were similar between morphotypes. Analysis of Hardy–Weinberg equilibrium revealed that, at the significance level of α 0.05, 5% of loci by sample combinations in the spotted-morph displayed significant deviations, whereas no deviations were registered at the significance level of α 0.001. In the non-spotted morphotype, 10% of loci deviated from

Table 1 Ballan wrasse microsatellites

a Division according phenotypic morphotypes										
Group	<i>N</i>	Size (cm)	Number of alleles	<i>Ar</i>	Number of private alleles	<i>Ho</i>	<i>uHe</i>	<i>F_{IS}</i>	No deviations Hardy–Weinberg equilibrium	Number of deviations LD
Spotted morph	41	30.1 ± 5.1	129	6.5	12	0.616 ± 0.050	0.619 ± 0.044	0.004 ± 0.027	1	11
Plain morph	48	25.8 ± 4.7	138	6.7	21	0.643 ± 0.039	0.656 ± 0.041	−0.002 ± 0.027	2	13
b Division according STRUCTURE clustering with a threshold of >0.50										
Group	<i>N</i>	Size (cm)	Number of alleles	<i>Ar</i>	Number of private alleles	<i>Ho</i>	<i>uHe</i>	<i>F_{IS}</i>	Number of deviations Hardy–Weinberg equilibrium	Number of deviations LD
STRUCTURE cluster I	44	29.3 ± 4.7	134	6.8	15	0.589 ± 0.047	0.587 ± 0.045	−0.015 ± 0.025	1	11
STRUCTURE cluster II	45	26.5 ± 5.6	135	6.8	16	0.671 ± 0.041	0.686 ± 0.039	0.007 ± 0.027	1	7

Summary statistics per grouping method (color morphotypes and STRUCTURE clustering) showing number of individuals, size (mean ± SD), total number of alleles, allelic richness, number of private alleles, observed heterozygosity (average ± SE), unbiased expected heterozygosity (average ± SE), and inbreeding coefficient (F_{IS}) (average ± SD), number of deviations from Hardy–Weinberg equilibrium at $P < 0.05$ and number of deviations from genotypic linkage disequilibrium (LD) at $P < 0.05$.

Hardy–Weinberg equilibrium at $\alpha 0.05$ and half of those at $\alpha 0.001$. Linkage disequilibrium (LD) was detected in the spotted morphs 11 times (5.8%) at $\alpha 0.05$ and 6 (3.2%) at $\alpha 0.001$, whereas in the non-spotted ones it was detected 13 times (6.8%) at $\alpha 0.05$ and 3 (1.6%) at $\alpha 0.001$.

Pairwise F_{ST} between morphotypes provided by the 20 microsatellite markers was found to be highly significant ($F_{ST} = 0.018$, $P < 0.0001$; $G'_{ST} = 0.050$, $P < 0.0001$) and the analysis of the allele frequencies between them revealed 10 alleles in 8 loci showing frequency differences ranging between 0.2 and 0.3 (Table 2). The pairwise F_{ST} per locus (Table 3) revealed 9 loci out of 20 showing an F_{ST} significantly different from zero (WrA103, WrA203, WrA111, WrA254, WrA113, WrA228, WrA261, WrA259, and WrB102), 8 of which coincided with the ones that showed the highest allele-frequency differences between the plain and spotted morphotypes. Factorial correspondence analyses further confirmed the separation between morphotypes, despite the low percentage of the total variation explained by the 2 first axes: 7.39% (Fig. 3).

The assessment of gene flow within and between morphotypes revealed an asymmetrical pattern. Hence, the probability of being non-migrant was 68%, whereas the gene flow between types was 32%.

Microsatellite analysis based upon clustering methods

BAPS reported one single genetic group (probability = 0.93) for a number of *Ks* ranging between 1 and 20. Likewise, STRUCTURE analyses blindly conducted on the 89 samples showed similar average likelihood at K1, K2 and K3, although with K1 showing the highest value (Fig. 4a). However, the lack of striking differences between likelihood at K1 and K2, combined with an independent metric (i.e. the existence of 2 clearly different morphotypes), encouraged us to explore the scenario provided by STRUCTURE at $K = 2$ to assess the frequency of each morphotype per cluster in order to investigate whether it coincided with the genetic data. Thus, first we clumped 10 runs of STRUCTURE at $K = 2$, and then, we set a threshold of inferred membership to cluster of >0.50 to distribute the individuals into the corresponding clusters. Despite the fact that at $K = 2$ STRUCTURE did not neatly divide the fish into spotted and non-spotted morphs based upon the genetic data alone, there were, nevertheless, significant differences in the frequencies of each phenotype in the 2 genetic clusters ($\chi^2 = 9.46$, $P = 0.002$). Hence, the distribution

Table 2 Ballan wrasse allele frequency

Locus	Allele	MORPHOTYPES		STRUCTURE K2	
		Spotted	Plain	Cluster I	Cluster II
WrA103	194	0.110	0.302	0.227	0.200
WrA103*	199*	0.317	0.083	0.273	0.111
WrA111	210	0.159	0.344	0.239	0.278
WrA111	222	0.220	0.052	0.170	0.089
WrA113*	197*	0.500	0.323	0.568	0.244
WrA203*	168*	0.549	0.281	0.523	0.289
WrA228*	167*	0.500	0.323	0.568	0.244
WrA254*	200*	0.793	0.576	0.791	0.568
WrA259*	195*	0.098	0.250	0.102	0.256
WrA261	234	0.256	0.083	0.159	0.167

Values for those alleles showing the highest difference in frequency (0.2-0.3) between morphotypes (spotted vs. plain) and genetic clusters according STRUCTURE at $K=2$ without priors (I vs. II) are depicted in boldface type. *These alleles matched both criteria.

Table 3 Ballan wrasse pairwise F_{ST} per locus between morphotypes (plain vs. spotted) and STRUCTURE K2 genetic clusters (I vs. II)

Locus	No alleles	MORPHOTYPES		STRUCTURE K2	
		F_{ST}	P -value	F_{ST}	P -value
WrA103*	6	0.0652	0.0000	0.0212	0.0178
WrA107	3	0.0135	0.1391	0.0545	0.0055
WrA111*	9	0.0456	0.0006	0.0168	0.0441
WrA112	13	0.0002	0.4169	0.0202	0.0111
WrA113*	9	0.0236	0.0110	0.0716	0.0000
WrA203*	16	0.0531	0.0000	0.0369	0.0022
WrA223	5	0.0031	0.2336	0.0053	0.2017
WrA224	5	0.0000	0.4577	0.0000	0.4954
WrA228*	9	0.0236	0.0113	0.0716	0.0000
WrA236	9	0.0000	0.6750	0.0049	0.1887
WrA237	3	0.0000	0.5866	0.0000	0.4352
WrA254*	5	0.0553	0.0055	0.0573	0.0050
WrA255	7	0.0036	0.2914	0.0000	0.8386
WrA256	10	0.0033	0.2679	0.0018	0.3290
WrA259*	8	0.0173	0.0432	0.0186	0.0409
WrA261*	6	0.0348	0.0146	0.0000	0.9804
WrB102	2	0.0339	0.0457	0.0539	0.0170
WrB212	9	0.0000	0.4617	0.0694	0.0000
WrB213	10	0.0017	0.2469	0.0082	0.0926
WrB215	6	0.0000	0.8548	0.0013	0.4322

Values in boldface type were significantly different from zero. *These loci showed the highest difference in allele frequencies between morphotypes.

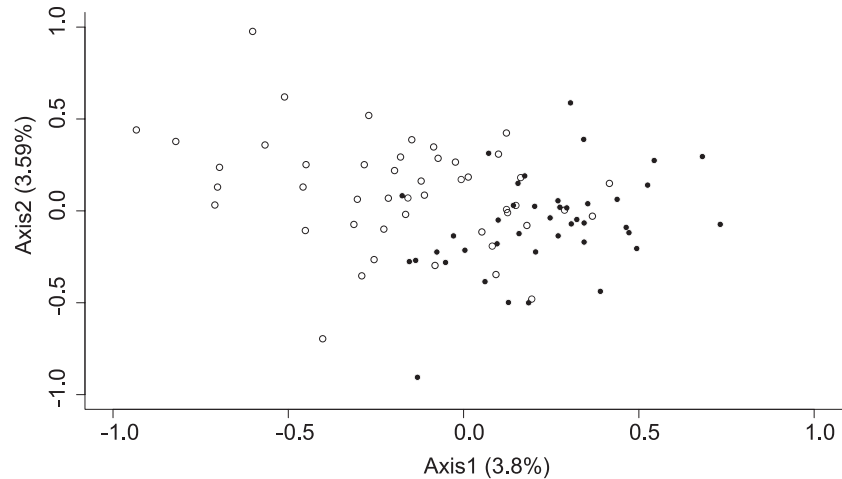


Figure 3 Two-dimensional representation of the Factorial Correspondence Analysis. Individuals are projected on the factor space defined by the similarity of their allelic states. White dots depict plain morphotype whereas black dots depict spotted individuals.

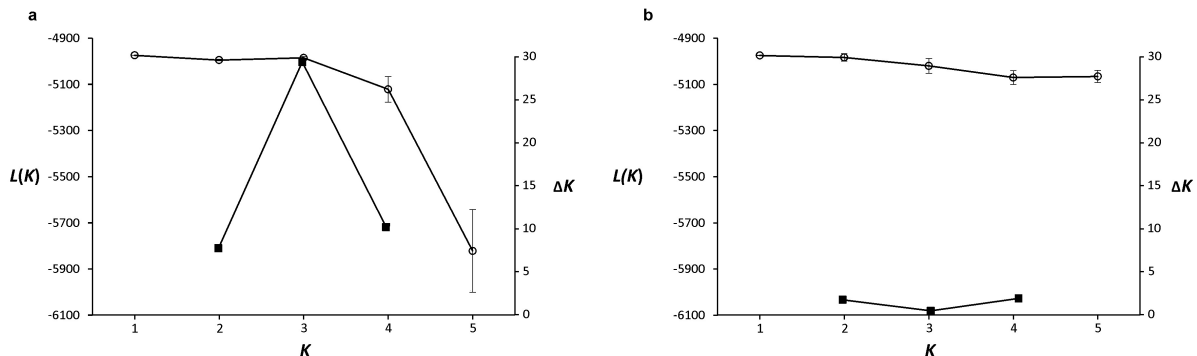


Figure 4 Plot of mean posterior probability values ($L(K) \pm SD$) per cluster (K) based on 10 replicates per K generated by STRUCTURE (circles), and ΔK analysis (squares) of $\text{LnP}(D)$ according to Evanno *et al.* (2005) for: (a) blind clustering approach and (b) using morphotypes as popinfo.

of fish into clusters was as follows: 28 spotted individuals (63.6%) were assigned to cluster I and 13 (28.9%) to cluster II, whereas 16 of the plain fish (36.4%) were assigned to cluster I and 32 (71.1%) to cluster II (STRUCTURE barplots in Fig. 5). Furthermore, individuals distributed between clusters I and II were even in number, and showed similar sizes, and almost identical number of total alleles, allelic richness and private alleles (Table 1b). Only H_o , UHe and F_{IS} took slightly lower values in cluster I than in II. Analysis of Hardy–Weinberg equilibrium revealed that, at the significance level of α 0.01, 5% of loci by sample combinations in cluster I displayed significant deviations; whereas no deviations were registered at the significance level of α 0.001. In cluster II, only 1 locus (5%) showed deviations from Hardy–Weinberg equilibrium at α 0.001. LD was detected in clus-

ter I 11 times (5.8%) at α 0.05 and 2 (1.1%) at α 0.001, whereas in cluster II it was detected 7 times (3.7%) at α 0.05 and 1 (0.5%) at α 0.001.

The genetic structure between clusters was highly significant ($F_{ST} = 0.025$, $P < 0.000$; $G'_{ST} = 0.067$, $P < 0.0001$), and this way of arranging the individuals revealed a slightly higher genetic differentiation than the division according to phenotypes ($F_{ST} = 0.018$, $P < 0.0001$; $G'_{ST} = 0.050$, $P < 0.0001$). Likewise, the analysis of allele frequencies revealed 6 alleles showing differences in frequencies between the 2 clusters ranging between 0.2 and 0.3 (Table 2), in the same direction and of similar magnitude as the ones between morphotypes. The latter point is important as it reveals that the 2 independent methods (sorting by morphotype and sorting by clustering) pointed to the same markers as separat-

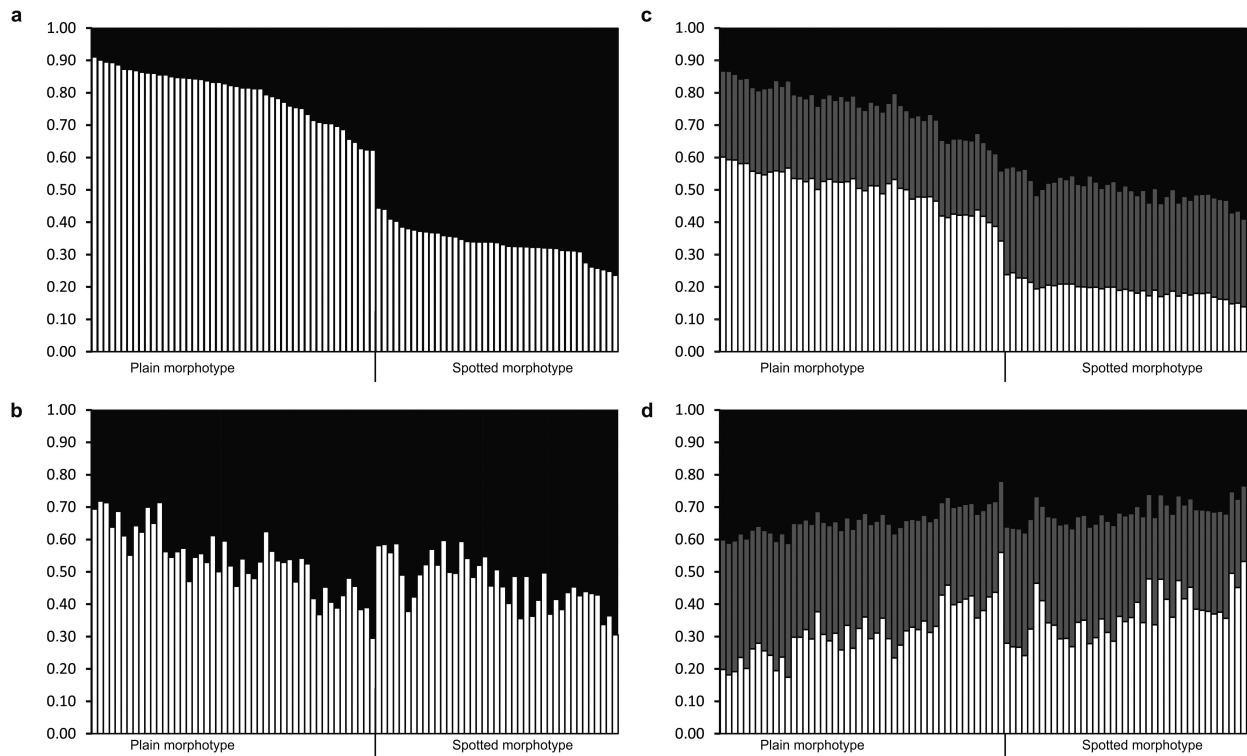


Figure 5 Inferred ancestry of plain and spotted morphotypes of ballan wrasse according STRUCTURE at $K = 2$ and $K3$ for: (a,c) clustering using morphotypes as popinfo; and (b,d) blind clustering approach. The order of the individuals has been kept identical in each of the 4 barplots.

ing the samples into 2 groups. Factorial correspondence analyses run on the total 20 markers confirmed the separation between clusters I and II and showed the same percentage of total variation explained by the 2 first axes as in the morphotypes analysis (7.39%).

When assisting STRUCTURE clustering with morphotypes as popinfo, $L(K)$ took very similar values at $K2$ and $K3$, and ΔK did not have enough grounds to discriminate between $K2$ and $K4$ either (Fig. 4b). However, and as expected, individuals were neatly divided between cluster I and II according to their phenotypes.

DISCUSSION

This represents the first genetic study of the sympatric color morphotypes of ballan wrasse. Highly significant genetic differences were observed between plain and spotted morphs sampled from the same region of north-west Spain ($F_{ST} = 0.018$ averaged over 20 loci, P

< 0.0001). In addition, blind clustering methods based upon genetic data alone, without the use of any kind of prior, sorted the samples into 2 genetically distinct groups ($F_{ST} = 0.025$ averaged over 20 loci, $P < 0.0001$), where the frequencies of the spotty and plain morphotypes were significantly different ($\chi^2 = 9.46$, $P = 0.002$). Thus, the combination of these 2 analytical approaches together with the asymmetric pattern of gene flow within and between morphs confirms that the genetic differentiation within this species appears to be associated with, but not necessarily completely explained by, the 2 main morphotypes spotty and plain.

The assessment of genetic differentiation between morphotypes of ballan wrasse conducted in the present study aimed not only to investigate potential disruptive selection between color patterns, but also to provide some insight about any putative taxonomic difference between them. In contrast with the popular belief, sig-

nificant homogeneity revealed by mitochondrial COI, with haplotype 1 clearly dominating in both groups, did not seem to provide strong support for the hypothesis of plain and spotted individuals belonging to 2 different species. However, and significantly, at the nuclear level, 40% of the studied loci showed large differences in allele frequencies between color morphotypes. Moreover, the loci flagged as displaying strong divergence between the 2 morphotypes, and the 2 genetic clusters which were associated with morphotype frequency differences, strongly coincided.

The degree of structure between morphs in ballan wrasse evokes a recurrent question on marine organisms; that is, the interpretation of low but still statistically significant levels of genetic differentiation. In an attempt to determine to what extent such observations reflected a real phenomenon and not the result of confounding factors (i.e. unrepresentative sampling, selection acting on the marker loci), a long-term study on coastal Atlantic cod combined genetics and extensive capture–mark–recapture to conclude that the low, and yet highly significant, F_{ST} (average 0.0037) was biologically meaningful as it did correspond to separate temporally persistent local populations (Knutsen *et al.* 2011). The degree of differentiation found in the present study was found to be 5-fold higher and matched, for instance, the one reported for color morphotypes of the rockfish *Sebastes oculatus* assessed with a suite of 24 microsatellites (Venerus *et al.* 2013). In this case, the genetic differentiation added to the bathymetric segregation between color morphotypes (“dark” and “light” fish) suggested the existence of speciation-by-depth in the absence of physical barriers, although the large number of individuals displaying high levels of admixture invoked incomplete reproductive barriers between color morphotypes. Likewise, significant genetic differences of the same magnitude ($F_{ST} = 0.032$) were found between littoral and profound morphotypes of the Arctic charr (*Salvelinus alpinus*) using microsatellite DNA analysis, with morphs showing strong reproductive isolation in time and space (Westgaard *et al.* 2004).

A previous study on spotty and plain morphotypes of ballan wrasse reported differences in growth trajectories, size at age, otolith length to body length, and mortality (Villegas-Ríos *et al.* 2013b), and also in investment in reproduction versus growth (Villegas-Ríos 2013). Based upon those results, which included extensive growth and maturation analyses, it was concluded that there was likely to exist large genetic differences between these 2 morphotypes. Clearly, while our genet-

ic data support this suggestion, there may be enough genetic overlap between the 2 morphotypes to indicate that it might not be as simple as spotty and plain forms being potential subspecies. It is, for example, possible that the 2 genetic sub-groups differ in their frequency of these morphotypes (i.e. one is primarily spotty and the other primarily plain). If this is the case, then some of the observed differences in traits between the spotty and plain morphotypes as reported by Villegas-Ríos *et al.* (2013b) may actually represent underestimates of the differences between the 2 genetic groups as there is no exact coincidence between morphotype and genetic grouping. Further studies using larger numbers of individuals and markers, and matching extensive life-history and trait data to the genetic data will be required to fully resolve this issue.

The accurate evaluation of the taxonomic status of ballan wrasse is key for the sustainable management of this species, which is commercially valuable throughout its distribution range for recreational and artisanal fishing. The small home range of *L. bergylta* together with its sedentary behavior highlights the potential use of small marine protected areas as a management tool to ensure a sustainable fishery (Villegas-Ríos *et al.* 2013c). In addition, the use of ballan wrasse as a biological cleaner fish for the control of sea lice on farmed salmonids can reduce the environmental impact of aquaculture (e.g. Skiftesvik *et al.* 2013), and has been reported to be an efficient alternative to the chemical treatment of sea lice, thus resulting in the initiation of dedicated breeding programs in Norway (Skiftesvik *et al.* 2013). Therefore, a thorough understanding of the within-species genetic structure associated with the morphotypes will be most valuable. Furthermore, the use of this species to control sea lice infestations on marine farms still primarily involves the capture and translocation of wild fish over long distances. Thus, the lack of knowledge about this species’ population genetic structure, the number of populations from which individuals are being harvested and mixed in the farms, or even the existence of local adaptations can have dramatic consequences if hybridization between translocated and local populations takes place and hampers the local evolutionary adaptations via maladaptive gene combinations or outbreeding depression (e.g. McGinnity *et al.* 2003). Likewise, overexploitation of wrasse in one area may leave a local population endangered or locally extinct if there is little gene flow among regions (Hutchinson 2008).

Finally, a remarkable feature is that Galician ballan wrasse analyzed here showed a substantially high-

er number of alleles than an almost 3-fold larger set of individuals sampled in the Norwegian coast (150 vs 115 alleles, respectively) (Quintela *et al.* 2014). Although this number of populations study does not allow any further assumptions to be made, we would like to point out the fact that also in the control region of mtDNA, a lower genetic diversity has been formerly reported for this species in the northernmost areas when comparing southern Norway and the British Isles, which was interpreted as the result of historical demographic events (population bottleneck followed by expansion) (D'Arcy *et al.* 2013). The same pattern has also been described for the corkwing wrasse (*Symphodus melops*), with south European populations showing twice as much allelic richness and average number of alleles per locus than the Scandinavian populations (Knutsen *et al.* 2013).

ACKNOWLEDGMENTS

Funding for this study was provided by the Norwegian Ministry of Trade and Fisheries. We are grateful to Raquel Mosquera, Belén Carro, Cristina Pérez Corbeira and Alfredo Veiga for their help with the collection of samples. We thank Anne Grete Sørvik Eide, Bjørghild Breistein Seliussen and Geir Dahle for their valuable assistance in the laboratory.

REFERENCES

- Antao T, Lopes A, Lopes R, Beja-Pereira A, Luikart G (2008). LOSITAN: A workbench to detect molecular adaptation based on a F_{ST} -outlier method. *BMC Bioinformatics* **9**, 323.
- Belkhir K, Borsa P, Chikhi L, Raufaste N, Bonhomme F (2004). GENETIX 4.05, logiciel sous Windows TM pour la génétique des populations.
- Corander J, Marttinen P (2006). Bayesian identification of admixture events using multilocus molecular markers. *Molecular Ecology* **15**, 2833–43.
- Corander J, Waldmann P, Marttinen P, Sillanpää MJ (2004). BAPS 2: Enhanced possibilities for the analysis of genetic population structure. *Bioinformatics* **20**, 2363–9.
- D'Arcy J, Mirimin L, Fitzgerald R (2013). Phylogeographic structure of a protogynous hermaphrodite species, the ballan wrasse *Labrus bergylta*, in Ireland, Scotland, and Norway, using mitochondrial DNA sequence data. *ICES Journal of Marine Science: Journal du Conseil* **70**, 685–93.
- D'Arcy J, Dunaevskaya E, Treasurer JW *et al.* (2012). Embryonic development in ballan wrasse *Labrus bergylta*. *Journal of Fish Biology* **81**, 1101–10.
- Earl DA, von Holdt BM (2012). STRUCTURE HARVESTER: A website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources* **4**, 359–61.
- Evanno G, Regnaut S, Goudet J (2005). Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study. *Molecular Ecology* **14**, 2611–20.
- Excoffier L, Laval G, Schneider S (2005). Arlequin ver. 3.0: An integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online* **1**, 47–50.
- Foll M, Gaggiotti O (2008). A genome-scan method to identify selected loci appropriate for both dominant and codominant markers: A Bayesian perspective. *Genetics* **180**, 977–93.
- Folmer O, Black M, Hoeh W, Lutz R, Vrijenhoek R (1994). DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. *Molecular Marine Biology and Biotechnology* **3**, 294–9.
- Gray SM, McKinnon JS (2007). Linking color polymorphism maintenance and speciation. *Trends in Ecology & Evolution* **22**, 71–9.
- Hoekstra, Hopi E, Drumm, Kristen E, Nachman, Michael W (2004). Ecological genetics of adaptive color polymorphism in pocket mice: Geographic variation in selected and neutral genes. *Evolution* **58**, 1329–41.
- Hutchinson WF (2008). The dangers of ignoring stock complexity in fishery management: The case of the North Sea cod. *Biology Letters* **4**, 693–5.
- Jakobsson M, Rosenberg NA (2007). CLUMPP: A cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics* **23**, 1801–6.
- Knutsen H, Jorde PE, Blanco González E, Robalo J, Albretsen J, Almada V (2013). Climate change and genetic structure of leading edge and rear end populations in a northwards shifting marine fish species, the corkwing wrasse (*Symphodus melops*). *PLoS ONE* **8**, e67492.
- Knutsen H, Olsen EM, Jorde PE, Espeland SH, André C, Stenseth NC (2011). Are low but statistically significant levels of genetic differentiation in marine fishes 'biologically meaningful'? A case study of coastal Atlantic cod. *Molecular Ecology* **20**, 768–83.

- McGinnity P, Prodöhl P, Ferguson A *et al.* (2003). Fitness reduction and potential extinction of wild populations of Atlantic salmon, *Salmo salar*, as a result of interactions with escaped farm salmon. *Proceedings of the Royal Society of London. Series B: Biological Sciences* **270**, 2443–50.
- Peakall R, Smouse PE (2006). GenA1Ex 6: Genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes* **6**, 288–95.
- Pritchard JK, Stephens M, Donnelly P (2000). Inference of population structure using multilocus genotype data. *Genetics* **155**, 945–59.
- Quintela M, Danielsen EA, Svåsand T, Knutsen H, Skiftesvik AB, Glover KA (2014). Isolation and characterization of twenty microsatellite loci for the ballan wrasse, *Labrus bergylta*. *Conservation Genetics Resources*, 1–4.
- Rousset F (2008). GENEPOP'007: A complete re-implementation of the genepop software for Windows and Linux. *Molecular Ecology Resources* **8**, 103–6.
- Skiftesvik AB, Bjelland RM, Durif CMF, Johansen IS, Browman HI (2013). Delousing of Atlantic salmon (*Salmo salar*) by cultured vs. wild ballan wrasse (*Labrus bergylta*). *Aquaculture* **402–3**, 113–8.
- Venerus LA, Ciancio JE, Riva-Rossi C, Gilbert-Horvath EA, Gosztonyi AE, Garza JC (2013). Genetic structure and different color morphotypes suggest the occurrence and bathymetric segregation of two incipient species of *Sebastes* off Argentina. *Naturwissenschaften* **100**, 645–58.
- Villegas-Ríos D (2013). Life-history and behaviour of *Labrus bergylta* in Galicia. *Department of Ecology and Marine Biodiversity*, PhD thesis, 228.
- Villegas-Ríos D, Alonso-Fernández A, Domínguez-Petit R, Saborido-Rey F (2013a). Intraspecific variability in reproductive patterns in the temperate hermaphrodite fish, *Labrus bergylta*. *Marine and Freshwater Research* **64**, 1156–68.
- Villegas-Ríos D, Alonso-Fernández A, Domínguez-Petit R, Saborido-Rey F (2014). Energy allocation and reproductive investment in a temperate protogynous hermaphrodite, the ballan wrasse *Labrus bergylta*. *Journal of Sea Research* **86**, 76–85.
- Villegas-Ríos D, Alonso-Fernández A, Fabeiro M, Bañón R, Saborido-Rey F (2013b). Demographic variation between colour patterns in a temperate protogynous hermaphrodite, the ballan wrasse *Labrus bergylta*. *PLoS ONE* **8**, e71591.
- Villegas-Ríos D, Alós J, March D, Palmer M, Mucientes G, Saborido-Rey F (2013c). Home range and diel behavior of the ballan wrasse, *Labrus bergylta*, determined by acoustic telemetry. *Journal of Sea Research* **80**, 61–70.
- Weir BS, Cockerham CC (1984). Estimating F-statistics for the analysis of population structure. *Evolution* **38**, 1358–70.
- Westgaard JI, Klemetsen A, Knudsen R (2004). Genetic differences between two sympatric morphs of Arctic charr confirmed by microsatellite DNA. *Journal of Fish Biology* **65**, 1185–91.
- Wilson GA, Rannala B (2003). Bayesian inference of recent migration rates using multilocus genotypes. *Genetics* **163**, 1177–91.

Cite this article as:

Quintela M, Danielsen E, Lopez L *et al.* (2016). Is the ballan wrasse (*Labrus bergylta*) two species? Genetic analysis reveals within-species divergence associated with plain and spotted morphotype frequencies. *Integrative Zoology* **11**, 162–72.