Genetic Diversity and Connectivity in *Maurolicus muelleri* in the Bay of Biscay Inferred from Thousands of SNP Markers

Naiara Rodriguez-Ezpeleta*, Paula Álvarez and Xabier Irigoien

1 Marine Research Division, AZTI, Sukarrieta, Spain, 2 Marine Research Division, AZTI, Pasaia, Spain

Mesopelagic fish are largely abundant poorly studied fish that are still intact, but which, due to their potentially great added value, will be imminently exploited by humans. Therefore, studies that provide information to anticipate the anthropogenic impact on this important resource are urgently needed. In particular, knowledge about their connectivity, potential adaptation and resilience are needed. This information can be obtained through the analysis of genome-wide markers which are now relatively easily and cost-efficiently discovered thanks to high-throughput sequencing technologies. Here, we have generated thousands of SNP markers in *Maurolicus muelleri*, based on the restriction-site associated DNA sequencing method, and preformed population connectivity and genetic diversity analyses in a subset of samples collected from the Bay of Biscay. Our study proves the method valid for obtaining genome-wide markers in this species and provides the first insights into the population genomics of *M. muelleri*. Importantly, the genomic resources developed here are made available for future studies and set the basics for additional endeavors on this issue.

**Keywords:** mesopelagic fish, high-throughput sequencing, RAD-seq, population connectivity, genetic diversity, *Maurolicus muelleri*

**INTRODUCTION**

The lack of physical barriers in the ocean and large populations sizes generally results in low genetic differentiation in marine fishes (Ward et al., 1994; Bradbury et al., 2008), which renders the task of inferring demographic patterns in this environment particularly difficult. Yet, over the last few years, an increasing number of studies based on high-throughput genetic data have provided evidences of fine scale population differentiation (e.g., Hess et al., 2013; Benestan et al., 2015; Rodriguez-Ezepeleta et al., 2016), challenging previous assumptions based on traditional methods (Hauser and Carvalho, 2008). In particular, the advent of the restriction site associated DNA sequencing (RAD-seq) (Baird et al., 2008) and related methods (Davey et al., 2011) has revolutionized the field of marine conservation genomics (Narum et al., 2013). This approach consists on subsampling putative homologous regions from the genome in several individuals with the aim of identifying and genotyping single nucleotide polymorphisms (SNPs). Interestingly, the method can be applied to organisms for which no prior genomic resources are available, and it is suitable to study both, neutral population structure and local adaptation (Allendorf et al., 2010).
In the context of marine management, RAD-seq is particularly relevant for poorly studied widely distributed species, as it can provide quick estimates of genetic diversity, population connectivity and adaptation more cost-effectively than relying on genome or transcriptome sequencing, or on non-sequencing based microsatellite or SNP typing.

Among the poorly studied marine organisms are mesopelagic fish, i.e., those inhabiting the deep scattering layer, that represent the largest fish biomass in the ocean (Kaartvedt et al., 2012; Irigoien et al., 2014) and play important roles in the marine ecosystem and global biogeochemical cycles (Klevjer et al., 2016). The large abundance of these fish make them particularly attractive for exploitation (St. John et al., 2016). Yet, the mesopelagic ecosystem is still largely unknown (Sutton, 2013), and the potential effects of newly introduced anthropogenic pressures in this realm should be anticipated so that sustainable management strategies for this valuable resource can be developed (St. John et al., 2016). Thus, there is an urgent need for establishing focused scientific surveys, developing appropriate sampling gear and generating additional biological data to booster knowledge in these species. In particular, developing genomic tools is foremost, as they can provide insights to understand population persistence, productivity and response to exploitation (Palumbi, 2003; Cowen et al., 2006).

Several studies have focused on deciphering species boundaries and population connectivity within mesopelagic fishes using genetic variants, but they are all based on allozymes (Gunawickrama and Naevdal, 2001; Kristoffersen and Salvanes, 2009), single mitochondrial markers (Kojima et al., 2009; Habib et al., 2012; Gordeeva, 2014) or a few microsatellites (Gordeeva, 2011; Van de Putte et al., 2012). To our knowledge, no studies based on genomic-wide SNP data have been published. Yet, simulated and empirical data based evidences support that high-throughput SNP data analyses provide more accurate population structure inferences than single or a few polymorphic marker based analyses (Haasl and Payseur, 2011; Rašić et al., 2014; Benestan et al., 2015; Szulkin et al., 2016). Here, in a first attempt to introduce genome-wide SNP discovery and genotyping to study mesopelagic fishes, we have focused on the Mueller's pearlside, *Maurolicus muelleri*.

Although initially thought to be distributed worldwide, previous reports of *M. muelleri* in the Northeast Atlantic (Gjøsætr and Kawaguchi, 1980), the South Atlantic (Hulley and Prosch, 1987), the Southeast Pacific (Robertson, 1976), and the coasts of Japan (Okiyama, 1971) are now believed to belong to different species of the genus *Maurolicus*. The genus was split into fifteen species based on morphological characters and geographic distribution (Parin and Kobyliansky, 1993), but recent studies have suggested an overestimate of species diversity in *Maurolicus* (Kim et al., 2008; Rees et al., 2017), illustrating a need for more in-depth phylogeographical analyses. In response to these needs, we have discovered and genotyped thousands of SNP markers in *M. muelleri*, which constitute the first genomic resource of a mesopelagic fish.

**MATERIALS AND METHODS**

**Tissue Sampling and DNA Extraction**

Specimens of *M. muelleri* were collected at several coastal locations of southern Bay of Biscay (Figure 1) through pelagic trawling on board the R/V Emma Bardán and Ramón FIGURE 1 | Map depicting the location where samples used for this study were obtained. Color indicates depth at which samples were caught, and boxplots depict size distribution of each catch.
Margalef in September 2016. Catches were collected between 106 and 212 m and during daytime. Fish were frozen at −20°C until DNA extraction. Genomic DNA was extracted from about 20 mg of tissue using the Wizard® Genomic DNA Purification Kit (Promega, Fitchburg, WI, United States) following manufacturer’s instructions for “Isolating Genomic DNA from Tissue Culture Cells and Animal Tissue.” Extracted DNA was suspended in Milli-Q water and concentration was determined with the Quant-iT™ dsDNA HS assay kit using a Qubit® 2.0 Fluorometer (Life Technologies). DNA integrity was assessed by electrophoresis, migrating about 100 ng of GelRed™-stained DNA on an agarose 1.0% gel.

RAD-Seq Library Preparation and Data Analysis
Restriction-site-associated DNA libraries of 94 individuals (10 to 20 per catch) were prepared following the methods of Etter et al. (2011). Briefly, starting DNA (ranging from 300 to 500 ng) was digested with the SbfI restriction enzyme and ligated to modified Illumina P1 adapters containing 5 bp unique barcodes. Pools of 30 or 32 individuals were sheared using the Covaris® M220 Focused-ultrasonicator™ Instrument (Life Technologies) and size selected to 300–500 pb by cutting agarose migrated DNA. After Illumina P2 adapter (including 5 pb index) ligation, each library was amplified using 14 PCR cycles. The three pools, each with one unique Illumina index, were combined and paired-end sequenced (100 pb) on an Illumina HiSeq2000.

RESULTS
RAD-Seq Genotyping
The number of RAD-seq reads passing quality filters per individual included in the final analyses ranges from 326 824 to 2 116 073 with an average of 974 008. The number of RAD-tags catalog were searched using sstacks. RAD loci found in at least 75% of the individuals were selected using populations and used to calculate (nucleotide diversity − \( \pi \), minor allele frequency − MAF, expected heterozygosity − \( H_e \), expected homozygosity − \( H_o \), and inbreeding coefficient − \( F_{IS} \)). Using PLINK version 1.07 (Purcell et al., 2007), SNPs with MAF larger than 0.05 and a genotyping rate larger than 0.9 were selected for population structure analyses. The obtained genotype dataset was exported to Structure and Genepop formats using PGDSPider version 2.0.8.3 (Lischer and Excoffier, 2012).

Genetic Diversity and Population Structure
\( F_{ST} \) values per group pairs were calculated following the Weir and Cockerham (1984) formulation as implemented in Genepop 4.3 (Rousset, 2008). Principal component analyses (PCA) were performed with the R package adegenet (Jombart and Ahmed, 2011) without any a priori grouping assumption. The percentage of appurtenance of each individual to each of the K (ranging from 1 to 4) hypothetical ancestral was calculated with the Bayesian clustering approach implemented in STRUCTURE (Pritchard et al., 2000) without any prior population assignment, based on the admixture model and a burn-in period of 100,000 iterations followed by 300,000 iterations from which estimates were obtained. Ten replicates for each K-value were performed and analyzed with CLUMPP (Jakobsson and Rosenberg, 2007) to identify common modes. Results were plotted using DISTRUCT (Rosenberg, 2004), and best K was identified according to the Evanno method (Evanno et al., 2005) as implemented in StructureHarvester (Earl and vonHoldt, 2012).

FIGURE 2 | Boxplot depicting first and third quartile, median and standard deviation of number of RAD-tags per individual (A) and relationship among number of RAD-tags and used reads (B) when the M parameter is set to 2 (red), 4 (green) or 8 (blue).
TABLE 1 | Number of tags present in at least 75% of the individuals, as well as number of positions remaining after each filtering step for each value of \( M \) used.

| \( M = 2 \) | \( M = 4 \) | \( M = 6 \) |
|---|---|---|
| Number of tags present in >75% of individuals | 22 624 | 24 179 | 24 504 |
| Number of positions | 2 031 164 | 2 166 128 | 2 190 124 |
| Number of variable positions (SNPs) | 132 782 | 180 398 | 207 080 |
| Number of SNPs with MAF > 0.05 | 13 852 | 19 925 | 22 331 |
| Number of SNPs genotyped in >90% of individuals | 1 625 | 2 350 | 2 409 |

obtained per individual ranges from about 30K to about 110K, with an average of 67 592, 66 359, and 65 852 for \( M \)-values of 2, 4, and 6, respectively (Figure 2A). Average coverage per individual ranges from 6 to 20\( \times \) (with an average of 11\( \times \)). Number of RAD-tags per individual increases with number of reads used for stack formation and does not reach a maximum value even when more than 100K reads are used (Figure 2B). This places the number of \( SbfI \) cut sites in *M. muelleri* above 32 000, which is larger than in other teleost species (Catchen et al., 2013a; Diaz-Arce et al., 2016; Rodríguez-Ezpeleta et al., 2016).

Above twenty thousand RAD-tags are present in at least 75% of the individuals, and comprise more than two million sites of

TABLE 2 | Average nucleotide diversity (\( \pi \)), inbreeding coefficient (\( F_{IS} \)), and observed (\( H_o \)) and expected (\( H_e \)) heterozygosity for each value of \( M \) when considering all positions included in tags present in at least 75% of the individuals (All), only variable positions within these tags (Variable) or only selected SNPs (Selected).

| \( M \) | \( \pi \) | \( F_{IS} \) | \( H_o \) | \( H_e \) |
|---|---|---|---|---|
| All | 0.0030 | 0.0105 | 0.0023 | 0.0030 |
| Variable | 0.0465 | 0.1613 | 0.0350 | 0.0462 |
| Selected | 0.1981 | 0.1458 | 0.1652 | 0.1880 |
| All | 0.0040 | 0.0135 | 0.0030 | 0.0040 |
| Variable | 0.0487 | 0.1632 | 0.0364 | 0.0484 |
| Selected | 0.1936 | 0.1537 | 0.1651 | 0.1925 |
| All | 0.0045 | 0.0163 | 0.0034 | 0.0045 |
| Variable | 0.0482 | 0.1741 | 0.0359 | 0.0479 |
| Selected | 0.1911 | 0.1543 | 0.1616 | 0.1900 |

FIGURE 3 | Radar plots showing average nucleotide diversity (\( \pi \)), inbreeding coefficient (\( F_{IS} \)), and expected (\( H_e \)) and observed (\( H_o \)) heterozygosity for each catch (colored lines) considering each value of \( M \) (2, 4, or 6) and for either, all positions included in tags present in at least 75% of the individuals (All), only variable positions within these tags (Variable) or only selected SNPs (Selected).
which more than a thousand are variable (Table 1). The number of tags present in more than 75% of the individuals increases with increasing M because more common loci can be found when these are composed by more alleles. Consequently, larger M-values results in more alleles and thus more variable positions (Catchen et al., 2013b; Rodriguez-Ezpeleta et al., 2016). Selecting only those positions with MAF larger than 0.05 and applying a stringent filtering of a minimum of 90% of the individuals being genotyped per SNP, results in three genotype datasets of 1 625, 2 350, and 2 409 SNPs for M-values 2, 4, and 6, respectively.

Genetic Diversity
The number of polymorphic sites contained in the tags present in at least 75% off the individuals is 6.5, 8.3, and 9.4% for M-values of 2, 4, and 6, respectively. When calculated considering all individuals a single group, overall nucleotide diversity (π), inbreeding coefficient (FIS), and expected (H_e) and observed heterozygosity (H_o) values are congruent across the three parameter sets used. Expectedly, using only variable sites or selecting those with at least 0.05 MAF increases nucleotide diversity and both, expected and observed heterozygosity (Table 2). For all set of parameters and position subsets used, observed heterozygosity is lower than expected both overall and when calculated per catch. Per catch, values of π and H_e are similar, whereas values of H_o and FIS differ (Figure 3). In general, catches with low observed heterozygosity are also those with highest inbreeding coefficient (9018, 9204, 9201) and vice versa (hauls 9007, 9219, 9234), although this does not hold for catches 9012 and 9029 with present large H_o, but medium FIS values with respect to the others. Similar to other studies (Rodriguez-Ezpeleta et al., 2016), we find that, for the four variables, absolute values depend on the set of parameters used for SNP discovery, but that relative values among groups are maintained.

Population Structure
In the Bayesian population structure analyses (Figure 4), all individuals display admixed representation of each of the hypothetical ancestral population, suggesting genetic connectivity within the area of study. Interestingly, catch 9018 shows an ancestry pattern that differs from the rest of the catches, yet, this difference is not visible in the PCA plots (Figure 5), where no differentiation among groups can be observed. In accordance with the hypothesis of high connectivity within the area of study, FST values between pairs of populations are low (Table 3), although, consistent with the pattern observed in the structure analyses, pairs including catch 9018 are those with highest FST values.

DISCUSSION
Restriction-site associated DNA sequencing constitutes an unprecedented opportunity for performing demographic inferences in species for which no prior genetic resources are available (Corander et al., 2013; Hess et al., 2013). Here, we have discovered and genotyped thousands of RAD-seq derived SNP markers in M. muelleri in a first attempt to use genome wide data to study diversity and connectivity in this mesopelagic species, which is particularly relevant in view of the imminent exploitation of this till now pristine marine resource.

We have demonstrated that the Sbf I restriction enzyme is as a good candidate for RAD-sequencing based SNP discovery in M. muelleri. Interestingly, we found that, although the average number of RAD-tags per individual is similar to that found in other species (Amores et al., 2011; Catchen et al., 2013b; Puebla et al., 2014; Diaz-Arce et al., 2016; Rodriguez-Ezpeleta et al., 2016), the number of RAD-tags obtained per individual increases with the number of sequencing reads produced; this suggests that most likely the genome of M. muelleri could have as many as 50 000 SbfI cut sites, which is higher than most fish species studied so far. The presence of a large number of cut sites, and therefore of loci, could suggest lower coverage per locus. Yet, although lower coverage than in other species is
obtained, all individuals had more than $6\times$ coverage and average was of $11\times$. Multiplexing less individuals per lane or sequencing the same pool in more than one lane would increase coverage accordingly. Yet, the large number of loci shared among 75% of the individuals and the above one thousand SNP markers obtained after filtering suggest that this increased sequencing would not be necessary.

In our analyses, observed heterozygosity is lower than expected. This could be indicator of (i) population differentiation within the area of study, referred to as the Wahlund effect (Wahlund, 1928), (ii) preferential mating with close relatives or inbreeding (Wright, 1921) or (iii) non-random sampling of members from a limited number of families (Robertson, 1965). Due to the large biomass of mesopelagic fish (Irigoien et al., 2014), it would seem unlikely that our samples, composed by a few individuals per location, contain relatives. Similarly, preferential mating with close relatives or inbreeding would be difficult also to explain. Thus, the natural explanation for the low observed heterozygosity obtained would be population differentiation within the area of study. Yet, we do not find clear evidences of population stratification except for catch 9018, which appears genetically differentiated in the STRUCTURE plots, but not in the PCA.

The fact that when separating individuals into catches, observed heterozygosity is also lower than expected, suggests that one of the three possible explanations should be also acting at the catch level. In this sense, the possibilities of inbreeding and non-random sampling members of the same family should also be considered. Indeed, the fact that size distributions are quite homogeneous within catches (Figure 1), supports previous observations that individuals are grouped according to their age (Staby et al., 2011), and claims for more studies linking the reproductive behavior of the species and the genetic diversity results obtained here, which should be further confirmed with additional data. Similarly, no explanation to the different genetic diversity estimates obtained per catch could be found. These differences could be simply due to a low sample size per catch or have a more biologically meaningful information, and thus, more analyses are needed to confirm either one.

Overall, we have validated the RAD-seq based SNP discovery method using the SbfI restriction enzyme for *M. muelleri*, which has resulted in the first genomic resources for this species that are moreover made available for future studies (raw sequences are available at NCBI SRA Bioproject PRJNA417219). This data will contribute to shed light on phylogeographic patterns within the species (Rodríguez-Ezpeleta et al., 2016), detect genetic

### TABLE 3 | $F_{ST}$ values per population pair.

|     | 9201 | 9201 | 9204 | 9007 | 9012 | 9219 | 9018 | 9020 |
|-----|------|------|------|------|------|------|------|------|
| 9204| 0.000|      |      |      |      |      |      |      |
| 9007| 0.000| 0.003|      |      |      |      |      |      |
| 9012| 0.000| 0.006| 0.000|      |      |      |      |      |
| 9219| 0.004| 0.011| 0.008| 0.003|      |      |      |      |
| 9018| 0.002| 0.008| 0.009| 0.010| 0.005|      |      |      |
| 9020| 0.001| 0.006| 0.006| 0.004| 0.007| 0.003|      |      |
| 9234| 0.002| 0.004| 0.000| 0.001| 0.002| 0.004| 0.003|      |

The three $M$ parameter values tested provided congruent outcomes; here, only those values corresponding to $M = 2$ are shown.
adaptation (Van Wyngaarden et al., 2017), and to resolve the phylogeny within the genus for species delimitation (Diaz-Arce et al., 2016).

AUTHOR CONTRIBUTIONS

NR-E designed the study, performed the analyses, interpreted the data, and wrote the manuscript. PA collected the samples, contributed to the interpretation of the data and revised the manuscript. XI contributed to the interpretation of the data and revised the manuscript.

REFERENCES

Allendorf, F. W., Hohenlohe, P. A., and Luikart, G. (2010). Genomics and the future of conservation genetics. Nat. Rev. Genet. 11, 697–709. doi: 10.1038/nrg2844

Amores, A., Catchen, J., Ferrara, A., Fontenot, Q., and Postlethwait, J. H. (2011). Genome evolution and meiotic maps by massively parallel DNA sequencing: spotted gar, an outgroup for the teleost genome duplication. Genetics 188, 799–808. doi: 10.1534/genetics.111.127324

Baird, N. A., Etter, P. D., Atwood, T. S., Currey, M. C., Shiver, A. L., Lewis, Z. A., et al. (2008). Rapid SNP discovery and genetic mapping using sequenced RAD markers. PLoS ONE 3:e3376. doi: 10.1371/journal.pone.0003376

Benestan, L., Gosselin, T., Perrier, C., Sainte-Marie, B., Rochette, R., and Bernatchez, L. (2015). RAD genotyping reveals fine-scale genetic structuring and provides powerful population assignment in a widely distributed marine species, the American lobster (Homarus americanus). Mol. Ecol. 24, 3299–3315. doi: 10.1111/mec.13245

Bradbury, I. R., Laurel, B., Snelgrove, P. V., Bentzen, P., and Campana, S. E. (2008). Global patterns in marine dispersal estimates: the influence of geography, taxonomic category and life history. Proc. Biol. Sci. 275, 1803–1809. doi: 10.1098/rspb.2008.0216

Catchen, J., Basmash, S., Wison, T., Currey, M., O’Brien, C., Yeates, Q., et al. (2013a). The population structure and recent colonization history of Oregon threespine stickleback determined using restriction-site associated DNA-sequencing. Mol. Ecol. 22, 2864–2883. doi: 10.1111/mec.12330

Catchen, J., Hohenlohe, P. A., Basmash, S., Amores, A., and Cresko, W. A. (2013b). Stacks: an analysis tool set for population genomics. Mol. Ecol. 22, 3124–3140. doi: 10.1111/mec.12354

Corander, J., Majander, K. K., Cheng, L., and Merila, J. (2013). High degree of cryptic population differentiation in the Baltic Sea herring Clupea harengus. Mol. Ecol. 22, 2931–2940. doi: 10.1111/mec.12174

Cowen, R. K., Paris, C. B., and Srinivasan, A. (2006). Scaling of connectivity in marine populations. Science 311, 522–527. doi: 10.1126/science.1122039

Davey, J. W., Hohenlohe, P. A., Etter, P. D., Boone, J. Q., Catchen, J. M., and Blaxter, M. L. (2011). Genome-wide genetic marker discovery and genotyping using next-generation sequencing. Nat. Rev. Genet. 12, 499–510. doi: 10.1038/nrg3012

Diaz-Arce, N., Arrizabalaga, H., Murua, H., Irigoien, X., and Rodriguez-Ezepeleta, N. (2016). RAD-seq derived genome-wide nuclear markers resolve the phylogeny of tunas. Mol. Phylogenet. Evol. 102, 202–207. doi: 10.1016/j.ympev.2016.06.002

Earl, D. A., and vonHoldt, B. M. (2012). STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conserv. Genet. Resour. 4, 359–361. doi: 10.1007/s12686-011-9548-7

Etter, P. D., Basmash, S., Hohenlohe, P. A., Johnson, E. A., and Cresko, W. A. (2011). SNP discovery and genotyping for evolutionary genetics using RAD sequencing. Methods Mol. Biol. 772, 157–178. doi: 10.1007/978-1-61779-228-1_9

Evanno, G., Regnaut, S., and Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol. Ecol. 14, 2611–2620. doi: 10.1111/j.1365-294X.2005.02553.x

Gosset, J., and Kawaguchi, K. (1980). A Review of the World Resources of Mesopelagic Fish. Rome: Food and Agriculture Organization, 1–151.

FUNDING

This project was supported by the Department of Agriculture, Fisheries and Food of the Basque Country.

ACKNOWLEDGMENTS

The authors wish to thank Naroa Aldanondo and Carlota Pérez for fish sampling and Itxaki Mendibil for preparing the RAD-seq libraries. This manuscript is contribution 833 from the Marine Research Division of AZTI.
Kristoffersen, J. B., and Salvanes, A. G. V. (2009). Distribution, growth, and population genetics of the glacier lanternfish (Benthosema glaciale) in the ocean. *Mar. Biol. Res.* 5, 596–604. doi: 10.1007/s12227-009-9247-y

Lischer, H. E., and Excoffier, L. (2012). PGDSpider: an automated data conversion tool for connecting population genetics and genomics programs. *Bioinformatics* 28, 298–299. doi: 10.1093/bioinformatics/bts462

Narum, S. R., Rueda, C. A., Davey, J. W., Miller, M. R., and Hohenlohe, P. A. (2013). Genotyping-by-sequencing in ecological and conservation genomics. *Mol. Ecol.* 22, 2841–2847. doi: 10.1111/mec.12350

Okiyama, M. (1971). Early life history of the gonostomatid fish, *Maurolicus muelleri* (Gmelin), in the Japan Sea. *Bull. Japan Sea Reg. Fish. Res. Lab.* 23, 21–53.

Palumbi, S. (2003). Population genetics, demographic connectivity, and the design of marine reserves. *Ecol. Appl.* 13, 146–158. doi: 10.1890/1051-0761(2003)013[0146:PGDCDT]2.0.CO2

Parin, V. N., and Kobylansky, S. G. (1993). Review for the genus *Maurolicus* (Stomopterygidae, Stomiiformes) with re-establishing validity of five species considered junior synonyms of *M. Mueller* and descriptions of nine species. *Biol. Ocean Fish. Squids* 128, 69–101.

Pritchard, J. K., Stephens, M., and Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics* 155, 945–959.

Puebla, O., Beringham, E., and McMullan, W. O. (2014). Genomic atlases of differentiation in coral reef fishes (*Acanthurus* spp. Serranidae). *Mol. Ecol.* 23, 5279–5303. doi: 10.1111/mec.12926

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A., Bender, D., et al. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* 81, 559–575. doi: 10.1086/519795

Rašić, G., Filipović, I., Weeks, A. R., and Hoffmann, A. A. (2014). Genome-wide SNP data lead to strong signals of geographic structure and relatedness patterns in the major arbovirus vector, *Aedes aegypti*. *BMC Genomics* 15:275. doi: 10.1186/1471-2164-15-275

Rees, D. J., Byrne, K., Thomas, L., Ferreira, M. A., Bender, D., et al. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* 81, 559–575. doi: 10.1086/519795

Rosenberg, N. A. (2004). *DISTRUCT*: a program for the graphical display of population structure. *Mol. Ecol. Notes* 4, 137–138. doi: 10.1046/j.1471-8286.2003.00566.x

Rousset, F. (2008). Genepop’007: a complete reimplementation of the Genepop software for Windows and Linux. *Mol. Ecol. Resour.* 8, 103–106. doi: 10.1111/j.1471-8286.2007.01931.x

St. John, M. A., Borja, A., Chust, G., Heath, M., Grigorov, I., Mariani, P., et al. (2016). A dark hole in our understanding of marine ecosystems and their services: perspectives from the mesopelagic community. *Front. Mar. Sci.* 3:31. doi: 10.3389/fmars.2016.00031

Staby, A., Rostad, A., and Kaartvedt, S. (2011). Vertical and horizontal movements and migration patterns. *Mar. Ecol. Prog. Ser.* 7, 298–309. doi: 10.3354/meps09363

Sutton, T. (2013). Vertical ecology of the pelagic ocean: classical patterns and new perspectives. *J. Fish Biol.* 83, 1508–1527. doi: 10.1111/jfb.12263

Szulkin, M., Gagnaire, P. A., Bierne, N., and Charmantier, A. (2016). Population genomic footprints of fine-scale differentiation between habitats in Mediterranean blue tits. *Mol. Ecol.* 25, 542–558. doi: 10.1111/mec.13486

Van de Putte, A. P., Van Houdt, J. K. J., Maes, G. E., Hellemans, B., Collins, M. A., and Volckaer, F. A. M. (2012). High genetic diversity and connectivity in a single species of marine fish from the Southern Ocean: the myctophid *Electrona antarctica*. *Deep Sea Res. II Top. Stud. Oceanogr.* 59(Suppl. C), 199–207. doi: 10.1016/j.dsr2.2011.05.011

Van Wyngaarden, M., Snellgrove, P. V., DiBacco, C., Hamilton, L. C., Rodríguez-Ezepeleta, N., Jeffery, N. W., et al. (2017). Identifying patterns of dispersal, connectivity and selection in the sea scallop, *Placopecten magellanicus*, using RADseq-derived SNPs. *Evol. Appl.* 10, 102–117. doi: 10.1111/evo.12432

Wahlund, S. (1928). Zusammensetzung von populationen und korrelationserscheinungen vom standpunkt der vererbungslehre ausbetrachtet. *Hereditas* 11, 65–106. doi: 10.1111/j.1601-5223.1928.tb02483.x

Ward, R. D., Woodmark, M., and Skibinski, D. (1994). A comparison of genetic diversity levels in marine, freshwater and anadromous fishes. *J. Fish Biol.* 44, 213–232. doi: 10.1111/j.1095-8649.1994.tb01200.x

Weir, B. S., and Cockerman, C. C. (1984). Estimating F-statistics for the analysis of population structure. *Evolution* 38, 1358–1370.

Wright, S. (1921). Systems of mating II. The effects of inbreeding on the genetic composition of a population. *Genetics* 6, 124–143.

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2017 Rodríguez-Ezepeleta, Álvarez and Irigoien. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.