Divergence by Depth in an Oceanic Fish

Peter Shum[1], Christophe Pampoulie[2], Carlotta Sacchi[3], and Stefano Mariani[1*]

[1]School of Environment & Life Sciences, University of Salford, Manchester, UK

[2]Marine Research Institute, Reykjavík, Iceland

[3]School of Biology & Environmental Science, University College Dublin, Dublin, Ireland

*Corresponding Author: s.mariani@salford.ac.uk

Introduction

Speciation phenomena in taxa diverging with gene flow, in the absence of obvious geographic barriers, remain a central focus in evolutionary biology (Bird et al., 2012; Smadja & Butlin, 2011; Fitzpatrick, Fordyce & Gavrilets, 2008). The classical model of allopatric speciation involves the evolution of reproductive isolation as a result of physical barriers that block gene flow in spatially separated populations; whereas populations diverging in sympatry lead to the formation of species from a single panmictic population which must exhibit strong divergent selection in order to overcome the homogenizing effects of gene flow (Gavrilets, 2003). Parapatric speciation represents an intermediate scenario of species formation whereby partial yet restricted contact zones exist between two populations with limited gene exchange (Gavrilets, 2003). Although exhaustive frameworks exist to identify and interpret speciation dynamics (Rettelbach et al., 2013; Bird et al., 2012; Smadja & Butlin, 2011), only recently has empirical attention been directed to the role of depth gradients in aquatic (Vonlathen et al., 2009) and especially oceanic biota (Roy, Hurlbut & Ruzzante, 2012; Bird et al., 2011; Brokovich et al., 2010; Hyde et al., 2008).

22 The strong physical gradients across depth layers in the ocean pose strong selective pressures in

23 aquatic organisms (Ingram, 2011; Somero, 1992), and one notable factor is the change in the light

24 environment (Warrant & Locket, 2004), which affects vision. Visual sensitivity in marine

25 vertebrates depends on the spectral tuning mechanism of the visual pigment (VP) (Yokoyama,

26 2000), which consists of an opsin protein (part of the largest family of G-protein-coupled

27 receptors) bound to a light-sensitive chromophore. Differently-charged amino acid (AA) residues

28 in the opsin will result in slightly different light absorbance by the photoreceptor cells

29 (Yokoyama, 2002).

30 The percomorph marine family of rockfishes (Sebastidae) have played a central role in the

31 understanding of depth-associated population divergence and speciation in the ocean (Ingram,

32 2011; Stefánsson et al., 2009a; Hyde et al., 2008; Alesandrini & Bernardi, 1999), and evidence

33 exists that the rhodopsin gene may have evolved in response to different light environments in the

34 main, ancient radiation of the genus, in the Pacific (Sivasundar & Palumbi, 2010). North-Atlantic

35 *Sebastes* have a much more recent history (Hyde & Vetter, 2007), with the four recognised extant

36 species having diversified during the Pleistocene (Bunke, Hanel, & Trautner, 2012). In particular,

37 the beaked redfish, *Sebastes mentella*, consists of two genetically distinguishable groups

38 (Stefánsson et al., 2009a; Pampoulie & Daníelsdóttir, 2008): a widely-distributed shallow-pelagic

39 (SP) form, found between 250 and 550m depth and a more circumscribed deep-pelagic (DP)

40 component, between 550 and 800m. However, doubts remain as to the forces at play and the time

41 scales associated with this divergence (Stefánsson et al., 2009b; Cadrin et al., 2010).

42 Here, we sought to investigate whether on-going processes of adaptation to different depth layers

43 may leave a signature of disruptive selection in the rhodopsin gene in a recently diversifying

44 *Sebastes* species. We also employed for the first time the mitochondrial DNA control region to

45 reconstruct historical demography and to further elucidate the evolutionary relationships between

46 'shallow' and 'deep' pelagic beaked redfish.

Material and Methods

47

Generation of molecular data

48

49 Archive samples were randomly selected from 25 shallow-pelagic (SP) (collected above 400m depth) and 25 deep-

50 pelagic (DP) (collected below 700m) *Sebastes mentella* from the Irminger Sea, south-west of Iceland previously

51 genotyped by Stéfansson et al. (2009b) (sample numbers 4&5 in the original article). DNA was isolated from gill

52 tissue that had been preserved in 96% EtOH using a modified salt extraction protocol (Miller et al., 1988) or the

53 DNeasy kit (Qiagen ©) following the manufacturer's protocol. The non-coding mitochondrial control region was

54 amplified by PCR using primers developed by Hyde & Vetter (2007); D-RF: 5′-CCT GAA AAT AGG AAC CAA

55 ATG CCA G-3′ and Thr-RF: 5′-GAG GAY AAA GCA CTT GAA TGA GC-3′. The primers by Chen et al. (2003);

56 Rh193: 5′-CNT ATG AAT AYC CTC AGT ACT ACC-3′ and Rh1039r: 5′-TGC TTG TTC ATG CAG ATG TAG A-3′

57 were used to amplify 744 bp of the intron-free rhodopsin gene in 10 shallow-pelagic *S. mentella*, 12 deep-pelagic *S.*

58 *mentella*, and 3 and 4 individuals of *S. marinus* and *S. viviparus* as outgroups. Reactions were carried out in 25μl

59 volumes containing 1x PCR buffer, 1 mM MgC12, 200μM dNTPs, 0.4 μM of each primer, 0.2 units *Taq* DNA

60 polymerase, with 1 μl of DNA template for mtDNA (4 μl for rhodopsin). Amplifications were performed in a

61 Biometra T3000 Thermocycler using the following temperature profiles: control region: 94ºC (2 min), 35 cycles of

62 [94ºC (30 sec), 59ºC (60 sec), 72 ºC (60 sec)], followed by 3 minutes at 72 ºC; for rhodopsin: 95 ºC (5 min), 37

63 cycles of [94 ºC (20 sec), 58 ºC (30 sec), 72 ºC (45 sec)], followed by 5 minutes at 72 ºC. A negative control was

64 included in all reactions. PCR products were subjected to electrophoresis through a 1% agarose gel 1X Tris–Borate–

65 EDTA Buffer, stained with SYBR green for visualisation via a UV-transilluminator and then purified through the

66 addition of exonuclease I and shrimp alkaline phosphatase to remove unincorporated primers and deoxynucleotides

67 in preparation to sequencing. Purified products were sequenced by Macrogen (Macrogen, Amsterdam, Europe;

68 http://dna.macrogen.com/eng/).

69 Data Analysis
70 *Genetic diversity and population differentiation*
71 The mtDNA control region was examined for nucleotide and haplotype diversity. This included the number of net

72 nucleotide substitutions per site between populations (D_a) which was calculated using DnaSP v5.10 (Librado &

73 Rozas, 2009). We estimated the level of genetic variation between populations calculating pairwise population F_{ST}

74   and Φst values performed in ARLEQUIN v3.5.1.2 (Excoffier & Lischer, 2010) with significant values tested by

75   5000 permutations.

76   Mismatch analysis was performed to examine the demographic history between the shallow-pelagic and deep-pelagic

77   *S. mentella* populations using ARLEQUIN and distributions were compared with a two-sample Kolmogorov-

78   Smirnov (K-S) test. For populations at stationary demographic equilibrium, theoretical and empirical studies show

79   that the mismatch distributions usually have multimodal, ragged or erratic distributions, and are typically smoother

80   or unimodal for populations that have undergone a recent expansion (Rogers & Harpending, 1992). To test the

81   goodness-of-fit distribution, we calculated the sum of squared deviations (SSD) and raggedness index (r) for a

82   stepwise expansion model for the data tested by Monte Carlo Markov Chain simulations (1000 steps) in

83   ARLEQUIN.

84   Haplotype genealogies for the *S. mentella* data set were constructed following a method described by Salzburger et

85   al. (2011) based on a maximum likelihood tree for mtDNA and rhodopsin genes sequences.

86   Data from a selected suite of 12 microsatellite loci previously used for genotyping by Stéfansson et al. (2009b) were

87   used to calculate pairwise genetic differentiation (Weir & Cockerham's F_{ST}, Hedricks G'_{ST} & Jost's D_{est}) between

88   populations with 9999 permutations carried out to obtain significance levels using GenAIEx 6.501 software (Peakall

89   & Smouse, 2006). Population structure was visualized by correspondence analysis (CA) using GENETIX 4.05

90   (Belkhir et al. 1996).

91   *Phylogenetic analysis and test for positive selection*

92   Maximum-likelihood (ML) analyses of the rhodopsin gene sequences were performed using PhyML 3.0 (Guindon &

93   Gascuel, 2003) under 1,000 replications; using Modeltest3.7 (Posada & Crandall, 1998), the model that best fit the

94   data was found to be F81+I (pinvar=0.9770). Trees generated from these results were used for a test for positive

95   selection at the rhodopsin gene, conducted using the Creevey-McInerney method (Creevey & McInerney, 2002)

96   implemented in CRANN (Creevey & McInerney, 2003). This test is a more sensitive tree-based analysis derived

97   from the relative ratio test (McDonald & Kreitmann, 1991). Given an appropriate rooted tree, the number of

98   synonymous and non-synonymous substitutions are calculated along each internal branch using the reconstructed

99   ancestral sequences. The method uses statistical tests for independence (χ^2 G-test or Fisher's exact test) to evaluate

100   whether the ratio between synonymous (Silent Invariable (SI) to Silent Variable (SV)) and non-synonymous

101   (Replacement Invariable (RI) to Replacement Variable (RV)) substitutions deviate from the expected value under the

102   neutral model. Where the G-test fails to produce a result, the Fisher's test is used and vice versa. Positive directional

103   selection is expected if there is a significantly higher number of RI substitutions or non-directional selection if

104   RV>RI. The test was performed using *S. alutus* as outgroup (Fig. 3) as it represents the closest common ancestor for

105   the Atlantic *Sebastes* spp. (Hyde & Vetter, 2007).

106   **Results and Discussion**

107   We recovered 16 mtDNA haplotypes, defined by 15 total variable sites across the two groups.

108   Haplotypes were almost completely segregated between depth layers, with eight haplotypes

109   found in the deep, never recovered in the shallow area, and resembling a starburst pattern, three

110   mutational steps away from the rest of the network (Fig. 1a). Of the 10 haplotypes found in the

111   shallow, eight were exclusive of this habitat, and only two individuals collected in the deep were

112   found to bear a 'shallow-type' sequence. The shallow-pelagic (SP) group exhibited much greater

113   diversity ($\hat{h}$=0.887±0.033, π=0.00504±0.00082) than the deep-pelagic (DP) group

114   ($\hat{h}$=0.543±0.119, π=0.00238±0.00071) (Table 1).

115   Partitioning of genetic variance between the populations showed highly significant and strong

116   population structure (Table. 2). The mismatch distributions of the two groups (Fig. 1b) were

117   significantly different (K-S test: D_{300}=0.3, p<<0.001), and confirmed what is visually apparent

118   from the haplotype network: a scenario of more recent and pronounced demographic and spatial

119   expansion in the deep-pelagic group compared to the shallow-pelagic (Table 3). Using Nei's (Nei,

120   1987) formula for divergence time: $T=D_a/2\mu$, where 2μ represents a general mtDNA evolutionary

121   rate, commonly assumed to be around 11% per million years for fish mtDNA control region

122 (Patarnello, Volckaert & Castilho, 2007), we find that the "deep" and "shallow" lineages split

123 over 44,000 years ago.

124 The rhodopsin gene was tested for signatures of positive selection using the Creevey-McInerney

125 method rooting the tree with *S. alutus* as an outgroup (Fig. 2). Values for the four substitution

126 variables, G-test, and *p*-values along each branch are presented in Table 4. Two branches

127 (numbers 26 and 27) showed significance at the *p*=0.05 level (Fig. 2). Branches 26 and 27 show

128 significant RI to RV deviations from neutrality, due to non-synonymous substitutions (Table 3),

129 suggesting that positive disruptive selection is acting on the rhodopsin gene for the clade and on

130 the internal branch leading to the shallow group. We observed a fixed non-synonymous AA

131 substitution within the transmembrane domain, which strongly discriminates the two groups

132 inhabiting shallow and deep environments (Fig. 3). The shallow-pelagic group exhibits a **G**TC at

133 position 119, which codes for Valine (L119V), while the deep-pelagic type displays an **A**TC,

134 coding for Isoleucine (L119I). Amino acid changes located in the transmembrane helical regions

135 have been known to be important for spectral tuning (Yokoyama et al., 2008), yet *in-vitro*

136 experimental spectral analyses of vertebrate rhodopsins suggest that amino acid substitutions at

137 site 119 have negligible effect on absorption spectra (Yokoyama et al., 2008). Nevertheless,

138 substitutions within the transmembrane protein domain III (helix-III), such as site 119, have been

139 shown to affect the decay rate of metarhodopsin II ("meta II"; Ou et al., 2011), which is an

140 intermediate of rhodopsin that binds and activates transducin, the visual G-protein (Smith, 2010).

141 Ou et al. (2011) have shown that an AA replacement L119C against the wild-type rhodopsin

142 resulted in shorter meta II lifetimes, suggesting more responsive structural alterations at the helix'

143 G-protein binding site.

144 Neither spectral nor conformational analyses have so far been conducted on shallow-pelagic and

145 deep-pelagic *Sebastes mentella*, but the AA variation observed here could underlie differential

146   hydrophobic activity and photoisomerization sensitivity (Ou et al., 2011) that could hold some

147   adaptive value, even without net change in wavelength of maximal absorption.

148   Sivasundar and Palumbi (2010) discovered a striking association between AA replacements along

149   the rhodopsin gene and inferred depth preference in many North Pacific *Sebastes*. Interestingly,

150   four North-Pacific *Sebastes* (*S. chlorostictus*, *S. elongates*, *S. aurora* and *S. melanostomus*)

151   typically associated with deeper waters were observed to exhibit the same AA replacement L119I

152   as detected in the deep-sea *S. mentella*. Similarly, one Pacific species (*S. diploproa*) exhibits the

153   AA replacement L119V, which is linked to a shift back to shallower waters, and mirrors the

154   polymorphism in the shallow-pelagic *S. mentella*.

155   Although larger sample sizes will be required in the future to test these patterns more robustly,

156   the implications of these findings are twofold, and have powerful resonance for both marine

157   evolution and fisheries management. First, mitochondrial variation between 'shallow' and 'deep'

158   *S. mentella* in the North Atlantic unveil a degree of historical divergence that previously

159   employed genetic markers either failed to detect (Bunke et al., 2012) or could not reliably frame

160   in a phylogeographic context (Stefánsson et al., 2009b).

161   The level of differentiation and haplotype sorting is such that evolutionary independence can be

162   broadly upheld for these two habitat-segregated lineages, and re-analysis of microsatellite data

163   confirm this picture (Fig. 4). In particular, the comparison of frequency-based indicators of

164   substructure (F_{ST} for mtDNA, and G'_{ST} and D_{est} to account for the hypervariability of

165   microsatellite loci) reveal values (Table 2) that match theoretical expectations under neutral

166   divergence, taking into account the four-fold strength of genetic drift at mitochondrial markers.

167   Interestingly, we also noticed two individuals with a 'shallow-type' mtDNA haplotype, which

168   were caught in the deep layer. One of these two (DP29, Fig. 2) also screened at the rhodopsin

169   locus, exhibits a sequence typical of the shallow layer, and its multilocus microsatellite genotype

170   also falls with the shallow group (Fig. 4), which can be interpreted as the occurrence of

171   individual movements along the water column during the life cycle (i.e. short-term "dives" into

172   the deep, by shallow-dwelling fish). Another deep-caught individual (DP1) also exhibits a

173   "shallow" haplotype, but a "deep-like" ATC rhodopsin sequence and an inconclusive multilocus

174   microsatellite genotype (Fig. 4). Collectively, this likely reflects the occurrence of introgressive

175   hybridisation between the two groups, as previously suggested by Pampoulie & Danielsdottir

176   (2008).

177   Furthermore, the stark pattern of depth-associated divergence at the rhodopsin gene is perhaps

178   even more surprising, were it not for the fact that comparable evolutionary genetic patterns have

179   recently been credited with a key role in the diversification of the more ancient Pacific *Sebastes*

180   group (Sivasundar & Palumbi, 2010). It has been hypothesized that fast-evolving markers will

181   allow to determine recent speciation events for closely related *Sebastes* spp. (Alesandrini &

182   Bernardi, 1999; Cadrin et al., 2010). The present data provide a snapshot of the evolutionary

183   mechanisms that may be at play in the young, species-poor, Atlantic *Sebastes* lineage, during its

184   initial phase of adaptive radiation, underpinned by positive selection at the rhodopsin gene.

185   Less than a decade ago, *S. mentella* was assumed to be panmictic in the North Atlantic, and the

186   rapidly increasing fishery pressure on these stocks did not recognise any possible substructure

187   until 2009 (Cadrin et al., 2010). These latest results dismiss the notion of panmixia in this oceanic

188   species, and, perhaps more intriguingly, open the possibility that the two 'shallow' and 'deep'

189   groups may represent two lineages experiencing adaptation towards divergent environmental

190   conditions. In the near future, it should be experimentally evaluated whether the amino acid

191   replacements at the 119 position actually produce detectable changes in retinal absorbance or

192   structural responsiveness, and whether more powerful molecular comparisons covering a wider

193   portion of the genome (e.g. SNP-based genome scans; transcriptomic approaches) will offer

194   further insights into the role of depth as a diversifying agent in the ocean.

195   **Acknowledgements**

196   We are grateful to the WKREDS workshop of the International Council for the Exploration of the

197   Sea (ICES) for inspiring this work. We are also indebted to Valerie Chosson for technical

198   assistance at the MRI, and Emma Teeling, Bruno Fonseca Simões and three reviewers, for the

199   constructive criticism offered.

References

200

201   Alesandrini S, Bernardi G. 1999. Ancient species flocks and recent speciation events: what can

202   rockfish teach us about cichlids (and vice versa)? *Journal of Molecular Evolution* 49:814–

203   818.

204   Belkhir K, Borsa P, Chikhi L, Raufaste N, & Bonhomme F. 1996. GENETIX 4.05, logiciel sous

205   Windows TM pour la génétique des populations. Laboratoire Génome, Populations,

206   Interactions, CNRS UMR 5000, Université de Montpellier II, Montpellier (France).

207   Bird CE, Fernandez-Silva I, Skillings DJ, Toonen RJ. (2012) Sympatric speciation in the post

208   "modern synthesis" era of evolutionary biology. *Evolutionary Biology* 39: 158–180.

209   Bird CE, Holland BS, Bowen BW, Toonen RJ. 2011. Diversification of sympatric broadcast-

210   spawning limpets (*Cellana* spp.) within the Hawaiian archipelago. *Molecular Ecology* 20:

211   2128–2141.

212   Brokovich E, Ben-Ari T, Kark S, Kiflawi M, Dishon G, Iluz D, Shashar N. 2010. Functional

213   changes of the visual system of the damselfish *Dascyllus marginatus* along its bathymetric

214   range. *Physiology and Behavior* 101: 413–421.

215   Bunke C, Hanel R, Trautner JH. 2012. Phylogenetic relationships among North Atlantic redfish

216   (genus *Sebastes*) as revealed by mitochondrial DNA sequence analysis. *Journal of Applied*

217   *Ichthyology* 29:82-92.

218   Cadrin SX, Bernreuther M, Daníelsdóttir AK, Hjörleifsson E, Johansen T, Kerr L, Kristinsson K,

219   Mariani S, Nedreaas K, Pampoulie C, Planque B, Reinert J, Saborido-Rey F, Sigurðsson T,

220   Stransky C. 2010. A Population structure of beaked redfish, *Sebastes mentella*: evidence of

221   divergence associated with different habitats. *ICES Journal of Marine Science* 67:1617–

222   1630.

223   Chen W-J, Bonillo C, Lecointre G. 2003. Repeatability of clades as a criterion of reliability: a

224   case study for molecular phylogeny of Acanthomorpha (Teleostei) with larger number of

225   taxa. *Molecular Phylogenetics and Evolution* 26: 262-288.

226   Creevey CJ, McInerney JO. 2002. An algorithm for detecting directional and non-directional

227   positive selection, neutrality and negative selection in protein coding DNA sequences. *Gene*

228   300:43-51.

229   Creevey CJ, McInerney JO. 2003. CRANN: detecting adaptive evolution in. *Bioinformatics*

230   19:1726.

231   Excoffier L, Lischer HE. 2010. Arlequin suite ver 3.5: a new series of programs to perform

232   population genetics analyses under Linux and Windows. *Molecular Ecology Resources* 10:

233   564–567.

234   Fitzpatrick BM, Fordyce JA, Gavrilets S. 2008. What, if anything, is sympatric speciation?

235   *Journal of Evolutionary Biology* 21:1452–1459.

236   Gavrilets S. 2003. Perspective: models of speciation: what have we learned in 40 years?

237   *Evolution* 57:2197-2215.

238   Guindon S, Dufayard JF, Anisimova M., Hordijk W, Gascuel O. 2010. PhyML 3.0: new algoriths,

239   methods and utilities. *Systematic Biology* 59:307-321.

240   Hyde JR, Kimbrell CA, Budrick JE, Lynn EA, Vetter RD. 2008. Cryptic speciation in the

241   vermilion rockfish (*Sebastes miniatus*) and the role of bathymetry in the speciation process.

242   *Molecular Ecology* 17:1122-1136.

243   Hyde J, Vetter RD. 2007. The origin, evolution, and diversification of rockfishes of the genus

244   *Sebastes* (Cuvier). *Molecular phylogenetics and evolution* 44:790-811.

245   Ingram T. 2011. Speciation along a depth gradient in a marine adaptive radiation. *Proceedings of*

246   *the Royal Society of London Series B* 278: 613–618.

247   Librado P, Rozas J. 2009. DnaSP v5: a software for comprehensive analysis of DNA

248   polymorphism data. *Bioinformatics* 25: 1451-1452.

249   McDonald JH, & Kreitman M. 1991. Adaptive protein evolution at the Adh locus in Drosophila.

250   Nature 351: 652-654.

251   Miller SA, Dykes DD, & Polesky H F. 1988. A simple salting out procedure for extracting DNA

252   from human nucleated cells. Nucleic Acids Research 16: 1215.

253   Nei M. 1987. *Molecular Evolutionary Genetics*. Columbia University Press, New York, NY.

254   Ou WB, Yi T, Kim JM, Khorana HG. 2011. The roles of transmembrane domain Helix-III during

255   rhodopsin photoactivation. *PLoS One* 6:e17398.

256   Pampoulie C. Daníelsdóttir AK. 2008. Resolving species identification problems in the genus

257   *Sebastes* using nuclear genetic markers. *Fisheries Research* 93:54-63.

258   Patarnello T, Volckaert F. Castilho R. 2007. Pillars of hercules: Is the atlantic-mediterranean

259   transition a phylogeographical break? *Molecular Ecology* 16:4426-4444.

260   Peakall R, Smouse PE. 2006. Genalex 6: genetic analysis in Excel. Population genetic software

261   for teaching and research. Molecular Ecology Notes 6: 288-295.

262   Posada D, Crandall KA. 1998. MODELTEST: testing the model of DNA substitution.

263   *Bioinformatics* 14:817-818.

264   Rettelbach A, Kopp M, Dieckmann U, & Hermisson J. 2013. Three Modes of Adaptive

265   Speciation in Spatially Structured Populations. *The American Naturalist* 182: E215-E234.

266   Rogers AR, & Harpending H. 1992. Population growth makes waves in the distribution of

267   pairwise genetic differences. *Molecular Biology and Evolution* 9: 552-569.

268   Roy D, Hurlbut TR, Ruzzante DE. 2012. Biocomplexity in a demersal exploited fish, white hake

269   (*Urophycis tenuis*): depth-related structure and inadequacy of current management

270   approaches. *Canadian Journal of Fisheries and Aquatic Sciences* 69:415-429.

271   Salzburger W, Ewing GB, von Haeseler A. 2011. The performance of phylogenetic algorithms in

272   estimating haplotype genealogies with migration. *Molecular Ecology* 20:1952-1963.

273   Sivasundar A, Palumbi SR. 2010. Parallel amino acid replacements in the rhodopsins of the

274   rockfishes (*Sebastes* spp.) associated with shifts in habitat depth. *Journal of Evolutionary*

275   *Biology* 23**:**1159-1169.

276   Smadja CM, & Butlin RK. 2011. A framework for comparing processes of speciation in the

277   presence of gene flow. *Molecular Ecology* 20:5123-5140.

278   Smith SO. 2010. Structure and activation of the visual pigment rhodopsin. *Annual Review of*

279   *Biophysics* 39:309-328.

280   Somero G N. 1992. Adaptations to high hydrostatic pressure, *Annual review of physiology*

281   54:557–577.

282   Stefánsson MO, Reinert J, Sigurðsson Þ, Kristinsson K, Nedreaas K, Pampoulie C. 2009a. Depth

283   as a potential driver of genetic structure of *Sebastes*. *ICES Journal of Marine Science*

284   66:680-690.

285   Stefánsson MÖ, Sigurðsson T, Pampoulie C, Daníelsdóttir AK, Thorgilsson B, Ragnarsdóttir R,

286   Gíslason D, Coughlan J, Cross TF, Bernatchez L. 2009b. Pleistocene genetic legacy suggests

287   incipient species of *Sebastes mentella* in the Irminger Sea. *Heredity* 102:514-524.

288 Vonlanthen P, Roy D, Hudson AG, Largiader CR, Bittner D, Seehausen O. 2009. Divergence
289   along a steep ecological gradient in lake whitefish (*Coregonus* sp.). *Journal of Evolutionary*
290   *Biology* 22:498–514.

291 Warrant EJ, Locket N. 2004. Vision in the deep sea. *Biological Reviews* 79:671-712.

292 Yokoyama S. 2000. Molecular evolution of vertebrate visual pigments. *Progress in Retinal and*
293   *Eye Researh* 19:385-419.

294 Yokoyama S. 2002. Molecular evolution of color vision in vertebrates. *Gene* 300:69-78.

295 Yokoyama S, Tada T, Zhang H, Britt L. 2008. Elucidation of phenotypic adaptations: Molecular
296   analyses of dim-light vision proteins in vertebrates. *Proceedings of the National Academy of*
297   *science of the United State of America* 105:13480–13485.

298   **Table 1.** Summary of mtDNA control region molecular

299   diversity: H = unique haplotypes; n = number of individuals; S

Population	H/n	S	$\hat{h} \pm$ SD	$\pi \pm$ SD	D_T	F_S	F^*	D^*
(SP)	8/25	10	0.887±0.033	0.00504±0.00082	-1.08590	-3.806	-2.30037	-2.35314
(DP)	6/25	8	0.543±0.119	0.00238±0.00071	-1.77639	-4.717	-1.49744	-1.08199

300   = Segregating sites; $\hat{h}$ = haplotype diversity; π = nucleotide

301   diversity (both with associated standard deviations, SD); D_T =

302   Tajima‟s D; F_S = Fu's F_S statistic; F^* = Fu and Li's F test; D^* =

303   Fu and Li's D test. SP = Shallow Pelagic; DP = Deep Pelagic.

304   **Table 2.** Analysis of fixation/differentiation indices for mtDNA and microsatellite data between
305   shallow-pelagic (SP) and deep-pelagic (DP) *S mentella*.

Marker	Group	Fixation/Differentiation index	estimate	p
mtDNA	SP vs DP	F_{ST}	0.636	<0.001
		Φ_{ST}	0.273	<0.001
Microsatellites	SP vs DP	F_{ST}	0.031	0.001
		G'_{ST}	0.135	0.001
		D_{est}	0.121	0.001

306 **Table 3.** Mismatch distribution parameter estimates for mtDNA control region

Population	Mismatch Distribution				
	τ	$\theta 0$	$\theta 1$	SSD *p*-value±SD	r *p*-value±SD
(SP)	1.9	0.0000	99,999	DE 0.007±0.18	0.30±0.21
				SE 0.007±0.002	0.27±0.28
(DP)	0.1	0.0000	99,999	DE 0.274±0.18	0.00±0.21
				SE 0.004±0.002	0.67±0.28

307 τ = tau; $\theta 0$ = theta 0;

308 $\theta 1$ = theta 1; SSD = sum of squared deviations; r = raggedness

309 statistic; DE: demographic expansion, SE: spatial expansion; (SP) = Shallow pelagic; (DP) = Deep-

310 pelagic.

311 **Table 4.** Creevey-McInerney positive selection analysis on *Sebastes* rhodopsin sequences

312 outgrouped with *S. alutus* (GenBank: EF212407.1), G-value p<0.05 for Fisher's[†] and G-Test*

Branch no.	RI	RV	SI	SV	G-value
rhodopsin					
0	0	2	0	0	0.00
1	0	2	0	0	0.00
2	0	2	0	0	0.00
3	0	5	0	0	0.00
4	0	5	0	0	0.00
5	0	5	0	0	0.00
6	0	5	0	0	0.00
7	0	5	0	0	0.00
8	0	5	0	0	0.00
9	0	5	0	0	0.00
10	0	5	0	0	0.00
11	0	8	0	0	0.00
12	0	8	0	0	0.00
13	0	8	0	0	0.00
14	0	8	0	0	0.00
15	0	8	0	0	0.00
16	0	11	1	0	2.01
17	0	0	1	0	0.00
18	0	9	0	0	0.00
19	0	11	0	0	0.00
20	0	12	0	0	0.00
21	0	12	0	0	2.2
22	0	12	1	0	2.2
23	0	12	1	0	2.2
24	0	12	1	0	2.2
25	0	12	1	0	2.2
26	1	12	2	0	5.5[†]
27	8	24	5	1	6.76*

314   **Figure 1**. *Sebastes mentella* genealogies for mtDNA (n=50; 25 SP + 25 DP) and rhodopsin

315   (n=22; 10 SP + 12 DP), and mtDNA mismatch distributions. (A) Haplotype network for the

316   shallow (red) and deep (blue) groups for mtDNA (i) and rhodopsin (ii). The size of each circle

317   represents the proportion of haplotypes. The lengths of the connecting lines reflect the number of

318   mutations between haplotypes. (B) Mismatch distributions from the mtDNA sequences of

319   shallow (i) and deep (ii) groups, respectively from above and below 550 metres depth

320   respectively. Dotted lines (Up/low bound.) represent the 95% boundaries of the simulated

321   distributions.

322   **Figure 2.** Creevey–McInerney analysis of *Sebastes* rhodopsin. Rhodopsin reveals significant

323   positive selection (*) at two nodes (26, 27).

324   **Figure 3.** Chromatograms illustrating the non-synonymous A/G mutation on the rhodopsin gene,

325   which discriminates between "Deep-Pelagic" (a: above) and "Shallow-Pelagic" (b: below)

326   *Sebastes mentella*.

327   **Figure 4.** Correspondence analysis based on microsatellite data. Each circle represent an

328   individual, red and blue refer to the shallow-pelagic (SP) and deep-pelagic (DP) groups

329   respectively.

Figure 1

Comprehensive image of mtDNA and Rhodopsin genetic divergence and Mismatch Distributions

Figure 1. S. mentella genealogies for mtDNA (n=50; 25 SP + 25 DP) and rhodopsin (n=22; 10 SP + 12 DP), and mtDNA mismatch distributions. (A) Haplotype network for the shallow (red) and deep (blue) groups for mtDNA (i) and rhodopsin (ii). The size of each circle represents the proportion of haplotypes. The lengths of the connecting lines reflect the number of mutations between haplotypes. (B) Mismatch distributions from the mtDNA sequences of shallow (i) and deep (ii) groups, respectively from above and below 550 metres depth respectively. Dotted lines (Up/low bound.) represent the 95% boundaries of the simulated distributions.

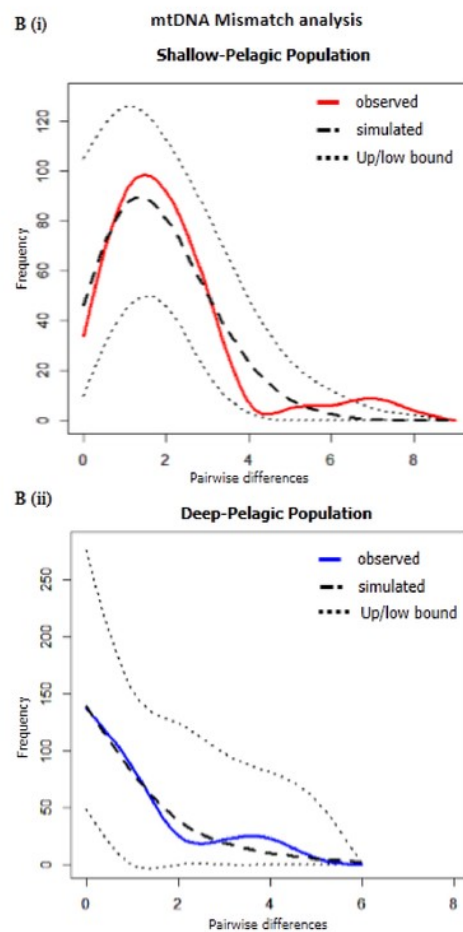
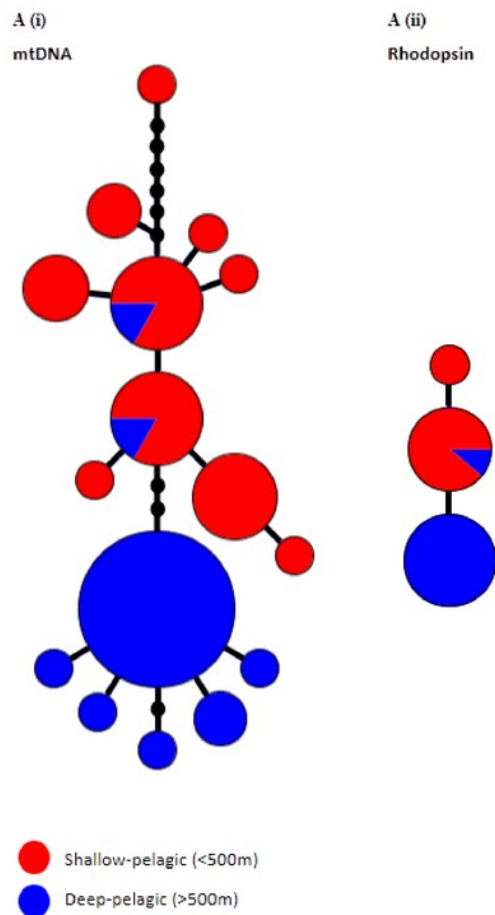



Figure 2

CRANN test tree

Creevey–McInerney analysis of Sebastes rhodopsin. Rhodopsin reveals significant positive selection (*) at two nodes (26, 27).

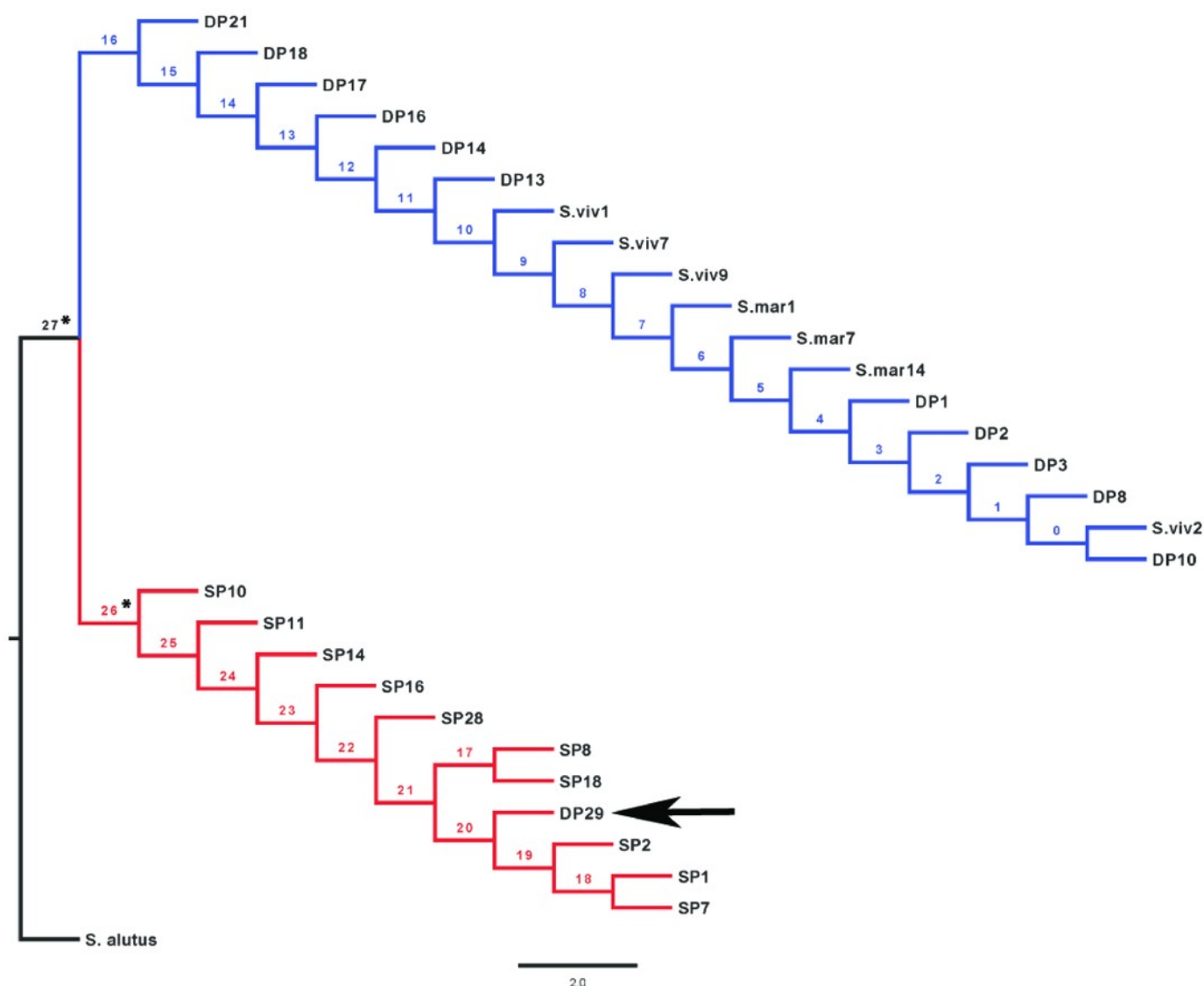



Figure 3

Example of non-synonymous base substitution

Figure 3. Chromatograms illustrating the non-synonymous A/G mutation on the rhodopsin gene, which discriminates between "Deep-Pelagic" (a: above) and "Shallow-Pelagic" (b: below) Sebastes mentella.

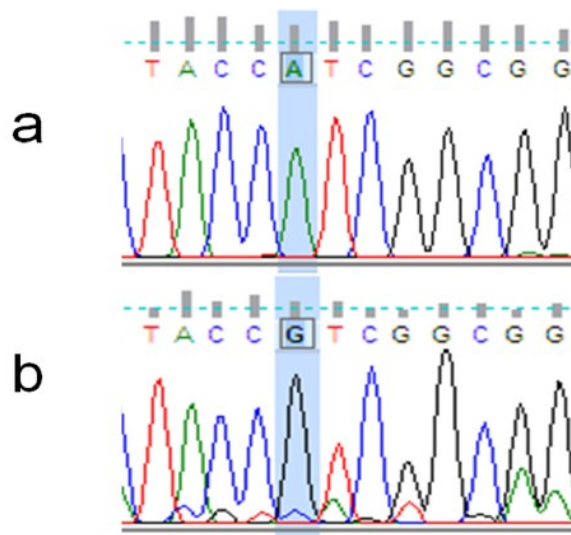



Figure 4

Ordination of microsatellite genotypes

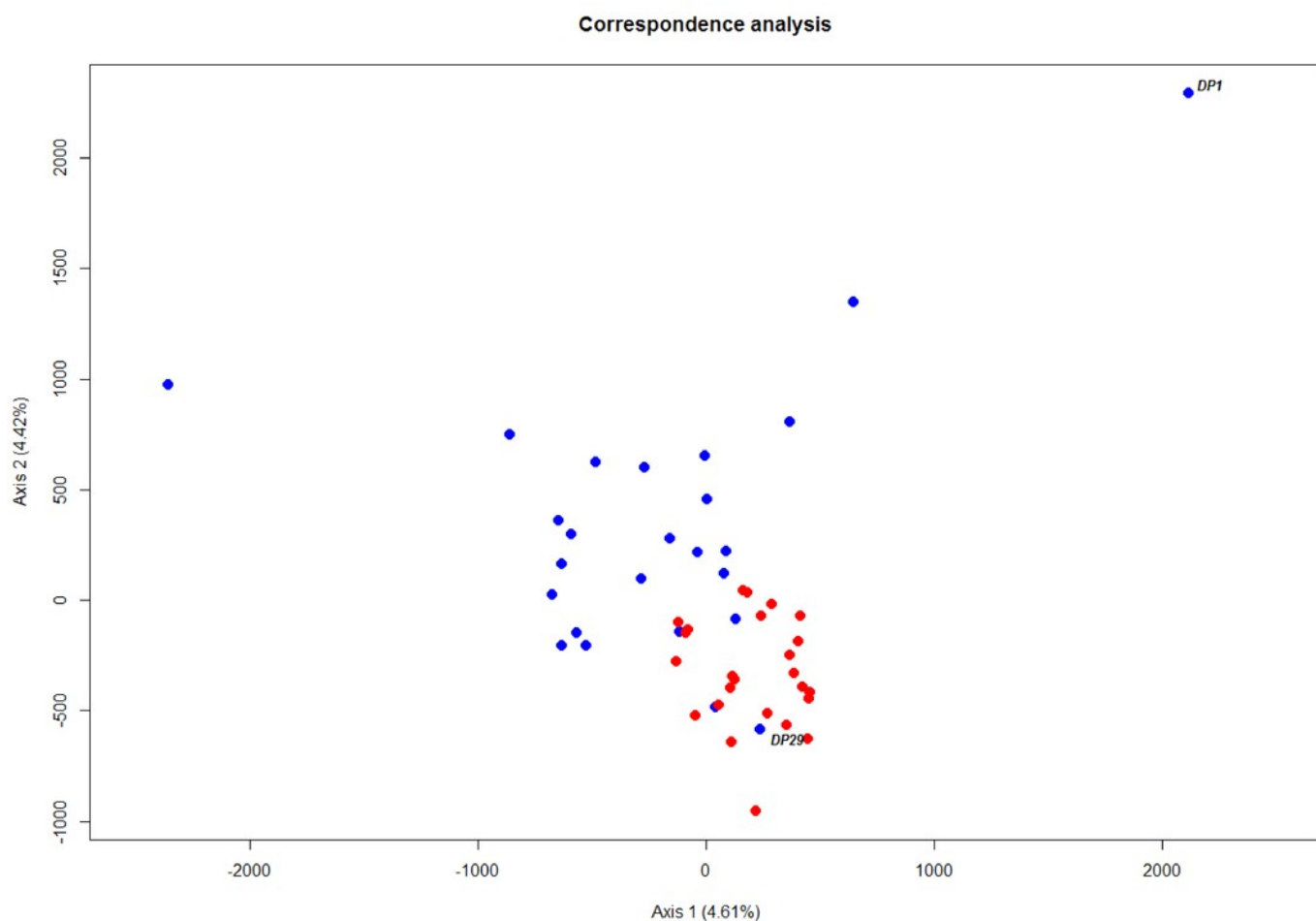
Figure 4. Correspondence analysis based on microsatellite data. Each circle represent an individual, red and blue refer to the shallow-pelagic (SP) and deep-pelagic (DP) groups respectively.

Summary of mtDNA control region molecular

Table 1. Summary of mtDNA control region molecular diversity: H = unique haplotypes; n = number of individuals; S = Segregating sites; FS = Fu's FS statistic; DT = Tajima"s D ; F* = Fu and Li's F test; D* = Fu and Li's D test. SP = Shallow Pelagic; DP = Deep Pelagic.

Population	H/n	S	$\hat{h} \pm SD$	$? \pm SD$	D_T	F_S	F*	D*
(SP)	8/25	10	0.887±0.033	0.00504±0.00082	-1.08590	-3.806	-2.30037	-2.35314
(DP)	6/25	8	0.543±0.119	0.00238±0.00071	-1.77639	-4.717	-1.49744	-1.08199

Analysis of fixation/differentiation indices for mtDNA and microsatellite data between shallow-pelagic (SP) and deep-pelagic (DP) S mentella.

Table 2. Analysis of fixation/differentiation indices for mtDNA and microsatellite data between shallow-pelagic (SP) and deep-pelagic (DP) S. mentella.

Marker	Group	Fixation/Differentiation index	estimate	p
mtDNA	SP vs DP	F_{ST}	0.636	<0.001
		Φ_{ST}	0.273	<0.001
Microsatellites	SP vs DP	F_{ST}	0.031	0.001
		G'_{ST}	0.135	0.001
		D_{est}	0.121	0.001

Mismatch distribution parameter estimates for mtDNA control region

Table 3. Mismatch distribution parameter estimates for mtDNA control region

Population	Mismatch Distribution				
	τ	θ_0	θ_1	SSD *p*-value±SD	r *p*-value±SD
(SP)	1.9	0.0000	99,999	DE 0.007±0.18	0.30±0.21
				SE 0.007±0.002	0.27±0.28
(DP)	0.1	0.0000	99,999	DE 0.274±0.18	0.00±0.21
				SE 0.004±0.002	0.67±0.28

τ = tau; θ_0 = theta 0; θ_1 = theta 1; SSD = sum of squared deviations; r = raggedness

statistic; DE: demographic expansion, SE: spatial expansion; (SP) = Shallow pelagic; (DP) = Deep-pelagic.

Creevey-McInerney positive selection analysis on Sebastes rhodopsin sequences outgrouped with S. alutus (GenBank: EF212407.1), G-value p<0.05 for Fisher's† and G-Test*

Table 4. Creevey-McInerney positive selection analysis on Sebastes rhodopsin sequences outgrouped with S. alutus (GenBank: EF212407.1), G-value p<0.05 for Fisher's† and G-Test*

Branch no. rhodopsin	RI	RV	SI	SV	G-value
0	0	2	0	0	0.00
1	0	2	0	0	0.00
2	0	2	0	0	0.00
3	0	5	0	0	0.00
4	0	5	0	0	0.00
5	0	5	0	0	0.00
6	0	5	0	0	0.00
7	0	5	0	0	0.00
8	0	5	0	0	0.00
9	0	5	0	0	0.00
10	0	5	0	0	0.00
11	0	8	0	0	0.00
12	0	8	0	0	0.00
13	0	8	0	0	0.00
14	0	8	0	0	0.00
15	0	8	0	0	0.00
16	0	11	1	0	2.01
17	0	0	1	0	0.00
18	0	9	0	0	0.00
19	0	11	0	0	0.00
20	0	12	0	0	0.00
21	0	12	0	0	2.2
22	0	12	1	0	2.2
23	0	12	1	0	2.2
24	0	12	1	0	2.2

25	0	12	1	0	2.2
26	1	12	2	0	$5.5^{\dagger}$
27	8	24	5	1	6.76*