

Recent speciation and hybridization in Icelandic deep-sea isopods: an integrative approach using genomics and proteomics

Eva Paulus¹, Saskia Brix², Annabelle Siebert², Pedro Martínez Arbizu³, Sven Rossel², Janna Peters⁴, Jörundur Svavarsson⁵, and Martin Schwentner⁶

¹University of Groningen

²Senckenberg am Meer

³Senckenberg Research Institute

⁴Senckenberg am Meer Deutsches Zentrum für Marine Biodiversitätsforschung

⁵University of Iceland

⁶Naturhistorisches Museum Wien

April 6, 2021

Abstract

The isopod species *Haploniscus bicuspis* (G.O. Sars, 1877) shows circum-Icelandic distribution in a wide range of environmental conditions and along well-known geographic barriers, such as the Greenland-Iceland-Faroe (GIF) Ridge. We wanted to explore population genetics, phylogeography and cryptic speciation as well as to investigate whether previously described, but unaccepted subspecies have any merit. Using the same set of specimens, we combined mitochondrial COI sequences, thousands of nuclear loci (ddRAD), and proteomic profiles, plus selected morphological characters using Confocal Laser Scanning Microscopy (CLSM). Five divergent genetic lineages were identified by COI and ddRAD, two south and three north of the GIF Ridge. Assignment of populations to the three northern lineages varied and detailed analyses revealed hybridization and gene flow between them, suggesting a single northern species with a complex phylogeographic history. No apparent hybridization was observed among lineages south of the Ridge, inferring the existence of two more species. Differences in proteomic profiles between the three putative species were minimal, implying an ongoing or recent speciation process. Population differentiation was high, even among closely associated populations, and higher in mitochondrial COI than nuclear ddRAD loci. Gene flow is apparently male-biased, leading to hybrid zones and instances of complete exchange of the local nuclear genome through immigrating males. This study did not confirm the existence of subspecies defined by male characters, which probably characterize different male developmental stages present in all species.

Recent speciation and hybridization in Icelandic deep-sea isopods: an integrative approach using genomics and proteomics

Running title: **Recent isopod speciation in Icelandic waters**

Eva Paulus^{1,2}, Saskia Brix^{2*}, Annabelle Siebert^{2,3}, Pedro Martínez Arbizu⁴, Sven Rossel⁴, Janna Peters⁴, Jörundur Svavarsson⁵, Martin Schwentner^{3,6}

¹ *University of Groningen, 9712 CP Groningen, Netherlands*

² *Senckenberg am Meer, German Centre for Marine Biodiversity Research (DZMB) c/o Biocenter Grindel, Center of Natural History (CeNak), Universität Hamburg, Martin-Luther-King-Platz 3, 20146 Hamburg, Germany*

³ Center of Natural History (CeNak), Universität Hamburg, Martin-Luther-King-Platz 3, 20146 Hamburg, Germany

⁴ Senckenberg am Meer, German Centre for Marine Biodiversity Research (DZMB), Südstrand 44, 26382 Wilhelmshaven

⁵Department of Life and Environmental Sciences, University of Iceland, Askja—Natural Science Building, Sturlugata 7, 102 Reykjavík, Iceland

⁶ Naturhistorisches Museum Wien, Burgring 7, 1010 Vienna, Austria

*corresponding author: Saskia Brix, saskia.brix-elsig@senckenberg.de

Abstract

The crustacean marine isopod species *Haploniscus bicuspis* (G.O. Sars, 1877) shows circum-Icelandic distribution in a wide range of environmental conditions and along well-known geographic barriers, such as the Greenland-Iceland-Faeroe (GIF) Ridge. We wanted to explore population genetics, phylogeography and cryptic speciation as well as to investigate whether previously described, but unaccepted subspecies have any merit. Using the same set of specimens, we combined mitochondrial COI sequences, thousands of nuclear loci (ddRAD), and proteomic profiles, plus selected morphological characters using Confocal Laser Scanning Microscopy (CLSM). Five divergent genetic lineages were identified by COI and ddRAD, two south and three north of the GIF Ridge. Assignment of populations to the three northern lineages varied and detailed analyses revealed hybridization and gene flow between them, suggesting a single northern species with a complex phylogeographic history. No apparent hybridization was observed among lineages south of the Ridge, inferring the existence of two more species. Differences in proteomic profiles between the three putative species were minimal, implying an ongoing or recent speciation process. Population differentiation was high, even among closely associated populations, and higher in mitochondrial COI than nuclear ddRAD loci. Gene flow is apparently male-biased, leading to hybrid zones and instances of complete exchange of the local nuclear genome through immigrating males. This study did not confirm the existence of subspecies defined by male characters, which probably characterize different male developmental stages.

Keywords

Haploniscidae, *Haploniscus bicuspis*, MALDI-TOF, cryptic species, ddRAD, DNA barcoding

Introduction

Deep-sea benthos fosters a high diversity of invertebrates with peracarid crustaceans like Isopoda being particularly species-rich (Wilson & Ahyong, 2015; Brandt et al., 2018). The underlying factors that shaped this diversity and the potential role of past climatic changes or physical barriers, such as deep-sea ridges, are not well understood. Similarly, fundamental population genetic studies are scarce for the deep sea (Taylor & Roterman, 2017). One problem is the limited availability of large-scale datasets in terms of individual numbers and geographic scale. The deep-sea crustacean isopod species *Haploniscus bicuspis* (G. O. Sars, 1877) offers to be an ideal surrogate, as it is one of the few isopod species with circum-Icelandic distribution and large numbers of samples available from the BIOICE (Benthic Invertebrates of Icelandic Waters) and IceAGE (Icelandic marine Animals: Genetics and Ecology) projects (Brix et al., 2014; Meißner, Brix, Halanych, & Jazdzewska, 2018).

The complex interactions of highly diverse water masses make Icelandic waters an interesting location to study the distribution and speciation processes of deep-sea taxa. While the warmer and more saline North

Atlantic occur south of Iceland, the colder waters of the Nordic Seas and the Arctic Ocean shape the environment to the North. Steep gradients in temperature and salinity, various sediment substrates, differences in food availability, and several shallow and deep ridges create a unique and diverse environment. This favors high biodiversity through a patchwork of different habitats within relatively small scales (Meißner et al., 2014; Brökeland & Svavarsson, 2017; Jochumsen, Schnurr, & Quadfasel, 2016; Ostmann, Schnurr, & Martínez Arbizu, 2014).

For other isopod (Brix & Svavarsson, 2010) and amphipod (Weisshappel, 2000, 2001; Dauvin et al., 2012) crustaceans, as well as many other benthic deep-sea animals, the Greenland-Iceland-Faeroe (GIF) Ridge is an important barrier that limits their distribution. The GIF Ridge has a saddle depth of about 480 meters between the Faeroe Islands and Iceland in the south-east, and 620 meters between Greenland and Iceland in the north-west of Iceland (Hansen & Østerhus, 2000). Near-bottom temperatures range from 12 °C in the North Atlantic south of the GIF Ridge, to -0.9 °C deep north of the GIF Ridge (Schnurr et al. 2014; Jochumsen et al., 2016). Iceland is located in an area especially susceptible to climatic change (Hanna, Jónsson, & Box, 2006) and hence is undergoing rapid changes in terms of species distribution and composition (e.g. Arnason, 2007; D’Alba, Monaghan, & Nager, 2010; Pecl et al., 2017). Naturally, the waters around Iceland have been influenced by past climatic changes. During the Last Glacial Maximum, the ice sheet extended beyond the Icelandic landmass and within reach of the shelf break at around 300 meters depth (Patton, Hubbard, Bradwell, & Schomacker, 2017). Around 15 ka BP, the ice sheet broke apart abruptly due to rising sea levels (Geirsdóttir, Miller, Axford, & Ólafsdóttir, 2009; Nordahl & Ingólfsson, 2015). Both the southern (i.e., Iceland-Faeroe Ridge) and the northern parts of the GIF Ridge (i.e. Denmark Strait between Iceland and Greenland) were likely influenced by spreading and contracting ice sheets, potentially limiting the distribution of benthic marine animals.

The GIF Ridge and associated ecological differences strongly affect the observed species compositions with most benthic deep-sea species (e.g., Isopoda, Amphipoda, Tanaidacea) being confined to one side of the GIF Ridge (e.g., Hansen, 1908; Negoescu & Svavarsson, 1997; Svavarsson, Strömberg, & Brattegard, 1993; Weisshappel, 2001; Gudmundsson, 1998; Stransky & Svavarsson, 2006; Brix, Stransky, et al., 2018; Brix, Lörz, et al., 2018; Jakiel, Stępień, & Błażewicz, 2018).

An interesting exception is the widespread asellote isopod species *Haploniscus bicuspis*. It occurs in all water masses and in various sampled depths (300-2900 m) around Iceland, whereas the majority of congeneric species are confined to the North Atlantic Ocean (except for *H. angustus* Lincoln, 1985; Brökeland & Svavarsson, 2017). This makes *H. bicuspis* an ideal surrogate to study patterns of genetic diversity, population genetics and the phylogeographic history of a benthic deep-sea isopod species. It also offers the possibility to assess proteomic differences using MALDI-TOF MS (Matrix-Assisted Laser Desorption/Ionization Time-of-Flight Mass Spectrometry) associated with ecological differentiation or potential cryptic speciation. Species identification with MALDI-TOF MS is well-established in microbiology for determination of bacteria, viruses, and fungi (Nachtigall, Pereira, Trofymchuk, & Santos, 2020; Singhal, Kumar, Kanaujia, & Virdi, 2015) and several proof-of-concept studies supported its general applicability for species delimitation in marine crustaceans (Riccardi et al., 2012; Laakmann et al., 2013; Bode et al., 2017; Rossel & Martínez Arbizu, 2019), yet to our knowledge no studies on peracarids have been performed so far. Understanding the taxonomic resolution of proteomic profiling in invertebrates is subject of ongoing research.

Haploniscus bicuspis was first described from specimens collected near Norway (Figure 1). Subsequently, Wolff (1962) described two subspecies based on the shape of the male antennula and pleopod 1 (see Figure 2): *Haploniscus bicuspis bicuspis* (including the Norwegian specimens) and *H. bicuspis tepidus* (from the Reykjanes Ridge south-west of Iceland). The latter features a narrower second segment of the antennula and a laterally rounded distal part of the pleopod 1 (in contrast to the presence of a distinct corner in *H. b. bicuspis*; Figure 2). Unfortunately, the type series of *H. bicuspis* comprises no adult male and both subspecies of Wolff (1962) currently hold the status “unaccepted” in WoRMS (Boyko et al., accessed August 2020). It is currently unknown whether the morphological variation described by Wolff suggests cryptic speciation or intraspecific variability.

Asellote isopods are particularly interesting for population genetic studies, as a pelagic larval stage is absent in their life cycle. Like all Peracarida, they hatch and rear juveniles in a brooding pouch (marsupium), and the juveniles largely resemble the adult morphology and lifestyle upon release (Lincoln, 1985). This limits their dispersal abilities compared to animals with free-swimming pelagic life stages. In general, locomotion and dispersal abilities in deep-sea asellote isopods depend on the adult stage (Brix et al., 2020). In the abyssal Pacific, Haploniscidae showed a mean species range of 183 km and a maximum range of 1,310 km, with 83 % of the species ($n = 24$) present in a single area only (Brix et al. 2020). These distributional ranges are much lower than for the swimming isopod families Desmosomatidae and Munnopsidae. As a typical walking haploniscid isopod, the body plan of *Haploniscus* does not show adaptations specific to swimming or burrowing, and the animals are found in close association with the sediment surface (Brix et al., 2020, Thiel and Haye, 2006). This suggests comparably poor dispersal capabilities of the adults as well.

This study aims to explore the genetic diversity within *Haploniscus bicuspis*, focusing on aspects of potential cryptic diversity, population genetics and its phylogeographic history. The latter is particularly interesting, as it may reveal historic glacial refugia and postglacial (re-)colonization for a deep-sea species, and thus potential population genetic consequences of past climatic changes. To do so, we combined analyses of mitochondrial COI sequences with thousands of nuclear ddRAD loci. We further assessed proteome-level differences between populations and putatively revealed cryptic species.

Methods

All *H. bicuspis* specimens examined were sampled aboard the research vessels RV Meteor (M85/3), RV Poseidon (POS456) and RV Maria S. Merian (MSM75) during the IceAGE (2011), IceAGE2 (2013) and IceAGE_RR (2018) expeditions. The specimens were sorted at the laboratories of the DZMB (German Centre for Marine Biodiversity Research) and deposited in the collections of the Center of Natural History (Hamburg, Germany; see Supplementary Table S1). Specimens were collected using an epibenthic sledge (EBS, see Brenke, 2005), except for samples from stations 23 and 63, which were collected using a Van Veen grab.

Following type material as loan from the Museum in Copenhagen was morphologically compared with the newly collected IceAGE specimens:

A dry vial apparently containing 1 dried specimen and labelled as ZMUC-CRU-5817 from Ingolf st. 78; The slide portion of ZMUC-CRU-5817; a dry vial containing 2 microvials of *H. b. bicuspis* from Ingolf st. 112; a dry vial containing 1 microvial of *H. b. bicuspis* from Ingolf st. 139; a vial with about 20 specimens in alcohol labelled as ZMUC-CRU-346, but also with the following label from Institut für Polarökologie, Kiel (not Zoological Museum, Copenhagen): *Haploniscus bicuspis*, Projekt H21-5, Station 333, Gerät EBS, Sammler A. Brandt.; the non-slide portion of ZMUC-CRU-5817 (NHMD-83661) consists of 1 syntype of *H. b. tepidus* in alcohol; the 2 dried syntypes of this subspecies from Ingolf st. 78 labelled as NHMD-272194; a dry vial containing 1 microvial of *H. b. bicuspis* from Ingolf st. 104; a dry vial containing 1 microvial of *H. b. bicuspis* from Ingolf st. 125; ZMUC-CRU-346 (NHMD-78194) with ZMUC supplementary label (Meteor st. 333, Kolbeinsey Ridge, Iceland, 67°56.5' N, 18°02.4' W).

Selected males were visualized using CLSM to assess the shape of the pleopod 1. Specimens were stained with the fluorescent markers Congo Red and Acid Fuchsin (1:1), following the methodology outlined in Michels and Büntzow (2010). For scanning, specimens were transferred to glycerin. A Leica TCS SPV with a Leica DM5000 B microscope and DPSS laser (10 mW, 561 nm) was used to perform the CLSM. Scans were recorded with the LAS AF 2.2. software.

DNA extractions were performed using the mid-sections of the animals, leaving the cephalothorax and the pleon intact for subsequent morphological analyses and as partial vouchers in the collections. The animals were dissected carefully using a micro scissor and the gut was removed to avoid contaminations. If possible, the same animal was used for proteomics, COI barcoding and ddRAD, transferring half of the dissected tissue in a vial for genetics and the other half into a vial for proteomics. DNA was extracted using the Marine Animal Tissue Genomic DNA Extraction Kit (Neo Biotech) or the Genomic DNA from tissue kit

(Macherey-Nagel). DNA was eluted in 70 μ L elution buffer. Chelex (Chelating Ion Exchange Resin, Bio-Rad) was used for some of the animals, for which only the COI fragment was sequenced (for protocol see Jennings, Golovan, & Brix, 2020).

COI

1 μ L DNA was PCR amplified using PuReTaq Ready-To-Go (RTG) PCR Beads (GE Healthcare) with 1 μ L of either dgLCO (GGT CAA CAA ATC ATA AAG AYA TYG G; Meyer, 2003)/dgHCO (TAA ACT TCA GGG TGA CCA AAR AAY CA; Meyer, 2003) or LCOJJ (CHACWAAAYCATAAAGATATYGG; Astrin & Stüben, 2008)/HCOJJ (AWACTTCVGGRTGVCCAAARAATCA; Astrin & Stüben 2008) primers and 22 μ L nuclease-free water. The PCR ran at 94 °C for 5 minutes, followed by 38 cycles of denaturation at 94 °C for 45 seconds, annealing at 45 °C for 50 seconds, and elongation at 72 °C for 1 minute. This was followed by the final elongation at 72 °C for 5 minutes. The resulting fragment length was assessed by gel electrophoresis using 1.5% TAE gels. Excess primers were removed using ExoSAP. The PCR products and the corresponding primers were sent to Macrogen or Eurofins for bidirectional sequencing.

The resulting sequences were individually checked for their quality using GENEIOUS Prime version 2019.2.3 and the forward and reverse sequences were assembled using the *de novo* assembly function. An alignment was created using MUSCLE (v 3.8.425, Edgar 2004). To assess the potential presence of cryptic diversity within *H. bicuspis*, we employed two species delimitation methods: generalized mixed Yule coalescent model (GMYC; Pons et al., 2006) and Automatic Barcode Gap Discovery (ABGD; Puillandre, Lambert, Brouillet, & Achaz, 2012). GMYC was run in R using the single threshold method (Reid & Carstens, 2012). The GMYC method uses an ultrametric tree as input to delimit clusters of putative species. The ultrametric tree was generated with BEAST2 (Bouckaert et al., 2019), running the analyses for 30 million generations, employing the Yule model and including every observed COI haplotype once. ABGD uses the so-called barcode gap, which corresponds to the gap between intra- and interspecific genetic distances, to delimit putative species. We used the web version of ABGD ([see https://bioinfo.mnhn.fr/abi/public/abgd/abgdweb.html](https://bioinfo.mnhn.fr/abi/public/abgd/abgdweb.html)) with standard settings except for a relative gap width of 0.5 and 100 steps and the uncorrected *p*-distances, pre-calculated with MEGA-X version 10.0.5 (Kumar, Stecher, Li, Knyaz, & Tamura, 2018). An unrooted phylogenetic network was calculated with SplitsTree (Huson & Bryant, 2006) to visualize the divergence between putative species (henceforth called lineages).

To assess the distribution of genetic diversity and the phylogeographic history, a median-joining haplotype network was generated using the program Network v 10.0 (see fluxus-engineering.com; Bandelt, Forster, & Röhl, 1999) and redrawn using InkScape. Furthermore, for each population (i.e., sampling station) nucleotide diversity (π) and gene diversity (*h*) were calculated with Arlequin 3.5.2.2 (Excoffier & Lischer, 2010). The demographic parameters Tajima's *D* and Fu's *F*, which assess deviations from neutrality, were also assessed. Genetic differentiation between populations was evaluated using pairwise F_{ST} , calculated only among populations of the same lineage and among the very closely associated lineages I_{COI}-III_{COI} (see below). For Tajima's *D*, Fu's *F*, and F_{ST} only populations with at least five individuals present were included.

ddRAD

Based on DNA yield, a subset of samples was selected for ddRAD sequencing, only including samples with at least 30 ng DNA. DNA concentration was measured with a Qubit 3.0 (Invitrogen). If samples had a lower starting concentration than 150 ng, they were concentrated via drying. All samples were brought to 24 μ L. The protocol described by Peterson et al. (2012) was mostly adhered to, with a few modifications (see also Schwentner & Lörz, 2021; Franchini, Monné Parera, Kautt, & Meyer, 2017).

Samples were grouped in batches of eight by starting DNA concentration. Restriction digestion was carried out at 37 °C for 4 hours using 3 μ L fastdigest buffer, 1.5 μ L fastdigest MspI and fastdigest EcoRI enzymes each (Thermo Fisher Scientific). Digested DNA was cleaned using 1.5x volume of magnetic beads and eluted in 21 μ L H₂O (AmpliClean Cleanup Kit, Nimagen). MspI and EcoRI adapters were ligated to the 19 μ L

digested DNA using 3 μ l of each adaptor, 3 μ l of 10x T4 ligase buffer, and 2 μ l T4 ligase (1-2 Weiss units). 16 EcoRI and 8 MspI adapter variations with unique barcodes (MspI adapters with four additional random nucleotides, as described in (Franchini et al., 2017) were used to provide a unique barcode combination for each sample within each batch of eight samples (Supplementary Table S1). Ligation commenced at 22 °C for 1 hour and was heat terminated at 65 °C for 10 minutes. Afterwards, the eight samples of each batch were pooled, and the pools were cleaned up using 1.5x volume of magnetic beads, eluting in 30 μ l of TE buffer. DNA fragments of 300 base pairs (bp) were selected with a BluePippin (Sage Science, +/- 30 bp allowed, “tight” setting) using a 1.5% agarose gel cassette with DF marker R2.

To reduce PCR amplification biases, four PCR replicates were run for each pool after size selection. PCRs comprised 5 μ l 2x Kappa HS HIFI mix, 0.3 μ l of each primer and 4.4 μ l library. Forward and reverse primers included eight different 8 bp indices (Supplementary Table S1) and were combined to add a unique index combination for each pool. The two-step PCR program ran at 95 °C for 3 minutes, followed by 17 to 19 cycles of 98 °C for 20 seconds, 72 °C for 25 seconds, and final elongation at 72 °C for 1 minute. The four replicates of each library were combined and cleaned up using 1x volume of magnetic beads, eluting in 18 μ l H₂O. DNA concentration was measured using a Qubit and mean size assessed using a TapeStation (Agilent). All samples were pooled at equal molarity and sent off to Macrogen (South Korea) for sequencing on one Illumina HiSeq4000 lane (100 bp, paired end).

The reads were pre-demultiplexed by indices by Macrogen upon delivery. Potential PCR duplicates were removed using the ‘clone_filter’ function of STACKS (Rochette, Rivera-Colón, & Catchen, 2019) and the demultiplexing by barcodes was performed with the ‘process_radtags’ function. Assembly of loci was performed using ipyrad (Eaton & Overcast, 2020). Parameters were optimized in multiple test runs. The crucial parameters in the final analyses were set to: #14 [clust_threshold] = 0.91, #20 [max_Hs_consens] = 0.1, and #22 [max_SNPs_locus] = 0.15 (see Supplemental File S1 for full set of parameters). To further optimize the number of retrieved loci, two different runs were performed: one including all specimens and one including all specimens previously delimited into lineages I_{COI}-III_{COI}. The latter run was performed to optimize loci recovery for this set of genetically and geographically closely associated lineages. In all analyses, a ‘populations’ file was included (defining lineages retrieved by COI) and requiring at least 50% of specimens of each population to be represented for a locus to be retained. Specimens with less than 50% of retrieved loci were removed after initial test runs.

To assess if the nuclear ddRAD data support the lineages differentiated by COI phylogenetic networks, principal component (PCA), Structure and coancestry analyses were performed on each of the three ddRAD data sets. Unrooted phylogenetic networks were computed with SplitsTree. PCA and Structure analyses were run via Python scripts, closely following the workflows outlined on the ipyrad homepage (<https://ipyrad.readthedocs.io/en/latest/API-analysis/index.html>; visited March 10th, 2020) using the .snps.hdf5 ipyrad output files. A minimum coverage of 80% was set for each locus. For PCA all available SNPs were included, but for Structure only one SNP of each locus was included to reduce artefacts by linkage. Structure analyses ran for 500,000 generations, with a burn-in of 50,000 generations, for k=2 to k=8 and with five replicates each. In each replicate, one SNP was randomly selected per locus. The replicates were summarized with CLUMPP (Jakobsson & Rosenberg, 2007). The best-fitting k was identified based on ad hoc posterior probability models of [Pr(X|K)] (Pritchard, Stephens, & Donnelly, 2000) and deltaK (Evanno, Regnaut, & Goudet, 2005) using the web version of STRUCTURE HARVESTER (see <http://taylor0.biology.ucla.edu/structureHarvester/>, accessed July 23rd, 2020; Earl & vonHoldt, 2012). The output of the best-fitting k was plotted with the online version of StructurePlot v2 (<http://omicsspeaks.com/strplot2/>; Ramasamy, Ramasamy, Bindroo, & Naik, 2014) sorting individuals by similarity. Coancestry analyses used RADpainter and fineRADstructure (Malinsky, Trucchi, Lawson, & Falush, 2018) closely following the proposed usage (<https://cichlid.gurdon.cam.ac.uk/fineRADstructure.html>, accessed 01.12.2020). The results were plotted using the provided R script FinestructureLibrary.R (<https://github.com/millanek/fineRADstructure/blob/master/FinestructureLibrary.R>, accessed 01.12.2020). One great advantage of the coancestry analysis compared to Structure is the utilization of the complete nuclear haplotypes instead of a single SNP per locus. The vcf files from ipyrad were used as input.

Key population genetic and demographic parameters were calculated from the vcf files generated by ipyrad. Nucleotide diversity, Tajima's D, the inbreeding coefficient F_{IS} and pairwise F_{ST} values were calculated for each locus with VCFtools (Danecek et al., 2011) and then averaged for each lineage and population. Nucleotide diversities were corrected by the total number of nucleotides. Heterozygosity per site per individual was reported by ipyrad and summarized (mean) for each lineage and population. To assess demographic changes over time, extended Bayesian skyline plots (Figure 5) were calculated with BEAST2 for lineages I-III jointly as these probably constitute a single species with inter-lineage hybridization (see discussion). Three runs were performed, one using the COI data set, and two runs each using 200 randomly chosen ddRAD loci (non-overlapping among runs), selecting only loci with 5-10 SNPs each. Loci were retrieved from the *snpsmap ipyrad outputfile. Among loci, the site and clock models were linked, but the tree models unlinked. The HKY model was selected with four gamma categories, empirical frequencies and kappa 2.0. A strict clock with a rate of 1 was enforced, as no suitable substitution rates are available (a few crustacean COI substitution rates have been published, their applicability to deep-sea Isopoda is highly questionable). The "Coalescent Extended Bayesian Skyline" prior was selected for each locus. The weights for "indicatorSampler.alltrees" and "indicatorScaler.alltrees" were set to 5000, of "EBSPupDownOperator.alltrees" to 3000 and of "bitflip.alltrees" to 10000. The MCMC chain was run for 100×10^6 generations, sampling every 5000th for EBSP. The output was analyzed with the EBSPAnalyser included in the BEAST2 package, discarding the first 25 % as burn-in. The final data including the 95% highest probability density intervals was plotted in R.

MALDI-TOF MS

For measurements using MALDI-TOF MS, the same tissue was used from a subset of the genetically studied individuals, always including a single pereopod with the attached muscles. Tissue was incubated in 5 μ l of a matrix solution containing α -Cyano-4-hydroxycinnamic acid (HCCA) as a saturated solution in 50% acetonitrile, 47.5% molecular grade water and 2.5% trifluoroacetic acid. After 5 minutes of incubation, 1.5 μ l of the extract solution was applied to one spot for crystallization on a target plate. Measurements were carried out on a Microflex LT/SH System (Bruker Daltonics), employing the flexControl 3.4 (Bruker Daltonics) software. Measured mass range was set from 2k to 20k Dalton. For peak evaluation, mass peak range from 2k to 10k Dalton was analyzed using a centroid peak detection algorithm, a signal-to-noise threshold of 2, and a minimum intensity threshold of 600, with a peak resolution higher than 400. The Proteins/Oligonucleotide method was employed for fuzzy control with a maximal resolution ten times above the threshold. For a sum spectrum, 200 satisfactory shots were summed up. Three mass spectra were measured for each specimen. Quality control by eye was carried out and mass spectra of inferior quality were discarded. In the following analyses, only specimens with a respective COI sequence were used. For comparison of inter- and intraspecific variance spectra of the congener species *H. foresti* (n = 9), *H. angustus* (n = 4) and *H. hamatus* (n = 5) were used (see Supplemental Table S1).

Data processing was carried out in R (R Core Team, 2020) using R packages MALDIquant (Gibb & Strimmer, 2012) and MALDIquantForeign (Gibb, 2015). Protein mass spectra were trimmed to an identical range from 2,000 to 20,000 m/z and smoothed with the Savitzky-Golay method (Savitzky & Golay, 1964). The baseline was removed based on the SNIP baseline estimation method (Ryan, Clayton, Griffin, Sie, & Cousens, 1988) using 15 iterations. Mass spectra were normalized using the TIC method implemented in MALDIquant. Noise estimation was carried out with a signal to noise ratio (SNR) of 7. Repeated peak binning was carried out with a tolerance of 0.002 in a strict approach and resulting bins were aligned using R package MALDIrppa (Palarea-Albaladejo, McLean, Wright, & Smith, 2018). For the resulting intensity matrix, missing values were interpolated from the corresponding spectrum. All signals below a SNR of 1.75 were assumed to be below detection limit and set to zero in the final peak matrix. This matrix was Hellinger transformed (Legendre & Gallagher, 2001) for further use. Intra- and interspecific Euclidean distances were calculated using 'vegdist' from R package vegan (Oksanen et al., 2013). To test group differentiation for classification approaches and to assess mass peak importance for group differentiation a Random Forest (RF, Breiman, 2001) analysis was carried out using R package randomForest (Liaw & Wiener, 2002, ntree = 2000, mtry = 35). Significant deviation from random of the observed model errors was calculated with the function MVSF.test from package

RFTools (<https://github.com/pmartinezarbizu/RFtools>) (Rossel & Martínez Arbizu, 2018). Significance of differences were tested using the distance-based multivariate analysis of variance (W*d) developed by Hamini et al. (2019).

Species distribution models (SDM)

We modelled the distribution of the three main groups retrieved by the genetic analyses (I-III, IV and V) based on major environmental factors. Random Forest classification models (Breiman, 2001) were calculated in R (R Core Team, 2020) using the package randomForest (Liaw & Wiener, 2002). The models were based on 2000 random trees and 1/3 of all variables randomly sampled at each split (Liaw and Wiener, 2002).

Two types of models were calculated: a.) a multiclass model in which all 3 groups (I-III, IV and V) are present and the model decides on the most probable class, and b.) a separate model for each group in which the model computes the probability of presence and absence for each group separately. For type b), a presence/absence matrix was produced for the groups I-III_{RAD}, IV_{RAD} and V_{RAD}. To avoid bias towards the absence class (which is most common), each tree was calculated with the same number of absence (randomly chosen) and presence. Predictor layers including the major environmental forces structuring the area (depth, bottom water salinity and temperature, bottom water oxygen, and particulate organic carbon flux (POC)) were downloaded from the Global Marine Environmental Dataset (GMED) (<http://gmed.auckland.ac.nz>). Probability of occurrence was predicted using the resulting models on a dataset containing 88,822 geographic locations in the study area, after excluding locations situated on land or those lacking values in one or more predictor variables. Significant deviation from random of the observed model errors was calculated with the function MVSF.test from package RFTools (<https://github.com/pmartinezarbizu/RFtools>).

Results

COI and ddRAD datasets

211 COI sequences of *H. bicuspis* were successfully sequenced. The final alignment was 530 base pairs (bp) long, of which 132 bp were parsimony-informative. Stop codons or indels were not observed. 52 of 78 sequenced ddRAD libraries were of sufficient quality for further analyses. The two ddRAD datasets comprised 1723 and 2993 loci with 18,999 and 21,697 bp, and the vast majority of loci yielded one or more SNPs (Table 1).

Genetic diversity and species delimitation

The data from COI and ddRAD revealed deep genetic divergences within *H. bicuspis* that might correspond to multiple speciation events. In COI, *H. bicuspis* was split into several lineages (= putative species), with GMYC suggesting seven and ABGD four lineages based on threshold values between 0.8 - 2.1 % uncorrected *p*-distance. Based on the distance distribution, we decided to work with the following five lineages: I_{COI} (east to south-east of Iceland; including station 1159 which was separated by GMYC), II_{COI} (east of Iceland, close to the shelf; I_{COI} and II_{COI} were combined in ABGD), III_{COI} (north of Iceland), IV_{COI} (south and south-west of Iceland; split into two syntopic lineages by GMYC), and V_{COI} (west and south-west of Iceland) (Figure 1, Table 2). None of these five lineages were syntopically recorded, and most lineages were clearly geographically separated. Only lineages IV_{COI} and V_{COI} occurred sympatrically along the Reykjanes Ridge. Lineage V_{COI} was genetically the most divergent lineage with uncorrected *p*-distances of 4.9 - 7.3 % to the other lineages.

In the nuclear ddRAD data five lineages could also be delimited, however, the assignment of populations north of the GIF Ridge differed in some important aspects (Figure 3 Table 3). To better discern these lineages north of the GIF Ridge, separate ipyrad analyses were performed for lineages I-III, including PCA, Structure, and coancestry analyses. I_{RAD} includes only south-eastern populations along the GIF Ridge towards the Faeroe Islands. Stations 1159 and 1172, which were also assigned to I_{COI}, clustered with the near-shelf station 1194 into lineage II_{RAD}. The latter station had formed II_{COI} together with station 1219, however, station 1219 clustered in ddRAD with the northern stations (corresponding to III_{COI}) into III_{RAD}.

As a consequence, individuals from several populations were assigned to different lineages depending on the studied markers. This was most notable at station 1219, whose individuals were assigned to II_{COI} and III_{RAD}, respectively. However, Structure analyses assigned a relatively large fraction of ~34 % of their nuclear genome to I_{RAD} (nearly as much as to III_{RAD}), but less than 3 % to II_{RAD} (Figure 3), suggesting hybridization among all three lineages. In the fineRADstructure analyses of all lineages this was not as evident, however, in the separate analyses of lineages I_{RAD}-III_{RAD}, individuals from station 1219 exhibited the highest shared coancestry with III_{RAD} followed by I_{RAD}. This putative hybridization might explain why I_{RAD} and III_{RAD} were not differentiated in the PCA that included all individuals, but only in the analyses focusing on lineages I-III (Figure 3). South of the GIF, lineages IV_{RAD} and V_{RAD} were delimited identically to IV_{COI} and V_{COI}, respectively, without any indication of hybridization or gene flow among these two lineages or with any of the other lineages (Figure 3).

Only two fully matured males were observed (stations 880 and 873; corresponding to lineages I_{COI} and I_{RAD}). Both featured the wider second antennular segment typical for *H. b. bicuspis*, as described by Wolff (1962). Pleopod 1 of one specimen had the two distinct corners of *H. b. bicuspis*, while the other had a more widely rounded corner, indicative of *H. b. tepidus* (Figure 2). The latter specimen thus features a mix of characters of both subspecies. All other males were juvenile or preparatory, thus pleopod morphologies were non-informative. Their antennula were mostly resembling *H. b. bicuspis*. Our results indicate a developmental explanation for the different male types described by Wolff (1962).

Proteomic differentiation

In total, 96 *H. bicuspis* specimens were used in the MALDI-TOF MS analysis. Based on the COI sequences, 56 specimens were assigned to *H. bicuspis* I, 7 to *H. bicuspis* II, 11 to *H. bicuspis* III, 9 to *H. bicuspis* IV and 13 to *H. bicuspis* V. Intra-specific Euclidean distances (pooling all lineages for *H. bicuspis*) ranged from 0.56 (10% quantile) to 0.83 (90% quantile), while inter-congener distances ranged from 0.93 (10% quantile) to 1.05 (90% quantile) with very little overlap. Lineage-specific distances were all in the same range with no distinct differences, and a small tendency was found with a higher inter-lineage distance of I-III vs IV than the intra-lineage distances of I-III (Figure 3). The Robust Distance-based Multivariate Analysis of Variance (Wd*) (Hamidi et al. 2018) revealed significant differences between lineages I-III, IV, and V ($p < 0.001$). The principal component analysis (PCA) (Figure 3) of the processed data shows two major groups. One comprises the specimens belonging to *H. bicuspis* group I-III (Figure 1, blue) from the North of Iceland, and the other includes specimens from sampling sites South of Iceland (Figure 1, beige and green). Based on classification votes, the TSNE plot constrained to predefined groups in a RF classification model (Figures 6-10) supports a differentiation of these regional groups based on proteome data. In a classification approach, 98.6% of group *H. bicuspis* I-III would be identified correctly with one specimen (ZMH K-58552) being classified as *H. bicuspis* IV. Similarly, one specimen of *H. bicuspis* IV (ZMH K-58494) would be classified as *H. bicuspis* I-III (11.1%), and three as *H. bicuspis* V (33.3%) (ZMH K-58496, ZMH K-58577 and ZMH K-58579). Of *H. bicuspis* V, 92.3% would be classified correctly with one specimen (ZMH K-58521) being assigned to IV. By investigating the most important variables within the RF model based on the highest decrease in Gini Index, peaks were identified that show group specific behavior.

Haploniscus bicuspis IV and V show shifts (in the range of 40-50 Daltons) in larger proteins compared to *H. bicuspis* I-III, while *H. bicuspis* V and IV were mainly separated by the relative expression of proteins with m/z of 2400, 4407 and 2680 (Figure 4). Whereas the mere presence or absence of these peaks would probably not be sufficient to distinguish between groups, relative peak intensities differ consistently between groups.

Phylogeography and Population Genetics

In most populations, genetic diversities were not pronounced (Table 4; Figure 2). A few COI haplotypes are particularly common, and the majority of the other haplotypes differ by one or two mutations from one of the common haplotypes in each population and lineage (Figure 2). Most populations feature only a few COI haplotypes. This is the case especially within lineages I_{COI}, II_{COI} and III_{COI} as well as the populations of IV_{COI} and V_{COI} along the Reykjanes Ridge. A notable exception is population 1072 (lineage V_{COI}) to

the west of Iceland, where nine haplotypes were observed. Nucleotide diversities were roughly an order of magnitude larger in the ddRAD data compared to COI (Table 4). Statistics differed slightly between the two datasets for populations and lineages in I_{RAD}, II_{RAD}, and III_{RAD}, but not extensively (see Table 4). Therefore, we will focus only on the analyses of the dataset which included all five lineages.

Observed heterozygosity across all ddRAD sites was close to 0.1 for all lineages and populations (Table 4). The level of inbreeding differed among lineages. Within lineages I_{RAD}-III_{RAD} and IV_{RAD}, the inbreeding coefficient F_{IS} was [?] 0.14 for each population, but 0.18 for V_{RAD} (Table 4). Due to the Wahlund effect, F_{IS} increased when populations were grouped into lineages I_{RAD}, II_{RAD} and III_{RAD}, or all together (up to 0.27).

Most lineages appear to be geographically restricted and separated from each other. Lineage I_{COI} is the most widely distributed lineage ranging from the south-east to the north-east of Iceland around a distance of 900 km. However, the population farthest into the Nordic Seas (station 1159) is well-differentiated from the others with at least five mutations separating the observed COI haplotype from all other haplotypes (Figure 2). Among and within all lineages, populations are strongly differentiated from each other. In the nuclear ddRAD data, population differentiation within lineages is not well resolved in the PCA and Structure analyses (Figure 3), however, the coancestry analyses with fineRADstructure grouped most individuals by populations within lineages, suggesting low but detectable levels of population differentiation (Figure 3). Population differentiation is particularly strong in COI, where the vast majority of populations are differentiated by significant F_{ST} values larger than 0.7, often over 0.9 (Table 5). High and significant F_{ST} values were also observed among several geographically close populations (<100 km apart), e.g. station 879 compared to 880 or 881. A few instances of geographically distant but genetically similar populations in COI were observed as well: 869 compared to 873 (210 km; lineage I_{COI}), 880 and 881 compared to 1172 (~450 km; lineage I_{COI}), and to a lesser degree 83 compared to 1072 (431 km; lineage V_{COI}) (Table 5). ddRAD F_{ST} values were lower with all values within and among populations of lineages I_{RAD}-III_{RAD} being [?] 0.29 (Table 5). Isolation-by-distance was observed in COI, though only a relatively small fraction of the genetic differentiation was explained by geographic distance (correlation coefficient $r_{Y1} = 0.56$; $p = 0.001$; determination of genetic differentiation by geographic distance = 31.3 %), but not in the ddRAD data (correlation coefficient $r_{Y1} = 0.36$; $p = 0.15$; determination of genetic differentiation by geographic distance = 13.2 %). The high levels of genetic differentiation are already observed at low geographic distances, which may explain the low correspondence between genetic differentiation and geographic distance.

The discordance between mitochondrial and nuclear data impacts the inferred geographic distribution of genetic lineages. It is noteworthy that I_{COI} extends much farther to the north-east than I_{RAD}, whereas II_{COI} is much more restricted to the near shelf of Iceland than II_{RAD}, which extends farther to the east into the Arctic Ocean. The difference at station 1219 is the most pronounced, which has been assigned to II_{COI}, but in ddRAD appears to be a hybrid between I_{RAD} and III_{RAD}, (Figure 2), which otherwise occur farther south or north, respectively.

Demographic parameters (Tajima's D and Fu's F_s) are slightly negative for most populations and lineages in COI. Notable exceptions are IV_{COI} (Tajima's D and Fu's F_s positive) and I_{COI}-III_{COI} jointly analyzed (Fu's F_s is positive). In the nuclear data set, most single populations had positive values close to zero (thus also lineages IV_{RAD} and V_{RAD}), however, lineages I_{RAD}, II_{RAD}, III_{RAD} and I_{RAD}-III_{RAD} each had negative values ranging from -0.13 to -0.59. The three EBSP analyses yielded slightly diverging outcomes concerning the long-term trends. While in COI a slightly decreasing population size over time was suggested, one of the ddRAD-based EBSPs suggested long-term stability and the other a temporal increase in population size. However, all three analyses suggest that population sizes were strongly reduced shortly before the present followed by a rapid recovery of population sizes, likely an expansion following a bottleneck (Figure 5). It is difficult to estimate when this bottleneck may have occurred due to the lack of a conclusive specific substitution rate for isopods.

Species Distribution models

The multivariate structure test performed with the MSVF.test function on the multiclass model containing all three groups reveals that all classes have an observed error, which is significantly lower ($p < 0.05$) than random error (null model). This means that there is significant multivariate structure to differentiate between the three lineages using the chosen predictor variables. The multiclass model had an out-of-bag training error of 4.35%. All 161 instances of lineage I-III were correctly classified (class error = 0). Lineage IV had an error of 6.6% in which one instance out of 15 was wrongly assigned to lineage V. Lineage V had an error of 25.8%, in which eight instances out of 31 were wrongly classified as lineage IV.

The prediction of the winner class of the model is depicted (Figure 6). Lineage I-III is distributed in the northern part of Iceland, from the shelf regions to the abyssal plains, having its southern distribution barrier along the GIF Ridge with an additional predicted occurrence south of the GIF Ridge. Lineage IV is present in the upper slope and shelf regions south of the GIF Ridge and extends into the Reykjanes Ridge at depths of around 1000 m. It is possible that this lineage extends into the deeper areas of the North Sea and further to the north along the coast of Norway. Lineage V extends from the abyssal plains of the North Atlantic to the north until a depth of around 1000 m to 2000 m, and is absent in the upper slope and shelf regions. Small disjunct areas are located on the west coast of Norway. The error rates of the presence/absence models were 0%, 6.7%, and 2.4% for lineages I-III, IV and V, respectively. The predictions for the single group models are depicted in Figures 7-9. Lineage I-III (Figure 7) shows basically the same predicted distribution as in the multiclass model. Lineage IV, however, displays an important extension into the abyssal areas in the North Atlantic, but with low probability of 0.5 to 0.6. On the other hand, lineage V shows an extension into the slope areas south of Iceland up to the 500 m isobath. The single lineage models show considerable overlap in the probability of occurrence between lineages IV and V in the slope areas. In this model, lineage V is not predicted to be present in the west coast of Norway.

The niche partitioning of the groups within the water masses is depicted in the temperature-depth plot (Figure 10). Lineage I-III occupies the northern water mass characterized by low temperature across all depths. This lineage is mainly distributed at bottom water temperatures below 2 degC. Lineage V is mainly distributed in southern water masses at temperatures between 2 degC and 3 degC and depths below 1,000 m. Lineage IV is mainly distributed at bottom water temperatures above 3 degC and depth above 3,000 m, with its core distribution above 1,000 m depth.

Discussion

Cryptic diversity and species distributions

Assessing biodiversity is essential to understand evolutionary and ecological processes and to observe changes in the distribution of species as a consequence of climate change and other anthropogenic stressors. A fundamental unit of biodiversity are species, though intraspecific genetic diversity is also considered a crucial factor (Błażewicz, Józwiak, Menot, & Pabis, 2019; Laikre, 2010; Brix et al., 2020). Delimiting species can be challenging in cases where differences between putative species are small and intraspecific variability and interspecific variation are not clearly demarcated (Kaiser et al., 2021). In the case of *Haploniscus bicuspis*, several genetically divergent lineages were recovered, of which some probably constitute separate, morphologically cryptic species. Commonly observed intraspecific COI distances within deep-sea isopods were usually below 6% (e.g., Brix, Riehl, & Leese, 2011; Brix et al., 2018a, 2018b; Havermans et al., 2013; Brandt, Brix, Held, & Kihara, 2014; Bober, Riehl, Henne, & Brandt, 2018), but intraspecific distances of more than 8% have been reported (e.g., Riehl, Lins, & Brandt, 2018). Conversely, interspecific COI distances as low as 4.6% were reported (Riehl et al., 2018), possibly suggesting that some of the larger intraspecific distances may be due to cryptic diversity, but commonly exceed 10% (e.g., Havermans et al., 2013; Brandt et al., 2014; Brix et al., 2018a, 2018b). In the following, we will discuss whether the five lineages might constitute different species.

Lineages I, II and III ($I_{\text{COI-III}_{\text{COI}}}$ and $I_{\text{RAD-III}_{\text{RAD}}}$) are consistently delimited from lineages IV and V in COI and nuclear loci, and with a few exceptions also in the proteomic data. However, these three lineages are not consistently delimited from each other. There is a clear discordance between the mitochondrial

and nuclear data, with individuals being assigned to different lineages (among I, II and III) based on the respective analyses. Our data also strongly suggests multiple instances of inter-lineage hybridization and gene flow most notable at station 1219 (see Figure 1) where all three lineages appear to have hybridized. The genetic distances in COI among these lineages are also smaller than among most other deep-sea isopods. We therefore suggest that these three lineages constitute a single species (referred to as *H. bicuspis*I-III in the following).

Lineages IV (IV_{COI} and IV_{RAD}) and V (V_{COI} and V_{RAD}) are consistently differentiated from each other and from *H. bicuspis* I-III in COI and nuclear loci. COI p -distances between lineage V and all other lineages exceed commonly observed intraspecific distances within deep-sea isopod species. For lineage IV, COI p -distances are lower in comparison to *H. bicuspis* I-III than commonly observed among deep-sea isopods. Furthermore, the p -distances are comparable to those observed within *H. bicuspis* I-III. However, in the nuclear loci, lineage IV is the most divergent. No instances of hybridization have been inferred for lineages IV or V. Distribution modelling suggests that their habitats are ecologically differentiated from those of *H. bicuspis* I-III and at least partially differentiated from each other. We therefore tentatively suggest that lineage IV and V constitute two morphologically cryptic species within *H. bicuspis* (termed *H. bicuspis* IV and *H. bicuspis* V in the following).

Similar patterns of cryptic diversity have been observed for other Icelandic Isopoda, which were previously assumed to be widespread species (Brix et al., 2014; Jennings, Brix, Bober, Svavarsson, & Driskell, 2018). The geographic distribution of the different putative *H. bicuspis* species fit well into the distribution patterns of other benthic deep-sea isopods around Iceland (Brix et al., 2018). Our discovery that the waters around Iceland alone may be inhabited by three cryptic species within *H. bicuspis* strongly questions the occurrence of *H. bicuspis* in the South Atlantic Ocean (Brökeland and Wägele, 2004; Menzies, 1962). The presence of further cryptic diversity in this presumably widespread species is highly likely. We deliberately did not perform molecular clock analyses to date the differentiation of the herein observed species, because evolutionary rates for deep-sea crustaceans are still lacking. A recent study suggests that the commonly applied rate of 1.4% per million years (Knowlton & Weigt, 1998) is too conservative and rather suggest a K2P divergence of 5-5.2% per million years for arctic marine invertebrates (Loeza-Quintana et al., 2019). If such rates are applicable to deep-sea benthic peracarids, such as *Haploniscus*, the divergence between the three putative species started roughly half a million to a million years ago. This is relatively young and suggests recent and putatively ongoing speciation processes.

The distribution of the three species appears to be largely governed by water masses and associated ecological parameters. In particular, the GIF Ridge represents a crucial barrier separating species. Whether the GIF Ridge is a physical barrier that hinders dispersal or just separates water masses and thereby shapes species distributions remains unanswered. The GIF Ridge has been discussed as an isolation barrier for many isopod species in the North Atlantic, for instance anthuridean isopods (Negoescu & Svavarsson, 1997), valviferan isopods (Stransky & Svavarsson, 2006), and desmosomatid and nannoniscoid isopods (Brix & Svavarsson, 2010). *Haploniscus bicuspis* sampled in this study were found between 316 m depth at station 1136 and 2422 m depth at station 1172, therefore they can easily cross the saddle depth of the GIF Ridge. Previously, *H. bicuspis* have been collected in the Arctic, even in shallow waters of 198 m depth (Brandt, 1997).

Ocean temperature is highly variable around Iceland. North of the GIF Ridge, bottom water temperature can be as low as -0.9 °C, and it only becomes warmer as it reaches the shelf (up to about 3 °C, Jochumsen et al., 2016). In contrast, south of the GIF Ridge, the North Atlantic Water can reach up to 10.5 °C (Hansen & Østerhus, 2000). Another environmental variable that may limit species distribution is the sediment structure, which was previously found to be important in peracarid crustacean distribution around Iceland (Stransky & Svavarsson, 2010). Ostmann et al. (2014) found that sediment characteristics vary in the surrounding waters of Iceland, with coarser sand found around the Reykjanes Ridge and more silt and clay found in the deep-sea east of the GIF Ridge. Oxygen concentration may also influence migration and gene flow between isopods, especially in the deep sea. Expanding oxygen minimum zones may have contributed to allopatric speciation in the past (White, 1988).

The distribution models provide interesting hypotheses that should be tested in future studies. These include the hypothesized distribution of *H. bicuspis* IV (and partly V) along the Norwegian coast, the occurrence of *H. bicuspis* I-III south of the GIF Ridge off the coast of Greenland, and whether *H. bicuspis* IV, V, or both occur in the abyssal plains south of Iceland. The former would require dispersal along a narrow habitable corridor along the north-west of Great Britain, whereas the expansion of *H. bicuspis* I-III south of the GIF Ridge would require crossing a region with unfavorable ecological conditions. This could have been mediated by transport via the north to south overflow of arctic deep-water. Whether *H. bicuspis* IV, V, or both occur on the abyssal plains of the North Atlantic south of Iceland cannot be answered with the available data. The ecological conditions appear suitable for both (possibly slightly better for IV), and sympatric occurrences of these two species were already observed for the Reykjanes Ridge. By examining additional populations of *H. bicuspis* genetically, future studies may shed light on these questions.

A comparison of proteomic profiles revealed differences between specimens from the northern (I-III) and southern (IV and V) species, and also between the two southern species IV and V. However, differentiation between species IV and V was less distinct. Differences between species are very small and occur mainly in a few relatively low-expressed proteins, which show differences in presence and relative intensities. Similarly, high similarity in proteomic patterns and only a small number of differentially expressed proteins were observed in cryptic mosquito species (Müller et al., 2013; Dieme et al., 2014). Like *H. bicuspis*, peaks in the mosquito *Anopheles gambiae* Giles, 1902 species complex were shifted by only a few Daltons, potentially reflecting few amino acid substitutions within proteins or minor post-transcriptional modifications (Müller et al., 2013). These small shifts in masses of specific proteins may thus infer polymorphisms of these molecules in *H. bicuspis*. Differences between the putative species were at least partly caused by a prevalence of a certain allele in the respective putative species. Despite apparent gene flow interruptions, differences between IV and V were less pronounced. Proteomic profiles depend on physiological responses to the environment, e.g., variations in proteomic profiles varied with season and habitat in ticks (Karger, Bettin, Gethmann, & Klaus, 2019). Thus, it remains unanswered whether the higher similarity of putative species IV and V results from very similar environmental impacts and thus comparable selection pressures in the past causing a similar physiological response. We interpret the high similarity between the different *H. bicuspis* species found in this study as a reflection of recent or ongoing speciation processes, as reported in fishes (Takács et al., 2014, Maasz et al., 2020). It is likely that young species are both morphologically and physiologically cryptic, revealing the limit of proteomic fingerprinting for species identification in such instances.

Haplonicus bicuspis was first described from the Norwegian Sea (Sars, 1877) north of the GIF Ridge and a few hundred kilometers east of our sampled area, but no genetic information is currently available for Norwegian representatives. Given the species' distribution around Iceland, we suggest that *H. bicuspis* I-III represents the true *Haplonicus bicuspis* (Sars, 1877) because the known and modelled distribution of *H. bicuspis* I-III extends well into the waters east of Iceland towards the Norwegian Sea and would correspond to the subspecies *H. b. bicuspis* described by Wolf (1962). However, Wolff differentiated the two subspecies based on the morphology of pleopod I, and both variants were found in *H. bicuspis* I-III (unfortunately no adult males were available for the other two putative species). In juvenile males the respective section of the male pleopod I is smooth and narrow and widens as they mature into preparatory males (see also Wolff 1962). It is plausible that the different pleopod I shapes assigned to the two subspecies are in fact successive developmental stages, with the distinct corner ascribed to *H. b. bicuspis* being the later stage. Even though pleopod I morphology might not be useful to distinguish Wolff's subspecies, *H. b. tepidus* might still represent a valid species (*H. tepidus*). Whether this corresponds to *H. bicuspis* IV or *H. bicuspis* V cannot be answered with the data currently available, as *H. b. tepidus* was described from the Reykjanes Ridge.

Phylogeography and population genetics

Historically, glacial cycles probably had a strong influence on species distributions. Most of Iceland's shelf was covered by an ice sheet during the Last Glacial Maximum (~25 ka), which broke up around 15 ka due to a northward shift of the polar front and rising sea levels (reviewed by Geirsdóttir et al., 2009). At that

time, today's pattern of currents was established. The near-shelf populations of *H. bicuspis* (represented by stations 1019, 1194 and 1219) were likely colonized following the break-up of the Icelandic ice shield. Deep-sea populations farther off the coast may have also been affected by changing temperatures, currents, etc. The observed demographic history for *H. bicuspis* I-III suggests a relatively recent decline in population sizes followed by a rapid recovery, though we do not have any age estimates (Figure 5). This overall pattern might reflect population declines and local extinctions during the Last Glacial Maximum. This was then followed by population expansions and recolonizations from unaffected deep-sea regions farther off the coast during the Holocene, which were scarcely sampled in our study. Colonization from less affected regions might also explain the rapid recovery of population sizes.

Our data suggests a complex colonization history of the near-shelf regions north of the GIF Ridge for all three lineages observed within *H. bicuspis* I-III. The consistently observed differences between the mitochondrial COI and nuclear ddRAD data imply differing migration behavior between males and females. The maternally inherited COI exhibited consistently higher levels of genetic differentiation between populations, which suggests males migrate more actively and potentially over wider distances, while females appear to be rather stationary. This is surprising, as *Haploniscus bicuspis* does not exhibit a strong sexual dimorphism in swimming structures as observed in other isopod families (e.g., Brix et al., 2020; Bober et al., 2018; Hessler, 1970; Riehl, Wilson, & Hessler, 2012). However, adult males of *H. bicuspis* have stouter second antennae with many more sensory sensilla and aesthetascs than females or juvenile and preparatory males. This difference was not described by Brökeland & Wägele (2004), as they did not have adult males available in the type series. The aesthetascs are hypothesized to have a chemosensory function, possibly enabling males to detect females across large distances. This suggests that males have a more active lifestyle, roaming around in search for females, which would explain the inferred sex-specific migration patterns and the different distribution patterns of COI and ddRAD lineages. As stated in Brix et al. (2020), locomotion of the adult stages does influence migration patterns and distance for Pacific isopod families, but also a “male behavior” is reflected in the Pacific data for Haploniscidae.

The ‘central’ region north of the GIF Ridge around stations 1194 and 1219 was probably first colonized by lineage II, possibly originating from close-by populations in the Norwegian Sea to the east or north-east of Iceland. Subsequently, males from northern (lineage III) and southern (lineage I) populations migrated into the area around station 1219, largely replacing the local population. The result is a local population at station 1219 that comprises a hybrid nuclear genome of lineages I and III, but the mitochondrial genome of lineage II. These newly arriving males must have outnumbered the local (male) population or were better adapted to the local environmental conditions, or both. We propose a similar scenario for the north-eastern populations around stations 1159 and 1172. These were probably first colonized by lineage I, possibly from the central or southern Norwegian Sea, with subsequent male-biased migration and introgression from lineage II, resulting in populations with lineage I mitochondrial and lineage II nuclear genomes. The reverse, a female-biased migration and introgression, is highly unlikely in both cases due to the overall higher rates of population differentiation in COI (suggesting lower female migration rates) and because in such a scenario these females should have contributed to the nuclear genome as well.

Overall, our results show that genetic differentiation between populations is usually high, even at comparably low geographic distances, suggesting low dispersal and gene flow rates. This is also supported by elevated inbreeding coefficients in some populations, implying that reproduction occurs frequently among closely related individuals. The above-described hybridization scenarios appear to represent rare events, which probably occurred at a time when local population densities were still low in the early phase of recolonization. The strongly male-biased migration behavior is noteworthy, with the only apparent sexual dimorphism in sensory structures. If such patterns hold true for other asellote isopods, population differentiation inferred from mitochondrial markers like COI might systematically underestimate dispersal and gene flow rates.

Haploniscus bicuspis IV is the only putative species with consistently positive values in Tajima's D and Fu's Fs, suggesting that it is currently undergoing a genetic bottleneck. At least in COI, this putative species appears closer related *H. bicuspis* I-III, which occurs exclusively in colder arctic waters. It is possible that

H. bicuspis IV is also better adapted to colder temperatures and is negatively affected by the northward movement of the Arctic front throughout the Holocene. The ranges of *H. bicuspis* IV and V meet along the Reykjanes Ridge, and we found *H. bicuspis* V at most stations, but with less genetic diversity among those found at the Reykjanes Ridge stations. Whether both species coexist due to the topological, structural, and environmental diversity of the Reykjanes Ridge (German et al., 1994; Taylor et al., in review) or whether *H. bicuspis* V is currently replacing *H. bicuspis* IV is an unanswered and interesting question.

Conclusion

Iceland’s marine environment is comparable to a “natural laboratory” that has shaped species’ distribution depending on a variety of factors. *Haplonicus bicuspis* is an Icelandic example of a benthic deep-sea species that was previously assumed to occur in all water masses around Iceland, but we have now found that there are cryptic species found within. *Haplonicus bicuspis* I-III is indicated as the “true *bicuspis*” North of Iceland. Whether species IV or V corresponds to *H. tepidus* needs to be evaluated, and they require a thorough redescription after the initial distinction by Wolff (1962).

Speciation processes in these benthic deep-sea species were probably driven by the interaction of geographic separation via the GIF Ridge with the associated ecological differences and past climatic changes. Genetic bottlenecks, potentially resulting from ice ages, were followed by population expansions into previously glaciated regions. Migration is male-biased, which results in hybrid zones and the complete exchange of the local nuclear genome in a few instances. Population genetics indicate mostly high levels of inbreeding and population differentiation, despite such cases of extensive male-biased dispersal. Such patterns may be typical for deep-sea benthic peracarid species, which are a crucial component of the fauna.

Acknowledgements

Financial support for the expeditions was given from the German Science Foundation to SB (BR3843/4-1,5-1 and MerMet17-06) with support of the shiptime on RV Sonne by the BMBF (Bundesministerium für Bildung und Forschung, Germany). This work was also supported by the DFG initiative 1991 ‘Taxon-omics’ [grant number RE2808/3-1 and RE2808/3-2]. We are happy about the excellent support of the sampling during the expeditions by the crews of the research vessels RV Meteor, RV Poseidon, RV Maria Sybilla Merian and RV Sonne. The processing of the samples would not have been possible without the technical support of Karen Jeskulke, Nicole Gatzemeier and Antje Fischer during the sorting and data management at the DZMB. Kathrin Phillipps-Bussau provided the final museum numbers from the CeNak. Thanks to Michael Sheridan for checking our English grammar. Thanks to Jørgen Olesen and the Museum in Copenhagen for allowing access to the Wolff material of *H. b. bicuspis* and *H. b. tepidus*. Terue Kihara, Karlotta Kürzel and Alex Kienecke are thanked for making it possible to complete the CLSM pictures in Wilhelmshaven to illustrate the male developmental stages. This is publication 77 of the Senckenberg am Meer Molecular Laboratory, XX of Confocal Laboratory and 13 of Proteomics Laboratory.

Data Accessibility and Benefit-Sharing Statement

For all DNA sequences, Genbank accessions numbers are available (XXXXX so far – submitted, see supplementary table1!). Final DNA sequence assembly will be uploaded to DRYAD as well as the Random Forest input files and proteomic fingerprints. A doi will be created for the BoLD project containing all information about sampling locations and COI data as well as morphological identifications.

Author Contributions

SB organized the expeditions and sampling. SB and MS devised the project and supervised the theses of AS and EP. Species identification was performed by SB and JS. The COI data was generated by AS and EP, the ddRAD data by EP, the morphological data by AS and SB; the MALDI-TOF data by EP, SB and SR. The genetic data was analyzed by MS and EP, the MALDI-TOF data by SR and JP. PMA conducted the distribution modelling. MS, EP and SB drafted the first version of the manuscript, all authors contributed significantly to the final version.

Figure Captions

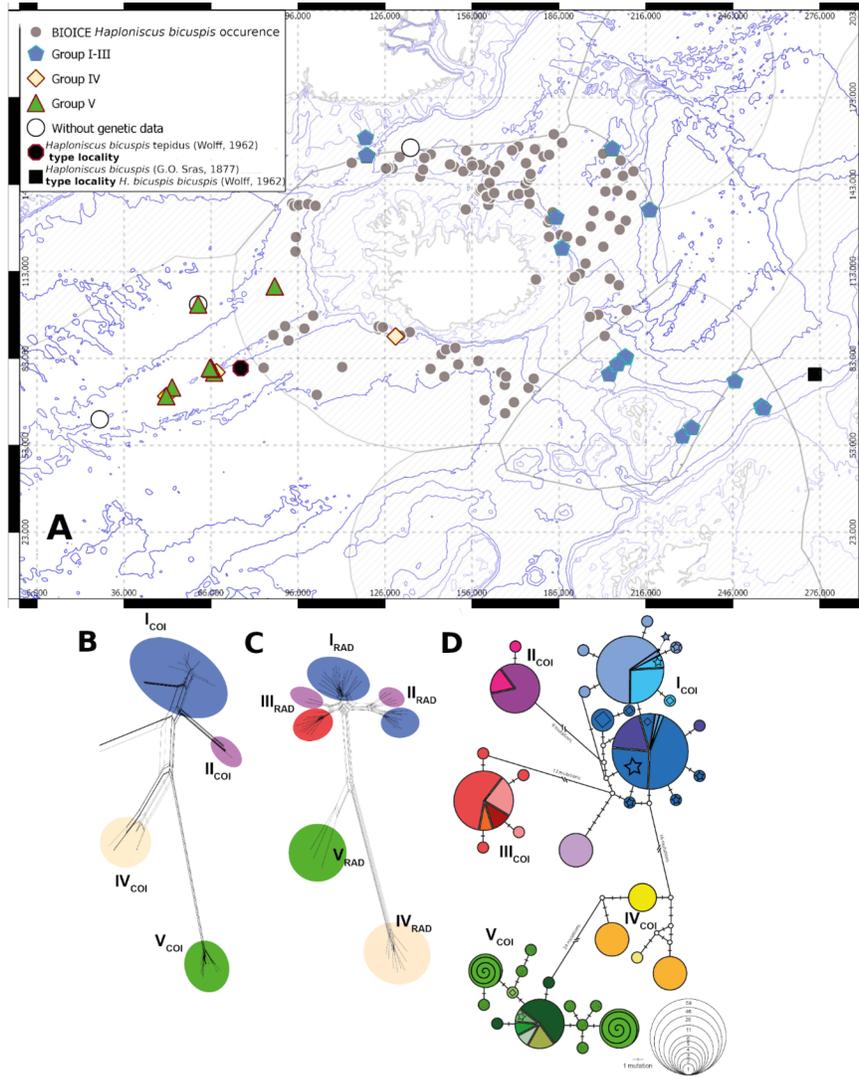


Figure 1 : A) Distribution map of *Haploniscus bicuspidus* sampled around Iceland from BIOICE, IceAGE and literature data. Black symbols refer to the type locality by Sars (1877) and Wolff (1962). B)-D) Unrooted genetic networks. B) COI haplotype network calculated with SplitsTree, C) network based on nuclear ddRAD loci calculated with SplitsTree and D) COI haplotype network calculated with Network. In B) and C), colors indicate lineages delimited by COI (highlighting the conflicts in the nuclear dataset). In D), colors represent single stations (see Figure A). Individuals are not shown to improve the readability of the figure, please refer to Supplemental Figure S1 for that.

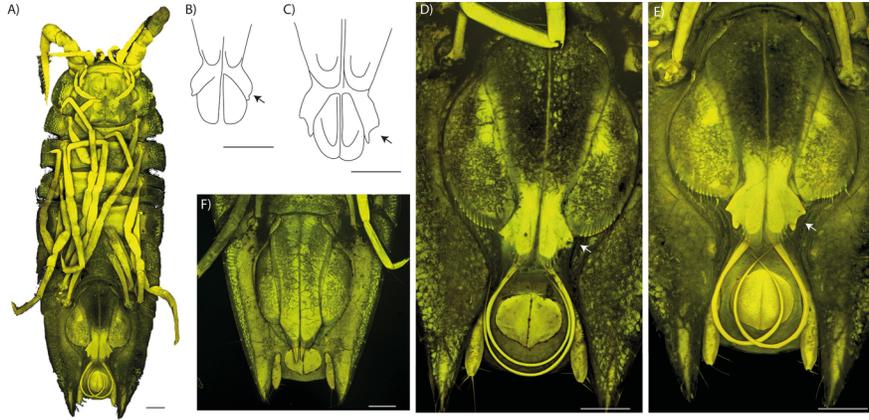


Figure 2 : Morphology of male pleopod 1 for (sub)species delimitation. A) ventral view of complete adult male (ZMH K-60856, lineage I; head facing upwards), B) male pleopod 1 of *H. bicuspis tepidus* , C) male pleopod 1 of *H. bicuspis bicuspis*, D) and E) pleopod 1 of adult males (ZMH K-60196 and ZMH K-60856, both lineage I) and F) pleopod 1 of juvenile male (ZMH K-60066, lineage I). C) and D) redrawn after Wolf (1962). Arrows indicate the distal part of pleopod 1, which is supposed to show species-specific differences (laterally rounded in *H. b. tepidus*, distinct corner in *H. b. bicuspis*) . A), D), E) and F) were taken with a CLSM. Scale bar = 100 μm

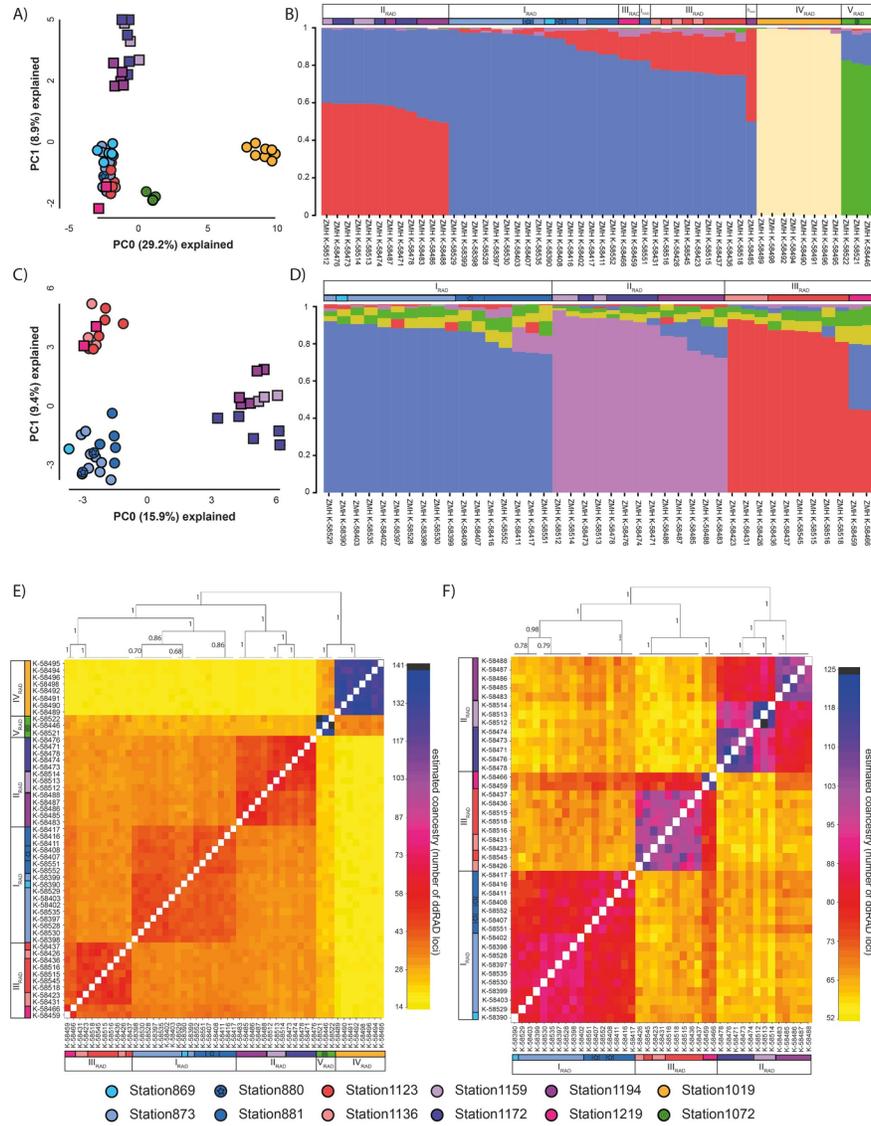


Figure 3: PCA, Structure and coancestry analyses. A), C) PCA plots, B), D) Structure plots and E), F) clustered fineRADstructure coancestry matrix based on nuclear ddRAD data. A), B) and E) includes all *H. bicuspis* individuals; C), D) and F) include only individuals assigned to lineages I-III. Colors represent single stations (see Figure 1).

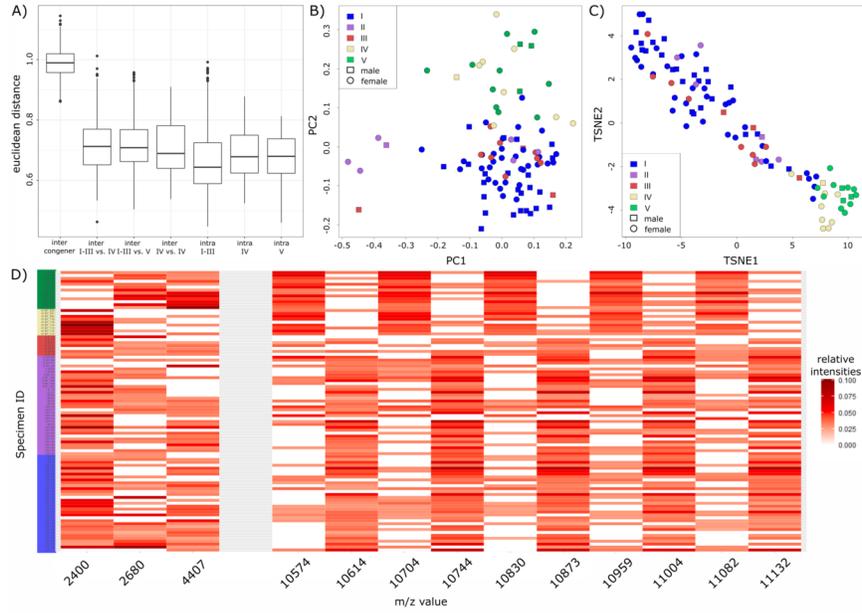


Figure 4: A) Within and between group euclidean distances based on MALDI-TOF mass spectra. B) PCA of hellinger-transformed mass spectra. C) TSNE plot of RF model generated with three classes (I-III, IV and V). Specimens in B) and C) are colored according to the respective groups. Symbols indicate the respective sex. D) Relative intensities of peaks which are important for group differentiation. A mass shift between I-III and IV-V was observed in peaks >10kDa.

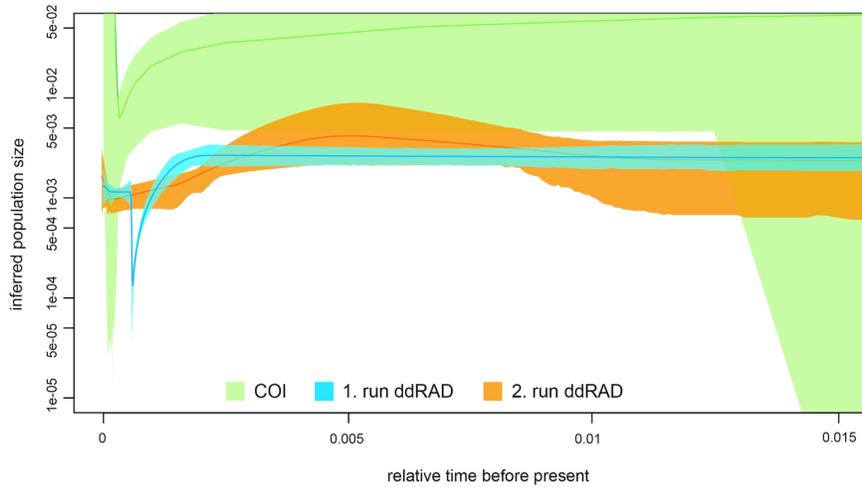


Figure 5: Extended Bayesian Skyline Plots (EBSPs) of *H. bicuspis* I-III. EBSPs were generated for the COI data set as well as for two nuclear ddRAD-based data sets. For the latter, 200 random loci (non-overlapping between runs) with 5-10 SNPs each were selected. For each run, the 95 % highest density probability intervals are shown.

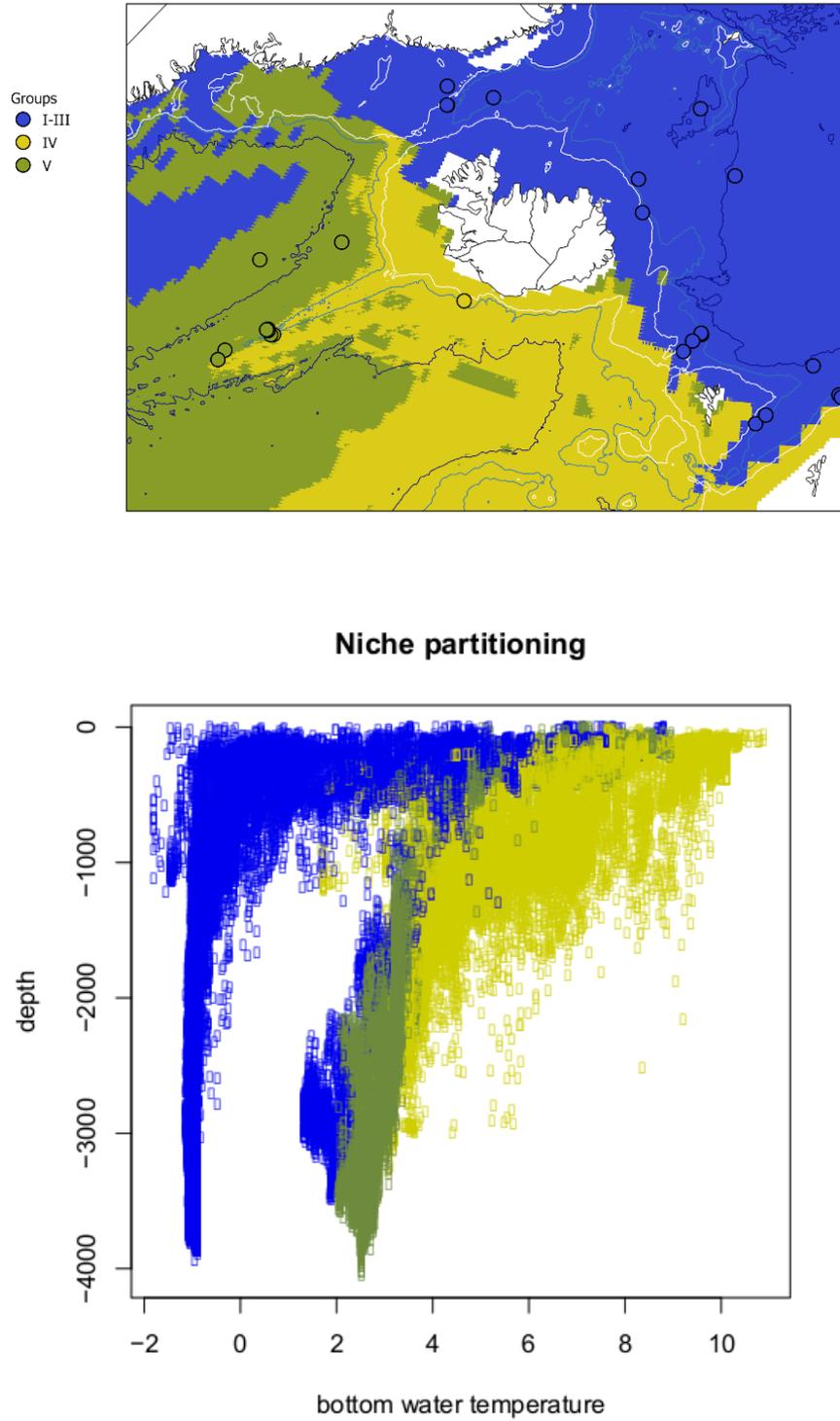


Figure 6: A) MAP/ winner The prediction of the winner class of the model; B) temperature and depth plots

Tables

Table 1. Overview of ddRAD analyses. The ddRAD data was analyzed once including all individuals and once including only individuals assigned to lineages I_{RAD}-III_{RAD}

Table 2. COI pairwise uncorrected p -distances in percent.

Table 3. Uncorrected p -distances of nuclear ddRAD loci in percent. Values are based on the analyses including all individuals.

Table 4: Population genetic and demographic parameters for each station and lineage. Only stations for which at least five individuals for the respective marker system (COI or ddRAD) were available are included (with the exception for V_{RAD}, as here only three individuals were available). For the respective lineages, all available individuals were included; also those from stations not shown here (see also Supplementary Table S1).

*, the assignment of individuals to the respective lineages differs among markers (see also Supplementary Table S1)

Table 5: Pairwise population differentiation based on F_{ST} . COI based F_{ST} values above and ddRAD based values below the diagonal. Only populations with at least five individuals available for the respective marker (COI, ddRAD) are included.

Supplementary Materials

Supplementary Table S1: List of studied specimens. For each specimen the collection number at the Center of Natural History (Hamburg, ZMH-K number), all GenBank and Bold registration numbers, detailed sampling and geographic information as well as index and barcode sequences used in ddRAD are provided.

Supplementary Figure S1: Unrooted genetic networks. A) COI haplotype network calculated with SplitsTree and B) network based on nuclear ddRAD loci calculated with SplitsTree. Colors indicate lineages delimited by COI (highlighting the conflicts in the nuclear dataset). All numbers correspond to the ZMH K- collection numbers (see Supplementary Table S1 for collection and locality details of each individual).

Supplementary File S1: Parameters of the ipyrad analyses.

Supplementary File S2: Map of stations in colours indicated in the ddRAD analyses

Probably Supplementary Figures:

Figure 7: Species Distribution Model MAP Lineages I-III_probability;

Figure 8: Species Distribution Model MAP Lineage IV_ probability;

Figure 9: Species Distribution Model MAP Lineage V_ probability,

Figure 10: Species Distribution Model errors

References

Arnason, R. (2007). Climate change and fisheries: Assessing the economic impact in Iceland and Greenland. *Natural Resource Modeling* . <https://doi.org/10.1111/j.1939-7445.2007.tb00205.x>

Astrin, J. J., & Stüben, P. E. (2008). Phylogeny in cryptic weevils: Molecules, morphology and new genera of western Palaearctic Cryptorhynchinae (Coleoptera: Curculionidae). *Invertebrate Systematics* . <https://doi.org/10.1071/IS07057>

Bandelt, H. J., Forster, P., & Röhl, A. (1999). Median-joining networks for inferring intraspecific phylogenies. *Molecular Biology and Evolution* . <https://doi.org/10.1093/oxfordjournals.molbev.a026036>

- Błażewicz, M., Józwiak, P., Menot, L., & Pabis, K. (2019). High species richness and unique composition of the tanaidacean communities associated with five areas in the Pacific polymetallic nodule fields. *Progress in Oceanography* . <https://doi.org/10.1016/j.pocean.2019.102141>
- Bober, S., Riehl, T., Henne, S., & Brandt, A. (2018). New Macrostylidae (Isopoda) from the Northwest Pacific Basin described by means of integrative taxonomy with reference to geographical barriers in the abyss. *Zoological Journal of the Linnean Society* . <https://doi.org/10.1093/zoolinnean/zlx042>
- Bode, M., Laakmann, S., Kaiser, P., Hagen, W., Auel, H., & Cornils, A. (2017). Unravelling diversity of deep-sea copepods using integrated morphological and molecular techniques. *Journal of Plankton Research* . <https://doi.org/10.1093/plankt/fbx031>
- Bouckaert, R., Vaughan, T. G., Barido-Sottani, J., Duchêne, S., Fourment, M., Gavryushkina, A., ... Drummond, A. J. (2019). BEAST 2.5: An advanced software platform for Bayesian evolutionary analysis. *PLoS Computational Biology* . <https://doi.org/10.1371/journal.pcbi.1006650>
- Boyko, C. B., Bruce, N. L., Hadfield, K. A., Merrin, K. L., Ota, Y., Poore, G. C. B., ... Wilson, G. D. F. (2008 onwards). World Marine, Freshwater and Terrestrial Isopod Crustaceans database. *Haploniscus bicuspis* (G. O. Sars, 1877). Accessed through: World Register of Marine Species at: <http://www.Marinespecies.Org/Aphia.Php?P=taxdetails&id=118656> on 2020-09-04.
- Brandt, A., Frutos, I., Bober, S., Brix, S., Brenke, N., Guggolz, T., ... Linse, K. (2018). Composition of abyssal macrofauna along the Vema Fracture Zone and the hadal Puerto Rico Trench, northern tropical Atlantic. *Deep-Sea Research Part II: Topical Studies in Oceanography* . <https://doi.org/10.1016/j.dsr2.2017.07.014>
- Brandt, A. (1997). Biodiversity of peracarid crustaceans (Malacostraca) from the shelf down to the deep Arctic Ocean. *Biodiversity and Conservation* . <https://doi.org/10.1023/A:1018318604032>
- Brandt, A., Błażewicz-Paszkowycz, M., Bamber, R. N., Mühlenthal-Siegel, U., Malyutina, M. V., Kaiser, S., ... Havermans, C. (2012). Are there widespread peracarid species in the deep sea (Crustacea: Malacostraca)? *Polish Polar Research* . <https://doi.org/10.2478/v10183-012-0012-5>
- Brandt, A., Brix, S., Held, C., & Kihara, T. C. (2014). Molecular differentiation in sympatry despite morphological stasis: deep-sea *Atlantoserolis* Wägele, 1994 and *Glabroserolis* Menzies, 1962 from the south-west Atlantic (Crustacea: Isopoda: Serolidae). *Zoological Journal of the Linnean Society* , 172(2), 318–359. <https://doi.org/10.1111/zoj.12178>
- Breiman, L. (2001). Random forests. *Machine Learning* . <https://doi.org/10.1023/A:1010933404324>
- Brenke, N. (2005). An epibenthic sledge for operations on marine soft bottom and bedrock. *Marine Technology Society Journal* . <https://doi.org/10.4031/002533205787444015>
- Brix, S., Lörz, A. N., Jażdżewska, A. M., Hughes, L., Tandberg, A. H. S., Pabis, K., ... Campean, A. J. (2018). Amphipod family distributions around Iceland. *ZooKeys* . <https://doi.org/10.3897/zookeys.731.19854>
- Brix, S., Meißner, K., Stransky, B., Halanych, K. M., Jennings, R. M., Kocot, K. M., & Svavarsson, J. (2014). The IceAGE project - A follow up of BIOICE. *Polish Polar Research* . <https://doi.org/10.2478/popore-2014-0010>
- Brix, S., Osborn, K., Kaiser, S., Truskey, S., Schnurr, S., Brenke, N., ... Martinez Abizu, P. (2020). Adult life strategy affects distribution patterns in abyssal isopods – implications for conservation in Pacific nodule areas. *Biogeosciences* , 17, 6163–6184. <https://doi.org/https://doi.org/10.5194/bg-17-6163-2020>
- Brix, S., Riehl, T., & Leese, F. (2011). First genetic data for species of the genus *Haploniscus* Richardson, 1908 (Isopoda: Asellota: Haploniscidae) from neighbouring deep-sea basins in the South Atlantic. *Zootaxa* . <https://doi.org/10.11646/zootaxa.2838.1.5>

- Brix, S., Stransky, B., Maljutina, M., Pabis, K., Svavarsson, J., & Riehl, T. (2018). Distributional patterns of isopods (Crustacea) in Icelandic and adjacent waters. *Marine Biodiversity* . <https://doi.org/10.1007/s12526-018-0871-z>
- Brix, S., & Svavarsson, J. (2010). Distribution and diversity of desmosomatid and nannoniscid isopods (Crustacea) on the Greenland-Iceland-Faeroe Ridge. *Polar Biology* , 33, 515–530. <https://doi.org/10.1007/s00300-009-0729-8>
- Brökeland, W., & Svavarsson, J. (2017). Distribution of haploniscids (Isopoda, Asellota, Haploniscidae) in Icelandic waters, with description of *Haploniscus astraphes* n. sp. from the Iceland basin and the Southeast Atlantic Ocean. *Zootaxa* . <https://doi.org/10.11646/zootaxa.4231.3.1>
- Brökeland, W., & Wägele, J. W. (2004). Redescription of three species of *Haploniscus* Richardson, 1908 (Isopoda, Asellota, Haploniscidae) from the Angola Basin. *Organisms Diversity and Evolution* . <https://doi.org/10.1016/j.ode.2004.02.004>
- D’Alba, L., Monaghan, P., & Nager, R. G. (2010). Advances in laying date and increasing population size suggest positive responses to climate change in common eiders *Somateria mollissima* in Iceland. *IBIS* , 152, 19–28. <https://doi.org/https://doi.org/10.1111/j.1474-919X.2009.00978.x>
- Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., . . . Durbin, R. (2011). The variant call format and VCFtools. *Bioinformatics* . <https://doi.org/10.1093/bioinformatics/btr330>
- Dauvin, J. C., Alizier, S., Weppe, A., & Gudmundsson, G. (2012). Diversity and zoogeography of Icelandic deep-sea Ampeliscidae (Crustacea: Amphipoda). *Deep-Sea Research Part I: Oceanographic Research Papers* . <https://doi.org/10.1016/j.dsr.2012.04.013>
- Dieme, C., Yssouf, A., Vega-Rúa, A., Berenger, J. M., Failloux, A. B., Raoult, D., . . . Almeras, L. (2014). Accurate identification of Culicidae at aquatic developmental stages by MALDI-TOF MS profiling. *Parasites and Vectors* . <https://doi.org/10.1186/s13071-014-0544-0>
- Earl, D. A., & vonHoldt, B. M. (2012). STRUCTURE HARVESTER: A website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources* . <https://doi.org/10.1007/s12686-011-9548-7>
- Eaton, D. A. R., & Overcast, I. (2020). ipyrad: Interactive assembly and analysis of RADseq datasets. *Bioinformatics* . <https://doi.org/10.1093/bioinformatics/btz966>
- Edgar, R. C. (2004). MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research* . <https://doi.org/10.1093/nar/gkh340>
- Evanno, G., Regnaut, S., & Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study. *Molecular Ecology* . <https://doi.org/10.1111/j.1365-294X.2005.02553.x>
- Excoffier, L., & Lischer, H. E. L. (2010). Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources* . <https://doi.org/10.1111/j.1755-0998.2010.02847.x>
- Franchini, P., Monné Parera, D., Kautt, A. F., & Meyer, A. (2017). quaddRAD: a new high-multiplexing and PCR duplicate removal ddRAD protocol produces novel evolutionary insights in a nonradiating cichlid lineage. *Molecular Ecology* . <https://doi.org/10.1111/mec.14077>
- Geirsdóttir, Á., Miller, G. H., Axford, Y., & Ólafsdóttir, S. (2009). Holocene and latest Pleistocene climate and glacier fluctuations in Iceland. *Quaternary Science Reviews*. <https://doi.org/10.1016/j.quascirev.2009.03.013>
- German, C. R., Briem, J., Chin, C., Danielsen, M., Holland, S., James, R., . . . Rudnicki, M. D. (1994). Hydrothermal activity on the Reykjanes Ridge: the Steinahóll vent-field at 63°06’N. *Earth and Planetary*

Science Letters . [https://doi.org/10.1016/0012-821X\(94\)90098-1](https://doi.org/10.1016/0012-821X(94)90098-1)

Gibb, S. (2015). MALDIquantForeign: Import/Export routines for MALDIquant. A package for R. Retrieved from <https://cran.r-project.org/package=MALDIquantForeign>

Gibb, Sebastian, & Strimmer, K. (2012). Maldiquant: A versatile R package for the analysis of mass spectrometry data. *Bioinformatics* . <https://doi.org/10.1093/bioinformatics/bts447>

Gudmundsson, G. (1998). Distributional limits of *Pyrgo* species at the biogeographic boundaries of the Arctic and the North-Atlantic Boreal Regions. *Journal of Foraminiferal Research* .

Hanna, E., Jonsson, T., & Box, J. E. (2006). Recent changes in Icelandic climate. *Weather* , 61(1), 3–9. <https://doi.org/10.1256/wea.80.04>

Hansen, B., & Osterhus, S. (2000). North Atlantic-Nordic Seas exchanges. *Progress in Oceanography* . [https://doi.org/10.1016/S0079-6611\(99\)00052-X](https://doi.org/10.1016/S0079-6611(99)00052-X)

Hansen, H. J. (1908). Crustacea Malacostraca. In *The Danish Ingolf-Expedition*.

Havermans, C., Sonet, G., d’Udekem d’Acoz, C., Nagy, Z. T., Martin, P., Brix, S., ... Held, C. (2013). Genetic and morphological divergences in the cosmopolitan deep-sea amphipod *Eurythenes gryllus* reveal a diverse Abyss and a bipolar species. *PLoS ONE* . <https://doi.org/10.1371/journal.pone.0074218>

Hessler, R. R. (1970). The Desmosomatidae (Isopoda, Asellota) of the Gay Head-Bermuda Transect. University of California Press.

Huson, D. H., & Bryant, D. (2006). Application of phylogenetic networks in evolutionary studies. *Molecular Biology and Evolution* . <https://doi.org/10.1093/molbev/msj030>

Jakiel, A., Stepień, A., & Błażewicz, M. (2018). A tip of the iceberg—Pseudotanaididae (Tanaidacea) diversity in the North Atlantic. *Marine Biodiversity* . <https://doi.org/10.1007/s12526-018-0881-x>

Jakobsson, M., & Rosenberg, N. A. (2007). CLUMPP: A cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics* . <https://doi.org/10.1093/bioinformatics/btm233>

Jennings, R. M., Brix, S., Bober, S., Svavarsson, J., & Driskell, A. (2018). More diverse than expected: distributional patterns of *Oecidiobanchus* Hessler, 1970 (Isopoda, Asellota) on the Greenland-Iceland-Faeroe Ridge based on molecular markers. *Marine Biodiversity* . <https://doi.org/10.1007/s12526-018-0857-x>

Jennings, R. M., Golovan, O., & Brix, S. (2020). Integrative species delimitation of desmosomatid and nannoniscid isopods from the Kuril-Kamchatka trench, with description of a hadal species. *Progress in Oceanography* . <https://doi.org/10.1016/j.pocean.2019.102236>

Jochumsen, K., Schnurr, S. M., & Quadfasel, D. (2016). Bottom temperature and salinity distribution and its variability around Iceland. *Deep-Sea Research Part I: Oceanographic Research Papers* . <https://doi.org/10.1016/j.dsr.2016.02.009>

Kaiser, S., Kihara, T. C., Brix, S., Mohrbeck, I., Janssen, A., & Jennings, R. M. (2021). Species boundaries and phylogeographic patterns in new species of *Nannoniscus* (Janiroidea: Nannoniscidae) from the equatorial Pacific nodule province inferred from mtDNA and morphology. *Zoological Journal of the Linnean Society* . <https://doi.org/10.1093/zoolinnean/zlaa174>

Karger, A., Bettin, B., Gethmann, J. M., & Klaus, C. (2019). Whole animal matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) mass spectrometry of ticks—Are spectra of *Ixodes ricinus* nymphs influenced by environmental, spatial, and temporal factors? *PLoS ONE* , 14(1). <https://doi.org/https://doi.org/10.1371/journal.pone.0210590>

Knowlton, N., & Weigt, L. A. (1998). New dates and new rates for divergence across the Isthmus of Panama. *Proceedings of the Royal Society B: Biological Sciences* . <https://doi.org/10.1098/rspb.1998.0568>

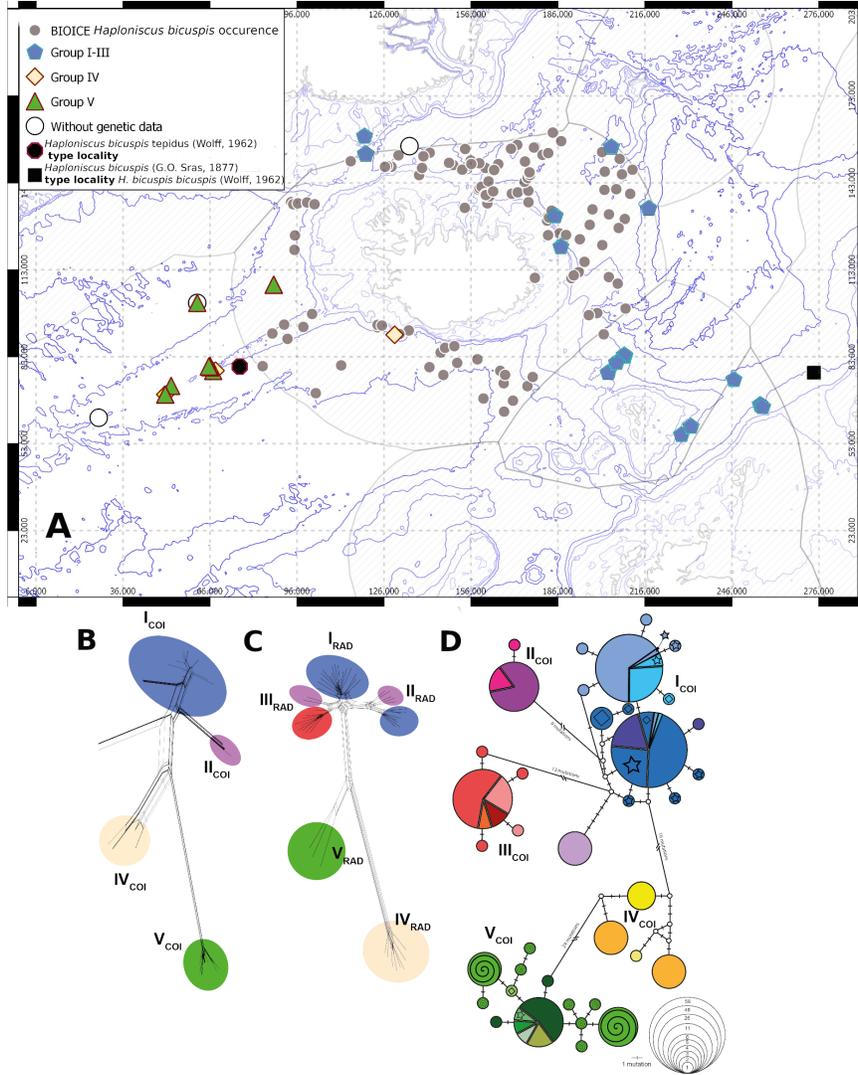
- Kumar, S., Stecher, G., Li, M., Knyaz, C., & Tamura, K. (2018). MEGA X: Molecular evolutionary genetics analysis across computing platforms. *Molecular Biology and Evolution* . <https://doi.org/10.1093/molbev/msy096>
- Laakmann, S., Gerdt, G., Erler, R., Knebelberger, T., Martínez Arbizu, P., & Raupach, M. J. (2013). Comparison of molecular species identification for North Sea calanoid copepods (Crustacea) using proteome fingerprints and DNA sequences. *Molecular Ecology Resources* . <https://doi.org/10.1111/1755-0998.12139>
- Laikre, L. (2010). Genetic diversity is overlooked in international conservation policy implementation. *Conservation Genetics* . <https://doi.org/10.1007/s10592-009-0037-4>
- Legendre, P., & Gallagher, E. D. (2001). Ecologically meaningful transformations for ordination of species data. *Oecologia* . <https://doi.org/10.1007/s004420100716>
- Liaw, A., & Wiener, M. (2002). Classification and Regression by randomForest. R News.
- Lincoln, R. J. (1985). Deep-sea asellote isopods of the north-east Atlantic: the family Haploniscidae. *Journal of Natural History* . <https://doi.org/10.1080/00222938500770411>
- Loeza-Quintana, T., Carr, C. M., Khan, T., Bhatt, Y. A., Lyon, S. P., Hebert, P. D. N., & Adamowicz, S. J. (2019). Recalibrating the molecular clock for Arctic marine invertebrates based on DNA barcodes. *Genome* . <https://doi.org/10.1139/gen-2018-0107>
- Maasz, G., Zrínyi, Z., Fodor, I., Boross, N., Vitál, Z., Sipos, D. I. K., ... Takács, P. (2020). Testing the applicability of MALDI-TOF MS as an alternative stock identification method in a cryptic species complex. *Molecules* . <https://doi.org/10.3390/molecules25143214>
- Malinsky, M., Trucchi, E., Lawson, D. J., & Falush, D. (2018). RADpainter and fineRADstructure: Population Inference from RADseq Data. *Molecular Biology and Evolution* . <https://doi.org/10.1093/molbev/msy023>
- Meißner, K., Brix, S., Halanych, K. M., & Jazdzewska, A. M. (2018). Preface—biodiversity of Icelandic waters. *Marine Biodiversity* . <https://doi.org/10.1007/s12526-018-0884-7>
- Meißner, K., Fiorentino, D., Schnurr, S., Martínez Arbizu, P., Huettmann, F., Holst, S., ... Svavarsson, J. (2014). Distribution of benthic marine invertebrates at northern latitudes - An evaluation applying multi-algorithm species distribution models. *Journal of Sea Research* . <https://doi.org/10.1016/j.seares.2013.05.007>
- Meyer, C. P. (2003). Molecular systematics of cowries (Gastropoda: Cypraeidae) and diversification patterns in the tropics. *Biological Journal of the Linnean Society* . <https://doi.org/10.1046/j.1095-8312.2003.00197.x>
- Michels, J., & Büntzow, M. (2010). Assessment of Congo red as a fluorescence marker for the exoskeleton of small crustaceans and the cuticle of polychaetes. *Journal of Microscopy* . <https://doi.org/10.1111/j.1365-2818.2009.03360.x>
- Müller, P., Pflüger, V., Wittwer, M., Ziegler, D., Chandre, F., Simard, F., & Lengeler, C. (2013). Identification of cryptic *Anopheles* mosquito species by Molecular Protein Profiling. *PLoS ONE* . <https://doi.org/10.1371/journal.pone.0057486>
- Nachtigall, F. M., Pereira, A., Trofymchuk, O. S., & Santos, L. S. (2020). Detection of SARS-CoV-2 in nasal swabs using MALDI-MS. *Nature Biotechnology* . <https://doi.org/10.1038/s41587-020-0644-7>
- Negoescu, I., & Svavarsson, J. (1997). Anthurideans (Crustacea, Isopoda) from the North Atlantic and the Arctic Ocean. *Sarsia* . <https://doi.org/10.1080/00364827.1997.10413650>
- Nordahl, H., & Ingólfsson, Ó. (2015). Collapse of the Icelandic ice sheet controlled by sea-level rise? *Arktos* . <https://doi.org/10.1007/s41063-015-0020-x>
- Oksanen, J., Blanchet, F. G., Kindt, R., Legendre, P., Minchin, P. R., & O'Hara, R. B. (2013). Package vegan. R Packag Ver 2.

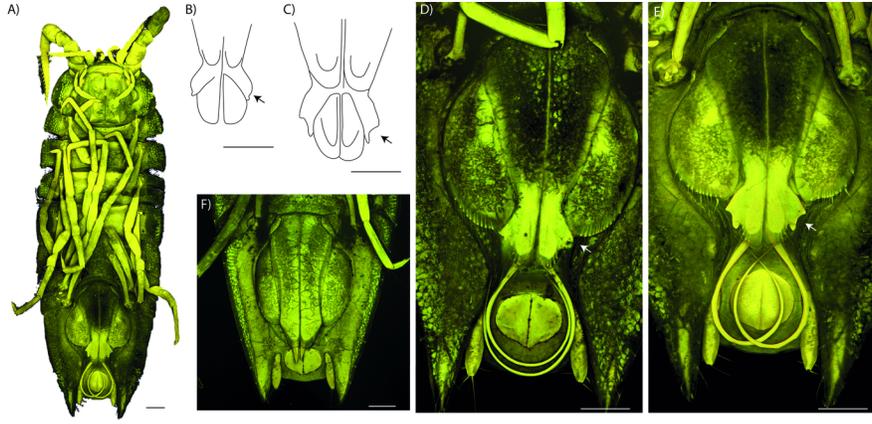
- Ostmann, A., Schnurr, S., & Martínez Arbizu, P. (2014). Marine environment around Iceland: Hydrography, sediments and first predictive models of Icelandic deep-sea sediment characteristics. *Polish Polar Research* . <https://doi.org/10.2478/popore-2014-0021>
- Palarea-Albaladejo, J., McLean, K., Wright, F., & Smith, D. (2018). MALDIrppa: quality control and robust analysis for mass spectrometry data. *Bioinformatics* , 34(3), 522–523. <https://doi.org/http://dx.doi.org/10.1093/bioinformatics/btx628>
- Patton, H., Hubbard, A., Bradwell, T., & Schomacker, A. (2017). The configuration, sensitivity and rapid retreat of the Late Weichselian Icelandic ice sheet. *Earth-Science Reviews* . <https://doi.org/10.1016/j.earscirev.2017.02.001>
- Pecl, G. T., Araújo, M. B., Bell, J. D., Blanchard, J., Bonebrake, T. C., Chen, I. C., ... Williams, S. E. (2017). Biodiversity redistribution under climate change: Impacts on ecosystems and human well-being. *Science* . <https://doi.org/10.1126/science.aai9214>
- Pons, J., Barraclough, T. G., Gomez-Zurita, J., Cardoso, A., Duran, D. P., Hazell, S., ... Vogler, A. P. (2006). Sequence-based species delimitation for the DNA taxonomy of undescribed insects. *Systematic Biology* . <https://doi.org/10.1080/10635150600852011>
- Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics* .
- Puillandre, N., Lambert, A., Brouillet, S., & Achaz, G. (2012). ABGD, Automatic Barcode Gap Discovery for primary species delimitation. *Molecular Ecology* . <https://doi.org/10.1111/j.1365-294X.2011.05239.x>
- Ramasamy, R. K., Ramasamy, S., Bindroo, B. B., & Naik, V. G. (2014). STRUCTURE PLOT: A program for drawing elegant STRUCTURE bar plots in user friendly interface. *SpringerPlus* . <https://doi.org/10.1186/2193-1801-3-431>
- R Core Team (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.
- Reid, N. M., & Carstens, B. C. (2012). Phylogenetic estimation error can decrease the accuracy of species delimitation: A Bayesian implementation of the general mixed Yule-coalescent model. *BMC Evolutionary Biology*. <https://doi.org/10.1186/1471-2148-12-196>
- Riccardi, N., Lucini, L., Bengali, C., Welker, M., Wicht, B., & Tonolla, M. (2012). Potential of matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI TOF MS) for the identification of freshwater zooplankton: a pilot study with three *Eudiaptomus* (Copepoda: Diaptomidae) species. *Journal of Plankton Research* , 34(6), 484–492. <https://doi.org/https://doi.org/10.1093/plankt/fbs022>
- Riehl, T., Lins, L., & Brandt, A. (2018). The effects of depth, distance, and the Mid-Atlantic Ridge on genetic differentiation of abyssal and hadal isopods (Macrostylidae). *Deep-Sea Research Part II: Topical Studies in Oceanography* . <https://doi.org/10.1016/j.dsr2.2017.10.005>
- Riehl, T., Wilson, G. D. F., & Hessler, R. R. (2012). New Macrostylidae Hansen, 1916 (Crustacea: Isopoda) from the Gay Head-Bermuda transect with special consideration of sexual dimorphism. *Zootaxa* . <https://doi.org/10.11646/zootaxa.3277.1.1>
- Rochette, N. C., Rivera-Colón, A. G., & Catchen, J. M. (2019). Stacks 2: Analytical methods for paired-end sequencing improve RADseq-based population genomics. *Molecular Ecology* . <https://doi.org/10.1111/mec.15253>
- Rossel, S., & Martínez Arbizu, P. (2018). Automatic specimen identification of harpacticoids (Crustacea: Copepoda) using Random Forest and MALDI-TOF mass spectra, including a post hoc test for false positive discovery. *Methods in Ecology and Evolution* . <https://doi.org/10.1111/2041-210X.13000>

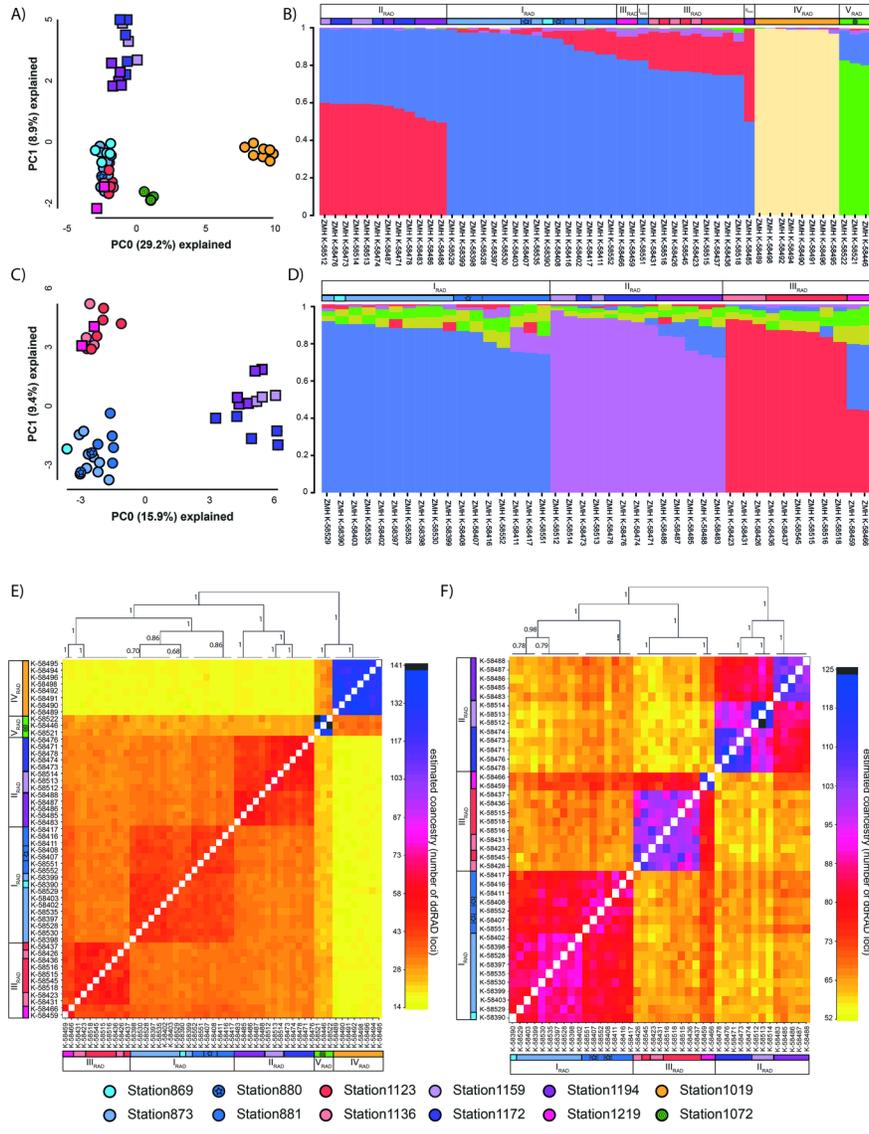
- Ryan, C. G., Clayton, E., Griffin, W. L., Sie, S. H., & Cousens, D. R. (1988). SNIP, a statistics-sensitive background treatment for the quantitative analysis of PIXE spectra in geoscience applications. *Nuclear Inst. and Methods in Physics Research*, B. [https://doi.org/10.1016/0168-583X\(88\)90063-8](https://doi.org/10.1016/0168-583X(88)90063-8)
- Sars, G. O. (1877). Prodrömus descriptionis crustaceorum et pycnogonidarum, quae in expeditione norvegica anno 1876, observavit. *Mathematik og Naturvidenskab*, 2, 337–368.
- Savitzky, A., & Golay, M. J. E. (1964). Smoothing and differentiation of data by simplified least squares procedures. *Analytical Chemistry*. <https://doi.org/10.1021/ac60214a047>
- Schnurr, S., Brandt, A., Brix, S., Fiorentino, D., Malyutina, M., & Svavarsson, J. (2014). Composition and distribution of selected munnopsid genera (Crustacea, Isopoda, Asellota) in Icelandic waters. *Deep-Sea Research Part I: Oceanographic Research Papers*. <https://doi.org/10.1016/j.dsr.2013.11.004>
- Schwentner, M., & Lörz, A. (2021). Population genetics of cold-water coral associated Pleustidae (Crustacea, Amphipoda) reveals cryptic diversity and recent expansion off Iceland. *Marine Ecology*, 42, e12625. <https://doi.org/https://doi.org/10.1111/maec.12625>
- Singhal, N., Kumar, M., Kanaujia, P. K., & Viridi, J. S. (2015). MALDI-TOF mass spectrometry: An emerging technology for microbial identification and diagnosis. *Frontiers in Microbiology*. <https://doi.org/10.3389/fmicb.2015.00791>
- Stransky, B., & Svavarsson, J. (2006). *Astacilla boreaphilis* sp. nov. (Crustacea: Isopoda: Valvifera) from shallow and deep North Atlantic waters. *Zootaxa*. <https://doi.org/10.11646/zootaxa.1259.1.1>
- Svavarsson, J., Stromberg, J.-O., & Brattegard, T. (1993). The Deep-Sea asellote (Isopoda, Crustacea) fauna of the Northern Seas: species composition, distributional patterns and origin. *Journal of Biogeography*. <https://doi.org/10.2307/2845725>
- Takacs, P., Bihari, P., Eros, T., Specziar, A., Szivak, I., Biro, P., & Csoma, E. (2014). Genetic heterogeneity reveals on-going speciation and cryptic taxonomic diversity of stream-dwelling gudgeons (Teleostei, Cyprinidae) in the Middle Danubian hydrosystem (Hungary). *PLoS ONE*. <https://doi.org/10.1371/journal.pone.0097278>
- Taylor, M. L., & Roterman, C. N. (2017). Invertebrate population genetics across Earth’s largest habitat: The deep-sea floor. *Molecular Ecology*. <https://doi.org/10.1111/mec.14237>
- Taylor, J. Devey, C. (...) & S. Brix: The geological and faunal composition of Steinhall vent sites and adjacent benthos communities, Reykjanes Ridge, Iceland. In review in *Frontiers in Marine Science*; since 09/2019
- Thiel, M., & Haye, P. A. (2006). The ecology of rafting in the marine environment. III. Biogeographical and evolutionary consequences. *Oceanography and Marine Biology*.
- Weissshappel, J. B. (2000). Distribution and diversity of the hyperbenthic amphipod family Eusiridae in the different seas around the Greenland-Iceland-Faeroe-Ridge. *Sarsia*. <https://doi.org/10.1080/00364827.2000.10414575>
- Weissshappel, J. B. (2001). Distribution and diversity of the hyperbenthic amphipod family Calliopiidae in the different seas around the Greenland-Iceland-Faeroe-Ridge. *Sarsia*. <https://doi.org/10.1080/00364827.2001.10420469>
- White, B. N. (1988). Oceanic anoxic events and allopatric speciation in the deep sea. *Biological Oceanography*, 5(4), 243–259.
- Wilson, G. D. F., & Ahyong, S. T. (2015). *The Natural History of the Crustacea* 2.
- Wolff, T. (1962). *The Systematics and Biology of bathyal and abyssal Isopoda Asellota (Galathea Report)*. Copenhagen: Danish Science Press.

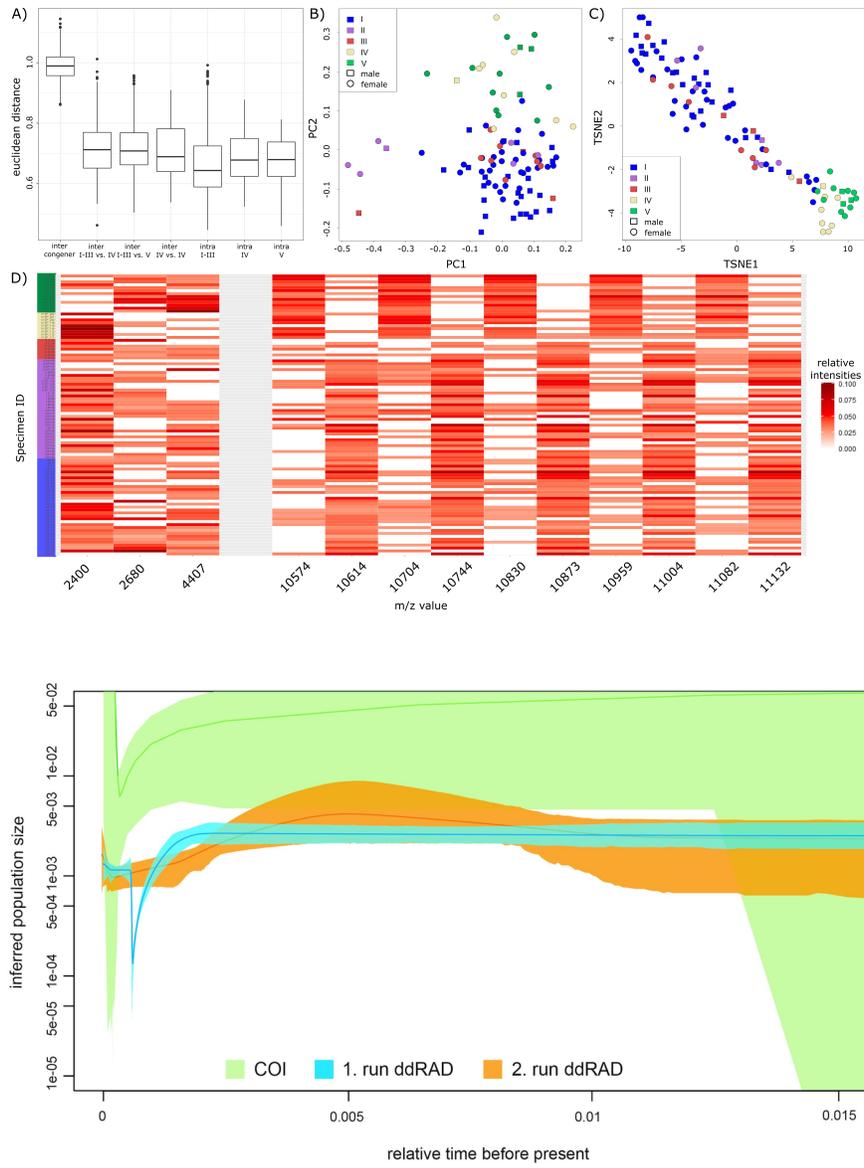
Hosted file

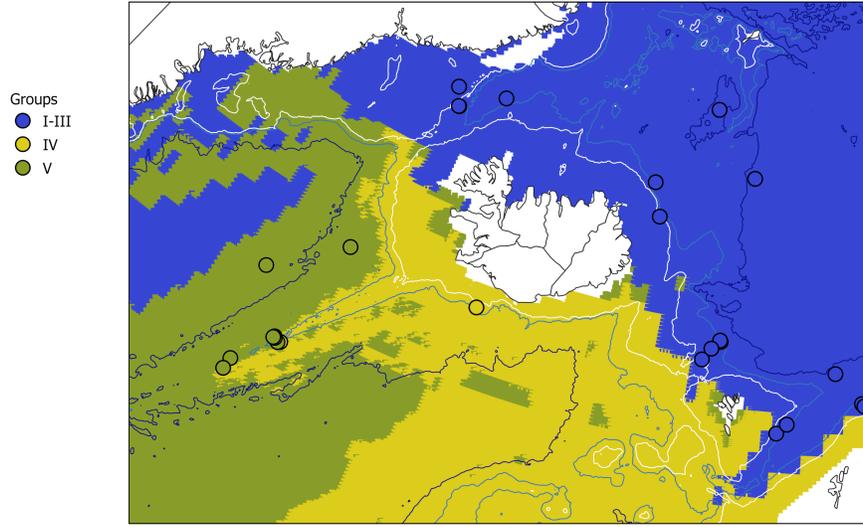
tables_2021-02-18.pdf available at <https://authorea.com/users/406160/articles/516929-recent-speciation-and-hybridization-in-icelandic-deep-sea-isopods-an-integrative-approach-using-genomics-and-proteomics>



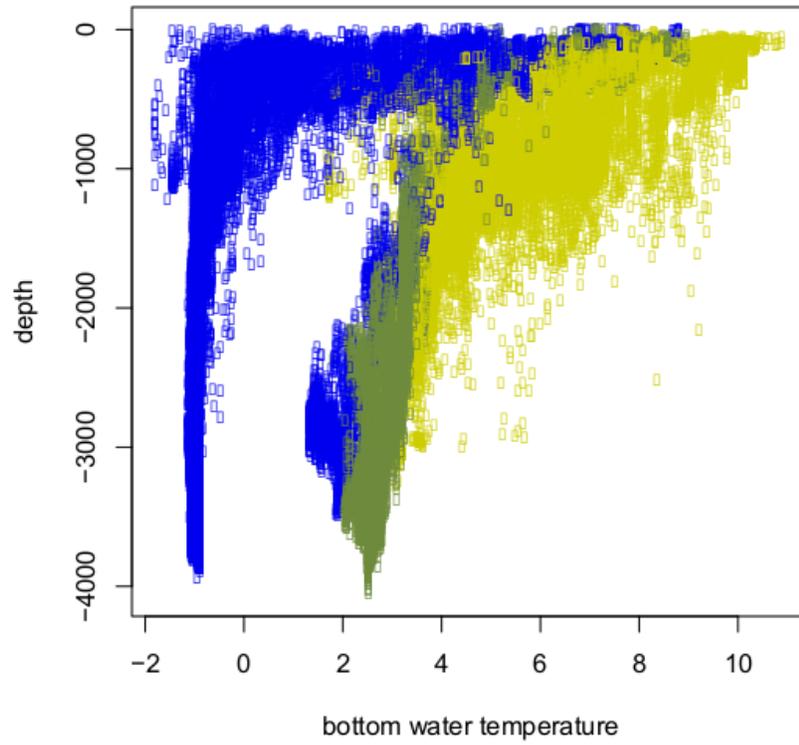


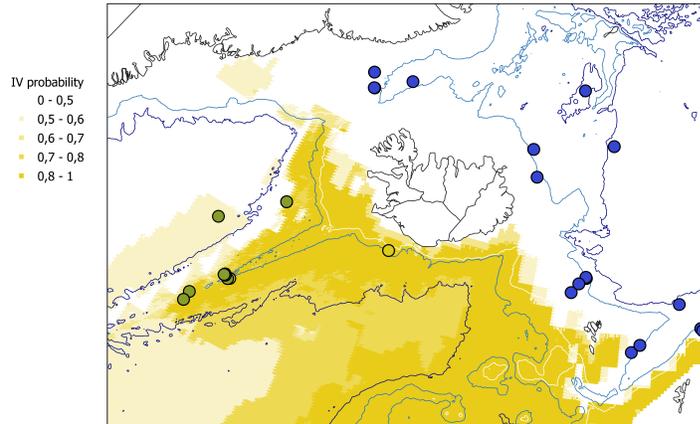
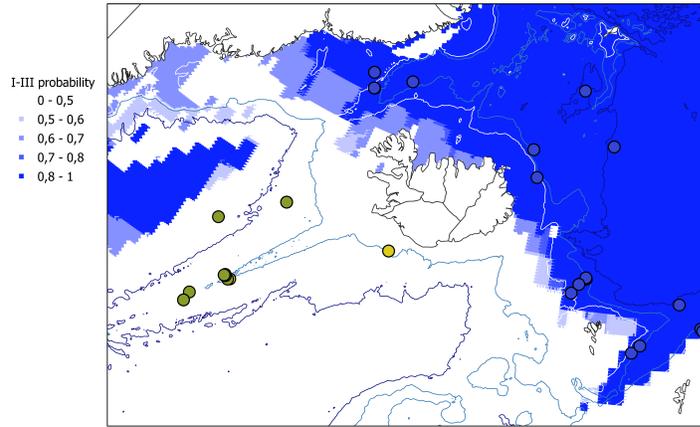


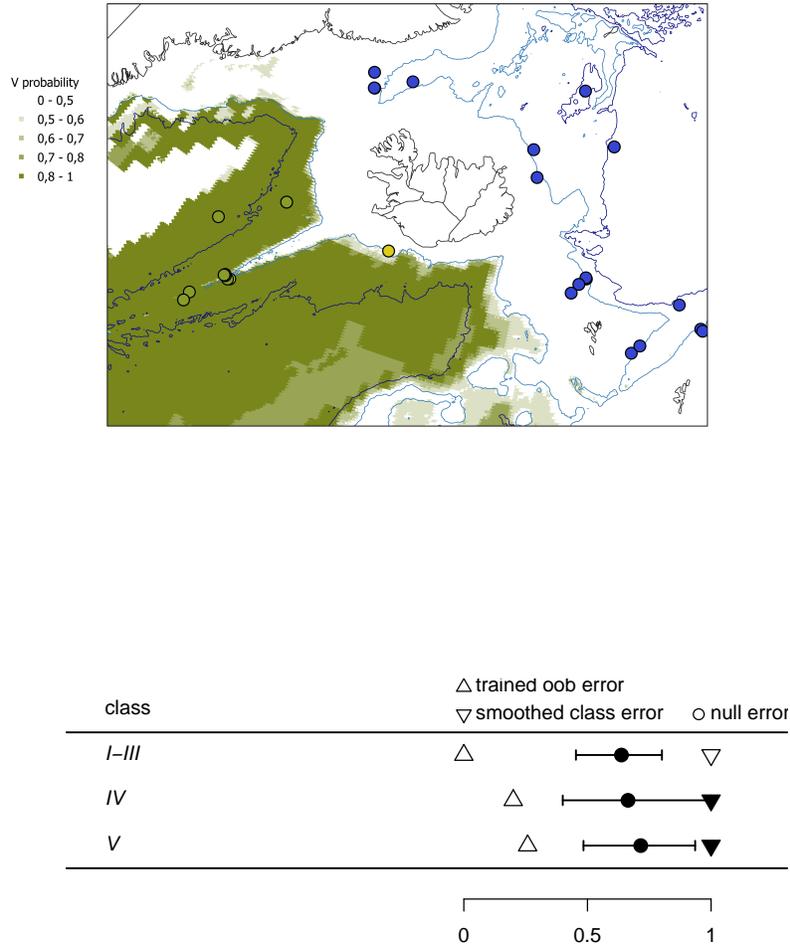


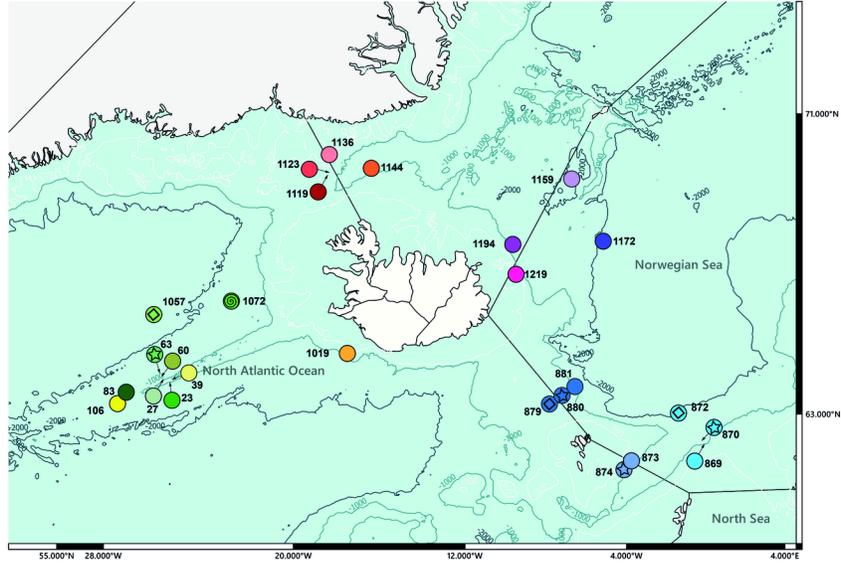


Niche partitioning









Hosted file

Table Supplement.pdf available at <https://authorea.com/users/406160/articles/516929-recent-speciation-and-hybridization-in-icelandic-deep-sea-isopods-an-integrative-approach-using-genomics-and-proteomics>

