SNPs for population genetic analysis and species delimitation of Greenland Nephtheidae corals (Octocorallia: Alcyonacea)

Student number: [redacted]
Supervisor: Dr Chris Yesson, ZSL
References formatted in the style of: Deep Sea Research Part I: Oceanographic Research Papers
Word count: 5999

MRes Biodiversity, Conservation, and Evolution
BIOS0013: Biodiversity research project 1
SNPs for population genetic analysis and species delimitation of Greenland Nephtheidae corals (Octocorallia: Alcyonacea)

Abstract

The Nephtheidae are a widely distributed family of Alcyonacean corals, found in cold waters at depths below 50m. In the Northwest Atlantic, nephtheid species commonly form dense communities harbouring high biodiversity, which are recognised as vulnerable marine ecosystems. Increased pressure from deep-sea fisheries around Greenland has been shown to have significant negative effects on nephtheid diversity and biomass. However, the underlying genetic structure of these populations is unknown. This study presents 22 novel single nucleotide polymorphism (SNP) markers which were used to assess the population structure of Nephtheidae corals across the east, south, and west coasts of Greenland. Population genetic patterns across the area were analysed using discriminate analysis of principle components (DAPC). Significant genetic differentiation ($F_{st}$) was found between all regional groups analysed, with no significant pattern of isolation by distance (IBD) or isolation by depth. Genetic cluster assignment appears to be species specific, and suggests that these SNP markers can be used for species delimitation within this family which has traditionally represented a taxonomic challenge. Hypotheses are presented to explain variation in $F_{st}$ and IBD results, and potential species-specific patterns of gene flow. The genetic tools developed here provide a foundation for further studies of nephtheid species. Understanding the population structure, reproductive behaviour, and dispersal capacity of deep-sea species is crucial to the designation of suitable management strategies around Greenland. These results represent an important step in elucidating the population structure of these habitat-forming species, and determining how they can best be protected from increasing anthropogenic threats.

Key words: cold-water coral, Greenland, Nephtheidae, SNP, population structure, VME
1. Introduction

Octocorallia (Cnidaria: Anthozoa) are the most diverse group of corals and represent a significant component of marine benthic ecosystems (Cairns, 2007; Roberts et al., 2009). There are currently over 3400 extant species of octocoral recognised (Williams and Cairns, 2015), however, the taxonomy of the group remains unresolved. Octocorals have a worldwide distribution, with an estimated 74% of their diversity found at depths exceeding 50m (Roberts and Cairns, 2014). Within the deep-sea habitats that they occupy (e.g. continental slopes, seamounts), octocorals represent significant ecosystem engineers, acting as important structural components of the benthos (Williams and Cairns, 2015). The single- or multi-species assemblages formed by colonies of cold-water octocorals are commonly referred to as “coral gardens”, “coral fields”, and “coral forests” (Bullimore et al., 2013; Neves et al., 2014). These gardens are not limited to octocorals, and may also contain corals from the Antipatharia, Scleractinia, and Stylasteridae families (Freiwald et al., 2004).

The complex biogenic habitats formed by cold-water corals support highly diverse biological communities of both fish and invertebrate species (Buhl-Mortensen et al., 2010; Baillon et al., 2012; Henry et al., 2013). Furthermore, the importance of cold-water coral species as biogenic substrates has been shown to increase with depth (down to 2500m) due to declining food supply and homogeneity of geological substrate (Buhl-Mortensen et al., 2010). However, the nature of the relationships between corals and their associated fauna are difficult to determine. While mutualistic relationships between invertebrates and octocorals have been documented, for example with ophiuroids (Girard et al., 2016), many species are found to use coral garden habitats facultatively (Edinger et al., 2007). The ecology of the habitat provided by octocorals is therefore non-uniform and dependant on the inhabiting species.

Associations between cold-water corals and commercially important species makes these coral species particularly vulnerable to anthropogenic impacts, such as fishing. Due to the depletion of shallower stocks and improved fishing technology since the 1970s, the trawl fishing industry has progressively been exploiting deeper waters (Davies et al., 2007). Cold-water corals are acutely vulnerable to bottom-trawling, as contact with these devices can decimate colonies through the complete removal of individuals with their associated fauna.
Octocorals have been shown to have slow growth rates, with longevities reaching hundreds of years and linear growth rates being as low as < 1 cm per year in some species (Sherwood and Edinger, 2009). These characteristics imply that the impacts from trawling damage can potentially persist for centuries before significant recovery can be observed (Freiwald et al., 2004; Althaus et al., 2009). These impacts are likely to be exacerbated in the coming years, as climate-induced changes in the availability of suitable habitats become more pronounced (Morato et al., 2020).

In response to concerns regarding long-term fishing impacts, the United Nations General Assembly (UNGA) have passed resolutions 61/105 and 64/72, pertaining to the dedication and protection of vulnerable marine ecosystems (VMEs) (Rodgers and Gianni, 2010). Due to the ecological and socio-economic value of cold-water coral gardens, they fall under the guidelines established by the Food and Agriculture Organisation of the United Nations (FAO) to be considered as VMEs (FAO, 2009). These ecosystems have also been acknowledged as VMEs by institutions including the North East Atlantic Fisheries Commission (NEAFC) (NEAFC, 2014; Buhl-Mortensen et al., 2019). Coupled with mounting evidence of declines in deep-sea coral populations impacted by fishing (Devine et al., 2019), it has become clear that research is needed into these populations in order to successfully designate VMEs. Exploited populations are likely characterised by reduced resilience and low genetic diversity, highlighting the need to understand the connectivity between populations, to identify areas of locally adapted genotypes (Valero et al., 2001).

The significant logistical constraints of sampling cold-water corals at depth often results in low sample numbers which are unconducive to population genetics studies. Consequently, alternative methods of population analysis have been demonstrated using traditional DNA barcoding to examine the connectivity of species (Herrera et al., 2012). However, the use of microsatellite markers has been successfully shown in a number of small-scale population studies on octocorals (Quattrini et al., 2015; Wright et al., 2015; Holland et al., 2017). For example, twelve microsatellite loci were developed for six populations and 83 samples of the species *Narella versluysi* (Alcyonacea: Primnoidae) from the Bay of Biscay to examine gene flow between colonies (Yesson et al., 2018a). However, the accuracy of using limited numbers of microsatellite markers to reflect genome-wide diversity is contentious due to the
ascertainment bias created by selecting only the most polymorphic markers (Väli et al., 2008; Glover et al., 2010). With the rapid development of high-throughput sequencing techniques, single nucleotide polymorphisms (SNPs) have emerged as ideal population genetic markers. These techniques can allow for the identification of potentially thousands of SNPs, with their codominant inheritance making them model candidates for analysing population patterns (Davey et al., 2011; Schlötterer et al., 2014). While SNPs have been successfully used to study genetic connectivity in shallow water, tropical corals (Drury et al., 2016; Devlin-Durante and Baums, 2017), they have yet to be widely applied to cold-water corals. Successful genotyping of the deep-sea gorgonian *Swiftia simplex* (Alcyonacea: Plexauridae) allowed for the development of 1,145 SNPs to assess putative population structure across the eastern Pacific (Everett et al., 2016). This suggests that SNPs can be effectively applied to this class to examine genetic flow between populations.

Greenland has an Exclusive Economic Zone (EEZ) of over 2.2 million km², however, currently has very few marine protected areas (MPAs). Areas that have been designated are exclusively inshore waters, with none proposed to directly protect VMEs (UNEP-WCMC, 2019). There is a significant knowledge gap regarding the distribution of VMEs in the North Atlantic, particularly within the Greenland EEZ. This is reflected by the lack of management in place to protect VMEs. Fisheries presently account for 80-95% of Greenland’s export income (Jacobsen, 2018), with the majority of this hailing from deep-sea fisheries in West Greenland. These fisheries rely on trawling targeting the Northern shrimp (*Pandalus borealis*) and Greenland halibut (*Reinhardtius hippoglossoides*) (Bowering and Nedreaas, 2000). Benthic surveys using a drop camera have been conducted across the West Greenland continental slope to assess the impacts of trawling on epibenthic communities in the area (Yesson et al., 2015; Yesson et al., 2017). These have demonstrated that trawling has a significant negative effect on sessile erect organisms, such as octocorals. More recently, benthic sleds have been used to identify a new VME composed of a cold-water coral garden, located within the Davis Strait (Long et al., in review).

A common feature of the benthic communities around Greenland are coral gardens dominated by octocorals from the Nephtheidae Gray, 1862 family, frequently known as “cauliflower corals” (Buhl-Mortensen et al., 2019; Long et al., in review). Members of this
family display highly variable morphology, often presenting with transitional forms of typical distinguishing characteristics (e.g. habitus, sclerite type). This morphological plasticity has resulted in a large number of species descriptions published within the last century – many of which are believed to be overlapping. The most recent taxonomical re-description of the North Atlantic Nephtheidae found three genera to be common in the area, and partitioned these based on polyp-type: Duva and Drifa have non-retractile polyps, whereas Gersemia have retractable polyps (Jensen, 2003). However, molecular phylogenetics suggest that Gersemia do not belong in the Nephtheidae, and in fact should be placed within the family Alcyoniidae (McFadden and van Ofwegen, 2013). Phylogenetic species delimitation of nephtheids presents a significant challenge, due to limited variation among anthozoan mitochondrial genomes, interspecific hybridisation, and shared ancestral polymorphisms (Forsman et al., 2009; Herrera and Shank, 2016). This represents a small example of the taxonomic difficulties present within the Nephtheidae, and how a complete taxonomic revision is necessary. Four species of nephtheid are currently known from around Greenland: Duva florida (Rathke, 1806), Drifa flavescens (Danielsson, 1887), Drifa groenlandica (Molander, 1915), and Drifa glomerata (Verrill, 1869).

While common around Greenland, cauliflower coral gardens are also observed in eastern Canada, and Northwest and Southeast Iceland (Buhl-Mortensen et al., 2019). To this date, there have been no published population genetics studies of the nephtheids in the North Atlantic, and limited studies on their reproductive biology. Dispersal is a key stage in the life history of coral species, and exerts a strong influence on the recovery potential of species, and patterns of genetic connectivity between populations (Jones et al., 2007; Baco et al., 2016). The study of population genetics in the deep sea allows for the assessment of connectivity and dispersal where traditional ecological studies are challenging. Advances in our knowledge of genetic connectivity between nephtheids is pivotal to the designation of protected areas in Greenland, in a manner which maximises the protection of biodiversity.

This study aims to assess genetic connectivity and population structure among nephtheid populations, covering a geographically extensive range across the west, east, and south of Greenland. Multiple novel SNP markers have been developed and used to evaluate the regional genetic structure of four nephtheid species, across a depth range of 58-1457 metres.
Specifically, the following questions were addressed: (i) do nephtheids show significant regional genetic structure around Greenland, indicative of a departure from panmixia; and (ii) can SNPs be utilised to detect species-specific patterns within this family. This work acts an important step in understanding gene flow around Greenland, in order to inform spatial management that is effective in protecting areas of high diversity and vulnerability.

2. Materials and methods

2.1. Sample collection

All nephtheid samples were caught between 2007 and 2017 as bycatch from stock assessment trawls by the Greenland Institute for Natural Resources (GINR), and the Department of Fisheries and Oceans Canada (DFO). The samples were gathered across a broad geographic distribution covering East and West Greenland, Baffin Bay, and East Canada (Fig. 1.). Samples were collected between 50 – 1500 metres, and after identification to the lowest possible taxonomic level, were labelled and stored in 70% ethanol. Prior to analysis, tissue samples were stored at 5°C, while DNA samples were either stored short-term at -20°C, or long-term at -80°C. Every sample labelled as Nephtheidae was included in analysis, also including two samples that had been identified as Gersemia sp. in order to identify any non-Nephtheidae species following analysis. After collation, samples were categorised as belonging to four regions based on location and surrounding oceanography. These were identified as: Baffin Bay, East Greenland, South Greenland, and Davis Strait (Fig. 1.). Details of each sample, including trawl ID and location can be found in Table 1.
Figure 1: Geographic distribution of Nephtheidae samples and major currents around Greenland. Samples are colour coordinated according to which geographic region they were assigned to. Blue arrows indicate Arctic origin water and red arrows indicate Atlantic origin water. BC = Baffin Current; LC = Labrador Current; IC = Irminger Current; WGC = West Greenland Current; EGC = East Greenland Current. Oceanographic information is from Daniault et al. (2011) and Bacon et al. (2014).
Table 1: Details for all samples included in analyses. Where possible, species identification was undertaken using examination of morphological characteristics.

<table>
<thead>
<tr>
<th>Sample ID</th>
<th>Trawl ID</th>
<th>Date</th>
<th>Longitude</th>
<th>Latitude</th>
<th>Depth (m)</th>
<th>Genus</th>
<th>Species</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gr16</td>
<td>20115136</td>
<td>22/08/11</td>
<td>-31.14112</td>
<td>65.44628</td>
<td>541</td>
<td>Drifa</td>
<td>groenlandica</td>
</tr>
<tr>
<td>Gr23</td>
<td>20115152</td>
<td>24/08/11</td>
<td>-33.15118</td>
<td>65.32718</td>
<td>668</td>
<td>Drifa</td>
<td>groenlandica</td>
</tr>
<tr>
<td>Gr26</td>
<td>20115128</td>
<td>21/08/11</td>
<td>-29.66123</td>
<td>65.53380</td>
<td>636</td>
<td>Drifa</td>
<td>groenlandica</td>
</tr>
<tr>
<td>Gr27</td>
<td>20115152</td>
<td>24/08/11</td>
<td>-33.15118</td>
<td>65.32718</td>
<td>668</td>
<td>Duva</td>
<td>florida</td>
</tr>
<tr>
<td>Gr28</td>
<td>20115143</td>
<td>23/08/11</td>
<td>-32.34307</td>
<td>65.50063</td>
<td>502</td>
<td>Duva</td>
<td>florida</td>
</tr>
<tr>
<td>Gr29</td>
<td>20115121</td>
<td>20/08/11</td>
<td>-29.06910</td>
<td>66.32043</td>
<td>274</td>
<td>Duva</td>
<td>florida</td>
</tr>
<tr>
<td>Gr30</td>
<td>20115153</td>
<td>20/08/11</td>
<td>-33.55612</td>
<td>65.14970</td>
<td>908</td>
<td>Duva</td>
<td>florida</td>
</tr>
<tr>
<td>Gr35</td>
<td>20092043</td>
<td>14/07/09</td>
<td>-60.11863</td>
<td>71.92487</td>
<td>727</td>
<td>Duva</td>
<td>florida</td>
</tr>
<tr>
<td>Gr36</td>
<td>20092030</td>
<td>12/07/09</td>
<td>-57.57410</td>
<td>71.78098</td>
<td>251</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Gr38</td>
<td>20092030</td>
<td>12/07/09</td>
<td>-57.57410</td>
<td>71.78098</td>
<td>251</td>
<td>Duva</td>
<td>florida</td>
</tr>
<tr>
<td>Gr43</td>
<td>20115127</td>
<td>21/08/11</td>
<td>-29.71245</td>
<td>65.61493</td>
<td>358</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Gr44</td>
<td>20093093</td>
<td>02/08/09</td>
<td>-46.20028</td>
<td>60.06947</td>
<td>126</td>
<td>Drifa</td>
<td>glomerata</td>
</tr>
<tr>
<td>Gr48</td>
<td>20092027</td>
<td>12/07/09</td>
<td>-56.64242</td>
<td>71.52985</td>
<td>185</td>
<td>Gersema</td>
<td>-</td>
</tr>
<tr>
<td>Gr49</td>
<td>20092034</td>
<td>13/07/09</td>
<td>-57.24505</td>
<td>72.26392</td>
<td>349</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Gr50</td>
<td>20092032</td>
<td>12/07/09</td>
<td>-56.62347</td>
<td>72.00173</td>
<td>230</td>
<td>Drifa</td>
<td>glomerata</td>
</tr>
<tr>
<td>Gr51</td>
<td>20092034</td>
<td>13/07/09</td>
<td>-57.24505</td>
<td>72.26392</td>
<td>349</td>
<td>Duva</td>
<td>florida</td>
</tr>
<tr>
<td>Gr52</td>
<td>20092040</td>
<td>14/07/09</td>
<td>-60.65625</td>
<td>72.35338</td>
<td>586</td>
<td>Duva</td>
<td>florida</td>
</tr>
<tr>
<td>Gr53</td>
<td>20092032</td>
<td>12/07/09</td>
<td>-56.62347</td>
<td>72.00173</td>
<td>230</td>
<td>Duva</td>
<td>florida</td>
</tr>
<tr>
<td>Gr55</td>
<td>20092063</td>
<td>17/07/09</td>
<td>-59.13068</td>
<td>69.73115</td>
<td>729</td>
<td>Duva</td>
<td>florida</td>
</tr>
<tr>
<td>Gr57</td>
<td>20092027</td>
<td>12/07/09</td>
<td>-56.64242</td>
<td>71.52985</td>
<td>185</td>
<td>Drifa</td>
<td>glomerata</td>
</tr>
<tr>
<td>Gr60</td>
<td>20093051</td>
<td>29/07/09</td>
<td>-49.53243</td>
<td>61.03255</td>
<td>177</td>
<td>Drifa</td>
<td>-</td>
</tr>
<tr>
<td>Gr66</td>
<td>20093113</td>
<td>04/08/09</td>
<td>-44.65978</td>
<td>59.55430</td>
<td>160</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Gr72</td>
<td>20103015</td>
<td>19/07/10</td>
<td>-53.54332</td>
<td>64.05840</td>
<td>267</td>
<td>Drifa</td>
<td>-</td>
</tr>
<tr>
<td>Gr76</td>
<td>20093066</td>
<td>30/07/09</td>
<td>-48.05093</td>
<td>60.37002</td>
<td>133</td>
<td>Drifa</td>
<td>-</td>
</tr>
<tr>
<td>Gr79</td>
<td>20093065</td>
<td>30/07/09</td>
<td>-48.20120</td>
<td>60.32400</td>
<td>434</td>
<td>Drifa</td>
<td>-</td>
</tr>
<tr>
<td>Gr84</td>
<td>20093091</td>
<td>02/08/09</td>
<td>-46.49040</td>
<td>60.10323</td>
<td>160</td>
<td>Drifa</td>
<td>glomerata</td>
</tr>
<tr>
<td>Gr86</td>
<td>20093052</td>
<td>29/07/09</td>
<td>-49.12500</td>
<td>61.11000</td>
<td>58</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Gr87</td>
<td>20093076</td>
<td>01/08/09</td>
<td>-46.89288</td>
<td>60.14983</td>
<td>355</td>
<td>Duva</td>
<td>florida</td>
</tr>
<tr>
<td>Gr102</td>
<td>20105030</td>
<td>29/08/10</td>
<td>-35.91407</td>
<td>64.87347</td>
<td>183</td>
<td>Drifa</td>
<td>groenlandica</td>
</tr>
<tr>
<td>Gr104</td>
<td>20105021</td>
<td>28/08/10</td>
<td>-33.71192</td>
<td>65.46290</td>
<td>181</td>
<td>Drifa</td>
<td>groenlandica</td>
</tr>
<tr>
<td>Gr105</td>
<td>20102038</td>
<td>01/07/10</td>
<td>-57.02800</td>
<td>72.47810</td>
<td>238</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Gr117</td>
<td>20102028</td>
<td>30/06/10</td>
<td>-56.53478</td>
<td>71.37888</td>
<td>224</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Gr119</td>
<td>20102031</td>
<td>30/06/10</td>
<td>-56.80848</td>
<td>71.81735</td>
<td>183</td>
<td>Drifa</td>
<td>glomerata</td>
</tr>
<tr>
<td>Gr124</td>
<td>20094009</td>
<td>06/08/09</td>
<td>-41.75923</td>
<td>60.86378</td>
<td>239</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Gr126</td>
<td>20094006</td>
<td>05/08/09</td>
<td>-41.94387</td>
<td>60.25382</td>
<td>312</td>
<td>Duva</td>
<td>florida</td>
</tr>
<tr>
<td>Gr127</td>
<td>20094006</td>
<td>05/08/09</td>
<td>-41.94387</td>
<td>60.25382</td>
<td>312</td>
<td>Drifa</td>
<td>glomerata</td>
</tr>
<tr>
<td>Gr134</td>
<td>20102036</td>
<td>01/07/10</td>
<td>-56.53183</td>
<td>72.28758</td>
<td>239</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Gr139</td>
<td>20093059</td>
<td>30/07/09</td>
<td>-48.87985</td>
<td>60.80798</td>
<td>116</td>
<td>Drifa</td>
<td>groenlandica</td>
</tr>
<tr>
<td>Gr144</td>
<td>20093059</td>
<td>30/07/09</td>
<td>-48.87985</td>
<td>60.80798</td>
<td>116</td>
<td>Duva</td>
<td>florida</td>
</tr>
<tr>
<td>Gr146</td>
<td>20123018</td>
<td>27/07/12</td>
<td>-50.96157</td>
<td>62.45940</td>
<td>72</td>
<td>Drifa</td>
<td>glomerata</td>
</tr>
<tr>
<td>Gr147</td>
<td>20123018</td>
<td>27/07/12</td>
<td>-50.96157</td>
<td>62.45940</td>
<td>72</td>
<td>Drifa</td>
<td>glomerata</td>
</tr>
<tr>
<td>Gr151</td>
<td>20114085</td>
<td>11/08/11</td>
<td>-33.75907</td>
<td>65.45837</td>
<td>183</td>
<td>Drifa</td>
<td>groenlandica</td>
</tr>
<tr>
<td>Gr153</td>
<td>20114006</td>
<td>01/08/11</td>
<td>-41.83383</td>
<td>60.87335</td>
<td>191</td>
<td>Drifa</td>
<td>glomerata</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gr156</td>
<td>20112042</td>
<td>06/07/11</td>
<td>-57.53520</td>
<td>71.78945</td>
<td>249</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Gr157</td>
<td>20112017</td>
<td>04/07/11</td>
<td>-55.05793</td>
<td>70.28818</td>
<td>177</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Gr159</td>
<td>20123044</td>
<td>22/06/11</td>
<td>-48.24065</td>
<td>60.40865</td>
<td>140</td>
<td>Drifa flavescens</td>
<td></td>
</tr>
<tr>
<td>Gr161</td>
<td>20124139</td>
<td>19/08/12</td>
<td>-29.71967</td>
<td>65.61647</td>
<td>328</td>
<td>Drifa groenlandica</td>
<td></td>
</tr>
<tr>
<td>Gr162</td>
<td>20124138</td>
<td>19/08/12</td>
<td>-29.75365</td>
<td>65.66185</td>
<td>328</td>
<td>Drifa -</td>
<td></td>
</tr>
<tr>
<td>Gr166</td>
<td>20124048</td>
<td>08/08/12</td>
<td>-40.43758</td>
<td>62.42835</td>
<td>439</td>
<td>Drifa flavescens</td>
<td></td>
</tr>
<tr>
<td>Gr167</td>
<td>20124139</td>
<td>19/08/12</td>
<td>-29.71967</td>
<td>65.61647</td>
<td>328</td>
<td>Drifa groenlandica</td>
<td></td>
</tr>
<tr>
<td>Gr170</td>
<td>20124089</td>
<td>08/08/12</td>
<td>-36.06205</td>
<td>64.90307</td>
<td>182</td>
<td>Duva florida</td>
<td></td>
</tr>
<tr>
<td>Gr172</td>
<td>20124136</td>
<td>19/08/12</td>
<td>-30.16977</td>
<td>65.42887</td>
<td>670</td>
<td>Drifa groenlandica</td>
<td></td>
</tr>
<tr>
<td>Gr173</td>
<td>20124048</td>
<td>08/08/12</td>
<td>-40.43758</td>
<td>62.42835</td>
<td>439</td>
<td>Drifa -</td>
<td></td>
</tr>
<tr>
<td>Gr178</td>
<td>20115127</td>
<td>21/08/11</td>
<td>-29.71245</td>
<td>65.61493</td>
<td>358</td>
<td>Drifa -</td>
<td></td>
</tr>
<tr>
<td>Gr186</td>
<td>20115127</td>
<td>21/08/11</td>
<td>-29.71245</td>
<td>65.61493</td>
<td>358</td>
<td>Drifa -</td>
<td></td>
</tr>
<tr>
<td>Gr187</td>
<td>20115127</td>
<td>21/08/11</td>
<td>-29.71245</td>
<td>65.61493</td>
<td>358</td>
<td>Duva florida</td>
<td></td>
</tr>
<tr>
<td>Gr188</td>
<td>20115127</td>
<td>21/08/11</td>
<td>-29.71245</td>
<td>65.61493</td>
<td>358</td>
<td>Drifa groenlandica</td>
<td></td>
</tr>
<tr>
<td>Gr192</td>
<td>20115127</td>
<td>21/08/11</td>
<td>-29.71245</td>
<td>65.61493</td>
<td>358</td>
<td>Drifa glomerata</td>
<td></td>
</tr>
<tr>
<td>Gr193</td>
<td>20115127</td>
<td>21/08/11</td>
<td>-29.71245</td>
<td>65.61493</td>
<td>358</td>
<td>Drifa -</td>
<td></td>
</tr>
<tr>
<td>Gr199</td>
<td>20115127</td>
<td>21/08/11</td>
<td>-29.71245</td>
<td>65.61493</td>
<td>358</td>
<td>Duva florida</td>
<td></td>
</tr>
<tr>
<td>Gr201</td>
<td>20121008</td>
<td>08/06/12</td>
<td>-55.01735</td>
<td>64.51943</td>
<td>359</td>
<td>Drifa glomerata</td>
<td></td>
</tr>
<tr>
<td>4_1</td>
<td>20131004</td>
<td>07/06/13</td>
<td>-54.16773</td>
<td>64.28665</td>
<td>251</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>6_5</td>
<td>20131006</td>
<td>07/06/13</td>
<td>-54.16638</td>
<td>64.45218</td>
<td>159</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>6_7</td>
<td>20131006</td>
<td>07/06/13</td>
<td>-54.16638</td>
<td>64.45218</td>
<td>159</td>
<td>Drifa -</td>
<td></td>
</tr>
<tr>
<td>7_8</td>
<td>20131007</td>
<td>08/06/13</td>
<td>-55.10023</td>
<td>64.54900</td>
<td>422</td>
<td>Duva florida</td>
<td></td>
</tr>
<tr>
<td>7_9</td>
<td>20131007</td>
<td>08/06/13</td>
<td>-55.10023</td>
<td>64.54900</td>
<td>422</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>7_10</td>
<td>20131007</td>
<td>08/06/13</td>
<td>-55.10023</td>
<td>64.54900</td>
<td>422</td>
<td>Drifa -</td>
<td></td>
</tr>
<tr>
<td>7_11</td>
<td>20131007</td>
<td>08/06/13</td>
<td>-55.10023</td>
<td>64.54900</td>
<td>422</td>
<td>Drifa -</td>
<td></td>
</tr>
<tr>
<td>7_14</td>
<td>20131007</td>
<td>08/06/13</td>
<td>-55.10023</td>
<td>64.54900</td>
<td>422</td>
<td>Duva florida</td>
<td></td>
</tr>
<tr>
<td>8_18</td>
<td>20131008</td>
<td>08/06/13</td>
<td>-54.51762</td>
<td>64.63718</td>
<td>271</td>
<td>Duva florida</td>
<td></td>
</tr>
<tr>
<td>10_23</td>
<td>20131010</td>
<td>08/06/13</td>
<td>-54.81835</td>
<td>64.86972</td>
<td>317</td>
<td>Drifa -</td>
<td></td>
</tr>
<tr>
<td>10_24</td>
<td>20131010</td>
<td>08/06/13</td>
<td>-54.81835</td>
<td>64.86972</td>
<td>317</td>
<td>Duva florida</td>
<td></td>
</tr>
<tr>
<td>11_25</td>
<td>20131011</td>
<td>08/06/13</td>
<td>-54.12745</td>
<td>64.88562</td>
<td>134</td>
<td>Drifa -</td>
<td></td>
</tr>
<tr>
<td>11_26</td>
<td>20131011</td>
<td>08/06/13</td>
<td>-54.12745</td>
<td>64.88562</td>
<td>134</td>
<td>Duva florida</td>
<td></td>
</tr>
<tr>
<td>12_29</td>
<td>20131012</td>
<td>08/06/13</td>
<td>-53.55123</td>
<td>64.63690</td>
<td>162</td>
<td>Drifa flavescens</td>
<td></td>
</tr>
<tr>
<td>13_33</td>
<td>20131013</td>
<td>09/06/13</td>
<td>-53.15925</td>
<td>64.61333</td>
<td>269</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>34_55</td>
<td>20131034</td>
<td>12/06/13</td>
<td>-56.37055</td>
<td>67.61258</td>
<td>133</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>42_60</td>
<td>20131042</td>
<td>13/06/13</td>
<td>-57.78742</td>
<td>67.85757</td>
<td>292</td>
<td>Duva florida</td>
<td></td>
</tr>
<tr>
<td>66_80</td>
<td>20131066</td>
<td>17/06/13</td>
<td>-54.80725</td>
<td>69.30393</td>
<td>170</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>69_86</td>
<td>20131069</td>
<td>18/06/13</td>
<td>-57.03458</td>
<td>69.32832</td>
<td>199</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>27_1</td>
<td>20132027</td>
<td>01/07/13</td>
<td>-56.83083</td>
<td>71.97638</td>
<td>180</td>
<td>Gersemia -</td>
<td></td>
</tr>
<tr>
<td>27_2</td>
<td>20132027</td>
<td>01/07/13</td>
<td>-56.83083</td>
<td>71.97638</td>
<td>180</td>
<td>Drifa groenlandica</td>
<td></td>
</tr>
<tr>
<td>27_5</td>
<td>20132027</td>
<td>01/07/13</td>
<td>-56.83083</td>
<td>71.97638</td>
<td>180</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>51_2</td>
<td>20132051</td>
<td>05/07/13</td>
<td>-55.03875</td>
<td>69.52473</td>
<td>138</td>
<td>Drifa glomerata</td>
<td></td>
</tr>
<tr>
<td>52_5</td>
<td>20132052</td>
<td>05/07/13</td>
<td>-56.31003</td>
<td>69.52268</td>
<td>226</td>
<td>Drifa -</td>
<td></td>
</tr>
<tr>
<td>98</td>
<td>PA03036</td>
<td>2014</td>
<td>-49.3954</td>
<td>60.91993</td>
<td>115</td>
<td>Drifa -</td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>PA03039</td>
<td>2014</td>
<td>-48.17157</td>
<td>60.53655</td>
<td>132</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>101</td>
<td>PA03039</td>
<td>2014</td>
<td>-48.17157</td>
<td>60.53655</td>
<td>132</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>106</td>
<td>PA03052</td>
<td>2014</td>
<td>-46.37507</td>
<td>59.91207</td>
<td>315</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>
2.2. DNA extraction

DNA extraction of samples was undertaken at the ZSL Institute of Zoology (London) between 2009 and 2018 by C. Yesson and B. Neves using the Bioline Isolate II Genome DNA kit, following the manufacturer’s instructions. Total genomic DNA extraction was performed on 20 new samples required for analysis using QuickExtract™ DNA extraction solution (Lucigen), following manufacturer guidelines, and adapted for a digestion time of 20 minutes. Any excess tissue after digestion was discarded. Upon genotyping, it was found that samples extracted using QuickExtract™ had universally not been successful and so were discarded from further analysis. Following extraction, DNA concentration was assessed using a Nanodrop 8000 spectrophotometer (Thermofisher). Any extractions which did not meet the required genomic yield for genotyping (5 ng.µl⁻¹) were discarded and not used for analysis.

2.3. SNP discovery

Protocol for developing SNPs was adapted from Yesson et al. (2018b). Potential SNP targets were developed via examining data from shotgun sequence reads from the Illumina Miseq platform (https://www.illumina.com/systems/sequencing-platforms/miseq.html). Twelve nephtheid samples of the species Duva florida, Drifa flavescens, Drifa groenlandica, and Drifa glomerata were independently sequenced on the Miseq. Sequence quality assessment was
carried out using FastQC (v.09; Babraham Bioinformatics, Cambridge, UK). Sequences with mean quality scores below 20 were removed.

The short read sequences were mapped to a reference *Dendronephthya gigantea* whole genome sequence. The software Geneious Prime v.2020.0.5 was used to identify potential SNP loci based on a number of control criteria: the genome fragment must have at least 50 base pairs either side of the SNP locus to allow for primer development; a minimum of 4 x read coverage; and, the SNP is not near other potential SNPs. A range of SNPs were selected to include both intra- and interspecific variation in variable base pairs. Putative SNPs were BLAST matched using the NCBI database (https://blast.ncbi.nlm.nih.gov/Blast.cgi). Any sequences producing positive matches to other taxa were discarded, assuming contamination. A final subset of 24 SNPs were selected for genotyping, however, two of these failed amplification for the majority of samples, leaving 22 SNPs for analysis.

2.4. Genotyping

Genotyping was performed using a Kompetitive Allele Specific PCR (KASP) assay (Semagn *et al*., 2014) and undertaken by LGC genomics (https://www.biosearchtech.com/products/pcr-kits-and-reagents/genotyping-assays/kasp-genotyping-chemistry). KASP is a homogeneous, fluorescence-based genotyping technology, utilising a quantitative polymerase chain reaction (qPCR) process to test for SNP variants. The technology involves allele-specific oligo extension fluorescence resonance energy transfer (FRET) to generate a signal. DNA samples were placed into two 96-well plates with two control wells, at concentrations of over 5ng.µl⁻¹ at 40µl volumes before being sent for genotyping. Results were delivered as bi-allelic scoring for each of the 22 SNPs.

2.5. Statistical analyses

All statistical analyses were performed in the software R v.3.6.3 (https://cran.r-project.org/), and adapted from Yesson *et al*., (2018b). Linkage disequilibrium (LD) between SNP markers was assessed using the LD function of the R package ‘genetics’ (Warnes *et al*., 2012). A Bonferroni correction was applied to *P*-values to assess significance. This correction is a
conservative test to minimise the chance of Type I errors. Summary statistics including observed and expected heterozygosity (Ho and He, respectively), were calculated using the functions ‘basic.stats’ and ‘HWE.test’ from packages ‘hierfstat’ and ‘genetics’ respectively (Goudet and Jombart, 2015).

Structure occurring between regions was assessed using a pairwise fixation index (Fst) (Weir and Cockerham, 1984). Significance was measured based on 1,000 permutations using the ‘stamppFst’ function in the R package ‘StAMPP’ (Pembleton et al., 2013). Isolation by distance (IBD) was tested using a Mantel test. Genetic distances were measured by Fst, and geographic distances measured as average distance between regions over water, approximated with the distance tool in Quantum GIS. The mantel test was performed using the R package ‘adegenet’ and the function ‘mantel.randtest’ with 100,000 permutations.

Assessment of genetic patterns in the data was conducted using a Discriminant Analyses of Principle Components (DAPC) using the R package ‘adegenet’ (Jombart, 2008). The number of genetic clusters present was examined using the ‘find.cluster’ function of adegenet. This runs successive K-means for clustering and assesses the optimal number by reference to Bayesian information criterion. The automatic cluster selection procedure ‘diffNgroup’ was used with n.iter set to $10^7$ and n.start set at 10,000. Principle components to retain for analysis were selected based on a percentage of variance explained threshold (95%). The DAPC analysis was performed using the ‘dapc’ function. The optimal number of principle components to retain was assessed using the ‘optim.a.score’ function.
3. Results

3.1. SNP discovery and genotyping

A total of 105 samples from four regions were collected and genotyped (Table 1; Fig. 1.). The 22 SNP markers analysed are presented in Table 2, with observed and expected heterozygosity. Significant deviations from HWE were observed for 15 of the 22 markers (Table 2). Significant linkage (LD) was found between 105 out of 231 pairs of markers. The largest subsets of unlinked markers contained six SNPs, with two combinations of six unlinked markers, set 1: 4, 9, 10, 11, 19, 22; and set 2: 4, 10, 11, 17, 19, 22. Analyses were conducted on the complete dataset of 22 SNPs for samples with at least 15 non-null SNPs (n = 105), no samples failed to meet this criteria. Parallel analyses were conducted using the subset of unlinked markers with the least null alleles (set 1). This analysis produced similar results and clustering to that using the full dataset of 22 SNPs. Considering these results, and that a DAPC is not influenced by marker linkage (Jombart et al., 2010), the original set of SNPs was used for the analysis.

3.2. Regional population structure

Significant genetic differentiation ($F_{st}$, $P < 0.05$) was found between all regions (Table 3). Both the greatest genetic and geographic distances were found between Baffin Bay and East Greenland, for both the complete and unlinked datasets. However, there was not a significant pattern of isolation by distance shown in the dataset (complete dataset: $Z = 0.19$, $P = 0.33$), indicating that geographically distant regions did not show any greater genetic divergence than less distant regions (Fig.2.a). There was also no significant pattern of isolation by depth (complete dataset: $Z = 0.15$, $P = 0.38$) (Fig.2.b). A repeat analysis was conducted on the species for which there was the largest number of samples ($Duva florida$, n = 27). Pairwise $F_{st}$ values ranged from -0.034 to -0.006, with no significant genetic differentiation between regions observed ($P > 0.7$).
<table>
<thead>
<tr>
<th>SNP ID</th>
<th>Sequence</th>
<th>Variant</th>
<th>Null (%)</th>
<th>Ho</th>
<th>Hs</th>
<th>Ht</th>
<th>Dst</th>
<th>Dstp</th>
<th>Fis</th>
<th>Dest</th>
</tr>
</thead>
<tbody>
<tr>
<td>nephtheid_1</td>
<td>ATTTATCCCGAGAT[A/C]TCGGGATTTTTTGAA</td>
<td>A/C</td>
<td>18.103</td>
<td>0.056</td>
<td>0.075</td>
<td>1</td>
<td>0.116</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>nephtheid_2</td>
<td>GCTTTAGCAGTACT[A/T]CGATCGACTACGA</td>
<td>A/T</td>
<td>8.621</td>
<td>0.356</td>
<td>0.412</td>
<td>0.056</td>
<td>0.075</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>nephtheid_3</td>
<td>ATATAAGCCTACCA[A/C]GTGAGTATTTTCCTCC</td>
<td>A/C</td>
<td>25.862</td>
<td>0.412</td>
<td>0.476</td>
<td>0.056</td>
<td>0.075</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>nephtheid_4</td>
<td>TTTGAGCGTAACTT[C/T]AGCAACTTTCAAC</td>
<td>C/T</td>
<td>30.172</td>
<td>0.056</td>
<td>0.075</td>
<td>0.012</td>
<td>0.075</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>nephtheid_5</td>
<td>TTTGAGCGTAACTT[A/G]GAAACTTTCAAC</td>
<td>A/G</td>
<td>18.103</td>
<td>0.412</td>
<td>0.476</td>
<td>0.056</td>
<td>0.075</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>nephtheid_6</td>
<td>TTTTATCGCTTAACTT[A/G]CATGTTTAAAAATTTG</td>
<td>A/G</td>
<td>8.621</td>
<td>0.056</td>
<td>0.075</td>
<td>0.012</td>
<td>0.075</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>nephtheid_7</td>
<td>TTTGACCGTTCTGTC[A/G]ACGCTTCTTATCCA</td>
<td>A/G</td>
<td>8.621</td>
<td>0.056</td>
<td>0.075</td>
<td>0.012</td>
<td>0.075</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>nephtheid_8</td>
<td>TTTTATCGCTTCTT[A/G]ATGGACCGCGATG</td>
<td>A/G</td>
<td>8.621</td>
<td>0.056</td>
<td>0.075</td>
<td>0.012</td>
<td>0.075</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>nephtheid_9</td>
<td>TTTTATCGCTTCTT[A/G]ATGGACCGCAGATG</td>
<td>A/G</td>
<td>8.621</td>
<td>0.056</td>
<td>0.075</td>
<td>0.012</td>
<td>0.075</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>nephtheid_10</td>
<td>TTTTATCGCTTCTT[A/G]ATGGACCGCAGATG</td>
<td>A/G</td>
<td>8.621</td>
<td>0.056</td>
<td>0.075</td>
<td>0.012</td>
<td>0.075</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>nephtheid_11</td>
<td>TTTTATCGCTTCTT[A/G]ATGGACCGCAGATG</td>
<td>A/G</td>
<td>8.621</td>
<td>0.056</td>
<td>0.075</td>
<td>0.012</td>
<td>0.075</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>nephtheid_12</td>
<td>TTTTATCGCTTCTT[A/G]ATGGACCGCAGATG</td>
<td>A/G</td>
<td>8.621</td>
<td>0.056</td>
<td>0.075</td>
<td>0.012</td>
<td>0.075</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>nephtheid_13</td>
<td>TTTTATCGCTTCTT[A/G]ATGGACCGCAGATG</td>
<td>A/G</td>
<td>8.621</td>
<td>0.056</td>
<td>0.075</td>
<td>0.012</td>
<td>0.075</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>nephtheid_14</td>
<td>TTTTATCGCTTCTT[A/G]ATGGACCGCAGATG</td>
<td>A/G</td>
<td>8.621</td>
<td>0.056</td>
<td>0.075</td>
<td>0.012</td>
<td>0.075</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>nephtheid_15</td>
<td>TTTTATCGCTTCTT[A/G]ATGGACCGCAGATG</td>
<td>A/G</td>
<td>8.621</td>
<td>0.056</td>
<td>0.075</td>
<td>0.012</td>
<td>0.075</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>nephtheid_16</td>
<td>TTTTATCGCTTCTT[A/G]ATGGACCGCAGATG</td>
<td>A/G</td>
<td>8.621</td>
<td>0.056</td>
<td>0.075</td>
<td>0.012</td>
<td>0.075</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>nephtheid_17</td>
<td>TTTTATCGCTTCTT[A/G]ATGGACCGCAGATG</td>
<td>A/G</td>
<td>8.621</td>
<td>0.056</td>
<td>0.075</td>
<td>0.012</td>
<td>0.075</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>nephtheid_18</td>
<td>TTTTATCGCTTCTT[A/G]ATGGACCGCAGATG</td>
<td>A/G</td>
<td>8.621</td>
<td>0.056</td>
<td>0.075</td>
<td>0.012</td>
<td>0.075</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>nephtheid_19</td>
<td>TTTTATCGCTTCTT[A/G]ATGGACCGCAGATG</td>
<td>A/G</td>
<td>8.621</td>
<td>0.056</td>
<td>0.075</td>
<td>0.012</td>
<td>0.075</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>nephtheid_20</td>
<td>TTTTATCGCTTCTT[A/G]ATGGACCGCAGATG</td>
<td>A/G</td>
<td>8.621</td>
<td>0.056</td>
<td>0.075</td>
<td>0.012</td>
<td>0.075</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>nephtheid_21</td>
<td>TTTTATCGCTTCTT[A/G]ATGGACCGCAGATG</td>
<td>A/G</td>
<td>8.621</td>
<td>0.056</td>
<td>0.075</td>
<td>0.012</td>
<td>0.075</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>nephtheid_22</td>
<td>TTTTATCGCTTCTT[A/G]ATGGACCGCAGATG</td>
<td>A/G</td>
<td>8.621</td>
<td>0.056</td>
<td>0.075</td>
<td>0.012</td>
<td>0.075</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>nephtheid_23</td>
<td>TTTTATCGCTTCTT[A/G]ATGGACCGCAGATG</td>
<td>A/G</td>
<td>8.621</td>
<td>0.056</td>
<td>0.075</td>
<td>0.012</td>
<td>0.075</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>nephtheid_24</td>
<td>TTTTATCGCTTCTT[A/G]ATGGACCGCAGATG</td>
<td>A/G</td>
<td>8.621</td>
<td>0.056</td>
<td>0.075</td>
<td>0.012</td>
<td>0.075</td>
<td>1</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Table 2: SNP markers and genetic estimates for nephtheid samples.** Null = percentage of null alleles (n = 105); Ho = mean observed heterozygosity; Hs = mean gene diversity within population (expected heterozygosity); Ht = overall gene diversity; Dst = gene diversity among samples; Dstp = heterozygosity corrected Dst; Fis = inbreeding coefficient per overall loci; Dest = measure of population differentiation. Bold values are significant (P < 0.05).
Table 3: Regional genetic and geographic distances for nephtheid samples. Upper triangle shows approximate over-water distances in km. Lower triangle shows pairwise $F_{st}$ values (all $P < 0.05$).

<table>
<thead>
<tr>
<th>Region</th>
<th>Code</th>
<th>BB</th>
<th>DS</th>
<th>SG</th>
<th>EG</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baffin Bay</td>
<td>BB</td>
<td>-</td>
<td>852</td>
<td>2320</td>
<td>3045</td>
</tr>
<tr>
<td>Davis Strait</td>
<td>DS</td>
<td>0.0388</td>
<td>-</td>
<td>1931</td>
<td>2115</td>
</tr>
<tr>
<td>South Greenland</td>
<td>SG</td>
<td>0.0180</td>
<td>0.0197</td>
<td>-</td>
<td>1884</td>
</tr>
<tr>
<td>East Greenland</td>
<td>EG</td>
<td>0.0586</td>
<td>0.0182</td>
<td>0.0499</td>
<td>-</td>
</tr>
</tbody>
</table>

Figure 2: Comparison between genetic distance (pairwise $F_{st}$ values) and physical distance between depth bins/regions (m/km). Analysis showed no significant pattern of **a)** isolation by distance ($P > 0.05$) or **b)** isolation by depth ($P > 0.05$).
3.3. Cluster analysis

Cluster analysis selected six genetic clusters with some degree of site specificity, in concordance with the significant $F_{st}$ results (Table 3). There appeared to be an East/West Greenland split, with clusters 3 and 6 largely occurring in the Baffin Bay and Davis Strait regions, while cluster 2 was confined to East Greenland (Fig. 3). Cluster 2 was the only completely site specific cluster, occurring in East Greenland, and is characterised by low heterozygosity. However, other samples from the East Greenland region were assigned to clusters 3, 4, and 5. Cluster 6 was predominantly limited to Baffin Bay, with only two samples from the South Greenland region assigned to this cluster (Fig. 3; Table 4.a). The greatest genetic similarity was found between the Baffin Bay and South Greenland regions (Table 3). This can be seen in the geographic pattern of clustering, as cluster 5 in particular covers both of these regions (Fig. 3). Cluster 4 is spread across the Davis Strait, and East and South Greenland, but was not found in Baffin Bay. A similar pattern of clustering was produced using just unlinked markers, with a comparable East/West Greenland split observed.

Table 4: Assignment of samples to genetic clusters. a) Regionally, b) depth-bin, c) species. N = number of samples.

<table>
<thead>
<tr>
<th>Genetic cluster</th>
<th>a) Region</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baffin Bay</td>
<td>37</td>
<td>1</td>
<td>0</td>
<td>9</td>
<td>0</td>
<td>8</td>
<td>19</td>
</tr>
<tr>
<td>Davis Strait</td>
<td>18</td>
<td>1</td>
<td>0</td>
<td>5</td>
<td>11</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>East Greenland</td>
<td>23</td>
<td>0</td>
<td>12</td>
<td>4</td>
<td>6</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>South Greenland</td>
<td>27</td>
<td>4</td>
<td>0</td>
<td>5</td>
<td>8</td>
<td>8</td>
<td>2</td>
</tr>
<tr>
<td>b) Depth (m)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0-200</td>
<td>36</td>
<td>3</td>
<td>3</td>
<td>4</td>
<td>7</td>
<td>12</td>
<td>7</td>
</tr>
<tr>
<td>200-600</td>
<td>47</td>
<td>3</td>
<td>5</td>
<td>12</td>
<td>15</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>600-800</td>
<td>14</td>
<td>0</td>
<td>4</td>
<td>5</td>
<td>2</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>800-1500</td>
<td>8</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>c) Species</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Duva florida</td>
<td>27</td>
<td>0</td>
<td>0</td>
<td>20</td>
<td>6</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Drifa glomerata</td>
<td>15</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>10</td>
<td>3</td>
</tr>
<tr>
<td>Drifa flavescens</td>
<td>3</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Drifa groenlandica</td>
<td>12</td>
<td>0</td>
<td>9</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>Gersemia</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>Unidentified</td>
<td>47</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>17</td>
<td>6</td>
<td>16</td>
</tr>
</tbody>
</table>
Cluster assignment showed no clear patterns with depth (Table 4.b), however it did appear to be largely species specific (Table 4.c). Cluster 2, which was geographically limited to East Greenland, is composed of *D. groenlandica* samples, with three unidentified samples. In comparison, cluster 3 shows the broadest geographic spread (Fig. 3), however appears to be species-specific to *D. florida* (Table 4.c). When cluster assignment was visualised against samples on an ultrametric phylogeny of nephtheid and *Gersemia spp.* sequences, the species-specificity of each cluster can be seen (Fig. 4). Based on this, clusters can be identified by species: cluster 1 contains *D. flavescens*; cluster 3 contains *D. florida*; cluster 5 contains *D. glomerata*; clusters 2 and 4 contain *D. groenlandica*; and cluster 6 contains *G. fruticosa*.

Figure 3: Geographic spread of genetic clusters. Each sample is colour coded with cluster assignment.
Figure 4: Genetic cluster assignment visualised on a maximum likelihood, ultrametric phylogeny of nephtheid and Gersemia spp. samples. Maximum likelihood phylogeny based on MutS and ITS1/2 mitochondrial DNA and nuclear DNA regions respectively. Triangles represent nodes collapsed for clarity. Branches are labelled by ZSL ID, morphological species ID, longitude, latitude, and depth (m). Samples analysed in the present study are highlighted with their genetic cluster assignment. Sequences were generated by Murphy (2014) and Ayers (2016). Phylogeny is courtesy of Yesson (Pers. Comm.).
The DAPC used two principle components and two discriminant functions, which accounted for 70% of the observed variance. The scatter plot (Fig. 5), shows the relative isolation of cluster 2 (East Greenland sub-population). Clusters 3, 4, and 6 show much closer affinities. Cluster 5, while geographically close to clusters 3 and 4 (Fig. 3), appears more genetically isolated than the other clusters (Fig. 5). Cluster assignment probabilities were typically high (76/105 higher than 90%).

Figure 5: Visualisation of discriminant analysis of principle components (DAPC) of nephtheid samples. Each point represents an individual. Ellipses show 75% confidence interval around each genetic cluster. Proportion of conserved variance is 70%, with axis 1 (x) accounting for 39% and axis 2 (y) accounting for 31%.
4. Discussion

In this study, SNP markers were developed for the first time for nephtheid corals, with 22 SNP loci successfully identified and genotyped across 105 individuals. The application of these markers has revealed both significant species-specific patterns and regional differentiation between Nephtheidae populations around Greenland. These markers therefore represent a useful tool for analysing genetic connectivity in these important, habitat-forming species.

4.1. Regional patterns of differentiation

The presence of significant regional population structure within the nephtheid data presented here is consistent with patterns seen in other deep-sea coral species, such as: *Lophelia pertusa* (Dahl et al., 2012), *Callogorgia delta* (Quattrini et al., 2015), and *Paragorgia arborea* (Herrera et al., 2012). The scleractinian species *L. pertusa* shows significant geographic partitioning within the North Atlantic, with grouping corresponding to ocean regions (Morrison et al., 2011). Different oceanographic processes and current regimes can create barriers to gene flow between regions, enhancing genetic differentiation. The low but significant pairwise genetic differentiation found between regional-scale nephtheid populations may suggest restricted gene flow between sampling locations. The circulation around Greenland displays complicated dynamics (Hopkins, 1991), mediated by the combined influence of warmer Atlantic water from the Irminger current (IC), and colder Arctic-derived water from the East Greenland and Baffin Currents (Bacon et al., 2014). Both the greatest genetic and geographic distances were found between the Baffin Bay and East Greenland regions. Baffin Bay is separated from the Labrador Sea and Southwest Greenland by the Greenland-Canada Ridge underlying the Davis Strait (Riis-Carstensen, 1948). These conditions have been shown to result in partial genetic isolation of populations of Greenland halibut (*Reinhardtius hippoglossoides*) across a similar study area (Knutsen et al., 2007). Compared with the sessile nephtheids, Greenland halibut are a highly mobile species with a large effective population size – characteristics which would suggest panmixia across a broad geographic range. The observed genetic structure of *R. hippoglossoides* sub-populations implies that the current systems around Greenland may act as larval retention systems and barriers to gene flow.
Despite the significant $F_{st}$ results, the lack of a significant pattern of IBD within the data indicates some level of connectivity and should not be overlooked. A significant pattern of IBD is expected if dispersal distances are lower than the range of study, assuming spatially homogenous gene flow and a balance between loss of alleles due to genetic drift, and replacement of alleles by gene flow (Bradbury and Bentzen, 2007). Within population studies of corals, significant IBD has been attributed to constrained larval dispersal limiting gene flow in both shallow- and deep-water species (Maier et al., 2005; Morrison et al., 2011). However, when utilising IBD slopes as a proxy of dispersal distance, the extent and duration of disturbance and disequilibrium within and between populations must be evaluated (Taylor and Roterman, 2017). In their meta-analysis of deep-sea IBD slopes, Baco et al. (2016) found that estimates of dispersal distance differed significantly between species in relation to taxonomic and life-history factors. The persistence of non-equilibrium factors in an environment impacts individual $F_{st}$ values, as population structure can only be used to infer current levels of connectivity where genetic drift and migration are in equilibrium. Therefore, a non-significant IBD result does not necessarily indicate panmixia across a region. If recent disturbance has triggered a range expansion or recolonization event, genetic similarity between subpopulations may be the result of shared ancestral polymorphisms (Slatkin, 1993). Low but significant $F_{st}$ values, as found among the nephtheids presented here, may not reflect moderate but restricted gene flow, but instead indicate a pattern of low-level gene flow coupled with shared ancestral polymorphisms (Marko and Hart, 2011). It has therefore been suggested that genealogical methods, such as isolation by migration (Hey and Nielsen, 2004), should be more widely used for population genetics studies where equilibrium cannot be assumed (Anderson et al., 2010), and so should be considered in future work on nephtheids.

When looking at other invertebrate species in the Northwest Atlantic, studies have shown high levels of gene flow around Greenland and East Canada, with little population differentiation. Microsatellite analysis of the snow crab (*Chionoecetes opilio*) revealed a notable absence of genetic structure along West Greenland down to the South Labrador Sea at depths of 20-500m (Puebla et al., 2008). This was hypothesised to be caused by surface circulation facilitating connectivity due to the East and West Greenland currents. Ribergaard
et al. (2004) applied a circulation model to study particle transport on the West Greenland shelf, which revealed strong northward transport of particles over long distances. This may explain the low genetic differentiation seen between the South Greenland and Baffin Bay regions. However, high connectivity from Southwest Greenland northwards to Baffin Bay is at odds with the higher differentiation found between the Davis Strait and Baffin Bay regions. The finding of regional genetic differentiation is also inconsistent with findings of limited population structure and panmixia across large geographic distances in the octocoral species: *Funiculina quadrangularis* (Wright et al., 2015), *Swiftia simplex* (Everett et al., 2016), and *Narella versluysi* (Yesson et al., 2018a). In these examples, prevailing current regimes were proposed to support genetic connectivity between distant populations, a pattern which would be expected around Greenland due to the connection seen in the predominant oceanographic systems.

The potential disparity in IBD and $F_{st}$ results could be explained by a taxonomic skew in regional samples, as well as limited sampling. Samples used for this study were not caught in species-specific surveys, therefore, there is not an even species spread across each region. Octocorals have been shown to display species-specific patterns in dispersal, which are then reflected in contrasting patterns of population structure and gene flow. Microsatellite analysis of the temperate octocoral species *Eunicella verrucosa* and *Alycyonium digitatum* revealed vastly different levels of population structure across the same study area (Holland et al., 2017). Differences in gene flow between populations were hypothesised to derive from differences in reproductive biology and life-history (Jenkins et al., 2018). The analysis of regional differentiation utilising just data from *Duva florida* samples presented here, found no significant genetic structure. This may indicate that analyses of genetic variance ($F_{st}$) were clouded by unequal taxonomic representation among regions. However, the low sample size of the *D. florida* analysis is unlikely to provide a clear signal, and as such, should be interpreted with caution. Increasing sampling of all nephtheid species therefore represents a key goal and should be the focus of future research efforts.

The large time frame across which samples were caught (2007-2017) may also be a confounding factor impacting genetic results. Genetic structure has been shown to vary temporally between coral populations (Morrison et al., 2011), and as such, temporal variation
in nephtheid population dynamics may be responsible for the significant regional differentiation seen. The longevity of nephtheid species is not currently known, however deep-water sea pens (Pennatulacea) have been shown to have decadal longevity (Neves et al., 2015). Should nephtheids have similar longevity, then the 10 year difference between sample collections could have significantly impacted the genetic results. This also prompts questions as to whether DNA degradation of older samples may have occurred, as has been evidenced to occur with prolonged ethanol storage of coral tissue (Gaither et al., 2011).

4.2. Inter-specific patterns of differentiation

Cluster analysis selected six genetic clusters which were largely species specific, with each selected cluster associated with a nephtheid species. This possibly confirms the hypothesis that regional patterns in differentiation were the result of taxonomic biases in sampling, rather than representing true genetic structure. Investigations into marine population structure have found that connectivity can vary greatly between taxa, including between closely related species over similar spatial scales (Bargelloni et al., 2005; Charrier et al., 2006). Differences in spawning times, rates of asexual reproduction, and effective population size have been hypothesised to cause variation in population structure among octocoral species occupying similar geographic ranges (Dahl et al., 2012; Holland et al., 2017). It is therefore within reason to assume that nephtheid species will not exhibit uniform patterns of gene flow, and may have species-specific population structure around Greenland. However, this hypothesis is difficult to verify due to the paucity of knowledge regarding the distributions, reproductive biology, and life-history of nephtheid species.

Reproduction in octocorals can be classified into two modes: brooding and broadcast spawning. Broadcast spawning produces a large number of oocytes, which are fertilised in the water column and can be widely dispersed, whereas larvae which have been brooded have a shorter dispersive phase (Sebens, 1983; Richmond and Hunter, 1990). The studies conducted to date about reproductive processes in nephtheids indicate that the species G. fruiticosa, D. florida, and D. glomerata are all brooders, with continuous production of planula larvae (Sun et al., 2009; Sun et al., 2010a; Sun et al., 2011). Environmental factors, including temperature, primary and secondary productivity, and depth have been found to impact patterns of larval
release. Peaks in planulation have been shown to coincide with seasonal fluctuations in primary productivity among colonies of *Drifa* spp. (Sun et al., 2010b). Differences in reproductive periodicity have potentially significant ramifications for patterns of genetic structure around Greenland, due to seasonal water mass formations and variations in the prevailing currents (Daniault et al., 2011; Bacon et al. 2014). Furthermore, depth has been shown to significantly influence reproduction in *D. glomerata*, as seasonality of planulation was more pronounced in colonies from 100-200m than in those from > 200m depth (Sun et al. 2010a). The coupling of physical oceanographic models with genetic methods has recently been demonstrated as effective in examining population connectivity of deep-sea corals (Bracco et al., 2019), and should therefore be considered in future studies of nephtheids.

Research is further hampered by historical difficulties in both the morphological and phylogenetic taxonomy of nephtheids. The Anthozoa show particularly low rates of mitochondrial DNA (mtDNA) evolution (Shearer et al., 2002). As a result, the mitochondrial genes commonly used for phylogenetic analyses lack the resolution necessary for discriminating species within octocoral genera (McFadden et al., 2006). More recent phylogenetic analyses of the nephtheids have found limited success utilising combinations of the MutS mtDNA region, and ITS nuclear ribosomal DNA marker (McFadden and van Ofwegen, 2013; Murphy, 2014; Ayre, 2016). Results presented here suggest that SNP markers have the potential to be utilised as a species delimitation method among the Northwest Atlantic nephtheids. Each genetic cluster aligned clearly with species clades when assessed against a combined ITS/MutS phylogeny, implying that the identified SNPs could be used in future taxonomic work where morphological characters are unclear. A combination of nuclear genes and SNP markers have been used successfully to identify sympatric species in the octocoral genus *Ovabunda* (McFadden et al., 2017). High-throughput genomic DNA sequencing to detect species-specific SNPs has also been demonstrated as an effective means for species identification among closely related *Heliopora* coral species (Iguchi et al., 2019). Moreover, the use of qPCR represents a faster and cheaper method of species identification than the sequencing of long DNA regions (Andrews et al., 2016). These methods do, however, rely on a single nucleotide to be descriptive and definite across a whole species, and as such, random mutations could result in misidentification. While there are still numerous challenges
to overcome, the use of SNP data represents a promising new direction for the study of Nephtheidae taxonomy.

The only species which showed variation in cluster assignment was *D. groenlandica*, which phylogenetically aligned with both clusters 2 and 4. Cluster 2 was also the only cluster to be regionally isolated, found only in East Greenland. Jensen (2003) defined the species *D. groenlandica* as having a limited distribution, occurring exclusively in East Greenland. This fits with the present results, and suggests that there may be a true *D. groenlandica* species represented by cluster 2, and a closely related, undescribed *Drifa sp.* represented by cluster 4. However, this hypothesis requires further morphological analysis and biological evidence to ascertain. Cluster 2 also exhibits a significant heterozygote deficiency. This could be explained by spatially restricted gene flow and consequent self-recruitment, as has been demonstrated in shallow-water coral populations (Ridgway et al., 2001). However, this appears unlikely as clusters 3, 4, and 5 also occur in the same geographic area, and show no spatial restrictions in distribution. There is currently no published research on reproduction in *D. groenlandica*, and as such it can be hypothesised that this species has a specific pattern of reproduction which results in either inbreeding, or restricted dispersal of larvae and self-recruitment.

4.3. Conservation implications and future directions

The expansion of deep-sea fisheries around Greenland is having increasingly negative effects on nephtheid communities (Pauly et al., 2003), highlighting the need for designated reserves to conserve the biodiversity supported by these ecosystems. All regions sampled within this study contain a minimum of three nephtheid species, with each species potentially displaying unique patterns in population structure due to reproductive differences. The presence of significant genetic differentiation between regions indicates that the Greenland nephtheids cannot be considered as one panmictic population, and the designation of protected areas should reflect this. The results of large scale analyses of connectivity in the deep sea suggest reserve size and spacing should be adequate to ensure a significant level of self-recruitment across taxa, while allowing sufficient space to facilitate external recruitment (Palumbi, 2003; Baco et al., 2016). In order to successfully designate protected areas, further investigation is
needed to unravel the specific population genetic structure of each nephtheid species. The potential for variation in connectivity between species suggests that a network of protected areas would be beneficial for the maintenance of taxonomic diversity, rather than the disjointed conservation methods currently in place (UNEP-WCMC, 2019).

Population genetic theory suggests that the East Greenland population of *D. groenlandica* (cluster 2) may be particularly vulnerable to disturbance given its low levels of heterozygosity and genetic isolation. The apparent isolation of this population implies that disturbance could cause serious long-term consequences in diversity loss, as there may be no source of new recruits to recolonise the area should the population be damaged. However, the samples used in this study had a relatively narrow range on the east-coast of Greenland, whereas *D. groenlandica* has also been reported to occur around Iceland (Jensen, 2003). Ergo, sampling effort may not be truly representative of the diversity of the area. Research effort should be focused on gaining a greater understanding of the distribution, reproduction, and dispersal patterns of this species to inform a successful management plan.

4.4. Conclusion

This study reports the first SNP markers developed for nephtheid corals, as well as the first population genetics study of the group. These markers have shown promise as potential species-delimitation tools for a family with a history of taxonomic complications. There appears to be some level of genetic structuring between regions, however results may be clouded by a taxonomic skew between regions and temporal variation between samples. A combination of genetic estimates of connectivity, modelling of dispersal, and biological data on reproductive strategies should be the next steps for future research, and will improve the quality of inferences regarding structure among nephtheid populations. With continued anthropogenic threats to deep-sea habitats around Greenland, information regarding connectivity between populations is vital for the development of ecologically informed conservation of these vulnerable habitats. The integration of genetic, biological, and oceanographic data is necessary to unravel complicated inter-specific patterns of connectivity in the deep sea, and will be invaluable in protecting these diverse ecosystems.
Acknowledgements

I would like to offer many thanks to my supervisor Chris Yesson for his guidance and support throughout the completion of this study. Thanks are also due to Luigi Colin for his knowledge and assistance with coral DNA extraction protocols.

References:


UNEP-WCMC., 2019. Protected Area Profile for Greenland from the World Database of Protected Areas., [https://www.protectedplanet.net/country/GRL](https://www.protectedplanet.net/country/GRL) [Accessed: 03/03/2020].


