

# **Methods and results from remotely operated vehicle (ROV) survey PAC2017-030: Exploring high and low current areas in the Salish Sea**

Jessica Nephin, Sharon Jeffery, Mary Thiess, Stephanie Archer, Ian Murdock, Jonathan Boschen-Rose, and Sarah Dudas

Institute of Ocean Sciences  
Science Branch, Pacific Region  
Fisheries and Oceans Canada  
9860 West Saanich Road  
Sidney, British Columbia, V8L 4B2  
Canada

2020

**Canadian Technical Report of  
Fisheries and Aquatic Sciences 3405**



## **Canadian Technical Report of Fisheries and Aquatic Sciences**

Technical reports contain scientific and technical information that contributes to existing knowledge but which is not normally appropriate for primary literature. Technical reports are directed primarily toward a worldwide audience and have an international distribution. No restriction is placed on subject matter and the series reflects the broad interests and policies of Fisheries and Oceans Canada, namely, fisheries and aquatic sciences.

Technical reports may be cited as full publications. The correct citation appears above the abstract of each report. Each report is abstracted in the data base *Aquatic Sciences and Fisheries Abstracts*.

Technical reports are produced regionally but are numbered nationally. Requests for individual reports will be filled by the issuing establishment listed on the front cover and title page.

Numbers 1-456 in this series were issued as Technical Reports of the Fisheries Research Board of Canada. Numbers 457-714 were issued as Department of the Environment, Fisheries and Marine Service, Research and Development Directorate Technical Reports. Numbers 715-924 were issued as Department of Fisheries and Environment, Fisheries and Marine Service Technical Reports. The current series name was changed with report number 925.

## **Rapport technique canadien des sciences halieutiques et aquatiques**

Les rapports techniques contiennent des renseignements scientifiques et techniques qui constituent une contribution aux connaissances actuelles, mais qui ne sont pas normalement appropriés pour la publication dans un journal scientifique. Les rapports techniques sont destinés essentiellement à un public international et ils sont distribués à cet échelon. Il n'y a aucune restriction quant au sujet; de fait, la série reflète la vaste gamme des intérêts et des politiques de Pêches et Océans Canada, c'est-à-dire les sciences halieutiques et aquatiques.

Les rapports techniques peuvent être cités comme des publications à part entière. Le titre exact figure au-dessus du résumé de chaque rapport. Les rapports techniques sont résumés dans la base de données *Résumés des sciences aquatiques et halieutiques*.

Les rapports techniques sont produits à l'échelon régional, mais numérotés à l'échelon national. Les demandes de rapports seront satisfaites par l'établissement auteur dont le nom figure sur la couverture et la page du titre.

Les numéros 1 à 456 de cette série ont été publiés à titre de Rapports techniques de l'Office des recherches sur les pêcheries du Canada. Les numéros 457 à 714 sont parus à titre de Rapports techniques de la Direction générale de la recherche et du développement, Service des pêches et de la mer, ministère de l'Environnement. Les numéros 715 à 924 ont été publiés à titre de Rapports techniques du Service des pêches et de la mer, ministère des Pêches et de l'Environnement. Le nom actuel de la série a été établi lors de la parution du numéro 925.

Canadian Technical Report of  
Fisheries and Aquatic Sciences 3405

2020

METHODS AND RESULTS FROM REMOTELY OPERATED VEHICLE (ROV) SURVEY  
PAC2017-030: EXPLORING HIGH AND LOW CURRENT AREAS IN THE SALISH SEA

by

Jessica Nephin<sup>1</sup>, Sharon Jeffery<sup>1</sup>, Mary Thiess<sup>2</sup>, Stephanie Archer<sup>3</sup>, Ian Murdock<sup>4</sup>, Jonathan  
Boschen-Rose<sup>5</sup>, and Sarah Dudas<sup>6</sup>

<sup>1</sup>Institute of Ocean Sciences  
Fisheries and Oceans Canada, 9860 West Saanich Road  
Sidney, British Columbia, V8L 4B2, Canada  
<sup>2</sup> Parks Canada - Ontario & Waterways Region  
2155 Ashburnham Road, Peterborough, Ontario, K9J 6Z6, Canada  
<sup>3</sup> Louisiana Universities Marine Consortium  
8124 Highway 56, Chauvin, Louisiana, 70344, United States of America  
<sup>4</sup> Victoria, British Columbia, Canada  
<sup>5</sup> Southampton, United Kingdom  
<sup>6</sup>Pacific Biological Station  
Fisheries and Oceans Canada, 3190 Hammond Bay Road  
Nanaimo, British Columbia, V9T 6N7, Canada

© Her Majesty the Queen in Right of Canada, 2020  
Cat. No. Fs97-6/3405E-PDF ISBN 978-0-660-35991-5 ISSN 1488-5379

Correct citation for this publication:

Nephin, J., Jeffery, S., Thiess, M., Archer, S., Murdock, I., Boschen-Rose, J., and Dudas, S.  
2020. Methods and results from remotely operated vehicle (ROV) survey PAC2017-030:  
Exploring high and low current areas in the Salish Sea. Can. Tech. Rep. Fish. Aquat. Sci.  
3405: vi + 39 p.



## CONTENTS

<b>ABSTRACT</b>	<b>v</b>
<b>RÉSUMÉ</b>	<b>vi</b>
<b>1 INTRODUCTION</b>	<b>1</b>
<b>2 METHODS</b>	<b>2</b>
2.1 Survey design . . . . .	2
2.1.1 Bathymetry data . . . . .	2
2.1.2 Current speed data . . . . .	4
2.1.3 Selection of sampling areas . . . . .	5
2.2 Data collection . . . . .	6
2.3 Image annotation . . . . .	7
2.3.1 Image capture and georeferencing . . . . .	7
2.3.2 Image annotation . . . . .	8
2.4 Data analysis . . . . .	11
2.4.1 Data preparation . . . . .	11
2.4.2 Diversity and community analyses . . . . .	12
<b>3 RESULTS AND DISCUSSION</b>	<b>16</b>
3.1 Habitat differences . . . . .	16
3.2 Diversity and community differences . . . . .	18
<b>4 ACKNOWLEDGEMENTS</b>	<b>22</b>
<b>REFERENCES</b>	<b>23</b>
<b>Appendices</b>	<b>26</b>
<b>A Image capture tool</b>	<b>26</b>

<b>B R code for processing Biigle data</b>	<b>29</b>
<b>C Species identification guide</b>	<b>33</b>

## ABSTRACT

Nephin, J., Jeffery, S., Thiess, M., Archer, S., Murdock, I., Boschen-Rose, J., and Dudas, S. 2020. Methods and results from remotely operated vehicle (ROV) survey PAC2017-030: Exploring high and low current areas in the Salish Sea. Can. Tech. Rep. Fish. Aquat. Sci. 3405: vi + 39 p.

This report describes the survey design, image annotation methods, data analysis, and results from a remotely operated vehicle (ROV) survey on the CCGS Vector from September 27 to October 2, 2017. The main goal of the survey was to test the feasibility of using geospatial data proxies (such as tidal current) to identify areas with increased diversity and a higher abundance of corals and sponges. Criteria were developed to define high and low current areas using high resolution tidal current model data, multibeam bathymetry and bathymetric derivatives (slope and rugosity). The criteria were then applied using existing data layers in the Salish Sea region to locate areas suitable for testing. During this short survey, eight ROV video transects were completed within three study areas: Howe Sound, Gabriola Island and Sutil Channel. Still images were extracted from the ROV videos and were annotated and reviewed using the online annotation application, Biigle. High current areas were found to have a greater abundance of corals and some sponges, and differed from low current areas in overall community composition. However, there were no detectable difference between low and high current areas in terms of species richness or evenness. Substrate type and bottom depth, which were strongly correlated with tidal current speed in this study, were also identified as potential drivers of the observed differences in community patterns. This report presents some of the first evidence on the relationship between tidal current speed and benthic community structure from deeper habitats (80 to 200 m) in the Salish Sea. Increasing our understanding of environmental proxies and their correlative relationships with benthic communities is an important step towards improving our ability to find and map the distribution of these communities at larger scales, which will enable us to more rapidly locate and prioritize sensitive, biodiverse and unique areas for protection.

## RÉSUMÉ

Nephin, J., Jeffery, S., Thiess, M., Archer, S., Murdock, I., Boschen-Rose, J., and Dudas, S. 2020. Methods and results from remotely operated vehicle (ROV) survey PAC2017-030: Exploring high and low current areas in the Salish Sea. Can. Tech. Rep. Fish. Aquat. Sci. 3405: vi + 39 p.

Dans ce rapport, on décrit la conception, les méthodes d'annotation des images, l'analyse des données et les résultats d'un relevé par véhicule sous-marin téléguidé (ROV) sur le NGCC Vector du 27 septembre au 2 octobre 2017. Le relevé visait principalement à évaluer la possibilité d'utiliser des données géospatiales indirectes (comme le courant de marée) pour répertorier les zones présentant une diversité et une abondance accrues de coraux et d'éponges. Des critères ont été établis pour définir les zones de courant fort et faible au moyen de données de modèles de courants de marée à haute résolution, la bathymétrie multifaisceaux et les dérivés bathymétriques (pente et rugosité). Les critères ont ensuite été appliqués aux couches de données existantes dans la région de la mer des Salish afin de localiser les zones appropriées pour les tests. Au cours de ce bref relevé, le ROV a fait huit transects vidéo dans trois zones d'étude : la baie Howe, l'île Gabriola et le chenal Sutil. Des images fixes ont été extraites des vidéos captées par le ROV puis annotées et examinées à l'aide de l'application d'annotation en ligne, Biigle. Les zones à fort courant se sont révélées plus riches en coraux et en éponges et se distinguent des zones à faible courant par la composition globale de la communauté. Cependant, aucune différence n'a été détectée entre les zones à courant faible et les zones à courant fort sur les plans de la richesse ou de l'équitabilité des espèces. Il a aussi été déterminé que le type de substrat et la profondeur du fond, qui étaient fortement liés à la vitesse des courants de marée dans cette étude, sont des facteurs pouvant expliquer les différences observées dans les modèles de communautés. Dans le rapport, on présente quelques-unes des premières preuves du lien entre la vitesse des courants de marée et la structure des communautés benthiques des habitats profonds (80 à 200 m) de la mer des Salish. L'amélioration de notre compréhension des indicateurs environnementaux et de leurs corrélations avec les communautés benthiques est une étape importante vers l'amélioration de notre capacité à trouver et à cartographier la répartition de ces communautés à plus grande échelle, ce qui nous permettra de localiser plus rapidement et de classer par ordre de priorité les zones sensibles, biodiversifiées et exceptionnelles à protéger.

## 1 INTRODUCTION

Canada has made a commitment to conserve marine ecosystems by signing on to the Convention of Biological Diversity Strategic Plan for Biodiversity 2011-2020 ([www.cbd.int/sp](http://www.cbd.int/sp)). The Government of Canada developed the Marine Conservation Target initiative to meet that commitment and conserve 10% of our coastal and marine areas. Currently, Canada has exceeded that target, with approximately 14% of marine and coastal areas under some form of protection. By 2030, the Government of Canada aims to increase the total protected area to 30%.

When choosing areas for protection, the Government of Canada prioritizes areas of ecological and biological significance (Schram et al. 2019). Areas of high biodiversity or high species abundance are considered ecologically significant; however, locating areas that meet these criteria can be challenging. With one of the longest coastlines in the world, the task of prioritizing marine areas for protection in Canada can be challenging. While much work has been completed to study the BC coast, the majority of it remains unsampled or under-sampled. Environmental proxies, also known as surrogates, can be helpful for identifying areas of ecological and biological significance when biological data are scarce (Huang et al. 2011; Wildsmith et al. 2017; Sarkar et al. 2005; McArthur et al. 2010; Dutertre et al. 2013). However, the effective application of environmental proxies relies on a good understanding of the relationships between ecosystem components and their environment drivers.

Current speed has been shown to have both direct and indirect effects on benthic communities. Current speed has a direct and positive relationship with food availability, with higher currents supplying benthic invertebrates with a higher concentration of plankton, organic material and nutrients (Witman et al. 1993; Lesser et al. 1994; Britton-Simmons et al. 2009) that can lead to faster growth rates (Fabricius et al. 1995; Sebens 1984). The direction and speed of current also have a controlling influence on dispersal, settlement and recruitment of benthic invertebrates (Roughgarden et al. 1988; Palardy and Witman 2011; Crisp 1955). Current regulates the transport of larvae to new areas and affects how likely individuals are to settle and reproduce once there depending on the species' tolerance for water movement. Additionally, current has an indirect effect on benthic communities via the local substrate type, which can play an important role in structuring benthic communities (McArthur et al. 2010). Current speed is inversely related to sedimentation, with higher currents more effectively removing loose sediment from the seabed resulting in harder substrates (Warwick and Uncles 1980; Kostylev et al. 2001; Jumars 1993; Wildish and Kristmanson 1997), which many sessile benthic invertebrates, such as corals and sponges, require for settlement.

Experimental and observational studies from shallow nearshore habitats have found evidence of increased diversity of benthic invertebrate in areas with higher current (Baynes and Szmant 1989; Palardy and Witman 2011), and differences in benthic assemblages between high and low current areas (Warwick and Uncles 1980). Locally, a study from the Salish Sea region found greater diversity in benthic communities within higher current tidal passages when compared with lower current inlets (Elahi et al. 2014). In-situ observations of high current areas have recently been used to support the identification of Ecologically and Biologically Significant Areas (EBSAs) in nearshore areas of the Northern Shelf Bioregion (Rubidge et al. 2020). For these reasons, high tidal current speed was identified as a potential indicator of benthic species

diversity and coral and sponge aggregations.

ROV research survey PAC2017-030 took place on the CCGS Vector from September 27 to October 2, 2017 in the Salish Sea. The main goal of the survey was to evaluate the use of tidal current speed as a proxy for identifying benthic areas with higher biodiversity, abundance or uniqueness within the Salish Sea. The proxy was assessed by comparing benthic diversity and the abundance of corals and sponges between areas of high and low tidal current at depths of 80 to 200 m. It was expected that the patterns in diversity and community structure observed between high and low current at shallower depths would hold true in deeper habitats. Specifically, we predicted that high current areas would have greater benthic diversity, and coral and sponge abundance, when compared to low current areas. Secondary goals of this survey were to map a glass sponge reef in Lions Bay, and retrieve and deploy a hydrophone for passive acoustic monitoring of sponge reefs. However, results from the sponge reef sampling will not be covered in this report.

## **2 METHODS**

### **2.1 Survey design**

To meet the survey objectives, two contrasting sampling area types were defined: high current and low current, both with hard substrate. Hard bottom substrates were targeted based on the assumption that hard substrates (e.g., bedrock, boulder and cobble) were more suitable coral and sponge habitat than soft substrates. To identify high and low current sampling areas within the study areas (Sutil Channel, Gabriola Island and Howe Sound, see Figure 1), two data sources were used: bathymetry and bottom tidal current speed. Bathymetry was used to identify depths suitable for ROV operation and to identify possible areas of hard substrates using bathymetric derivatives (slope and rugosity) and bathymetric features (ridges and mounds) as proxies for hard substrate. The proxy was based on our assumption that ridges and mound features, as well as areas of relatively steeper slopes and higher rugosity, were more likely to be areas with hard substrate. Bottom tidal current speed was used to identify areas of relatively higher tidal current within the study area.

#### **2.1.1 Bathymetry data**

Multibeam bathymetry, sourced from Canadian Hydrographic Service (CHS) and Natural Resources Canada (NRCan) surveys, had a horizontal resolution of 5 metres. Bathymetry derivatives, slope, rugosity and broad and fine Benthic Position Index (BPI) were calculated from the bathymetry raster using python code provided in the ArcGIS Benthic Terrain Modeler (BTM 3.0) toolbox (Walbridge et al. 2018). Rugosity was calculated with the Vector Ruggedness Measure method (Sappington et al. 2007). Bathymetric features (Figure 2) were classified from slope, fine BPI and broad BPI layers using the classification dictionary shown in Table 1 which was roughly based on work from Lundblad et al. (2006).

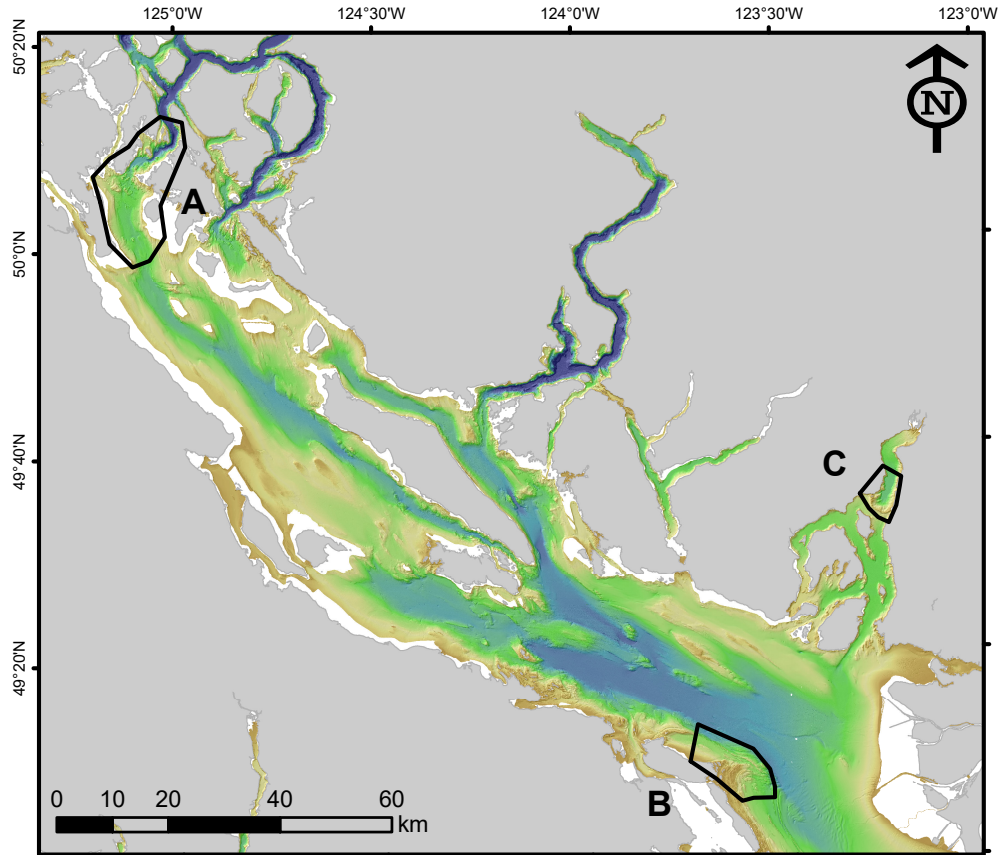


Figure 1. Study areas shown with the 5 m resolution multibeam bathymetry raster in the background. A = Sutil Channel, B = Gabriola Island and C = Howe Sound.

Table 1. A classification dictionary for categorizing bathymetric features from slope, fine BPI and broad BPI. Values represent the upper and lower limits for each category. Local crest, local crest in depression, and narrow crest categories were used to locate ridge and mount features.

<b>Bathymetric feature</b>	<b>Broad BPI lower</b>	<b>Broad BPI upper</b>	<b>Fine BPI lower</b>	<b>Fine BPI upper</b>	<b>Slope lower</b>	<b>Slope upper</b>
Narrow depression		-100		-100		
Local crest in depression		-100	100			
Broad depression		-100	-100	100		
Local depression	-100	100		-100		
Local crest	-100	100	100			
Flat	-100		-100	100		5
Open Slopes	-100		-100	100	5	80
Narrow crest		100	100			
Depressions on crests		100		-100		

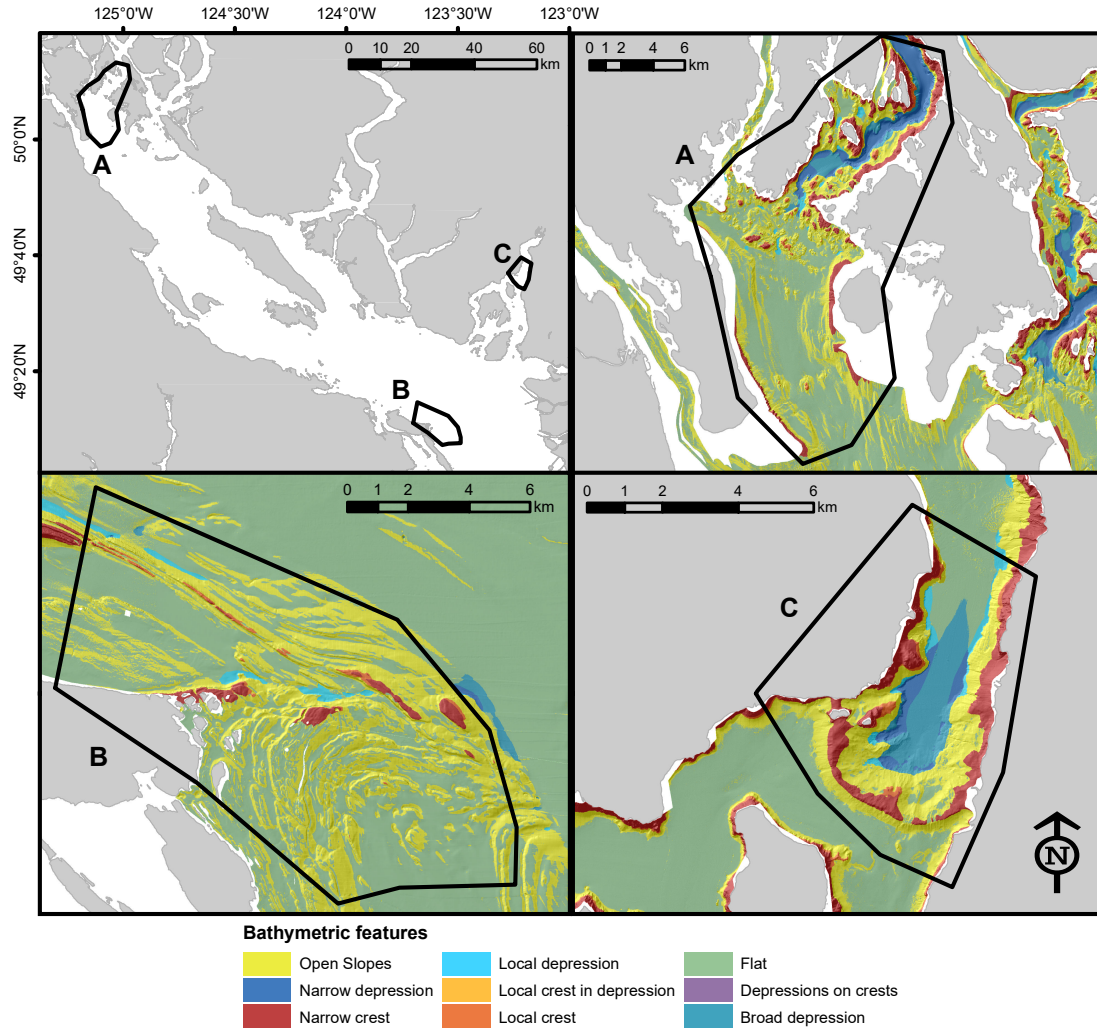


Figure 2. Classified bathymetry features (using Table 1) for each study area derived from 5 m resolution multibeam bathymetry. Local crest, local crest in depression, and narrow crest categories were used to locate ridge and mount features. A = Sutil Channel, B = Gabriola Island and C = Howe Sound.

### 2.1.2 Current speed data

Bottom tidal current speed was sourced from the SalishSeaCast model (Soontiens et al. 2016; Soontiens and Allen 2017). The region covered by the model includes the Strait of Georgia, Strait of Juan de Fuca, Johnstone Strait, Fraser River, and other connecting waterways. The model has a horizontal resolution of 440 by 500 metres and a non-uniform vertical resolution with 40 depth levels that are clustered near the surface (from 1 m width near surface to 27 m width near the seafloor). For more details on the SalishSeaCast model see [salishsea.eos.ubc.ca/nemo](http://salishsea.eos.ubc.ca/nemo) and [salishsea.eos.ubc.ca/erddap/index.html](http://salishsea.eos.ubc.ca/erddap/index.html) for data access.

Hourly zonal (u) and meridional (v) velocities from the year 2017 were downloaded from the



model version v17-02. Bottom level values were extracted using the provided mask layers for each variable. Velocities were horizontally shifted to the central nodes via linear interpolation and tidal current speed was calculated monthly, from the shifted u and v values using the root mean square:  $\sqrt{\text{mean}(u^2 + v^2)}$ . Current speed was then interpolated using spline with barriers (ESRI ArcMap 10.4) with a smoothed coastal high water line as a barrier and then smoothed using a circular 13-cell focal mean. See [gitlab.com/dfo-msea/environmental-layers/salishsea-nemo](https://gitlab.com/dfo-msea/environmental-layers/salishsea-nemo) for code that downloads the SalishSeaCast data and calculates and interpolates bottom tidal current speed.

### 2.1.3 Selection of sampling areas

Sampling was restricted to depths between 30 and 250 metres for safe ROV operation, and to areas that were predicted to have hard substrate. Areas of hard substrate were located through the use of bathymetric proxy variables, because a comprehensive substrate model of the area was not available at the time. Hard substrate was predicted to be located in areas where slope was greater than 5 degrees or ridge or mound features were present and where rugosity was relatively high (i.e., greater than 0.005).

Within the sampling areas constrained by depth and substrate, two sampling area types, high and low current, were differentiated based on the bottom tidal current speed layer. High current areas (Figure 3) were defined as areas with tidal current speeds greater than the 60<sup>th</sup> percentile of values in the Salish Sea model and low current areas as areas with tidal current speeds less than the 60<sup>th</sup> percentile of values.

Random start and end points for ROV transects were generated within the high current and low current sampling areas. For all three study areas (Howe Sound, Gabriola Island, and Sutil Channel), six random sampling points were generated in both high and current areas, avoiding known sponge reefs. Any random points within 200 m of land were then moved manually, until they were greater than 200 m from any coastline. This was requested by ROV operators as a best practice to ensure the vessel towing the ROV has adequate space to manoeuvre.

To create transect lines from each randomly selected sampling point, the random points were buffered by 500 m. Buffers were then converted into polyline features and points were generated every 100 m along the polyline buffers. From those generated points along the buffer, a single point was manually selected. The pair of points (the initial random point and the selected point along the buffer) would be used to delineate the transect line. The point along the buffer was selected based on three criteria: it was either downhill or uphill (perpendicular to depth contour lines) from the initial random point, at least 200 m away from land and within the sampling area type (high or low current), if possible. From the pair of points, the deeper point was labelled 'start' and the shallower point was labelled 'end' to ensure the ROV would move upslope when sampling. Each start and end point combination was assigned a transect number. Eight transects were completed during the survey: four in high current areas and four in low current areas (Figure 3).

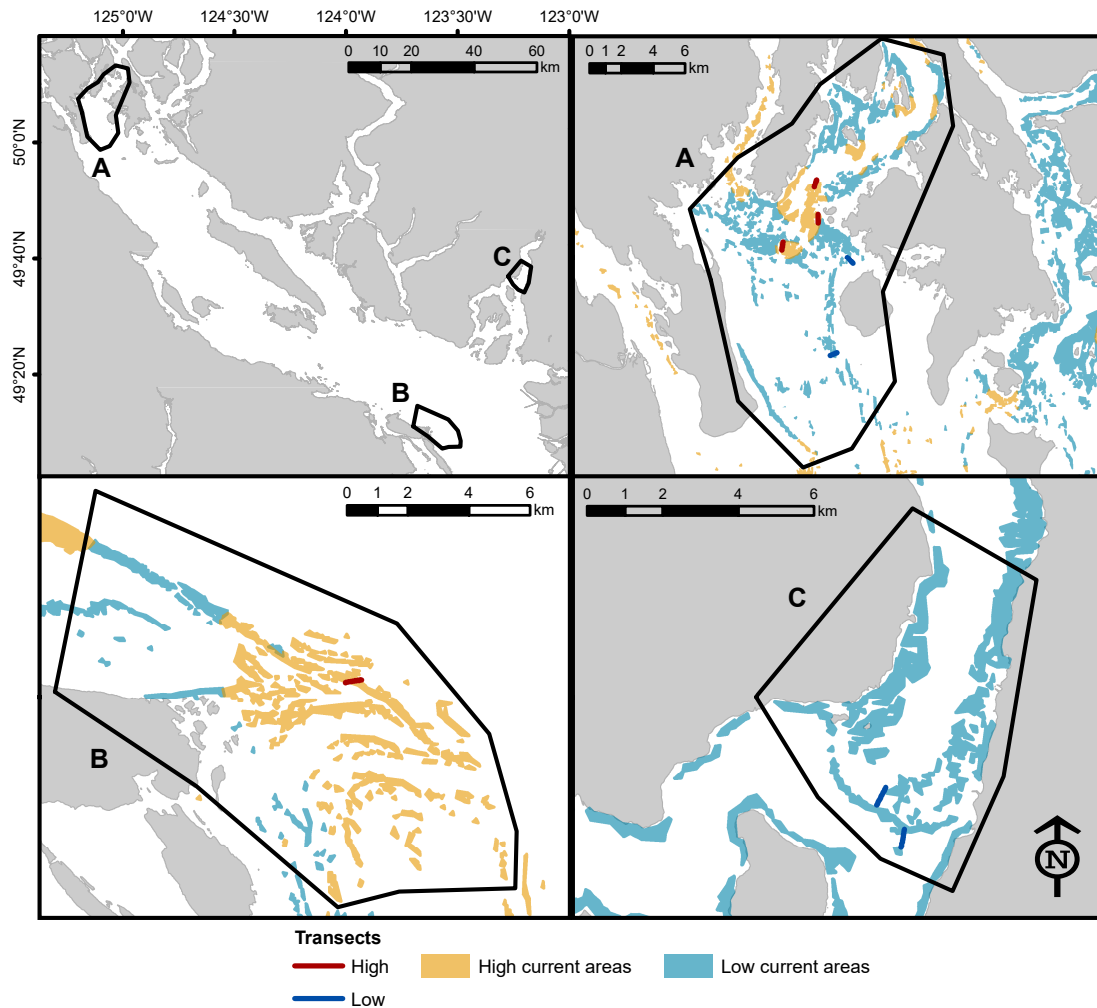


Figure 3. High and low current areas and the eight completed transects. A = Sutil Channel, B = Gabriola Island and C = Howe Sound.

## 2.2 Data collection

ROV operations were conducted from the Coast Guard vessel CCGS Vector using a Deep Ocean Engineering Phantom HD2+2 ROV (Figure 4), an upgraded HD2. The ROV was deployed from the Vector's starboard side A-frame. The ROV was towed using a 3/16 inch winch wire affixed to a 360 kg clump weight. The winch wire was taped to the ROV's 300 m fiber-optic tether upon deployment to eliminate the drag of the tether during ROV operations. On transect, the ROV was driven at speeds up to one knot, but more typically 0.5 knots or less. For high current transects the ROV was deployed only during or near slack current time windows based on the CHS predictions to ensure safe ROV operating conditions.

During transect operations, videos were recorded with an Insite Pacific MiniZeus high-definition (1080i) video camera. The camera tilt was controlled by ROV pilots, and was set to an oblique

angle relative to the seafloor (slightly forward and down, toward the seafloor). The ROV was driven in close proximity to the seafloor, with a target altitude of 0.5 to 1 m above the seafloor when on transect. The MiniZeus camera was typically operated with the zoom as far out as possible and adjusted as needed during operation. Forward illumination for the MiniZeus HD camera was provided by two DeepSea Power and Light SLS-5200 LED lights and two Remote Ocean Systems Q-LED II LED floodlights, both of which supplied a warm white illumination. The ROV was equipped with a set of parallel green scaling lasers. They were affixed to the bottom of the MiniZeus camera, spaced 10 cm apart and orientated to point in the middle of the camera's field of view.

The latitudinal and longitudinal position of the ROV was calculated using an EdgeTech Broadband Acoustic Tracking System, from an Applied Acoustic 300 series transponder. The ROV latitude and longitude values were smoothed and gaps were filled via linear interpolation where necessary using Hypack software. ROV tracking data exported from Hypack contained latitude, longitude, date, time and depth of the ROV at a 1 Hz resolution.

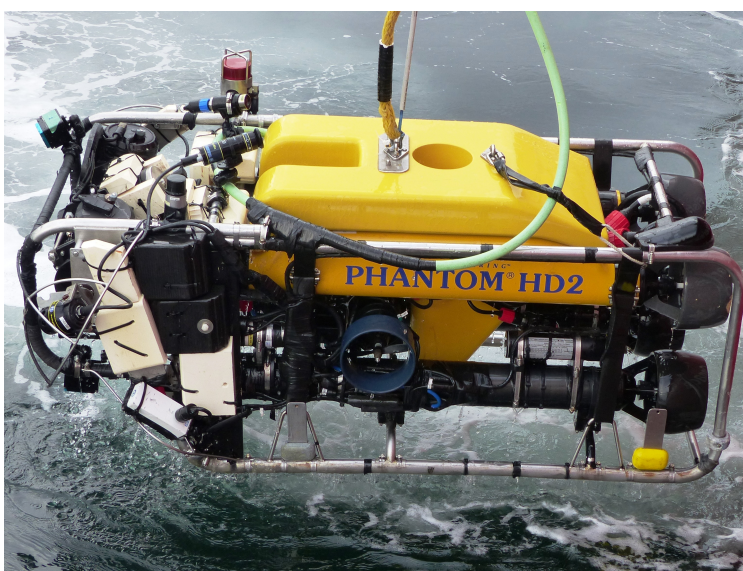


Figure 4. Phantom ROV after completing a transect, in operation off the CCGS Vector.

## 2.3 Image annotation

### 2.3.1 Image capture and georeferencing

Still images were captured from the video files every three seconds using an image capture tool (see Appendix A). The tool saved the images using a standardized naming convention (TransectName\_Date\_Time.jpeg). The images were then imported into Biigle (see Section 2.3.2). ROV tracking information was used to assign a spatial position to each image by matching the date and time of the image with the tracking data. Once matched, a comma separated value file containing the latitude and longitude of the ROV and the associated image filename was created and uploaded to Biigle (see below).

An image capture rate of three seconds was chosen as it allowed for some overlap between images. Overlap between images was required to minimize the risk that a species captured by the video would be missed in the images. A number of image capture rates were tested (from 2 to 10 seconds) and three seconds provided the optimal amount of overlap. However, the amount of overlap between images did vary with ROV speed and height of the ROV off the bottom. Although variations in ROV speed and height off the bottom were not available from the ROV sensor data for this survey, they were estimated by watching the videos. Based on that qualitative assessment, it was determined that the speed and height variations were unlikely to cause gaps between the images when using a three second image capture rate and thus it was unlikely that species captured by the video would be missed in the images.

### 2.3.2 Image annotation

The annotation of images captured from the ROV videos was completed in Biigle (<https://biigle.de/>), a web-based tool for collaborative annotation of marine images and videos (Langenkämper et al. 2017). Videos were viewed prior to annotation in Biigle and a list of all species visible in the videos was created. From that list, a species guide was developed for annotators to use as a reference (see Appendix C).

A total of 5,303 images were captured from the eight ROV transects. The images were divided into sets, with several sets of images within each transect. Two annotators then completed annotation tasks one set at a time. To allay annotator bias, image sets were randomized across transects and high low groups and annotation tasks were scheduled such that a single annotator did not annotate an entire transect on their own. The annotation task schedule also served to ensure that learning bias, if present, would not be confounded with either high or low current transect type.

Annotators were asked to apply appropriate image and annotation labels to each image. Image labels apply to the entire image (e.g., image quality) and annotation labels apply to a specific location within an image (e.g., to identify fish to species). In Biigle, label trees are used to store image and annotation labels. The image label tree used was based on previous ROV survey annotation projects and included labels that categorised the status of the ROV (on or off-transect), image quality, dominant and sub-dominant substrate, and documented the presence of dead reef-building sponges (Table 2). The annotation label tree used was developed from the list of species observed in the ROV video (Table 3). Additional annotation labels were used to label laser point marks, visible anthropogenic objects, and species that were unknown or notable. A notable label was used, in conjunction with a species label, when an annotator was uncertain in their identification.

The first step in image annotation was to categorise the status of the ROV. When the status of the ROV matched a label within the 'image not annotated' category (Table 2), for example ROV 'off bottom', no additional image or species labels were applied. Fourteen percent of all images (718 out of 5,303) were categorised as 'image not annotated'. For all other images, image quality, dominant and sub-dominant substrate and the presence of dead sponge were noted. Sub-dominant substrate labels were optional as they were not relevant or required when only a single substrate type was present. Items within the image were then annotated, beginning with laser point marks. Laser point marks were labelled to estimate the area of the

image using the known distance between the lasers (i.e., 10 cm). Finally, individual animals within an image were annotated by attaching the appropriate label to a point location, with the exception of reef-building sponges (Figure 5). Reef-building hexactinellida, *Aphrocallistes vastus*, and *Heterochone calyx* sponges were annotated using polygons to capture the area of the colony within the image. Additionally, individual zoanthids could not be counted because they were too numerous and difficult to distinguish from one another. When zoanthids were present in an image, they were labelled by placing one point annotation to mark their presence.

Table 2. Image labels used for annotating image-level qualities. Substrate labels were used to define both dominant and sub-dominant substrate. The Biigle label tree can be accessed here: <https://biigle.de/label-trees/195>.

Category	Label
Image not annotated	Close up (zoom) Investigation (collection/deployment) Off bottom Other technical issue Duplicate image (no movement)
Image quality	Excellent quality Average/good quality Poor quality
Substrate	Bedrock creviced Bedrock smooth Boulders Cobble Gravel Sand Mud
Reef-building sponges	Dead sponge(s) present

Once images were annotated, each image was reviewed by the other annotator. This quality control step was performed to limit differences between annotators and catch errors in identification. If an annotator disagreed with a label, they added the label they believed to be correct, as well as a ‘correction’ label to indicate that it had been corrected. After the initial review was completed by the annotators, a subsequent review by project leads and species experts occurred. During this final review phase, species previously labelled as ‘unknown’ were identified (where possible) and all annotations with ‘correction’ labels were re-examined and assigned the most appropriate label based on expert opinion. In addition, labels for species that were known to be difficult to identify (e.g., *Sebastes* spp.) were reviewed for consistency using the Biigle largo re-evaluation tool (see Biigle manual for details) and misidentification errors were corrected.

Table 3. Species labels used for annotation and their hierarchical structure. All labels in the hierarchy were available for annotation. Example images for all species can be found in Appendix C. The Biigle label tree can be accessed here: <https://biigle.de/label-trees/190>.

<p><b>Arthropoda</b></p> <ul style="list-style-type: none"> <li>↳ Decapoda           <ul style="list-style-type: none"> <li>↳ <i>Cancer</i> sp.</li> <li><i>Chionoecetes</i> sp.</li> <li>Decorator crab</li> <li>Hermit crab</li> <li>Lithode crab 1- like <i>Acantholithodes</i></li> <li>Lithode crab 2 - large box crab</li> <li><i>Lopholithodes foraminatus</i></li> <li><i>Munida quadrispina</i></li> <li><i>Pandalus platyceros</i></li> <li>Shrimp - unknown sp.</li> </ul> </li> </ul>	<p><b>Mollusca</b> . . . . . continued</p> <p>Gastropoda</p> <ul style="list-style-type: none"> <li>↳ <i>Calliostoma</i> sp.</li> <li><i>Fusitriton oregonensis</i></li> <li><i>Peltodoris lentiginosa</i></li> <li>Snail - unknown sp.</li> </ul>
<p><b>Asciacea</b></p> <ul style="list-style-type: none"> <li><i>Corella</i> sp.</li> </ul>	<p><b>Pisces</b></p> <ul style="list-style-type: none"> <li>↳ Agonidae</li> <li>Cottidae           <ul style="list-style-type: none"> <li>↳ <i>Icelinus</i> sp.</li> </ul> </li> <li><i>Cymatogaster aggregata</i></li> <li>Gadidae           <ul style="list-style-type: none"> <li>↳ <i>Gadus macrocephalus</i></li> </ul> </li> <li>Hexagrammidae           <ul style="list-style-type: none"> <li>↳ <i>Ophiodon elongatus</i></li> </ul> </li> <li><i>Hydrolagus colliei</i></li> <li><i>Lycodes pacificus</i></li> <li>Pleuronectidae           <ul style="list-style-type: none"> <li>↳ <i>Lyopsetta exilis</i></li> </ul> </li> <li><i>Porichthys notatus</i></li> <li><i>Raja</i> sp.           <ul style="list-style-type: none"> <li>↳ <i>Raja rhina</i></li> </ul> </li> <li><i>Sebastes</i> sp.           <ul style="list-style-type: none"> <li>↳ <i>Sebastes diploproa</i></li> <li><i>Sebastes elongatus</i></li> <li><i>Sebastes maliger</i></li> <li><i>Sebastes proriger</i></li> <li><i>Sebastes ruberrimus</i></li> </ul> </li> <li><i>Squalus suckleyi</i></li> </ul>
<p><b>Brachiopoda</b></p>	
<p><b>Cnidaria</b></p> <ul style="list-style-type: none"> <li>↳ Alcyonacea           <ul style="list-style-type: none"> <li>↳ <i>Anothothela pacifica</i></li> <li><i>Paragorgia pacifica</i></li> <li><i>Swiftia</i> sp.</li> </ul> </li> <li>Anthozoa           <ul style="list-style-type: none"> <li>↳ Anemone - unknown sp.</li> <li><i>Cribrinopsis fernaldi</i></li> <li><i>Metridium</i> sp.</li> <li><i>Pachycerianthus fimbriatus</i></li> </ul> </li> <li>Pennatulacea           <ul style="list-style-type: none"> <li>↳ <i>Halipteris</i> sp.</li> </ul> </li> </ul>	
<p><b>Echinodermata</b></p> <ul style="list-style-type: none"> <li>↳ Asteroidea           <ul style="list-style-type: none"> <li>↳ Asteroid 1 - long armed</li> <li>Asteroid 2</li> <li><i>Ceramaster</i> sp.</li> <li><i>Henricia</i> sp.</li> <li><i>Hippasteria phrygiana</i></li> <li><i>Mediaster aequalis</i></li> <li><i>Pteraster</i> sp.</li> </ul> </li> <li>Echinoidea           <ul style="list-style-type: none"> <li>↳ <i>Stronglyocentrotus franciscanus</i></li> </ul> </li> <li>Holothuroidea           <ul style="list-style-type: none"> <li>↳ <i>Psolus chitonoides</i></li> </ul> </li> <li>Ophiuroidea</li> </ul>	<p><b>Porifera</b></p> <ul style="list-style-type: none"> <li>↳ Demospongiae           <ul style="list-style-type: none"> <li>↳ Demosponge 1 - white columnar</li> <li>Demosponge 2 - dirty columnar</li> <li>Demosponge 3 - squat, round</li> <li>Demosponge 4 - <i>lophon</i> like</li> <li>Encrusting sponge - unknown</li> <li><i>Polymastia</i> sp.</li> <li>Yellow sponge</li> </ul> </li> <li>Hexactinellida           <ul style="list-style-type: none"> <li>↳ <i>Aphrocallistes vastus</i></li> <li><i>Heterochone calyx</i></li> <li><i>Rhabdocalyptus dawsoni</i></li> </ul> </li> </ul>
<p><b>Mollusca</b></p> <ul style="list-style-type: none"> <li>↳ Cephalopoda           <ul style="list-style-type: none"> <li>↳ <i>Rossia pacifica</i></li> </ul> </li> </ul>	<p><b>Worms</b></p> <ul style="list-style-type: none"> <li>↳ Feather duster</li> <li>Serpulid</li> <li>Spaghetti worm</li> </ul>
	<p><b>Zoantharia</b></p>



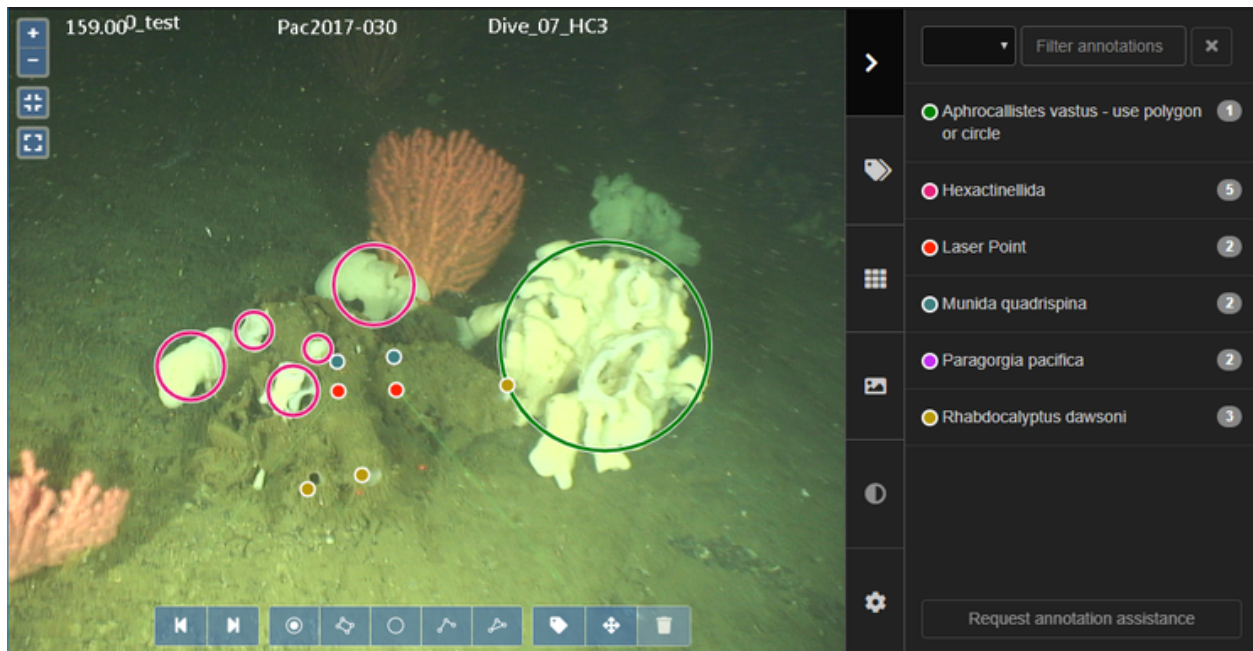


Figure 5. Annotation view in the Biigle web application showing point annotations and circle (areal) annotations for reef-building hexactinellid sponges.

## 2.4 Data analysis

### 2.4.1 Data preparation

After all images were annotated and reviewed, the image and annotation label data were exported from Biigle by downloading the comma separate value format of the image and annotation reports and the area report. These three reports were combined into one dataset that included image metadata (i.e., image filename, name of the primary annotator, image area in square metres, longitude and latitude), image labels, annotation labels, and annotation areas from reef-building sponge polygon annotations. Image area was included in the annotation report and was calculated automatically in Biigle for every image with laser point annotations using the known distance between lasers points. Data quality checks were performed by looking for missing longitude and latitude values, missing image labels (e.g., missing image quality or substrate type labels), missing laser point annotations and images with no annotations. All omissions were cleaned up by adding any missing labels in Biigle and then the reports were exported from Biigle again. See Appendix B for R code that was used to create the dataset from Biigle reports and perform data quality checks.

Species counts and areas (for reef-building sponges) were aggregated to the transect level by summing the counts and areas. Substrate observations were aggregated to the transect level by calculating the percent of images assigned to each substrate category. Dominant and sub-dominant substrate observations were grouped into 4 aggregate substrate categories for analysis: hard (bedrock to gravel), mixed (hard dominant), mixed (soft dominant) and soft (mud).

There were no observations of sand.

Before the transect-level data were used for analysis, sampling effort was calculated to effort-correct species counts and areas. Effort-correction was necessary because there were differences in sampling effort between transects. Sampling difference stemmed from variation in the surveyed transect lengths and ROV altitude (height of the ROV off-bottom) while surveying. Since there was overlap between images from using the three-second image capture rate, sampling effort could not simply be calculated by summing the field of view area of all images. A sampling effort index was instead estimated by multiplying the mean image area for each transect by the length of each transect, producing a relative measure of sampling effort among transects.

Transect length was calculated from the longitude and latitude coordinates of the images in R using the `gLength` function in the `rgeos` package (Bivand and Rundel 2019). The sampling effort index was then normalized by dividing sampling effort values by their maximum value so that sampling effort of the transect with the greatest effort was equal to one and all others were between zero and one. To apply the effort correction, species counts and areas were divided by the normalized sampling effort index. Sampling effort was relatively larger for high current transects compared to low current transects (Figure 6).

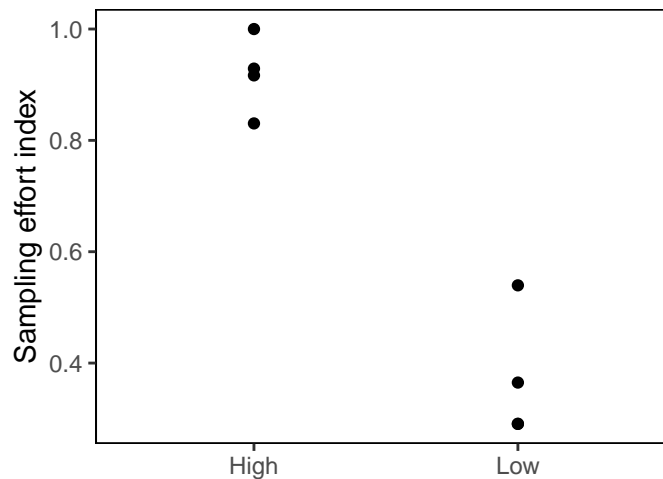


Figure 6. Sampling effort index for each transect in high and low current areas.

#### 2.4.2 Diversity and community analyses

Community analyses, performed using cluster and ordination methods, aim to group samples (i.e., transects) based on their shared species composition. Following the assumption that similar habitats will support similar groups of species, the identified groups of samples can subsequently be used to explore the underlying relationship between species composition and environmental gradients.

Prior to conducting diversity and community analyses, further refinement of the annotated



species-transect dataset was required (Table 3). Some annotated species required grouping or removal to ensure that the same species were not represented by different labels, at various taxonomic levels, across the consolidated dataset. After processing, the dataset used for analysis included 40 taxa (19 species and 21 species groups or higher-level taxa) of the original 77 labels. The processing steps were as follows:

1. When labels could not be consistently differentiated from each other, or when there were very few observations for a given label, taxonomically similar labels were grouped together. See Table 4 for a list of all labels that were grouped, the specific justification for grouping, and their new label names.
2. When labels could not be consistently differentiated from each other, but grouping was not appropriate because they were taxonomically or ecologically distinct, labels were removed from the analysis. This rule was applied in two cases. First, 'Shrimp-unknown' and *Munida quadrispina* labels were removed because image quality was too low overall to differentiate between these small organisms (e.g., shrimp was likely to be mislabelled as *Munida* and vice versa). The removal of 'Shrimp-unknown' and *Munida quadrispina* labels accounted for a large amount of data loss (38% of all individuals counted in the images). Second, taxa labelled as Pennatulacea were excluded because they had not been identified consistently between annotators. This resulted in the removal of very few labels (i.e., less than 5 Pennatulacea annotations were removed from the dataset, representing 0.03% of all individuals counted in the images).
3. When taxa were typically identified with a higher-level taxonomic label and were unlikely to represent the same taxa labelled using lower-level taxonomic labels (e.g., *Sebastes maliger* were unlikely to be labelled as Pisces), higher-level taxonomic groups were retained (e.g., Pisces).
4. When taxa were typically identified at a species level (e.g., *Sebastes maliger*), but were sometimes identified with a higher-level taxonomic label (e.g., *Sebastes*), the higher-level taxonomic labels were reviewed. If the image quality was not sufficient to identify the individual at the lower taxonomic level, the higher-level annotation label was removed. Decapoda, Alcyonacea, *Sebastes* and Anemone labels were removed from the analysis so the lower-level taxonomic (e.g., species) annotations within their respective taxonomic groups could be retained. The removal of those labels accounted for a very small amount of data loss (0.5% of all individuals counted in the images).

Univariate diversity measures were calculated from the cleaned species-transect dataset. Richness was measured by summing the number of species and higher-level taxa present in each transect. Evenness (Pielou's J) was calculated using the equation  $J = H / \log n$ , where  $H$  is the Shannon Index and  $n$  is species richness. The Shannon Index was calculated using the equation  $H = -\sum_{i=1}^n p_i \ln p_i$ , where  $p_i$  is the proportional abundance of species  $i$ . The rarefy function from the R package *vegan* (Oksanen et al. 2019) was used to compute rarefaction curves and the rarefied number of species for each transect. The minimum number of individuals observed on any transect (419 individuals from low current transect 7) was used to sub-sample the number of individuals in other transects for rarefaction.

Community analyses were performed in R with the *vegan* package (Oksanen et al. 2019). To represent dissimilarities between sites in 2-dimensional space, hierarchical cluster analysis

and non-metric multidimensional scaling (NMDS) were used. Dissimilarities were calculated based on effort-corrected species counts. For species where count data were not available (reef-building sponges and zoanthids), presence in the image was used as an alternative to counts. Presence values were summed by transect and effort-corrected following the method used for count data. Dissimilarities between transects (sites) were measured with Bray-Curtis distance for clustering and ordination analyses. A hierarchical clustering analysis was completed using the `hclust` function with the 'complete linkages' measure for determining clusters. NMDS analysis was completed with the `metaMDS` function. The `metaMDS` algorithm returned a convergent solution after 20 attempts. To visualise the difference between high and low current transects in ordination space, group ellipsoids, representing the minimum area that contains all group points, were computed using the `ordihull` function. To measure the difference between high and low current transects based on community composition, an analysis of similarities was completed with the `anosim` function using 999 permutations to test for significance.

To investigate the environmental drivers of site dissimilarities, the `envfit` function in the `vegan` R package was used. The `envfit` method tests for significant linear relationships between environmental variables and the ordination axes derived from the NMDS analysis. Significance was measured via permutation with the default 999 permutations completed. Environmental variables included in the analysis were the bathymetric derivatives and tidal current speed described in Section 2.1, as well as mean bottom depth and substrate type sourced from the ROV data. Substrate was represented as the percent of images per transect with hard dominant substrate (which ranged from 0 to 58%).

Table 4. Species groups used for community analysis including the justification for each grouping.

Group	Label used for analysis	Original annotation label	Justification
Arthropoda	Crab	<i>Cancer</i> sp. <i>Chionoecetes</i> sp. Decorator crab Hermit crab Lithode crab 1 ↳ ( <i>Acantholithodes</i> like) Lithode crab 2 ↳ (large box crab) <i>Lopholithodes foraminatus</i>	Image quality was too low to differentiate crabs that were typically small
Echinodermata	Asteroid spA	Asteroid 1 ↳ (long armed) <i>Henricia</i> sp.	Very similar to <i>Henricia</i> sp., likely the same species or group of species
	Asteroidea	Asteroid 2 <i>Ceramaster</i> sp. <i>Hippasteria phrygiana</i> <i>Mediaster aequalis</i> <i>Pteraster</i> sp.	Could not consistently differentiate the short-armed seastars
Mollusca	Cephalopoda	<i>Rossia pacifica</i>	Very few observations
	Gastropoda	<i>Calliostoma</i> sp. <i>Fusitriton oregonensis</i> Snail - unknown sp.	Image quality was too low to differentiate between gastropod species consistently
Pisces	Agonidae-Cottidae	Agonidae Cottidae <i>Icelinus</i> sp.	Could not consistently differentiate fish in these two families
	Pisces	<i>Porichthys notatus</i>	Not very visible in images and had very few observations
	Gadidae	<i>Gadus macrocephalus</i>	Very few observations
	Pleuronectidae	<i>Lyopsetta exilis</i>	Could not differentiate species and family level identifications
	<i>Raja</i> sp.	<i>Raja rhina</i>	Very few observations
Porifera	Demosponge spA	Demosponge 2 ↳ (dirty columnar)	Distinctive demosponge that was consistently identified
	Porifera	Demosponge 1 ↳ (white columnar) Demosponge 3 ↳ (squat and round) Demospongiae <i>Polymastia</i> sp.	Could not consistently differentiate most demosponges with other porifera
	Reef sponge	<i>Aphrocallistes vastus</i> <i>Heterochone calyx</i> Reef-building hexactinellida	Image quality was too low to consistently differentiate between types of reef-building sponge
Worms	Worms	Feather duster Spaghetti worm	Worms were not able to be differentiated (except for serpulids which remained a separate group)

### 3 RESULTS AND DISCUSSION

#### 3.1 Habitat differences

In addition to current speed there were several habitat differences between transects in high and low current areas. One unintended habitat difference, that was not suitably controlled for in the survey design, was depth (Figure 7). The mean depth from transects in high current areas was 166 m, while the mean depth from transects in low current areas was 108 m, a difference of 58 m. As a result, depth may have had a differential influence on the species community occurring at high and low current areas, as depth has been shown to be an important driver in the distribution of benthic communities on the BC coast (Rubidge et al. 2016). For example, Agonidae and Cottidae species tended to occur at shallower depths within the transects (mean depth of observations 92 m) and *Sebastes ruberrimus* tended to occur at deeper depths within the transects (mean depth of observations 148 m). Unfortunately, the influence of depth cannot be parsed out in the analysis as it is confounded with high and low current transect type.

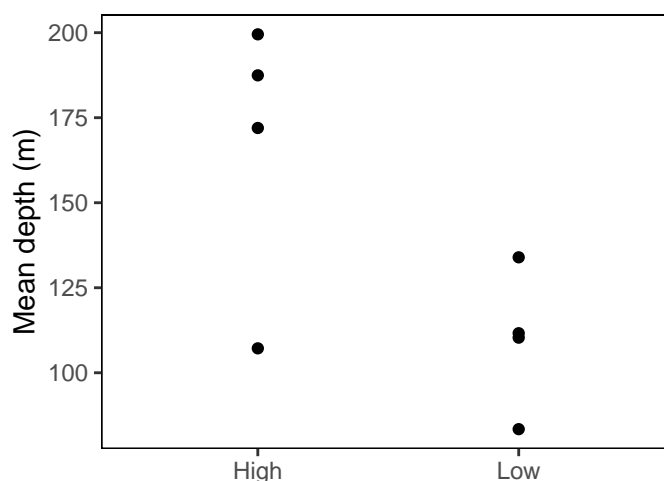


Figure 7. Mean ROV depth from transects in high and low current areas.

Substrate type also varied considerably between transects in high and low current areas (Figure 8). Transects in high current areas had a larger percentage of hard substrate with an average of 28% of images with rocky substrate from bedrock to gravel, compared with transects in low current areas that had only 1% of images on average with rocky substrates. Conversely, transects in low current areas had a much higher percentage of muddy substrate (86% of images from low current transects compared to 48% of images for high current transects). Substrate differences were not explicitly controlled for in the survey design and it is not unexpected that low current areas would have a greater accumulation of softer, finer substrates. However, given that both low and high current areas were defined as areas with relatively higher slope and rugosity and areas with ridges and mound features, the expectation was that a greater percentage of images with hard substrate would have been found, even in low current areas. Future studies will benefit from using a modelled substrate layer for the area rather than bathymetry based proxies

that did not effectively locate hard substrate.

In addition, the low percentage of hard substrate observed in low current areas may have been related to inaccurate placement of the transects within the low current sampling area (Figure 3 blue areas). All low current transects began within areas delineated as low current sampling areas, but did not consistently remain within the area throughout the transect due to requirements of safe ROV operation (i.e., maintaining 200 m distance from shore and transiting up-slope). In comparison, high current transects were more often fully contained within the high current sampling area (Figure 3 orange areas). In future surveys, it is recommended to develop a sampling design that ensures all transects are fully contained within the delineated sampling areas.

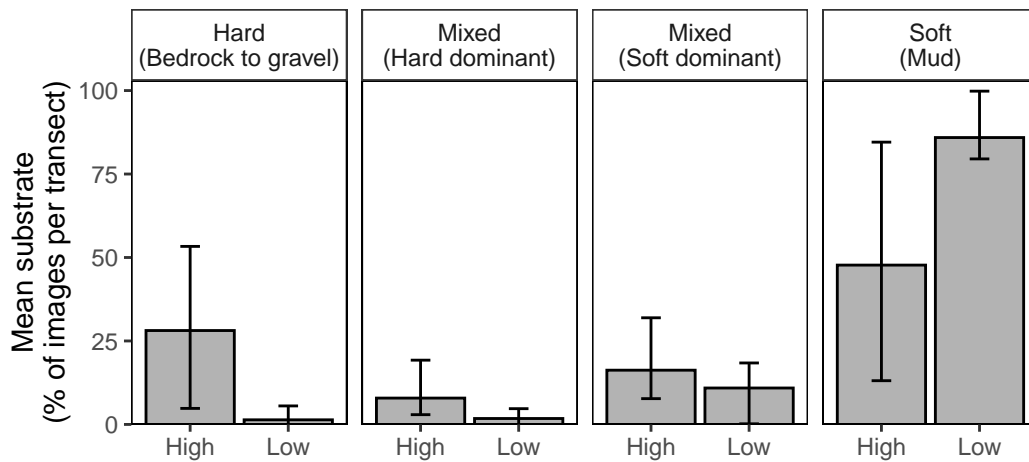


Figure 8. Mean percent of images with each substrate category from transects in high and low current areas. Error bars represent range between the minimum and maximum.

One expected habitat difference between high and low current areas was the areal coverage of reef-building hexactinellida sponges (Figure 9). Reef-building sponge area was observed to be greater in high current areas ( $p$ -value=0.066,  $df=3$ , from Welch t-test), with an average of 49.6  $m^2$  in high current transects and 0.5  $m^2$  in low current transects based on effort-corrected values. Sponges are known to occur in greater density in areas with moderate to high current speeds as a result of the increased food supply available to them (Ramiro-Sánchez et al. 2019; Rice et al. 1990; Rooper et al. 2014). However, the greater percentage of hard substrate in high current transects (Figure 8) could also have played a role in the increased areal coverage of reef-building sponges, confounding the relationship with current, as sponges are also found more often on hard substrates due to the increased number of possible attachment sites (Leys et al. 2004). An attempt to control for the effect of substrate type on the areal coverage of reef-building sponges was made. Unfortunately, substrate differences between low and high current transects were too large (Figure 8) to be controlled for, especially given the small sample size ( $n=4$ ).

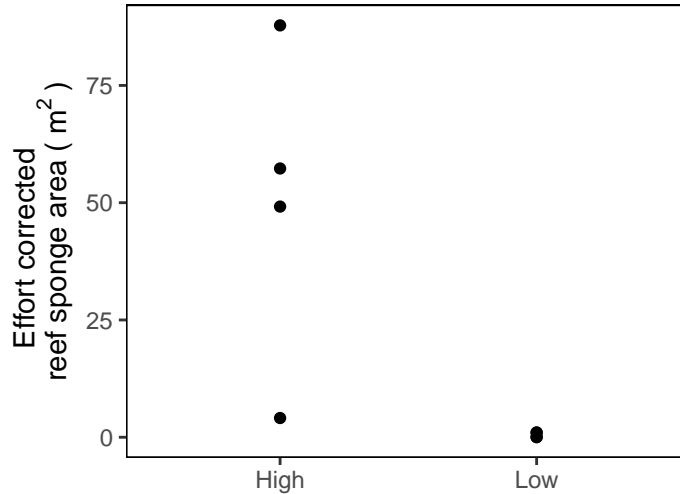


Figure 9. Effort corrected areal coverage of reef-building sponges from transects in high and low current areas. Sponge area was normalized by sampling effort estimated by transect distance multiplied by mean transect image area.

### 3.2 Diversity and community differences

No difference in diversity was observed between high and low current areas when assessed using the commonly reported univariate measures: species richness and evenness (Figure 10). Average richness and evenness were not significantly different between transects in high and low current areas (mean richness: 28 (high), 25 (low),  $p$ -value=0.4; mean evenness: 0.7 (high), 0.5 (low),  $p$ -value=0.2). Given the greater sampling effort in high current areas (Figure 6), one might have expected species richness in high current areas to be greater than in low current areas based on the sampling effort differences alone. However, there remained no difference between high and low current transects in species richness when evaluated with a rarefied number of species (mean rarefied richness: 23 (high), 21 (low),  $p$ -value=0.7). Rarefaction curves (not shown) were approaching their asymptote for all transects, except for low current transect 7 ('Low 7'). The rarefaction results indicate that sampling effort differences did not have a large effect on the species diversity comparison.

More samples (greater than  $n=4$  for each transect type), and a larger difference in tidal current speeds between areas defined as high and low, may be needed to detect an effect of current speed on benthic diversity. Kregting et al. (2016) found very little change in the benthic community structure across small changes in tidal current speed. Thus, it is possible that high current transects (defined as areas with current speed values greater than the 60<sup>th</sup> percentile of modelled values within the study area) did not have large enough tidal current speeds to sustain the expected higher number of species. The mean tidal current speed was 0.09 m/s across high current transects and 0.04 m/s across low current transects based on modelled values (see Section 2.1.2). Those current speeds are relatively low because they represent smoothed and interpolated yearly mean values from a circulation model with a 500 m horizontal resolution. In-situ maximum values would be greater in magnitude.

Detectable differences in community composition were observed between high and low current areas in this study (i.e., species or species groups that were abundant in transects from high current areas differed from those abundant in transects in low current areas). Taxa that were more abundant in low current transects included *Rhabdocalyptus dawsoni* boot sponges, *Pachyercianthus fimbriatus* tube-dwelling anemones, *Pandalus platyceros* spot prawns, pleuronectidae, gadidae and *Lycodes pacificus* fish, and *Halipteris* sea whips. Taxa that were more abundant in high current transects included reef-building hexactinellida sponges, *Swiftia*, *Paragorgia pacifica* and *Anthothela pacifica* corals, *Metridium* anemones, and asteroidea sea stars. As expected, corals and reef-building sponges were more abundant in high current transects (Figure 11). However, some sponges (*Rhabdocalyptus dawsoni*) were much more abundant in the low current transects.

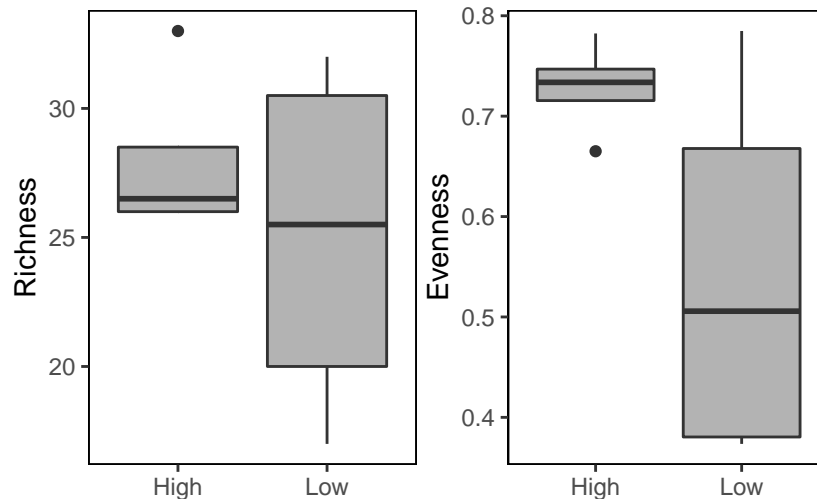


Figure 10. Boxplots of taxa richness and Pielou's evenness from transects in high and low current areas.

Differences in community composition between high and low current areas were also evident with an NMDS ordination (Figure 12), which showed transects from high and low current areas distinctly clustered. The ordination illustrates the relationship between transects in 2-dimensional space based on the abundance of species and species groups in each transect. The low stress value (0.067) suggests that there is little chance that the ordination is a misrepresentation of the measured dissimilarity between transects. Transects from high current areas were clustered more tightly together than transects from low current areas, indicating that transects in low current areas were more variable in their community composition. An analysis of similarity between high and low current transects (ANOSIM statistic  $R=0.4$ ,  $p\text{-value}=0.037$ ) reinforced the difference in community composition that was illustrated with the ordination.

The ordination shows a clear separation between high and low transects. However, high current transect 4 ('High 4') and low current transect 7 ('Low 7') appear closer to each other than any of their within-group transects. This was supported by the dendrogram (Figure 12), where transects High 4 and Low 7 were found to be more similar to each other than any other transects. High 4 may be more similar to low current transects because they shared a number of habitat characteristics. High 4 was the shallowest transect (mean depth 107 m), had the lowest percent

of hard substrate (8%), and had the lowest mean tidal current speed (0.07 m/s) among the high current transects. These findings provide additional evidence that the degree of difference between study comparison groups (high and low current in this case) should be maximized as much as possible within the study area to facilitate the detection of community and diversity difference between them.

Ordination axes, which represent the variability in community composition, were found to have linear relationships with the percent of hard substrate ( $r^2=0.74$ ,  $p\text{-value}=0.046$ ), mean depth ( $r^2=0.79$ ,  $p\text{-value}=0.014$ ) and with modelled bottom tidal current speed ( $r^2=0.61$ ,  $p\text{-value}=0.08$ ). Environmental vectors in Figure 12 represent the direction and strength of the relationship, thus increases in current speed, depth and percent of hard substrate were associated with the species composition found in high current transects. Unfortunately, the percent of hard substrate and depth were both correlated with current speed ( $\rho=0.65$  and  $0.71$ , respectively), thus their individual influences on community structure cannot be disentangled.

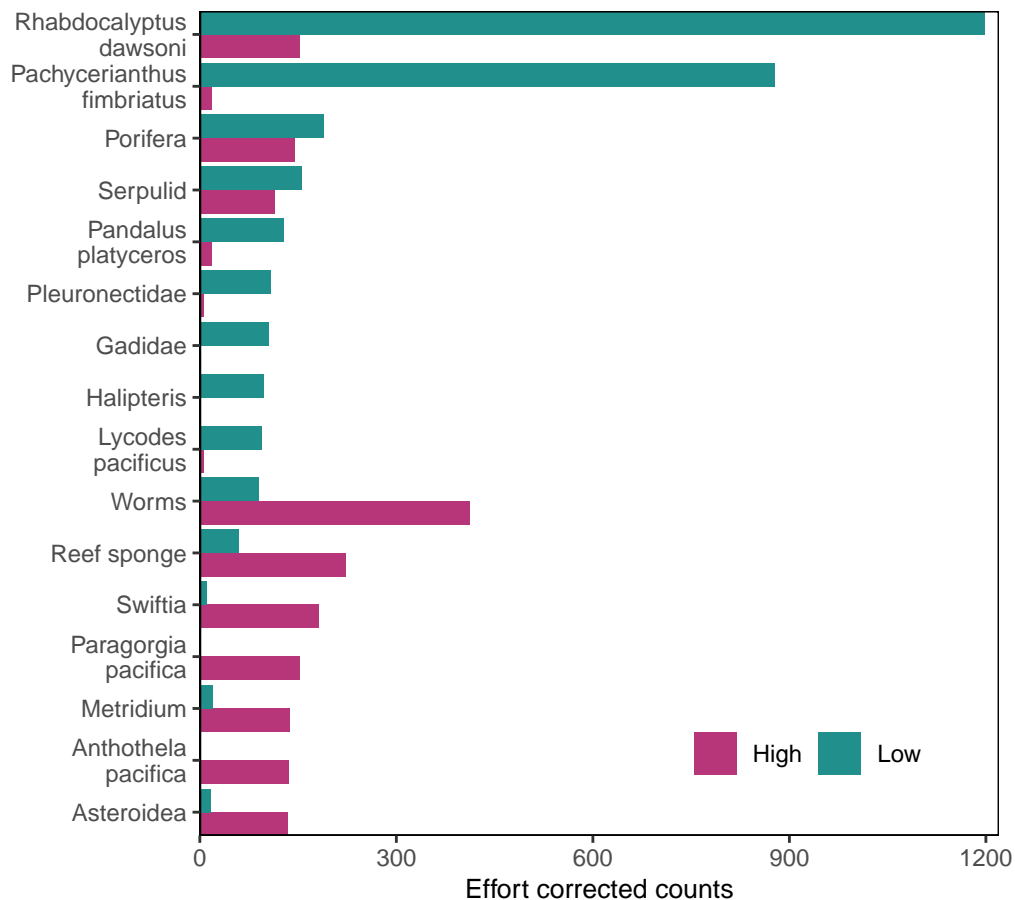


Figure 11. Mean effort-corrected counts from the ten most abundant species or higher-level taxa in transects from high and low current areas. Reef-building sponges were recorded as presence in each image and their counts represent the number of images in which they occurred. Counts for all other taxa represent the total number of individuals observed in transects. Species counts were corrected with the sampling effort index estimated by transect distance multiplied by mean transect image area.



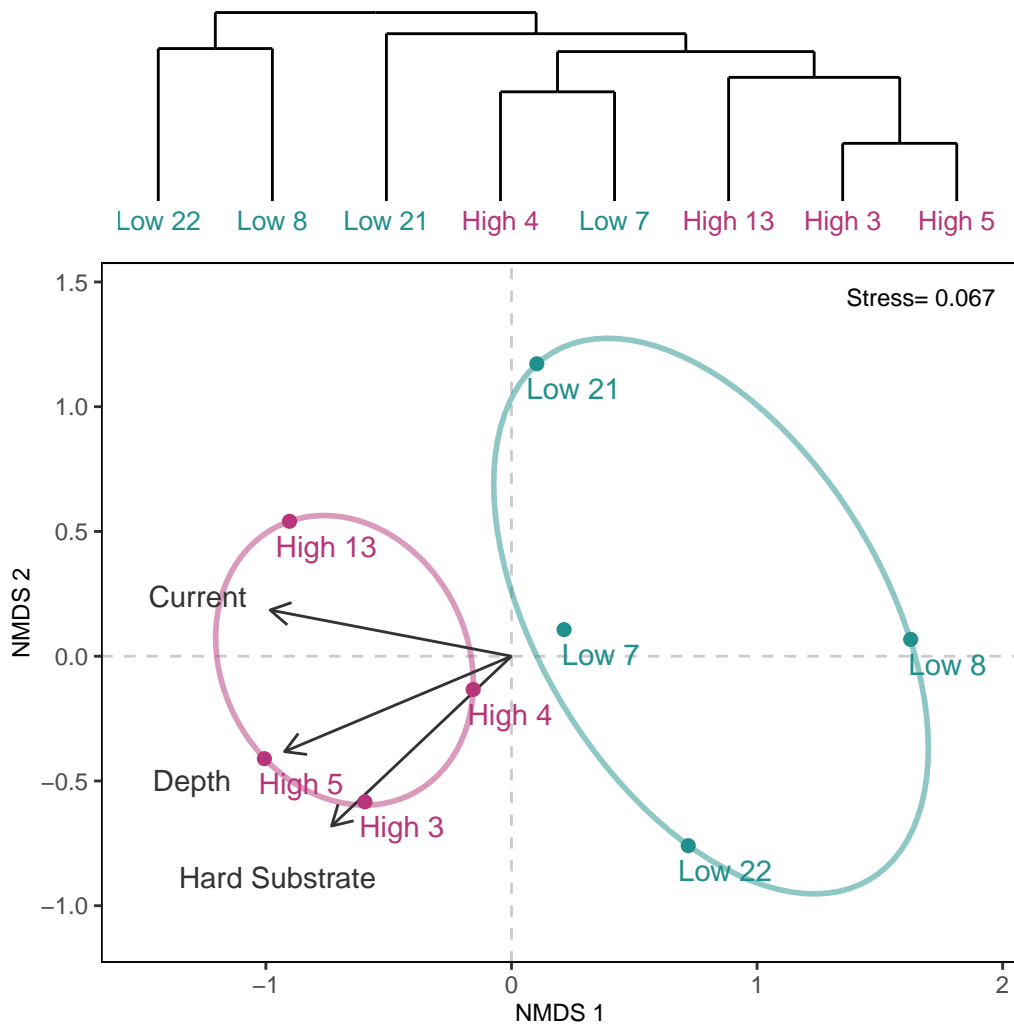


Figure 12. Cluster dendrogram and NMDS ordination using Bray-Curtis distance calculated from effort corrected species counts. Ordination showing transect (site) scores along with environmental variable vectors from fitted vectors with significant linear relationships with the ordination axes. Ellipses represent minimal area ellipsoids for high and low transects.

Based on studies at shallower depths in other regions (Baynes and Szman 1989; Warwick and Uncles 1980; Palardy and Witman 2011), it was expected that differences in species diversity and community structure would be observed between high and low current areas in the Salish Sea. This survey and subsequent data analysis provided some initial empirical evidence for the influence of tidal current speed on benthic community patterns in deeper waters (80 to 200 metres), and supports the recent proposal that, where information is sufficient, high tidal current areas should be considered Ecologically and Biologically Significant Areas (Rubidge et al. 2020). A larger sample size, and a larger difference between tidal current speeds of the study comparison groups, may be required to reliably detect the expected difference in benthic diversity.

Unfortunately, given the limitations of the study design (small sample size and comparison groups confounded by bottom depth and substrate type), it was not possible to disentangle the effect of tidal current speed on benthic community patterns from the effect of substrate and depth. Further studies are needed to confirm that tidal current speed is the primary driver of the increase in reef-building sponge abundance and change in benthic community structure observed here. However, that may be challenging as substrate and tidal current speed are physically related parameters (Kostylev et al. 2001; Wildish and Kristmanson 1997; Warwick and Uncles 1980), with softer sediments more frequently occurring in lower tidal current areas.

Future studies would benefit from a greater number of transects completed across a wider range of tidal current speeds to improve our understanding of the magnitude of tidal current speed required to produce shifts in benthic community structure and diversity. Additionally, in-situ measurements of bottom current speed would be valuable to ground-truth the circulation model data and provide a more accurate measure of the difference in current speed magnitude between high and low current areas. The ultimate goal would be to develop predictive relationships between environmental parameters (tidal current speed, substrate type, and bottom depth) and benthic community patterns and diversity to improve our knowledge of benthic communities distributions within the Salish Sea and other areas of the BC coast.

#### **4 ACKNOWLEDGEMENTS**

The authors would like to thank the following people and organisations for their contributions to this work. Katie Gale and Georgia Clyde for their thorough reviews of the manuscript. Wolfgang Carolsfeld for technical expertise and ROV piloting. Jonathan Martin for ROV piloting and developing the batch image capture tool. James Pegg and Ben Snow for providing ROV data support. Lily Burke for ROV field support. Cherisse DuPreez for ROV field support and for the introduction to the Biigle web-application used for image annotation. The Biigle development team for providing a free, open software for image annotation. Danny Kent for species review and identification. Tammy Norgard for providing support for image annotation. The captain and crew of the CCGS Vector for ROV deployment and field operations.

## REFERENCES

- Baynes, T. and Szmant, A. 1989. Effect of current on the sessile benthic community structure of an artificial reef. *Bulletin of Marine Science* 44. 545–566.
- Bivand, R. and Rundel, C. 2019. rgeos: Interface to Geometry Engine - Open Source ('GEOS'). Available from <https://cran.r-project.org/package=rgeos>.
- Britton-Simmons, K.H., Foley, G. and Okamoto, D. 2009. Spatial subsidy in the subtidal zone: Utilization of drift algae by a deep subtidal sea urchin. *Aquatic Biology* 5(3). 233–243. doi:[10.3354/ab00154](https://doi.org/10.3354/ab00154).
- Crisp, D.J. 1955. The behaviour of barnacle cyprids in relation to water movement over a surface. *Journal of Experimental Biology* 32(3). 569–590.
- Dutertre, M., Hamon, D., Chevalier, C. and Ehrhold, A. 2013. The use of the relationships between environmental factors and benthic macrofaunal distribution in the establishment of a baseline for coastal management. *ICES Journal of Marine Science* 70(2). 294–308. doi:[10.1093/icesjms/fss170](https://doi.org/10.1093/icesjms/fss170).
- Elahi, R., Dwyer, T.R. and Sebens, K.P. 2014. Mesoscale variability in oceanographic retention sets the abiotic stage for subtidal benthic diversity. *Marine Ecology Progress Series* 498. 117–132. doi:[10.3354/meps10642](https://doi.org/10.3354/meps10642).
- Fabricius, K.E., Genin, A. and Benayahu, Y. 1995. Flow-dependent herbivory and growth in zooxanthellae-free soft corals. *Limnology and Oceanography* 40(7). 1290–1301. doi:[10.4319/lo.1995.40.7.1290](https://doi.org/10.4319/lo.1995.40.7.1290).
- Huangn, Z., Brooke, B.P. and Harris, P.T. 2011. A new approach to mapping marine benthic habitats using physical environmental data. *Continental Shelf Research* 31(2 SUPPL.). 4–16. doi:[10.1016/j.csr.2010.03.012](https://doi.org/10.1016/j.csr.2010.03.012).
- Jumars, P. 1993. *Concepts in biological oceanography: An inter-disciplinary approach*. Oxford University Press, New York.
- Kostylev, V.E., Todd, B.J., Fader, G.B.J., Courtney, R.C., Cameron, G.D.M. and Pickrill, R.A. 2001. Benthic habitat mapping on the Scotian Shelf based on multibeam bathymetry, surficial geology and sea floor photographs. *Marine Ecology Progress Series* 219. 121–137. doi:[10.3354/meps219121](https://doi.org/10.3354/meps219121).
- Kregting, L., Elsaesser, B., Kennedy, R., Smyth, D., O'Carroll, J. and Savidge, G. 2016. Do changes in current flow as a result of arrays of tidal turbines have an effect on benthic communities? *PloS one* 11(8). e0161,279–e0161,279. doi:[10.1371/journal.pone.0161279](https://doi.org/10.1371/journal.pone.0161279).
- Langenkämper, D., Zurowietz, M., Schoening, T. and Nattkemper, T.W. 2017. BIIGLE 2.0 - Browsing and annotating large marine image collections. *Frontiers in Marine Science* 4. 83. doi:[10.3389/fmars.2017.00083](https://doi.org/10.3389/fmars.2017.00083).

- Lesser, M.P., Witman, J.D. and Sebens, K.P. 1994. Effects of flow and seston availability on scope for growth of benthic suspension-feeding invertebrates from the Gulf of Maine. *Biological Bulletin* 187(3). 319–335. doi:[10.2307/1542289](https://doi.org/10.2307/1542289).
- Leys, S.P., Wilson, K., Holeton, C., Reising, H.M. and Austin, W.C. 2004. Patterns of glass sponge (Porifera, Hexactinellida) distribution in coastal waters of British Columbia, Canada. *Marine Ecology Progress Series* 283. 133–149.
- Lundblad, E.R., Wright, D.J., Miller, J., Larkin, E.M., Rinehart, R., Naar, D.F., Donahue, B.T., Anderson, S.M. and Battista, T. 2006. A benthic terrain classification scheme for American Samoa. *Marine Geodesy* 29(2). 89–111. doi:[10.1080/01490410600738021](https://doi.org/10.1080/01490410600738021).
- McArthur, M.A., Brooke, B.P., Przeslawski, R., Ryan, D.A., Lucieer, V.L., Nichol, S., McCallum, A.W., Mellin, C., Cresswell, I.D. and Radke, L.C. 2010. On the use of abiotic surrogates to describe marine benthic biodiversity. *Estuarine, Coastal and Shelf Science* 88(1). 21–32. doi:[10.1016/j.ecss.2010.03.003](https://doi.org/10.1016/j.ecss.2010.03.003).
- Oksanen, J., Blanchet, F.G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., Minchin, P.R., O'Hara, R.B., Simpson, G.L., Solymos, P., Stevens, M.H.H., Szoecs, E. and Wagner, H. 2019. *vegan*: Community ecology package. Available from <https://cran.r-project.org/package=vegan>.
- Palardy, J.E. and Witman, J.D. 2011. Water flow drives biodiversity by mediating rarity in marine benthic communities. *Ecology Letters* 14(1). 63–68. doi:[10.1111/j.1461-0248.2010.01555.x](https://doi.org/10.1111/j.1461-0248.2010.01555.x).
- Ramiro-Sánchez, B., González-Irusta, J.M., Henry, L.A., Cleland, J., Yeo, I., Xavier, J.R., Carreiro-Silva, M., Sampaio, Í., Spearman, J., Victorero, L., Messing, C.G., Kazanidis, G., Roberts, J.M. and Murton, B. 2019. Characterization and mapping of a deep-sea sponge ground on the Tropic Seamount (Northeast Tropical Atlantic): Implications for spatial management in the high seas. *Frontiers in Marine Science* 6(May). 278. doi:[10.3389/fmars.2019.00278](https://doi.org/10.3389/fmars.2019.00278).
- Rice, A.L., Thurston, M.H. and New, A.L. 1990. Dense aggregations of a hexactinellid sponge, *Pheronema carpensteri*, in the Porcupine Seabight (northeast Atlantic Ocean), and possible causes. *Progress in Oceanography* 24(1). 179–196. doi:[10.1016/0079-6611\(90\)90029-2](https://doi.org/10.1016/0079-6611(90)90029-2).
- Rooper, C.N., Zimmermann, M., Prescott, M.M. and Hermann, A.J. 2014. Predictive models of coral and sponge distribution, abundance and diversity in bottom trawl surveys of the Aleutian Islands, Alaska. *Marine Ecology Progress Series* 503. 157–176. doi:[10.3354/meps10710](https://doi.org/10.3354/meps10710).
- Roughgarden, J., Gaines, S. and Possingham, H. 1988. Recruitment dynamics in complex life cycles. *Science* 241(4872). 1460–1466. doi:[10.1126/science.11538249](https://doi.org/10.1126/science.11538249).
- Rubidge, E., Gale, K.S.P., Curtis, J.M.R., McClelland, E., Feyrer, L., Bodtker, K. and Robb, C. 2016. Methodology of the Pacific marine ecological classification system and its application to the Northern and Southern Shelf Bioregions. DFO Canadian Science Advisory Secretariat Research Document 035. 124.
- Rubidge, E., Jeffery, S., Gregr, E.J., Gale, K.S.P. and Frid, A. 2020. Assessment of nearshore features in the Northern Shelf Bioregion against criteria for determining Ecologically and Biologically Significant Areas (EBSAs). DFO Canadian Science Advisory Secretariat Research Document 023. 63.

- Sappington, J.M., Longshore, K.M. and Thompson, D.B. 2007. Quantifying landscape ruggedness for animal habitat analysis: A case study using bighorn sheep in the Mojave Desert. *The Journal of Wildlife Management* 71(5). 1419–1426. doi:[10.2193/2005-723](https://doi.org/10.2193/2005-723).
- Sarkar, S., Justus, J., Fuller, T., Kelley, C., Garson, J. and Mayfield, M. 2005. Effectiveness of environmental surrogates for the selection of conservation area networks. *Conservation Biology* 19(3). 815–825. doi:[10.1111/j.1523-1739.2005.00236.x](https://doi.org/10.1111/j.1523-1739.2005.00236.x).
- Schram, C., Ladell, K., Mitchell, J. and Chute, C. 2019. From one to ten: Canada's approach to achieving marine conservation targets. *Aquatic Conservation: Marine and Freshwater Ecosystems* 29(S2). 170–180. doi:[10.1002/aqc.3133](https://doi.org/10.1002/aqc.3133).
- Sebens, K.P. 1984. Water flow and coral colony size: Interhabitat comparisons of the octocoral *Alcyonium siderium*. *Proceedings of the National Academy of Sciences* 81(17). 5473–5477. doi:[10.1073/pnas.81.17.5473](https://doi.org/10.1073/pnas.81.17.5473).
- Soontiens, N. and Allen, S.E. 2017. Modelling sensitivities to mixing and advection in a sill-basin estuarine system. *Ocean Modelling* 112. 17–32. doi:[10.1016/j.ocemod.2017.02.008](https://doi.org/10.1016/j.ocemod.2017.02.008).
- Soontiens, N., Allen, S.E., Latornell, D., Souëf, K.L., Machuca, I., Paquin, J.P., Lu, Y., Thompson, K. and Korabel, V. 2016. Storm surges in the Strait of Georgia simulated with a regional model. *Atmosphere-Ocean* 54(1). 1–21. doi:[10.1080/07055900.2015.1108899](https://doi.org/10.1080/07055900.2015.1108899).
- Walbridge, S., Slocum, N., Pobuda, M. and Wright, D.J. 2018. Unified geomorphological analysis workflows with Benthic Terrain Modeler. *Geosciences* 8(3). 94. doi:[10.3390/geosciences8030094](https://doi.org/10.3390/geosciences8030094).
- Warwick, R.M. and Uncles, R.J. 1980. Distribution of benthic macrofauna associations in the Bristol Channel in relation to tidal stress. *Marine Ecology Progress Series* 3. 97–103. doi:[10.3354/meps003097](https://doi.org/10.3354/meps003097).
- Wildish, D.J. and Kristmanson, D. 1997. *Benthic suspension feeders and flow*. Cambridge University press, New York.
- Wildsmith, M.D., Valesini, F.J. and Robinson, S.F. 2017. The value of enduring environmental surrogates as predictors of estuarine benthic macroinvertebrate assemblages. *Estuarine, Coastal and Shelf Science* 197. 159–172. doi:<https://doi.org/10.1016/j.ecss.2017.08.006>.
- Witman, J.D., Leichter, J.J., Genovese, S.J. and Brooks, D.A. 1993. Pulsed phytoplankton supply to the rocky subtidal zone: Influence of internal waves. *Proceedings of the National Academy of Sciences of the United States of America* 90(5). 1686–1690. doi:[10.1073/pnas.90.5.1686](https://doi.org/10.1073/pnas.90.5.1686).

## APPENDIX A Image capture tool

```
#####  
# Python Batch Image Capture Tool  
#  
# Originally developed by Jonathan Martin  
# January 2019  
# Adapted by Jessica Nephin (jessica.nephin@dfo-mpo.gc.ca)  
# March 2019  
#  
# Requirements:  
# Requires ffmpeg command line tool to capture images,  
# Download at https://ffmpeg.org/  
#  
# Description:  
# 1) Captures images at a specified frequency from videos listed  
# in an imported csv file.  
# 2) Renames the images by transect name, date and time for  
# uploading to Biigle.  
#  
# Instructions:  
# 1) Check that all required modules are installed.  
# 2) Modify inputs (e.g., rate of image capture).  
# 3) Run python script from the directory containing the videos.  
#####  
  
#####  
# Required modules  
#####  
  
import os  
from datetime import datetime  
from datetime import timedelta  
import subprocess  
import csv  
import re  
  
#####  
# Inputs  
#####  
  
# Rate of image capture in seconds  
stillsFrequency = 5  
  
# Full path to csv with video filenames,  
# file should contain video filename, start time and transect name,  
# without headings  
# column 1 = Filename (e.g., videofile.mov)  
# column 2 = Start datetime (e.g., yyyy/mm/dd hh:mm:ss)  
# column 3 = Transect name  
videofiles = "<path to directory>/<name of file>.csv"
```

```

# Directory to save images
imageDir = "<path to directory>/<name of directory for images>/"

#####
# Setup
#####

# Function to convert a csv file into dictionaries
def csv_dict(variables_file, col):
    # Load csv and create a dictionary
    with open(variables_file, mode="r") as infile:
        reader = csv.reader(infile)
        dict_list = {rows[0]:rows[col] for rows in reader}
        return dict_list

# Calls csv_dict,
# returns dictionary of video filenames and start times
startsDict = csv_dict(videofiles, 1)
transectsDict = csv_dict(videofiles, 2)

# Get video filenames
videoNames = [k for k in startsDict]

#####
# Loop through each video file and capture images
#####

# Loop through each video file
for video in videoNames:

    # Get actualStartTime
    actualStartTime = startsDict[video]

    # Get transect name
    transectName = transectsDict[video]

    # Create directory to save images if it doesn't exist
    if not os.path.exists(imageDir + transectName):
        os.makedirs(imageDir + transectName)

    # rename all image files with timestamps
    def renameImageFiles():
        global actualStartTime
        global stillsFrequency
        counter = 0

        #stillTime
        startTimeDateTime = datetime.strptime(
            actualStartTime, "%Y-%m-%d %H:%M:%S")

        # files from image grabs
        imageList = os.listdir(imageDir + transectName)

```

```

for fileName in imageList:

    # Elapsed seconds
    elapsedSecs = stillsFrequency * counter
    # Calculate elapsed time
    elapsedDatetime = (startTimeDateTime +
                       timedelta(seconds=elapsedSecs))
    # text string
    strDateTime = str(elapsedDatetime)
    # sub space for underscore
    nospaces = re.sub("\s+", "_", strDateTime)
    nocolons = re.sub(":", "", nospaces)
    fileDateTime = (transectName + "_" +
                    re.sub("-", "", nocolons))

    # rename
    orig = imageDir + transectName + "/" + fileName
    new = (imageDir + transectName + "/" +
           fileDateTime + ".jpeg")
    os.rename(orig, new)

    # add to counter
    counter += 1

# Get framerate of video
command = ("ffprobe -v error -select_streams v -of "
           "default=noprint_wrappers=1:nokey=1 "
           "-show_entries stream=r_frame_rate " + video)
FrameRate = str(subprocess.check_output(command) )
fr = FrameRate.split("/")
frNum = float(fr[0]) / float(fr[1])

# Image capture frequency calculated with framerate
FrameRateFreq = str(int(round( frNum * stillsFrequency )))

# generate stills, start at the beginning of video
path = imageDir + "/" + transectName + "/" + transectName
FFMCommand = ("ffmpeg -ss 0 -i " + video +
              " -vf \"select=not(mod(n,\" +
              FrameRateFreq + "))\" -vsync vfr -q:v 1 " +
              path + "_%03d.jpg")
subprocess.call(FFMCommand, shell = True)

#adds time stamp to captures and moves them to a named directory
renameImageFiles()

```



## APPENDIX B R code for processing Biigle data

```
#####  
# R code to combine Biigle reports and perform data quality checks  
#  
# Developed by Jessica Nephin (jessica.nephin@dfo-mpo.gc.ca)  
# September 2019  
#  
# Requirements:  
# Packages readxl and reshape2  
#  
# Description:  
# 1) Prepares CSV image and CSV annotation exports for merging  
# 2) Checks for missing or duplicate laser point annotations  
# 3) Merges CSV image, CSV annotation and Area reports  
# 4) Performs data quality checks on combined dataset  
#  
# Instructions:  
# 1) Check that all required packages are installed  
# 2) Export CSV image, CSV annotation and Area reports from Biigle  
#####  
  
#####  
# Load Biigle reports into an R workspace  
#####  
  
# load required package  
library(readxl)  
# Set working directory  
setwd("<Path to folder containing Biigle reports>")  
  
# Extract csv annotation report from .zip folder  
# Save files in a folder named "Annotation"  
# Load all annotation reports into the workspace  
files <- list.files(path = "Annotation", full.names = TRUE)  
annotations <- NULL  
for(f in files){  
  tmp <- read.csv(f, header= TRUE, stringsAsFactors = FALSE)  
  annotations <- rbind(annotations, tmp)  
}  
  
# Extract csv image report from .zip folder  
# Save files in a folder named "Image"  
# Load all image reports into the workspace  
files <- list.files(path = "Image", full.names = TRUE)  
images <- NULL  
for(f in files){  
  tmp <- read.csv(f, header= TRUE, stringsAsFactors = FALSE)  
  images <- rbind(images, tmp)  
}  
  
# If polygon annotations were made then,
```

```

# Extract the area report from .zip folder
# Save files in a folder named "Area"
# Load all area reports into the workspace
files <- list.files(path = "Area", full.names = TRUE)
area <- NULL
for(f in files){
  # Load only the desired columns and set column type
  tmp <- as.data.frame(
    read_xlsx(f, skip=1,
              col_types=c("text","skip","skip","skip","text",
                          "skip","text","skip","skip","numeric",
                          "skip","skip","numeric")))
  area <- rbind(area, tmp)
}

#####
# Clean image reports
#####

# load required package
library(reshape2)

# Get filenames for images labelled as 'Image Not Annotated'
notAnnotated <- images$filename[grep("Image Not Annotated",
                                     images$label_hierarchy)]

# Remove images which were not annotated
images <- images[which(!images$filename %in% notAnnotated),]

# Seperate label_hierarchy into 2 columns (e.g., Substrate > Mud)
image_labels <- colsplit(images$label_hierarchy, ">",
                          c("Category","Label"))

# Clean up 'Category' and 'Label' columns
image_labels$Category <- gsub(" |-", "", image_labels$Category)
image_labels$Label <- sub("*. _", "", image_labels$Label)
image_labels$Label <- sub(" [(]Dominant[)]", "", image_labels$Label)
image_labels$Label <- gsub("~\\s+|\\s+$", "", image_labels$Label)

# Add yes label for dead sponge category
image_labels$Label[image_labels$Category ==
                  "Deadsponge(s)present"] <- "yes"

# Bind 'image_labels' back with 'images'
images <- cbind(images, image_labels)

# Reshape 'images' so that each image label category is a column
images <- dcast(images, value.var = "Label",
                filename+image_id+longitude+latitude~Category)

# Convert latitude and longitude back to numeric
images$latitude <- as.numeric(images$latitude)
images$longitude <- as.numeric(images$longitude)

```

```

# Check results of reshape
# If dcast() returns numbers rather than text as expected then,
# there are duplicate image labels in any one category
head(images)

#####
# Check for laser point annotation errors
#####

# Subset dataframe to get only laser point annotations
laserpts <- annotations[annotations$label_name == "Laser Point",]

# Make sure all annotation shapes are points
table(laserpts$shape_name)
# If other annotation shapes exist, return their image filenames
unique(laserpts$filename[laserpts$shape_name != "Point"])

# Check number of 'laser point' annotations in each image
numpts <- aggregate(label_name~filename, length, data=laserpts)

# Images with only 1 laser point
unique(numpts$filename[numpts$label_name == 1])

# Images with more than 2 laser point
unique(numpts$filename[numpts$label_name > 2])

#####
# Clean annotation reports
#####

# Remove any annotations from images labelled 'Image Not Annotated'
# 'notAnnotated' object defined in previous Code Block
annotations <- annotations[!(annotations$filename %in%
                             notAnnotated),]

# Get just laser point annotations,
# and remove those records from 'annotations' dataframe
lasers <- annotations[annotations$label_name == "Laser Point",]
annotations <- annotations[annotations$label_name != "Laser Point",]

# Find images with only laser point annotations
# These are images with no species annotations, i.e. absence records
absences <- lasers[!(lasers$filename %in% annotations$filename),]

# Remove duplicate laser point records,
# store in dataframe as absence records
absences <- absences[!duplicated(absences$filename),]
absences$label_name <- "Empty"
absences$label_hierarchy <- "Empty"

# Combine annotation records with absence records

```

```

annotations <- rbind(annotations, absences)

#####
# Merge reports from Biigle into a single dataset
#####

# If polygon annotations were made then,
# Merge 'areas' with 'annotations'
annotations <- merge(
  annotations, area, all.x=TRUE,
  by.x=c("filename", "annotation_id", "label_name"),
  by.y= c("image_filename", "annotation_id", "label_names"))

# Merge 'images' and 'annotations' reports
dat <- merge(images, annotations, all=TRUE,
             by=c("filename", "image_id"))

#####
# Perform quality control checks on the combined dataset
#####

# check for missing values in longitude, latitude
unique(dat$filename[is.na(dat$latitude)])

# Remove any records with missing image longitude, latitude
dat <- dat[which(!dat$filename %in%
                unique(dat$filename[is.na(dat$latitude)])),]

# Check for images with no image quality labels
unique(dat$filename[is.na(dat$ImageQuality)])

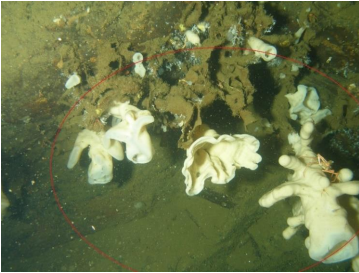
# Check for images with no dominant substrate label
unique(dat$filename[is.na(dat$SubstrateDominant)])

# Check for images with image labels but missing annotations
unique(dat$filename[is.na(dat$label_name)])

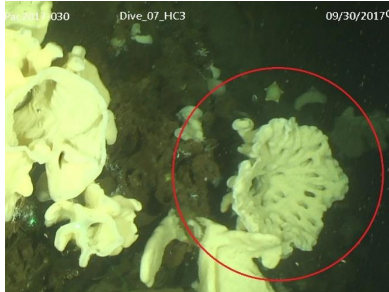
```

## APPENDIX C Species identification guide

**Aphrocallistes vastus**- Cream coloured glass sponge, variety of morphologies, never with holes



**Heterochone**- Cream coloured glass sponge, goblet shaped with numerous perforations



**Rhabdocalyptus dawsoni**- Tubular glass sponge that is always covered in silt. Can be oriented upright or lying sideways



**Yellow sponge**- Encrusting patches of yellow sponge often found amongst glass sponge aggregations. Looks yellow and crumbly



**Polymastia**- White encrusting demosponge found growing on rock with numerous projections protrude from the base. White/cream/yellow



**Demosponge 1**- short tubular sponge, white/yellow



**Demosponge 2**- Taller tubular sponge, dirty yellow



**Demosponge 3**- Squat, creamy yellow sponge





**Anemone**- Generic anemone. Use for any anemone that can't be ID'd to species



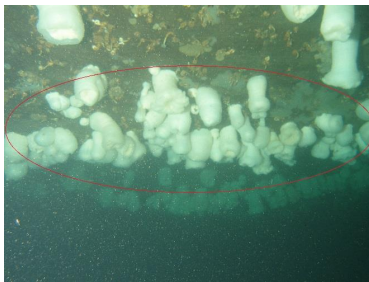
**Cribrinopsis fernaldi**- Squat, pink anemone with long, thin tentacles that are sometimes droopy



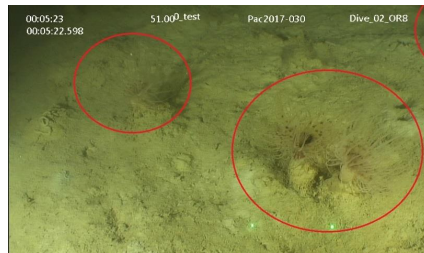
**Zoanthid**- Small, thick-walled anemone. Brown stalk with lighter tentacles. Stalk often covered with silt. Grows in aggregations



**Metridium**- Plumose anemone. Tall anemone with frilly tentacles. Can be orange or white. Tentacles are often retracted



**Pachycerianthus fimbriatus**- burrowing anemone. Only found in soft sediment. Banding on tentacles. Can be beige or dark purple



**Paragorgia pacifica**- Large, pink gorgonian fan coral. Tips of branches are bulbous. Will look fuzzy if tentacles are extended



**Swiftia torreyi**- Small, branching gorgonian coral. Much thinner branches than P. pacifica



**Anthothela pacifica**- white, creeping gorgonian coral



**Halipteris**- Large, white sea whip



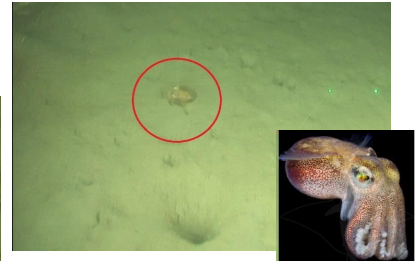
**Calliostoma**- Small snail with pointed top



**Fusitriton oregonensis**- Large, hairy snail



**Rossia pacifica**- Stubby squid. Small, benthic squid. Shaped more like an octopus than squid. Usually on mud



**Peltodoris lentiginosa**- Large white dorid nudibranch



**Asteroid 1**- Sea star with long arms (similar to a painted star – Orthasterias). Arm length: disc ratio = 6-10 : 1



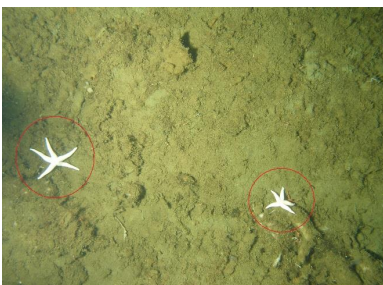
**Asteroid 2**- Orange sea star with short arms and webbing between them (similar to vermilion star – Mediaster). Arm length: disc ratio = 2-3 : 1



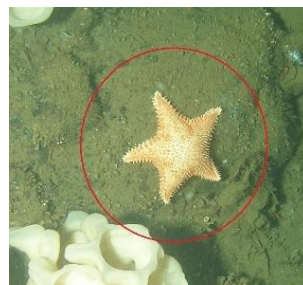
**Ceramaster**- Cookie cutter star. Very short arms, yellow/orange. Arm length: disc ratio = <2 : 1, may have ridges along outer margin, body may be flat or puffy



**Henricia sp.**- Small sea star with very long thin arms (Arm length: disc ratio = 2 : 1). Usually looks white



**Hippasterias phrygiana**- Large, orange sea star with spines. Arm length: disc ratio = 2 : 1



**Psolus chitonoides**- Orange, round, armoured, disc-shaped cucumber with red feeding tree





**Chionoecetes sp-** Large, long-legged spider crab with roundish body and small claws. Usually found on mud



**Cancer sp-** Classic Dungeness crab shape. May be seen as mating pairs gripping each other face to face



**Crab decorator-** Small, long legged crab with very small pincers. Sometimes decorated, often has claws up in the air, often clinging to erect features



**Crab lithode 1-** Low profile crab with only 6 walking legs, wide teardrop carapace, short pincer arms and larger claws



**Crab lithode 2-** Large box crab with high profile carapace, 6 thick walking legs and large pincers. Use this label when the species appears to be a box crab, but you can't tell for certain that it is *L. foraminatus*.



**Lopholithodes foraminatus-** Brown box crab. Large box crab with high profile carapace, 6 thick walking legs and large pincers.



**Pandalus platyceros-** Spot prawn. Large shrimp with white banding on legs



**Pandalus sp.-** Various shrimp species, all with a similar morphology to *P. platyceros*, but different colorations.



**Munida-** Squat lobster. Small, red crustacean with squat carapace. Often seen with long claws extended upwards and carapace inside a hole or crevice.



**Brachiopod-** Small, brown, attached to rock resembles clam shaped bivalve



**Corella sp.-** Small, clear tunicate, often in clusters





**Spagetti worm**- Long, thin tentacle like strands originating from a central indentation in the mud. Difficult to see



**Serpulid**- Calcareous tube worm. White, twisted, calcareous casings attached to rock



**Agonidae**- Poacher family. Bottom dwelling fish that sits on pectoral fins; scaly, rigid body; sometimes with whiskers. Unlikely any poacher will be identified to species.



**Cottidae**- Sculpin family. Bottom dwelling fish that sits on belly; colour is brown/black/white, often mottled. Use when sculpin cannot be identified to species. Less rigid and scaly than Agonidae.



**Icelinus sp**- Large sculpin (Cottidae) with a large, deep head, and blunt snout



**Lycodes pacificus**- Elongated, soft-bodied, creamy yellow fish, usually on mud, sometimes seen with tail curled.



**Porichthys notatus**- Midshipmen. Small soft-bodied fish, often with horizontal banding, benthic, large round head. Can be hard to distinguish from a sculpin (Cottidae).



**Gadus macrocephalus**- Large cod species with 1 whisker under bottom lip. Very rotund. (cod are distinguished by their 3 individual dorsal fins)



**Gadidae**- Species in the cod family, which are distinguished by their 3 individual dorsal fins. Often schooling and slender bodied. Use this label when the cod can't be identified as G. macrocephalus.



**Cymatogaster aggregata**- Schooling, silver fish. Laterally compressed, body tall in center and elongated at mouth and tail. Swims by flapping pectoral fins



**Sebastes sp.-** Rockfish family. Characterized by 2 part dorsal fin with spiny rays at anterior and soft rays at posterior. Use this label when the fish is known to be a rockfish, but it can't be identified to species.



**Sebastes diploproa-** Splitnose rockfish. Benthic rockfish, orange, can be banded, always has notch in top lip



**Sebastes elongatus-** Greenstripe rockfish. Elongated, benthic rockfish. Usually on mud. Has dark, broken horizontal bars along sides



**Sebastes maliger-** Quillback rockfish. Deepbodied rockfish with light vertical band behind eyes and light patch on dorsal



**Sebastes proriger-** Redstripe rockfish. Elongated body, short dorsal spines, light lateral line, body coloration is pink/orange/rusty and even.



**Sebastes ruberrimus-** Yelloweye rockfish. Large rockfish, orange or red, even coloration except light lateral line and eye. Juveniles are darker with 2 distinct horizontal white lines.



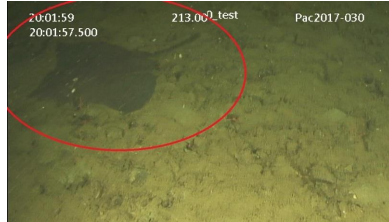
**Ophiodon elongates-** Lingcod. Member of the greenling family (Hexagrammidae). Large elongated body, large head, benthic, and tapered mouth



**Hydrolagus colliei**- Ratfish. Large pectoral fins, long tail, green eyes



**Raja sp.**- Large skate. Use this label with a skate cannot be identified to species.



**Squalus suckleyi**- Spiny dogfish. Elongate, sleek, grey bodied shark



**Pleuronectidae**- Flatfish family. Catch all label for any flatfish that cannot be identified to species (e.g. *L. exilis*)



**Lyopsetta exilis**- slender, elongated flatfish, often with brown spots or mottling

