Appendix S1

Data and reproducibility

A warming western boundary current increases the prevalence of commercially disruptive parasites in broadbill swordfish

Jessica A. Bolin, Karen J. Evans, David S. Schoeman, Claire M. Spillman, Thomas S. Moore II, Jason R. Hartog, Scott F. Cummins & Kylie L. Scales

Fisheries Oceanography

Table of Contents

Figure S1

Temporal frequency of swordfish sampled over the study period and associated infection metrics. Top row = prevalence data set. Bottom row = intensity data set. (Left): Sampling frequency of swordfish over the study period. No samples obtained in April, July or November 2020, and January 2022, due to COVID-19 disruptions. Further, no samples from November 2019 to February 2020 were collected in the intensity dataset. (Top right): Percentage of swordfish infected with *Kudoa musculoliquefaciens*; (bottom right): Median number of spores/g^{-1} of swordfish tissue.

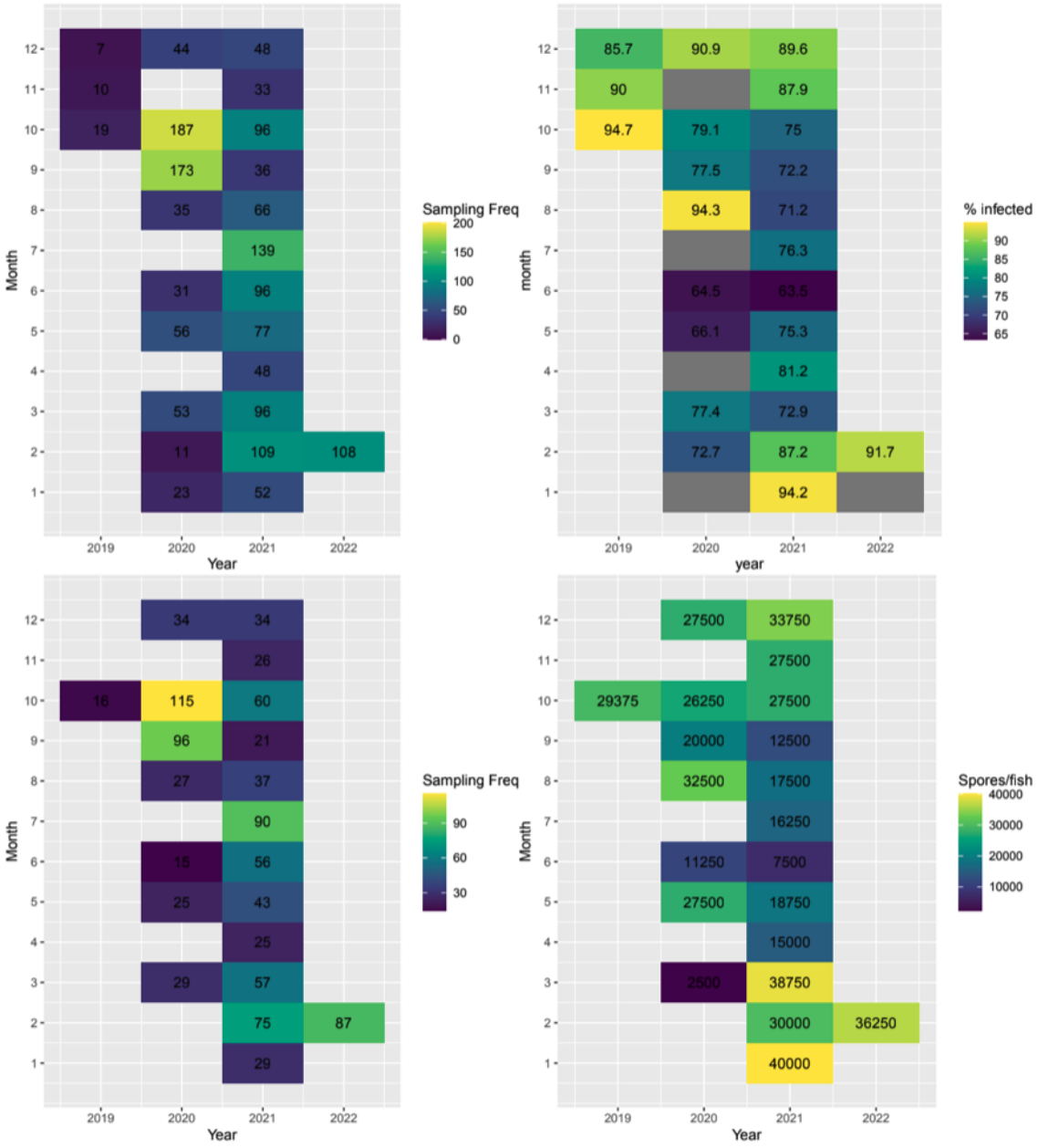
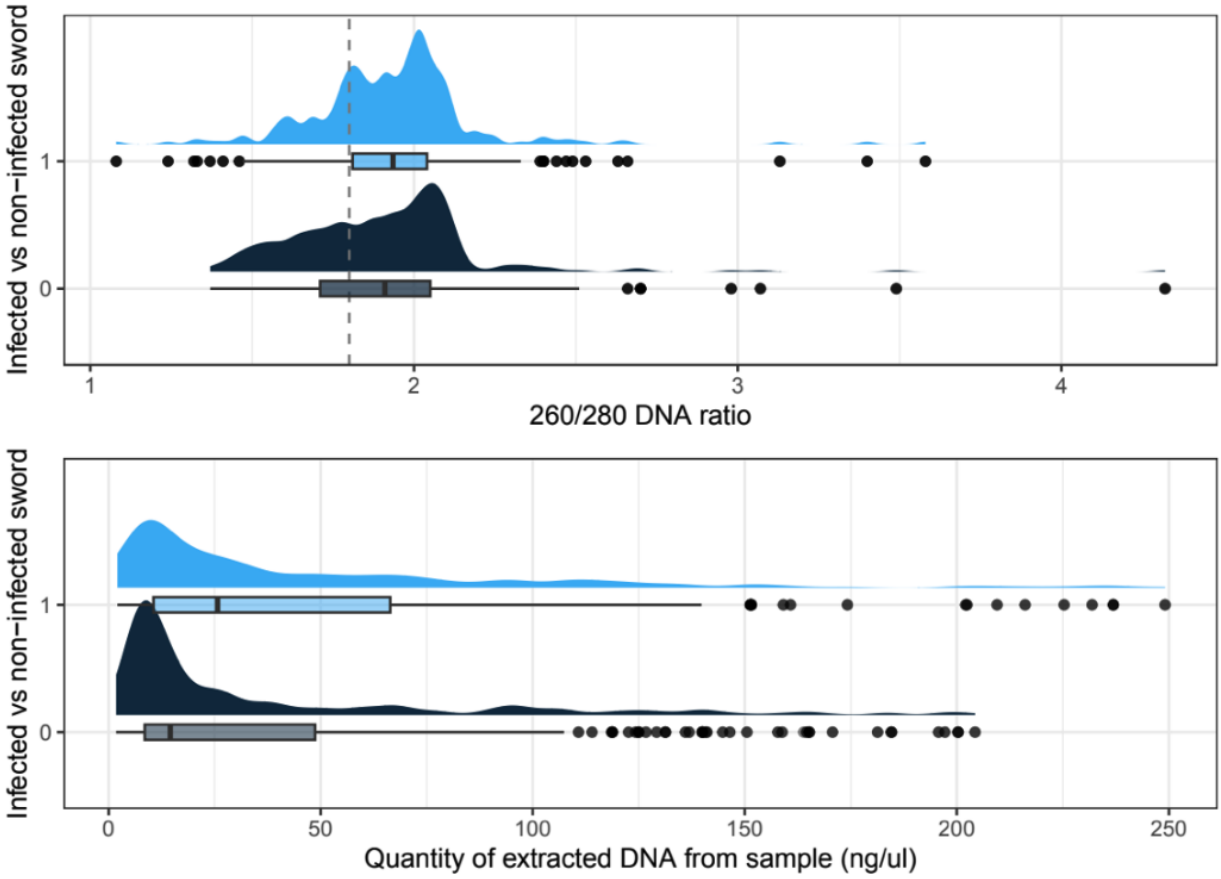


Figure S2

Distributions of the quantity and purity of swordfish samples subject to DNA extraction. Top: 260/280 nm (purity) ratios. A ratio of ~1.8 indicates the DNA is free of residual protein (dashed grey line). However, kudoid DNA could still be detected via PCR in samples with purity ratios lower than 1.8. Bottom: DNA yield extracted from swordfish samples (ng/µl).



Section S1

Machine Specifications

All analyses were run on a machine with the following specifications:

```
System:             Darwin
Release:            21.2.0
Machine:            x86_64
macOS Version:      12.1
CPU:                Intel(R) Core(TM) i7-8850H CPU @ 2.60GHz
Cores:              12
RAM (GB):           17.18
R Version:          R version 4.2.2 (2022-10-31)
Version Control:    git version 2.39.1
```

Fishing locations

Latitude and longitude coordinates of fishing locations were provided directly by Walker Seafoods Australia and 4 Seas Pty. Ltd, after returning to port from each fishing trip. Data cane in the form of either AL06 logbook sheets, or excel spreadsheets via eLogs. Fishing trip identifiers (i.e., operation) were renamed to randomly generated numbers between 0–100 prior to analysis. Therefore, it is not possible to identify the company or vessel associated with any one observation in the dataset.

Dynamic ocean variables

Metrics of ocean state (excluding chlorophyll-*a*) are from the Australian Community Climate Earth-System Simulator - Seasonal Version 2 (ACCESS-S2) reanalysis. ACCESS-S2 is the result of a collaboration between the Bureau of Meteorology and the UK Meteorological Office (UKMO). We downloaded `.nc` files at their native resolution (i.e., 0.25˚), and at daily and monthly frequencies via the NCI THREDDS data server. These files are natively on a tripolar grid, so we regridded them to a standard lat/lon format via the command line, using CDO. We downloaded monthly chlorophyll-a (OCI algorithm) from the Aqua MODIS satellite, via NOAA ERDDAP at their native 0.04˚resolution. Dataset ID: `erdMH1chlamday`.

Topography

We downloaded shapefiles of the position and total area of seamounts and guyots off Eastern

Australia from the global seafloor geomorphic features map, provided by Blue Habitats. See Harris et al. 2014 for a complete description of the dataset. We downloaded gridded global bathymetry from the General Bathymetric Chart of the Oceans (GEBCO) 2021 grid at 0.05 arc second resolution.

Table S1

R packages used for data munging and cleaning are primarily those comprising the `tidyverse`. Packages commonly used for certain tasks throughout the analysis (besides `base` and `stats`) are specified below.

Purpose	Packages	References
Downloading environmental data	`RCurl, rvest`	(Temple 2022, Wickham 2022)
Extracting environmental data	`terra, exactextractr, sf, raster, geosphere`	(Hijmans 2023, Baston 2022, Pebesma 2018, Hijmans 2022, Hijmans 2022)
Data exploration	`lattice, ggplot2, sp, gstat, ncf`	(Sarkar 2008, Wickham 2016, Pebesma & Bivand 2005, Bivand et al 2013, Pebesma 2004, Gräler et al 2016, Bjornstad 2022)
Modelling (mixed effects models)	`lme4, MASS, bbmle`	(Bates et al 2015, Venables & Ripley 2002, Bolker 2022)
Modelling (for model comparison)	`glmmTMB, MASS, rstanarm, glmmADMB, spaMM`	(Brooks et al 2017, Venables & Ripley 2002, Goodrich et al 2022, Brilleman et al 2018, Fournier et al 2012, Skaug et al 2016, Rousset and Ferdy 2014)
Model evaluation	`gstat, performance, DHARMa, car, lme4, lattice, ncf, dismo, AUC, PresenceAbsence`	(Pebesma 2004, Gräler et al 2016, Lüdecke et al 2021, Hartig 2022, Fox and Weisberg 2019, Bates et al 2015, Sarkar 2008, Bjornstad 2022, Hijmans et al 2022, Freeman & Moison 2008, Ballings & Van den Poel 2022)
Model predictions and visualisation	`ggplot2, lattice, sjPlot, effects, terra, metR, tmap, viridis, boot`	(Wickham 2016, Sarkar 2008, Lüdecke 2022, Fox and Weisberg 2019, Hijmans 2023, Campitelli 2021, Tennekes 2018, Garnier et al. 2021, Canty & Ripley 2021)

Table S2

Model covariate key containing all variables downloaded and/or derived, and identifiers used in statistical models.

Variable	Model ID	Units	Source	Other
Sea-surface temperature (mean, std. dev.)	`mean_SST`, `sd_SST`	˚C	ACCESS-S2	
SST anomaly (daily, monthly and seasonal)	`do_anom`, `mo_anom`, `seas_anom`	˚C	ACCESS-S2	Manually derived from SST climatology (1981–2018)
Climatological SST (monthly and seasonal; mean, std. dev.)	`mon_clim_sst`, `seas_clim_sst`, `mon_clim_sst_sd`, `seas_clim_sst_sd`	˚C	ACCESS-S2	Manually derived from SST climatology (1981–2018)
Degree heating days	`DHD`	day	ACCESS-S2	Manually derived from SST climatology (1981–2018)
Heating rate	`heat_rate`	˚C/day	ACCESS-S2	Manually derived from SST climatology (1981–2018)
Distance to EAC core	`dist_core`	km	ACCESS-S2	Manually derived using principal component's analysis. See Methods and Appendix S2.
Sea-surface salinity	`mean_SSS`	psu	ACCESS-S2	
Sea-surface height anomaly	`mean_ssh_corrected`	m	ACCESS-S2	
Meridional velocity	`mean_v`	m/s	ACCESS-S2	
Zonal velocity	`mean_u`	m/s	ACCESS-S2	
Eddy kinetic energy	`mean_eke`	cm^2/s^2	ACCESS-S2	$(u^2 + v^2) / 2$
Current speed	`mean_speed`	m/s	ACCESS-S2	$sqrt(u^2 + v^2)$
Depth (mean, std. dev.)	`mean_depth`, `sd_depth`	m	GEBCO	
Distance to seamount/guyot	`SeamountDistKM`	km	Blue Habitats	Manually derived
Mixed layer depth	`mean_mld1`	m	ACCESS-S2	
Heat content in the upper 300 m	`mean_hc300`	m	ACCESS-S2	

Chlorophyll-*a*	`mean_chla`	Mg m^{-3}	Aqua MODIS	OCI algorithm
Fishing trip identifier	`operation`	-	Direct from industry partners	Random intercept used in the models.

References

Ballings M, Van den Poel D (2022). AUC: Threshold Independent Performance Measures for Probabilistic Classifiers. R package version 0.3.2.

Baston (2022). exactextractr: Fast Extraction from Raster Datasets using Polygons. R package version 0.8.1.

Bates, D., Martin Maechler, Ben Bolker, Steve Walker (2015). Fitting Linear Mixed-Effects Models using lme4. Journal of Statistical Software, 67(1), 1-48.

Bivand, RS., Edzer Pebesma, Virgilio Gomez-Rubio, 2013. Applied spatial data analysis with R, Second edition. Springer, NY.

Bjornstad ON (2022). ncf: Spatial Covariance Functions. R package version 1.3-2.

Bolker B, R Development Core Team (2022). bbmle: Tools for General Maximum Likelihood Estimation. R package version 1.0.25.

Brilleman SL, Crowther MJ, Moreno-Betancur M, Buros Novik J & Wolfe R. Joint longitudinal and time-to-event models via Stan. StanCon 2018. 10-12 Jan 2018. Pacific Grove, CA, USA.

Brooks, ML., Kasper Kristensen, Koen J. van Benthem, Arni Magnusson, Casper W. Berg, Anders Nielsen, Hans J. Skaug, Martin Maechler and Benjamin M. Bolker (2017). glmmTMB Balances Speed and flexibility Among Packages for Zero-inflated Generalized Linear Mixed Modeling. The R Journal, 9(2), 378-400.

Campitelli, E., (2021). metR: Tools for Easier Analysis of Meteorological Fields. R package version 0.13.0.

Canty A and Brian Ripley (2021). boot: Bootstrap R (S-Plus) Functions. R package version 1.3-28.

Fournier DA, Skaug HJ, Ancheta J, Ianelli J, Magnusson A, Maunder M, Nielsen A, Sibert J (2012). “AD Model Builder: using automatic differentiation for statistical inference of Highly parameterized complex nonlinear models.” Optim. Methods Softw., 27, 233-249.

Fox, J. and Sanford Weisberg (2019). An {R} Companion to Applied Regression, Third Edition. Thousand Oaks CA: Sage.

Freeman, E. A. and Moisen, G. (2008). PresenceAbsence: An R Package for Presence-Absence Model Analysis. Journal of Statistical Software, 23(11):1-31.

Garnier, S., Noam Ross, Robert Rudis, Antônio P. Camargo, Marco Sciaini, and Cédric Scherer (2021). Rvision - Colorblind-Friendly Color Maps for R. R package version 0.6.2.

Goodrich B, Gabry J, Ali I & Brilleman S. (2022). rstanarm: Bayesian applied regression modeling via Stan. R package version 2.21.3.

Gräler, B., Edzer Pebesma and Gerard Heuvelink, 2016. Spatio-Temporal Interpolation using gstat. The R Journal 8(1), 204-218

Hartig F (2022). DHARMa: Residual Diagnostics for Hierarchical (Multi-Level / Mixed) Regression models. R package version 0.4.6.

Hijmans RJ, Phillips S, Leathwick J, Elith J (2022). dismo: Species Distribution Modeling. R package version 1.3-9.
Hijmans R (2022). raster: Geographic Data Analysis and Modeling. R package version 3.6-11.
Hijmans R (2022). geosphere: Spherical Trigonometry. R package version 1.5-18.
Hijmans R (2023). terra: Spatial Data Analysis. R package version 1.6-52.
Lüdecke et al., (2021). performance: An R Package for Assessment, Comparison and Testing of Statistical Models. Journal of Open Source Software, 6(60), 3139.
Lüdecke D (2022). sjPlot: Data Visualization for Statistics in Social Science. R package version 2.8.12.
Pebesma, E.J., 2004. Multivariable geostatistics in S: the gstat package. Computers & Geosciences, 30: 683-691.
Pebesma, E.J., R.S. Bivand, 2005. Classes and methods for spatial data in R. R News 5 (2).
Pebesma, E., 2018. Simple Features for R: Standardized Support for Spatial Vector Data. The R Journal 10 (1), 439-446.
Rousset, F. and Jean-Baptiste Ferdy (2014) Testing environmental and genetic effects in the presence of spatial autocorrelation. Ecography 37(8): 781-790.
Sarkar, Deepayan (2008) Lattice: Multivariate Data Visualization with R. Springer, New York. ISBN 978-0-387-75968-5
Skaug H, Fournier D, Bolker B, Magnusson A, Nielsen A (2016-01-19). Generalized Linear Mixed Models using 'AD Model Builder'. R package version 0.8.3.3.
Temple Lang D (2022). RCurl: General Network (HTTP/FTP/...) Client Interface for R. R package version 1.98-1.9.
Tennekes M (2018). "tmap: Thematic Maps in R." _Journal of Statistical Software_, *84*(6), 1-39.
Venables, W. N. & Ripley, B. D. (2002) Modern Applied Statistics with S. Fourth Edition. Springer, New York. ISBN 0-387-95457-0
Wickham. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York, 2016.
Wickham H (2022). _rvest: Easily Harvest (Scrape) Web Pages_. R package version 1.0.3.