Supplementary Material:

**Linking hunting weaponry to attack strategies in sailfish and striped marlin**

Hansen et al. 2019 *Proceedings of the Royal Society B: Biological Sciences*

Introduction

**Sailfish and Striped Marlin**

Striped marlin and sailfish are closely related1 (radiation event starting ~18M years ago2) and overlap considerably in their body length distributions with marlin being substantially heavier for a given length3. The ratios of rostra length to body length are almost identical (30% of body length in sailfish and 27-30% in striped marlin3, see also4, 5) and there is also considerable overlap in prey species. Both predators are highly opportunistic and known to have a large proportion of schooling prey fish in their diet based on stable isotope analysis6.

Methods

**Locating and video recording sailfish and striped marlin predation events**

For both species, sardine schools were discovered opportunistically, most often using feeding seabirds as a guide to the sardine school locations. Upon discovery, the sardine schools were filmed underwater by one to three researchers whilst snorkelling at approximately 2-10 m. Group-hunting sailfish predating upon seven different schools of sardines (*Sardinella aurita*) were filmed (63 min, CASIO EX-FH100, 240fps) in 2012. This was done 15-30 km offshore Northern Quintana Roo in the Gulf of Mexico (N 21° 28.3-41.15’, W 86° 38.41-41.30’). Group-hunting striped marlin were filmed (64 min, GoPro HERO4, 120 fps) 10-30 km offshore Baja California, Mexico (November 2018) (N 24° 54.52-48.5’, W 112° 34.46-23.51’) while predating upon three different schools of sardines (*Sardinops sagax caerula*). Both species were observed preying upon sardines of similar size (approximately 20 cm standard length).

**Markov Chain**

*State definitions*

The states were defined as follows: *approach,* the predator swam at cruising speed towards prey school*; dash*, the predator accelerated directly at prey school and dispersed it*; bill use,* the predator moved its rostrum at prey*; prey contact,* the predator made physical contact with prey*; open mouth* the predator opened its mouth*; ingest,* the predator captured and consumed a prey item*; leave,* the predator swam away from the prey school.

*Model selection*

An analysis based on the Bayesian information criterion (BIC) showed that a first-order model (where each state depends on the immediately preceding state) explained the data substantially better than a zeroth-order model (where the order of states does not contain any information) and a second-order model (where each state depends on the two immediately preceding it). Therefore, we used the first-order model for the main analysis. See below for details.

The Bayesian Information Criterion (BIC) can be used as a consistent estimator of the order *k* of a Markov chain as follows (see, for example, 7):

,

where *d* is the number of states of the Markov chain, the observed sequence of states, and the *k*th order maximum likelihood of the observed sequence.

The results in Table S1 show that for the attack sequences of both the sailfish and the marlin. The differences between the values of the BIC for the different values are very large. A difference of 10 or more is commonly regarded as sufficient for the model selection8.

For the computation of the BIC we concatenated the separate attack sequences of the sailfish and the marlin, respectively, merging the pairs of states *leave approach* into single states. (In the concatenated sequences, it does not make sense to regard *approach* and *leave* as separate states, because *leave* can only be followed by *approach*, and *approach* can only be preceded by *leave*.)

**Supplementary Table S1**. Values of the BIC for different orders of the Markov chains for the striped marlin and the sailfish. For both species, the BIC has its smallest value for *k* = 1.

|  |  |  |
| --- | --- | --- |
| Order of the Markov chain | BIC | |
| Striped Marlin  = 3050 | Sailfish  = 387 |
| 0 | 7190 | 1148 |
| 1 | 3911 | 800 |
| 2 | 4900 | 1662 |
|  |  |  |

**Rostrum use: isolated prey or school of prey**

Calculating exact distances between prey fish was not attempted in this analysis due to inherent difficulties quantifying these distances in 3-dimensional space, therefore only fish clearly isolated from the school by at least 5 body lengths (as judged by the naked eye) were qualitatively classified as isolated (see Supplementary Video 2 for an example).

**Morphological analysis**

*Rostra collection*

Rostra of sailfish and striped marlin were collected by observers embarking on French longliners operating from Réunion Island in the western Indian Ocean.

*Micro-CT details*

The full 55 mm of the rostrum tip was only available for one of the 10 tips, the rest were a fraction of this, the minimum being 30 cm (see Supplementary Table 2 for details). The micro-CTs were performed at the Julius-Wolf Institute, the Charité (Micro- CT Skyscan 1172, Skyscan, Kontich, Belgium), and the Natural History Museum, Berlin (Phoenix nanotome|s, General Electric, Wunstorf), and set to obtain a resolution of 2.5-9.2 μm. The dimension of 55 mm is a constraint of the micro-CT machine at the Julius-Wolf Institute. The data was reconstructed using Nrecon reconstruction package (Skyscan) and datos|x 2.0 (Phoenix GE). The reconstructed volumes were visualized to reveal the structure with CTvox and further processed using Fiji9.

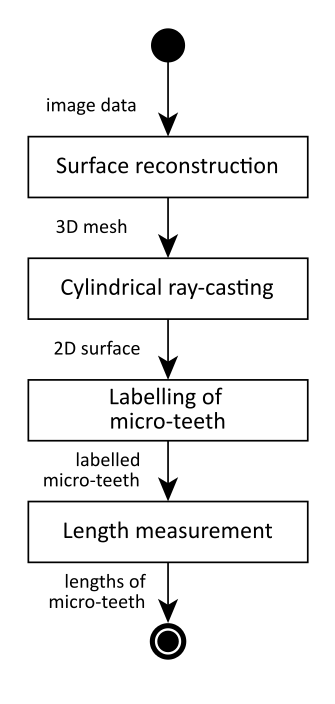
A more detailed investigation of the material was carried out using a scanning electron microscope (SEM) and energy dispersive X-ray spectroscopy (EDX). The SEM analysis confirmed that the caps of micro-teeth are composed of dense mineralised material and the tubules in the micro-teeth are approximately 0.5 µm. The EDX showed that denticle tips were rich in phosphorus and calcium with little carbon (i.e. organic material).

*Micro-tooth identification and quantification*

Image Processing and Labelling:

We developed an image processing and visualization software in Python (version 3.6), primarily based on the Visualization Toolkit10 (version 8.2), which is well suited to process and visualize large-scale 3D data. The automated image processing workflow (Fig S1) was run from the command line and partly from a graphical user interface.

We acquired the image data with two 3D X-ray microscopy systems, the Bruker Skyscan 1172 (6.87 µm resolution) and the GE phoenix nanotom|s (9.17 µm resolution). The 3D image data was stored in archives in Hierarchical Data Format11 (version 5) for convenient handling and as a uniform interface for the image processing workflow.



**Supplementary Fig S1.**Workflow of image processing and labelling.

Surface reconstruction:

The surface reconstruction step produced a 3D mesh and included Gaussian smoothing, noise removal and segmentation. Also, the roll axis of each bill was determined and the 3D mesh axis-aligned.

Cylindrical ray-casting:

The axis-aligned 3D mesh was converted into a two-dimensional regular grid (height map) using ray-casting, where the rays were orthogonal to the roll axis of the bill.

Labelling of micro-teeth:

To find the micro-teeth, the software identified local maxima performing a continuous wavelet transformation (using the Python module scipy.signal13) and applying a hill-climbing algorithm. All identified teeth were labelled with colour marks. The software enabled human observers to visualize the bills to inspect the labelling, if necessary, modify it, and determine the type (intact, broken, or re-growing) of each micro-tooth.

Length measurement:

To measure the length of a micro-tooth, the software calculated its local base plane and computed the perpendicular distance from its tip to the base plane (see Fig S2).



**Supplementary Fig S2.** A.) Scatterplot representing the length of every micro-tooth measured (mm) coloured by species (red: sailfish, blue: striped marlin). The x-axis shows the angular position of the micro-teeth in degrees from the roll-axis of the rostrum. 0°/360° is the dextral side, 90° is the dorsal side, 180° is the sinistral side and 270° is the ventral side. B.) micro-CT image of one sailfish rostrum tip showing lateral teeth at the top and bottom. C.) diagram of bill cross section looking towards head, showing angular positions in A.)

*Statistical analysis*

For each of the four response variables (tooth length, fraction broken, fraction growing, and density), we ran a series of 11 Bayesian mixed-level regression models using the brm function of the brms package14. All models included a group-level intercept for individuals (‘random intercept’). For each of the four response variables, we present the WAIC weights of all 11 models; and the summary statistics of the model with the lowest WAIC weight. Model convergence was evaluated with Rhat and with posterior predictive checks. In all situations, the simulated replicated data from the fitted model were in good agreement with the observed data.

Results

*Capture and rostrum use*



**Supplementary Fig S3.** Rostrum use, capture and success rate data for sailfish and striped marlin. A.) the average amount of bill use per attack sequence, B.) the average capture rate per attack sequence, C.) the average capture rate per minute, D.) the proportion of bill uses that were slashes, E.) the proportion of bill uses that were on isolated prey, and F.) the proportion of bill uses that were in a horizontal motion. Error bars in A, B and C show standard error.

*Morphological analysis*

*Model output and selection*

Tooth length:

|  |  |  |  |
| --- | --- | --- | --- |
|  | Population-level effects | Δ WAIC | se |
| 1 | species + s(angle, bs = "cc", by = species) + dist\*species | 0 | 0 |
| 2 | species + s(angle, bs = "cc", by = species) + dist | 28.94 | 12.55 |
| 3 | species + s(angle, bs = "cc") + dist\* species | 474.09 | 62.85 |
| 4 | species + s(angle, bs = "cc") + dist | 487.85 | 63.63 |
| 5 | s(angle, bs = "cc") + dist | 487.97 | 63.60 |
| 6 | species + dist | 10445.94 | 293.54 |
| 7 | species + s(angle, bs = "cc") | 536.20 | 64.96 |
| 8 | species | 10573.55 | 294.38 |
| 9 | s(angle, bs = "cc") | 535.16 | 65.06 |
| 10 | dist | 10445.46 | 293.53 |
| 11 | 1 | 10573.32 | 294.36 |

**Supplementary Table S3.** WAIC values of 11 models of tooth length. Best performing model is used as reference model (i.e., Δ WAIC = 0). Dist = distance.

**Summary best performing model (Model 1):**

Family: exgaussian

Links: mu = identity; sigma = identity; beta = identity

Formula: length ~ species + s(angle, bs = "cc", by = species) + dist \* species + (1 | id)

Data: DATA (Number of observations: 20365)

Samples: 3 chains, each with iter = 2000; warmup = 1000; thin = 1;

total post-warmup samples = 3000

Smooth Terms:

Estimate Est.Error l-95% CI u-95% CI Eff.Sample Rhat

sds(sanglespeciesSF\_1) 0.07 0.02 0.04 0.12 844 1.00

sds(sanglespeciesSM\_1) 0.08 0.03 0.05 0.15 710 1.00

Group-Level Effects:

~id (Number of levels: 10)

Estimate Est.Error l-95% CI u-95% CI Eff.Sample Rhat

sd(Intercept) 0.04 0.01 0.02 0.07 1097 1.00

Population-Level Effects:

Estimate Est.Error l-95% CI u-95% CI Eff.Sample Rhat

Intercept 0.29 0.02 0.25 0.32 1019 1.00

**speciesSM -0.04 0.03 -0.09 0.02 1272 1.00**

dist -0.00 0.00 -0.00 -0.00 2904 1.00

speciesSM:h 0.00 0.00 0.00 0.00 3158 1.00

Family Specific Parameters:

Estimate Est.Error l-95% CI u-95% CI Eff.Sample Rhat

sigma 0.07 0.00 0.07 0.08 3062 1.00

beta 0.06 0.00 0.06 0.07 3063 1.00

Fraction broken:

|  |  |  |  |
| --- | --- | --- | --- |
|  | Population-level effects | Δ WAIC | se |
| 1 | species + s(angle, bs = "cc", by = species) + s(dist, by =species) | 0 | 0 |
| 2 | species + s(angle, bs = "cc", by = species) + s(dist) | 30.89 | 12.93 |
| 3 | species + s(angle, bs = "cc") + s(dist\* species) | 176.83 | 26.90 |
| 4 | species + s(angle, bs = "cc") + s(dist) | 195.25 | 28.90 |
| 5 | s(angle, bs = "cc") + s(dist) | 195.19 | 28.90 |
| 6 | species + s(dist) | 1159.99 | 70.21 |
| 7 | species + s(angle, bs = "cc") | 596.28 | 50.73 |
| 8 | species | 1541.37 | 78.55 |
| 9 | s(angle, bs = "cc") | 595.06 | 50.72 |
| 10 | s(dist) | 1160.88 | 70.22 |
| 11 | 1 | 1541.01 | 78.55 |

**Supplementary Table S4.** WAIC values of 11 models of fraction broken. Best performing model is used as reference model (i.e., Δ WAIC = 0). Dist = distance.

**Summary best performing model (Model 1):**

Family: bernoulli

Links: mu = logit

Formula: broken ~ species + s(angle, bs = "cc", by = species) + s(dist, by = species) + (1 | id)

Data: DATA (Number of observations: 20365)

Samples: 3 chains, each with iter = 2000; warmup = 1000; thin = 1;

total post-warmup samples = 3000

Smooth Terms:

Estimate Est.Error l-95% CI u-95% CI Eff.Sample Rhat

sds(sanglespeciesSF\_1) 0.44 0.13 0.26 0.77 871 1.00

sds(sanglespeciesSM\_1) 0.45 0.15 0.26 0.84 787 1.00

sds(sdistspeciesSF\_1) 3.00 1.21 1.34 6.06 1132 1.01

sds(sdistspeciesSM\_1) 4.04 2.30 1.06 9.57 578 1.01

Group-Level Effects:

~id (Number of levels: 10)

Estimate Est.Error l-95% CI u-95% CI Eff.Sample Rhat

sd(Intercept) 0.63 0.19 0.37 1.10 1318 1.00

Population-Level Effects:

Estimate Est.Error l-95% CI u-95% CI Eff.Sample Rhat

Intercept -1.24 0.27 -1.75 -0.71 871 1.00

**speciesSM 0.90 0.43 0.04 1.78 955 1.01**

sdist:speciesSF\_1 -1.11 0.34 -1.85 -0.50 2228 1.00

sdist:speciesSM\_1 -1.35 0.83 -3.31 -0.11 1000 1.01

Fraction growing:

|  |  |  |  |
| --- | --- | --- | --- |
|  | Population-level effects | Δ WAIC | se |
| 1 | species + s(angle, bs = "cc",by = species) + s(dist, by = species) | 0 | 0 |
| 2 | species + s(angle, bs = "cc", by = species) + s(dist) | 10.33 | 7.46 |
| 3 | species + s(angle, bs = "cc") + s(dist, by = species) | 13.02 | 12.01 |
| 4 | species + s(angle, bs = "cc") + s(dist) | 23.09 | 13.74 |
| 5 | s(angle, bs = "cc") + s(dist) | 21.06 | 13.53 |
| 6 | species + s(dist) | 535.01 | 51.07 |
| 7 | species + s(angle, bs = "cc") | 60.64 | 19.23 |
| 8 | species | 574.69 | 51.16 |
| 9 | s(angle, bs = "cc") | 57.91 | 18.91 |
| 10 | s(dist) | 533.31 | 51.35 |
| 11 | 1 | 573.85 | 51.43 |

**Supplementary Table S5.** WAIC values of 11 models of fraction growing. Best performing model is used as reference model (i.e., Δ WAIC = 0). Dist = distance.

**Summary best performing model (Model 1):**

Family: bernoulli

Links: mu = logit

Formula: Growing ~ species + s(angle, bs = "cc", by = species) + s(dist, by = species) + (1 | id)

Data: DATA (Number of observations: 20365)

Samples: 3 chains, each with iter = 2000; warmup = 1000; thin = 1;

total post-warmup samples = 3000

Smooth Terms:

Estimate Est.Error l-95% CI u-95% CI Eff.Sample Rhat

sds(sanglespeciesSF\_1) 0.91 0.28 0.52 1.63 719 1.00

sds(sanglespeciesSM\_1) 0.60 0.43 0.08 1.74 557 1.00

sds(sdistspeciesSF\_1) 1.55 1.41 0.08 5.34 716 1.00

sds(sdistspeciesSM\_1) 10.08 6.39 2.61 26.80 960 1.00

Group-Level Effects:

~id (Number of levels: 10)

Estimate Est.Error l-95% CI u-95% CI Eff.Sample Rhat

sd(Intercept) 0.73 0.36 0.33 1.66 642 1.00

Population-Level Effects:

Estimate Est.Error l-95% CI u-95% CI Eff.Sample Rhat

Intercept -4.27 0.34 -4.95 -3.58 774 1.00

**speciesSM -2.23 0.69 -3.71 -0.98 1046 1.00**

sdist:speciesSF\_1 -0.28 0.32 -1.01 0.36 1427 1.00

sdist:speciesSM\_1 -1.75 2.20 -6.71 2.35 1238 1.00

Density:

|  |  |  |  |
| --- | --- | --- | --- |
|  | Population-level effects | Δ WAIC | se |
| 1 | species + zone\*species + s(section\*species) | 3.40 | 6.52 |
| 2 | species + zone\*species + s(section) | 0 | 0 |
| 3 | species + zone + s(section\* species) | 46.62 | 20.13 |
| 4 | species + zone + s(section) | 31.52 | 24.66 |
| 5 | zone + s(section) | 29.49 | 25.62 |
| 6 | species + zone | 161.86 | 29.55 |
| 7 | species + s(section) | 20.25 | 29.23 |
| 8 | species | 162.59 | 30.65 |
| 9 | s(section) | 19.86 | 29.61 |
| 10 | zone | 162.00 | 30.65 |
| 11 | 1 | 163.48 | 30.80 |

**Supplementary Table S6.** WAIC values of 11 models of tooth density. Best performing model is used as reference model (i.e., Δ WAIC = 0).

**Summary best performing model (Model 2):**

Family: exgaussian

Links: mu = identity; sigma = identity; beta = identity

Formula: density ~ species + zone + s(section) + zone \* species + (1 | id)

Data: DF (Number of observations: 375)

Samples: 3 chains, each with iter = 4000; warmup = 1000; thin = 1;

total post-warmup samples = 9000

Smooth Terms:

Estimate Est.Error l-95% CI u-95% CI Eff.Sample Rhat

sds(ssection\_1) 4.82 1.75 2.38 8.95 2981 1.00

Group-Level Effects:

~id (Number of levels: 11)

Estimate Est.Error l-95% CI u-95% CI Eff.Sample Rhat

sd(Intercept) 0.50 0.16 0.28 0.89 3085 1.00

Population-Level Effects:

Estimate Est.Error l-95% CI u-95% CI Eff.Sample Rhat

Intercept 2.34 0.24 1.84 2.81 3633 1.00

**speciesSM -0.11 0.38 -0.85 0.64 3565 1.00**

zonedorsal 1.53 0.29 0.99 2.14 4109 1.00

zonesinistral -0.10 0.15 -0.39 0.19 7536 1.00

zoneventral 1.00 0.18 0.65 1.36 5784 1.00

speciesSM:zonedorsal -2.31 0.34 -2.97 -1.66 5223 1.00

speciesSM:zonesinistral0.04 0.26 -0.47 0.54 6609 1.00

speciesSM:zoneventral -1.36 0.30 -1.93 -0.77 6753 1.00

ssection\_1 -1.26 0.76 -2.84 0.15 4079 1.00

Family Specific Parameters:

Estimate Est.Error l-95% CI u-95% CI Eff.Sample Rhat

sigma 0.49 0.08 0.35 0.66 4428 1.00

beta 1.18 0.11 0.97 1.38 5609 1.00

**References**

1. Collette BB, McDowell JR & Graves JE 2006. Phylogeny of recent billfishes (*Xiphioidei*). Bulletin of Marin Science 79: 455-468.
2. Santini F & Sorenson L 2013. First molecular timetree of billfishes (Istiophoriformes: Acanthomorpha) shows a Late Miocene radiation of marlins and allies. Italian Journal of Zoology. 80: 481-489.
3. Nakamura, I. 1983. Systematics of the billfishes (Xiphiidae and Istiophoridae). Publ. Seto Mar. Biol. Lab. 28: 255–396.
4. Domenici P, Wilson ADM, Kurvers RHJM, Marras S, Herbert-Read JE, Steffensen JF, Krause S, Viblanc PE, Couillaud P & Krause J 2014. How sailfish use their bills to capture schooling prey. Proceedings of the Royal Society London B 281: 20140444.
5. Fierstine H.L. & Voigt N.L. 1996 Use of rostra characters for identifying adult billfishes (Teleostei: Perciformes: *Istiophoridae* and *Xiphiidae*). Copeia 1996: 148-161.
6. Tsaia C-N, Chiangb W-C, Suna C-L, Shaoc K-T, Chend S-Y & Yeh S-Z 2015. Stomach content and stable isotope analysis of sailfish (*Istiophorus platypterus*) diet in eastern Taiwan waters. Fisheries Research 166: 39–46.
7. Csiszár I, Shields PC. The consistency of the BIC Markov order estimator. The Annals of Statistics. 2000; vol. 28(6), 1601–1619.
8. Guttorp P. Stochastic Modeling of Scientific Data, Chapman and Hall/CRC, 1995.
9. Schindelin, J., Arganda-Carreras, I. & Frise, E. et al. (2012), "Fiji: an open-source platform for biological-image analysis", Nature methods 9(7): 676-682) to produce 3D renderings, longitudinal and cross-sectional slices.
10. Schroeder W, Martin K, Lorensen B. 2006. The Visualization Toolkit (4th ed.). Kitware. ISBN 978-1-930934-19-1
11. The HDF Group. 1997-2019. Hierarchical Data Format, version 5. <http://www.hdfgroup.org/HDF5>/
12. Jones E, Oliphant E, Peterson P, et al. 2001 – 2019. SciPy: Open Source Scientific Tools for Python. <http://www.scipy.org/>
13. Bürkner, P.-C. (2017). brms: An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software, 80*(1), 1-28.