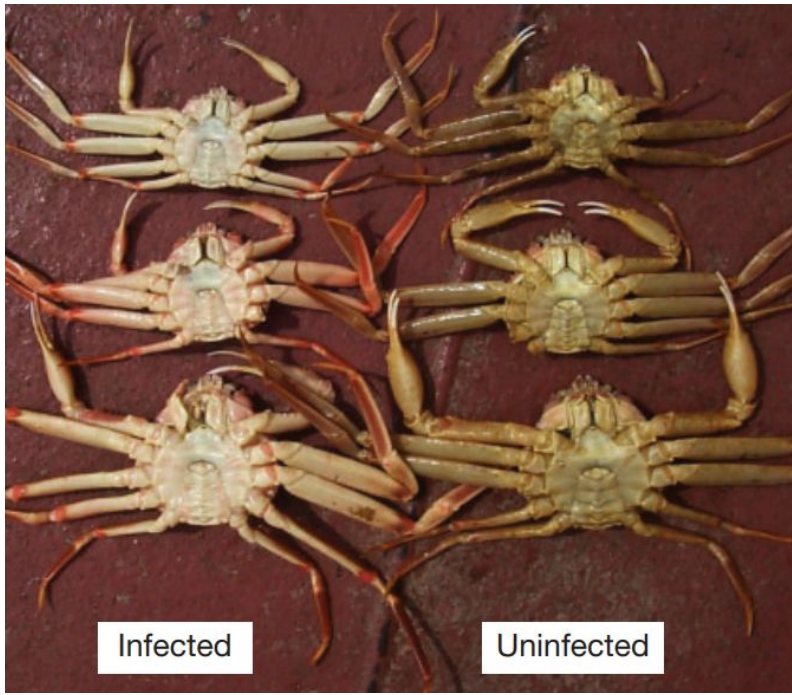
Modeling *Hematodinium sp.* infection in Alaskan Tanner crab

Aidan Coyle

Tanner crab and *Hematodinium sp.*

- Parasitic dinoflagellate
- Infection route unknown, likely waterborne
  - Are molting or wounded crabs more vulnerable?
  - Are crabs more vulnerable when small?
  - Is infection more likely at particular times of year?
- Infection != visual positive



Data Set

- 6 annual Tanner crab pot surveys in Southeast Alaska
  - 2007-2012
- Bernoulli distribution - visible infection by *Hematodinium*
- Chosen variables to examine

Continuous	**Categorical**
Year	*Location*
Carapace width	Sex
Julian day (Feb 1 = 32)	Shell condition
Pot depth	Black Mat disease
	Missing legs?

Cleaning, Data Checks, and Other Modeling Prep

- Removed all lines with NAs + data entry errors + irrelevant data
  - ~14,500 remaining lines
- Checked for correlation between variables
  - Continuous vs. continuous: Pearson's test
  - Categorical vs. categorical: Cramer's V test
  - Continuous vs. categorical: Spearman rank-order correlation
- Found correlation between sex and carapace width
- Will be building GLMM with multiple random effects

Model Building

- Problem: Model is too big to just run using Laplace approximation
  - 1: Scale all continuous variables
  - 2: Build Laplace models with one fixed effect + random effects
  - 3: Build small Laplace model using most important variables (step 2), slowly add in further variables with update()
- 2 “full models”: Tested with AIC (plus all intermediate models)
  - 1: Carapace width (all except sex and missing legs)
  - 2: Sex (all except carapace width and missing legs)

```
lap_CW_mod <- update(lap_four_mod, PARASITE_CODE ~
BLACKMAT_CODE + SHELL_CONDITION_CODE + DEPTH_SCALED +
DAY_SCALED + WIDTH_SCALED + (1 | LOCATION_CODE) + (1 |
s.YEAR))
```

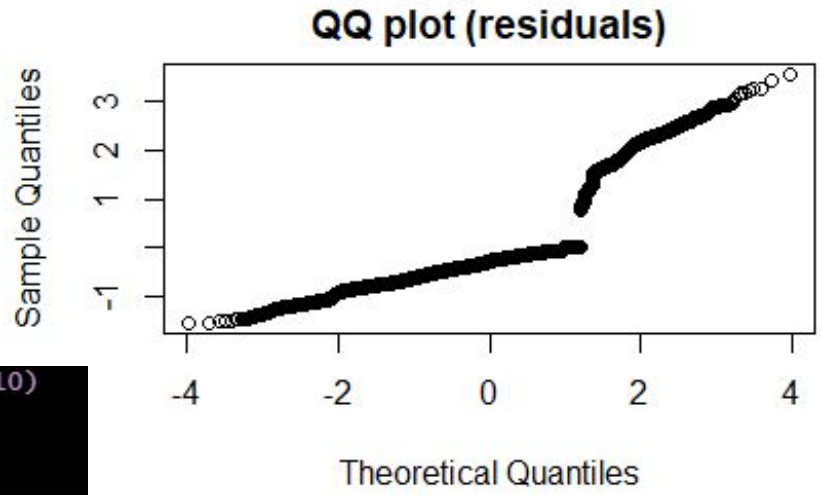
Model Selection and Results

- Used MuMIn::dredge()
  - Tests all combinations of fixed effects in model
- 2 models with weights > 0.001
  - 1 (wt = 0.72): Doesn't include Julian day
  - 2 (wt = 0.28): Includes Julian day
- Averaged with MuMIn::model.avg()

```
Model-averaged coefficients:
(full average)
                      Estimate Std. Error Adjusted SE z value Pr(>|z|)
(Intercept)          -10.65688   11.85864    11.85963   0.899    0.369
BLACKMAT_CODE.L      -10.12176   16.77164    16.77304   0.603    0.546
DEPTH_SCALED          -0.50887    0.05463     0.05464   9.313   <2e-16 ***
SHELL_CONDITION_CODE.L -1.54969    0.26261     0.26263   5.901   <2e-16 ***
SHELL_CONDITION_CODE.Q -1.68320    0.19912     0.19914   8.453   <2e-16 ***
SHELL_CONDITION_CODE.C  0.64983    0.10591     0.10592   6.135   <2e-16 ***
WIDTH_SCALED            0.20714    0.03107     0.03108   6.665   <2e-16 ***
DAY_SCALED             -0.00975    0.05200     0.05200   0.187    0.851
```

Model Diagnostics

- Chi-square goodness of fit test: PASSED
- Q-Q plot of residuals: OK
- Hosmer-Lemeshow test: FAILED
  - Variation that isn't accounted for



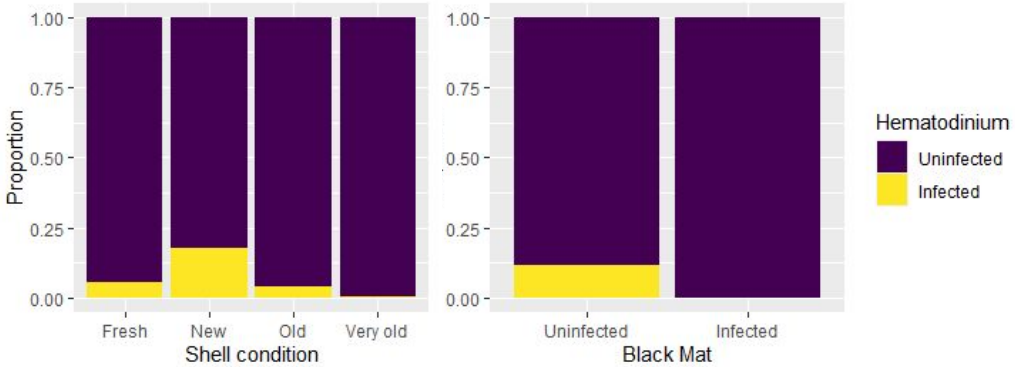

```
> logitgof(crabdat$PARASITE_CODE, fitted(first_mod), g = 10)

        Hosmer and Lemeshow test (binary model)

data:  crabdat$PARASITE_CODE, fitted(first_mod)
X-squared = 52.459, df = 8, p-value = 1.372e-08
```

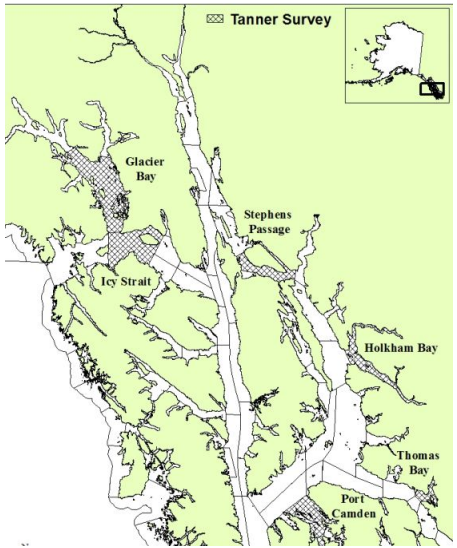
Discussion

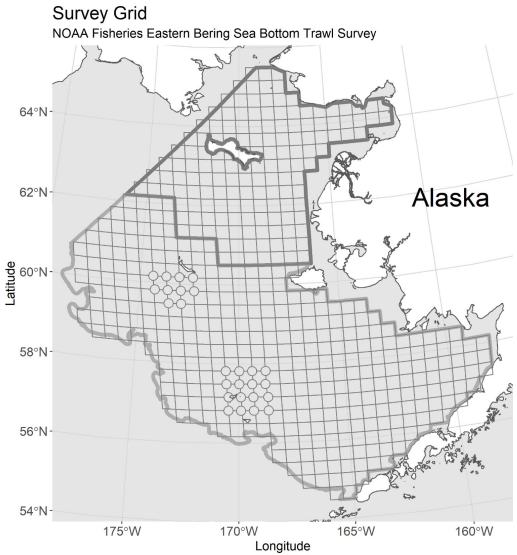
- Unaccounted variance - missing predictor(s)
- Incomplete model analysis:
  - Shell condition: Shows linkage to molt time (but not lost legs)
  - Depth: Movement into shallow waters post-infection
  - Carapace width: Larger crab more likely to be infected
  - Black mat: Not significant, but intriguing
    - 3rd-most important by AIC (after shell condition and carapace width)
    - Likely not enough sampled with Black Mat



Next Steps

- See if similar model works on NOAA EBS trawl data
- Possible future work
  - Expose freshly molted crab to *Hematodinium*
  - Investigate potential exclusion of *Hematodinium* by Black Mat fungus






Thank you!