**AFSC Pathobiology Historical Disease Data Rescue**

**Data Summary:** The original Pathobiology Access database holds collection data and results from histology, blood smears and PCR assays from fish, invertebrates, mammals, plankton, and environmental samples collected by the AFSC Pathobiology group from 1978 through 2013. In July, 2022 the database was subset for majid, cancrid and lithodid crustaceans only (BCS Master Data.csv) and moved to a github repository for collaborative data exploration. Preliminary plots were output to assess spatial coverage and sample sizes of bitter crab syndrome (BCS) collections (i.e. blood smear or PCR assay) by Large Marine Ecosystem. Summaries for each region are provided below along with any details on sampling protocol/design, if known. **NOTE:** Prior to 1988, only histology samples were collected (i.e. no BCS diagnosis available).

**GOA/SEAK Samples:**

Coverage: Majority *C. bairdi,* 1990/1993/2003-2012, Alaska Peninsula/Kodiak Island/Southeast AK

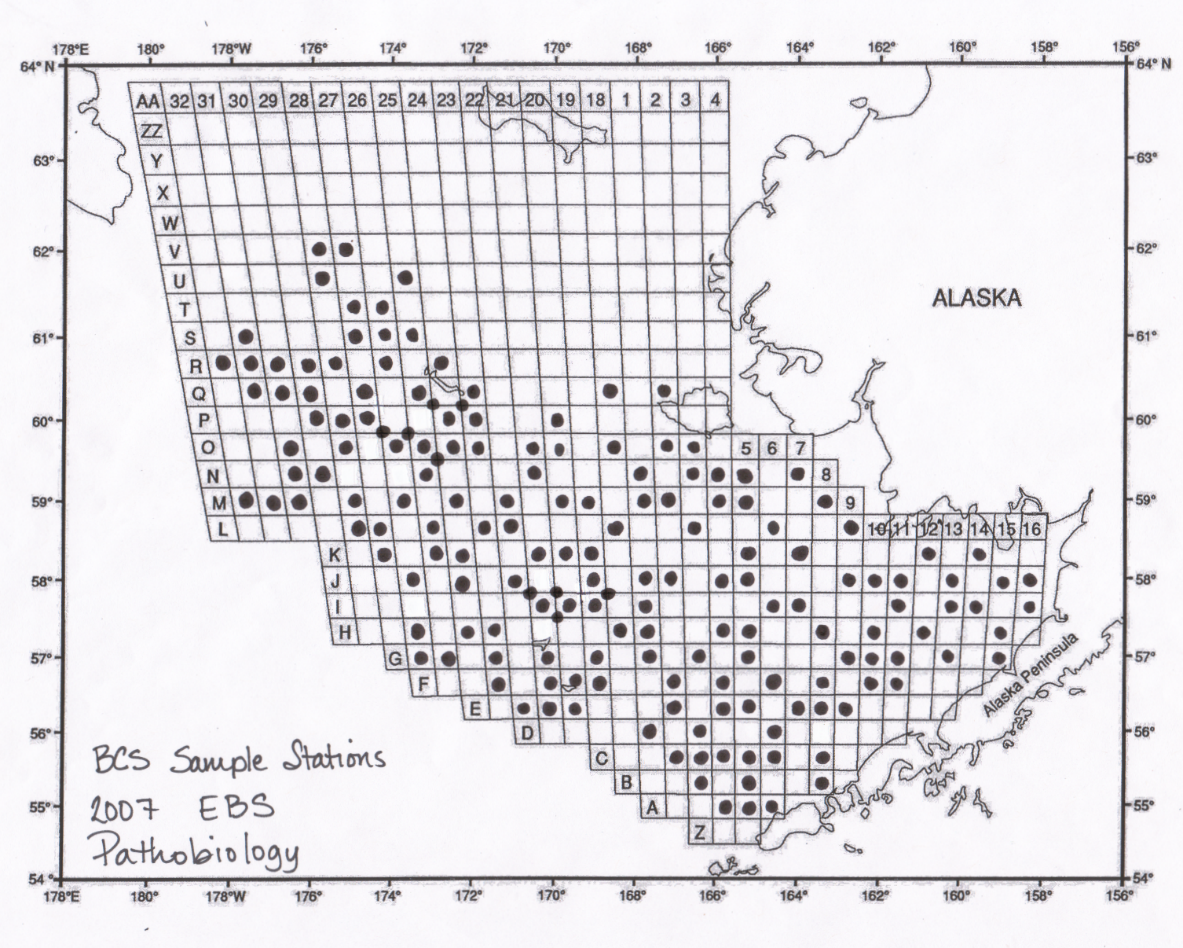
Sampling Design: SEAK samples (2003-2012) were collected on the ADF&G SE Tanner Crab survey; no sampling protocols were found. Alaska Peninsula samples appear to have been collected on the NMFS GOA BT survey (2003/2005) and the ADF&G large-mesh bottom trawl survey (2007-2010); no sampling protocols found

**Bering Sea Samples:**

Coverage: Primarily *C. bairdi and C. opilio*, Eastern Bering Sea shelf (1988-2013) and slope (1988/1991/2000/2014)

Sampling design: The majority of the samples were collected on the NMFS ESB bottom trawl survey. Special project sampling protocols exist for the EBS shelf bottom trawl survey (2007-2012) and Bering Sea slope survey (2014). The protocol designates to collect hemolymph from 10-20 crabs of the dominant *Chionoecetes* species at each station completed by the vessel with Pathobiology staff (goal to sample approximately half of the 250 pre-selected stations; see map below; ONE boat only). In 2009 and 2011 only, the protocol designates to collect hemolymph from the following species when available: *Paralithodes spp., Hyas spp., Pagurus spp., E. isenbeckii, Elassochirus spp., shrimp.* In 2014, BCS index sites for each species (10 stations per site) were adopted. These index site data are NOT included in the historical disease dataset.

*To follow up on: Did pre-2007 collections sample the same stations designated in protocols from 2007 onwards? 2014 slope data not in database? Also there appear to be a small number of samples from outside the NMSF survey, may need to filter these out*



**Arctic Samples:**

Coverage: Majority *C. opilio*, Northern Bering Sea (1988/1991/1994/2001/2005/2006/2010), Chukchi Sea (1990, 2012-2013) and Beaufort Sea (2008 only)

Sampling design: Chukchi Sea samples (2012-2013) and Beaufort Sea samples (2008) were collected on NMFS bottom trawl surveys via a stratified sampling plan; no sampling protocols found. The majority of Northern Bering Sea samples appear to have been sampled within the NMFS bottom trawl survey area but the cruise is unclear; no sampling protocols found.

*To follow up on: Some stations from the EBS shelf grid are assigned to the NBS LME. Need to index via station look-up table and re-assign if filtering by LME.*

Additional LME’s not included in initial data exploration due to small sample sizes/geographical relevance include the Aleutian Islands and Atlantic Coast/Newfoundland.

**Potential concerns of study designs detailed above for use in spatial models:**

  For regions/collections where sampling protocols were not located, it’s impossible to know what portion of the population was targeted and how sample sites were selected. While I have no reason to believe that sampling was non-random, index sites were designed in 2014 to target known areas of high prevalence….hence a bit of doubt.

2)      Sampling date is highly correlated with prevalence (i.e. as the disease progresses throughout the year, diagnostic assays and smears are more likely to pick up infections). Thus, infection status is confounded with sampling date within a given year, across years, and across regions where sampling dates differ.

3)      250 pre-selected stations within the EBS were only sampled with one boat. This means that stations may have been sampled in a non-random way, depending on how the vessels split up the grid within a given year?

4)      At or around 2003 in all datasets, there was a transition from blood smears to PCR assays to diagnose BCS infection status. While we have data in hand to compare the sensitivity of the two methods, prevalence rates between the two methods are likely not directly comparable as is.

5) Crab size is highly correlated with crab size (i.e. small, immature crab appear more likely to have BCS). With no size/maturity targets for sampling, if mature crab are sampled more frequently than immatures, prevalence will likely be underestimated.

6) Bias towards sampling most prevalent species in the catch? Although bairdi/opilio ranges differ, stations where they are both caught might non-randomly misrepresent one species