Kelp Wrack Metrics

Data collected at Bandon Beach from 2018-2021

David Bilderback, Oregon Institute of Marine Biology

Ross Whippo, Oregon Institute of Marine Biology

Created: 2022-04-01 ; (Updated: 2023-09-23)

Table of Contents

# Summary

Hi Dave, this is as far as I’ve gotten with the analysis so far and I thought you might like to take a look at it. I have a few questions about the data set itself and would appreciate feedback on the following issues:

I am also curious as to the status of the ‘mature sporophyte’ data set. Do you have additional entries for that as well? Would you like me to combine them into a single dataset? It’s fairly easy for me to put them together, and then take apart during analysis, as-needed.

Overall, this is a very impressive and exciting data set. I’m particularly interested in digging into the recruitment question that is represented in the final figure. It would be useful to generate a model that back-calculates other demographics of the local population, like spore release and settlement. Perhaps relating the sporophyll and recruitment data would be a neat way to do it (all within a single set of data, no less!).

I have other questions that came up over the course of this and previous analyses, and I will work to add more to this document. If you have any particular thoughts on what I’ve presented here, do let me know. I will send you updates as I progress in my analyses.

# Methods

## Field Surveys

Dave, can you provide me a comprehensive description of your field methods? I have what I pulled from the metadata below, but it would be great to have more detail about the specific location, the scope of the transects, what time of day they were conducted, any details you may have about how high on the tideline you looked for sporophytes, the general condition of the sporophytes (were they dessicated? fresh? both? neither? Did you treat them differently depending on their condition [beyond being ‘intact’]?), did you ever subset if there were too many on the beach, and any other pertinent information regarding your collections.

All intact, individual young sporophytes and clusters of young sporophytes with holdfasts, stipes, bulbs and blades were collected and brought to the lab. The diameter of some holdfasts were measured to the nearest mm. Because of the elasticity of the stipe, its length was measured to the nearest 0.5 cm. The diameter of the bulb was measured with a caliper to the nearest mm. The width of some of the widest blades were measured to the nearest mm.

## Data Analysis

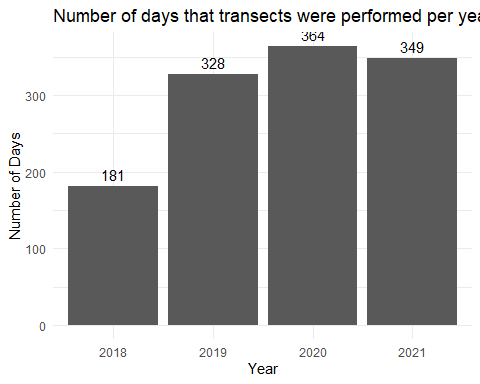
GOES HERE.

# Results

# Import all data and manage for analyses/visualizations  
  
# read in young dataset  
young\_sporophyte <- read\_csv("Data/young\_sporophyte\_data.csv",   
 col\_types = cols(Date = col\_character(),   
 Stipe = col\_number(),  
 Bulb = col\_number(),   
 Blade = col\_double(),   
 `Ho Fa` = col\_double(),   
 Single = col\_character(),   
 Group = col\_character()))  
  
# fix erroneous date  
young\_sporophyte$Date <- young\_sporophyte$Date %>%  
 recode("11/28/2028" = "11/28/2018",  
 "10/9/2028" = "10/9/2018",  
 "6/17/2028" = "6/17/2018",  
 "56/2/2021" = "5/26/2021")  
young\_sporophyte$Date <- mdy(young\_sporophyte$Date)  
# fix bulb diameter and blade width typos  
young\_sporophyte['Bulb'][young\_sporophyte['Bulb'] == 53.0] <- 5.3  
young\_sporophyte['Bulb'][young\_sporophyte['Bulb'] == 30.0] <- 3.0  
young\_sporophyte['Blade'][young\_sporophyte['Blade'] == 22.0] <- 2.2  
# fix substrate typos  
young\_sporophyte$Subst <- young\_sporophyte$Subst %>%  
 recode("NL" = "Nl",  
 "Ni" = "Nl",  
 "MY" = "My",  
 "R" = "Ro",  
 "B" = "Ba",  
 "Bs" = "Ba",  
 "Bo" = "Ba",  
 "Bl" = "Nl",  
 "LS" = "Ls",  
 "SS" = "Ss",  
 "Hy" = "Hf",  
 "Mp" = "Mu",  
 "Dl" = "Do")  
  
  
# fix CoSp1 typos  
young\_sporophyte$CoSp1 <- young\_sporophyte$CoSp1 %>%  
 recode("LS" = "Ls",  
 "Po" = "Pl",  
 "Cor" = "Co",  
 "Io" = "Is",  
 "Tu" = "none",  
 "ls" = "Ls",  
 "Ba" = "Bo",  
 "By" = "Bo")  
  
# substrate type dataset  
substrateData <- read\_csv("Data/substrateCodes.csv")  
  
# split columns add none value  
substrate <- substrateData %>%  
 separate(`Subst=Substrate`, into = c("Subst", "Substrate"), sep = "=") %>%  
 bind\_rows(c(Subst = "none", Substrate = "none"))  
  
  
  
# cospecies dataset  
cospeciesData <- read\_csv("Data/cospeciesCodes.csv")  
  
# split columns add none and duplicate columns  
cospecies <- cospeciesData %>%  
 separate(`Cosp = Cospecies`, into = c("Cosp", "Cospecies"), sep = " = ") %>%  
 bind\_rows(c(Cosp = "none", Cospecies = "none")) %>%  
 mutate(CoSp1 = Cosp,  
 CoSp2 = Cosp,  
 CoSp3 = Cosp)

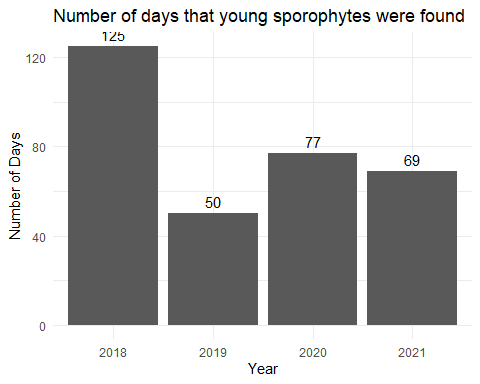
## For each of the four years, how many days was the beach transect surveyed for young sporophytes?

young\_sporophyte\_q1 <- young\_sporophyte %>%  
 separate(Date, c("year", "month", "day"), sep = "-") %>%  
 unite("month-day", month:day, remove = FALSE) %>%   
 select(year, "month-day") %>%  
 group\_by(year) %>%  
 summarise(days = length(unique(`month-day`)))  
  
ggplot(young\_sporophyte\_q1, aes(x = year, y = days)) +  
 geom\_col() +  
 theme\_minimal() +  
 labs(y = "Number of Days", x = "Year", title = "Number of days that transects were performed per year") +  
 geom\_text(aes(label = days), vjust = -0.5) +  
 labs(title = "Number of days that transects were performed per year")



## For each of the four years how many days were young sporophytes found on the transect?

young\_sporophyte\_q2 <- young\_sporophyte %>%  
 separate(Date, c("year", "month", "day"), sep = "-") %>%  
 unite("month-day", month:day, remove = FALSE) %>%   
 filter(!is.na(Stipe)) %>%  
 select(year, "month-day") %>%  
 group\_by(year) %>%  
 summarise(days = length(unique(`month-day`)))  
  
ggplot(young\_sporophyte\_q2, aes(x = year, y = days)) +  
 geom\_col() +  
 theme\_minimal() +  
 labs(y = "Number of Days", x = "Year", title = "Number of days that young sporophytes were found") +  
 geom\_text(aes(label = days), vjust = -0.5)



## Is there a relationship between stipe length and bulb diameter?

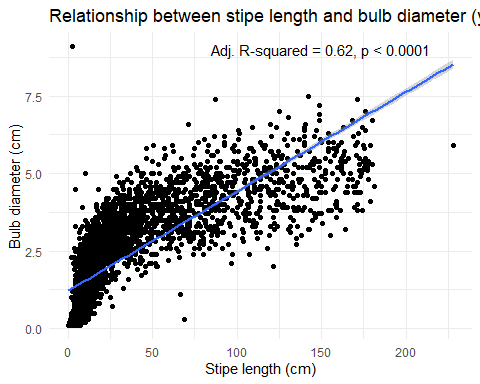
fit1 <- lm(Bulb ~ Stipe, data = young\_sporophyte)  
fit2 <- lm(Bulb~poly(Stipe,2,raw=TRUE), data=young\_sporophyte)  
fit3 <- lm(Bulb~poly(Stipe,3,raw=TRUE), data=young\_sporophyte)  
fit4 <- lm(Bulb~poly(Stipe,4,raw=TRUE), data=young\_sporophyte)  
fit5 <- lm(Bulb~poly(Stipe,5,raw=TRUE), data=young\_sporophyte)  
  
AIC(fit1, fit2, fit3, fit4, fit5)

## df AIC  
## fit1 3 13085.093  
## fit2 4 11202.472  
## fit3 5 10084.824  
## fit4 6 9605.674  
## fit5 7 9413.794

summary(fit1)

##   
## Call:  
## lm(formula = Bulb ~ Stipe, data = young\_sporophyte)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -3.2761 -0.7953 -0.0011 0.6495 7.7899   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 1.2204165 0.0169020 72.2 <2e-16 \*\*\*  
## Stipe 0.0320471 0.0003605 88.9 <2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.9352 on 4836 degrees of freedom  
## (2122 observations deleted due to missingness)  
## Multiple R-squared: 0.6204, Adjusted R-squared: 0.6203   
## F-statistic: 7903 on 1 and 4836 DF, p-value: < 2.2e-16

ggplot(young\_sporophyte, aes(x = Stipe, y = Bulb, na.rm = TRUE)) +  
 geom\_point() +  
 geom\_smooth(method = "lm") +  
 theme\_minimal() +  
 labs(y = "Bulb diameter (cm)", x = "Stipe length (cm)", title = "Relationship between stipe length and bulb diameter (young sporophytes)") +   
 annotate("text", x = 150, y=9, label = "Adj. R-squared = 0.62, p < 0.0001")



## Is there a relationship between bulb diameter and widest blade width?

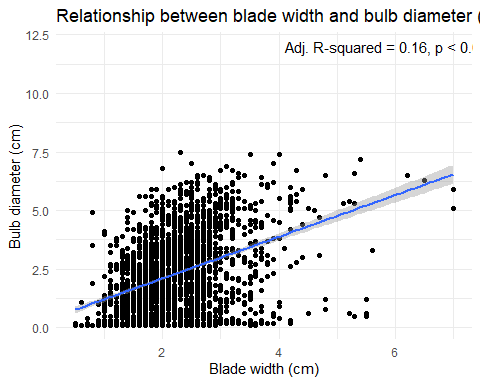
fit1 <- lm(Bulb ~ Blade, data = young\_sporophyte)  
fit2 <- lm(Bulb~poly(Blade,2,raw=TRUE), data=young\_sporophyte)  
fit3 <- lm(Bulb~poly(Blade,3,raw=TRUE), data=young\_sporophyte)  
fit4 <- lm(Bulb~poly(Blade,4,raw=TRUE), data=young\_sporophyte)  
fit5 <- lm(Bulb~poly(Blade,5,raw=TRUE), data=young\_sporophyte)  
  
AIC(fit1, fit2, fit3, fit4, fit5)

## df AIC  
## fit1 3 9127.266  
## fit2 4 9090.172  
## fit3 5 9081.677  
## fit4 6 9080.884  
## fit5 7 9081.510

summary(fit1)

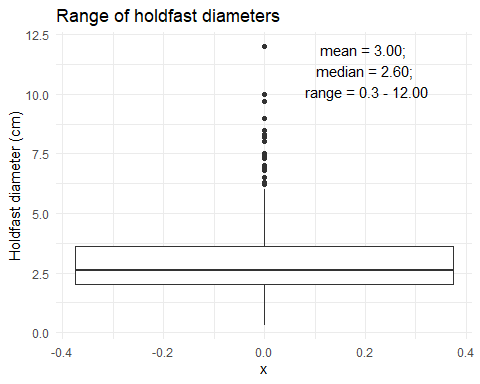
##   
## Call:  
## lm(formula = Bulb ~ Blade, data = young\_sporophyte)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -4.7133 -1.2024 -0.1246 1.1309 5.1310   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 0.32473 0.09518 3.412 0.000656 \*\*\*  
## Blade 0.88882 0.04144 21.447 < 2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 1.515 on 2485 degrees of freedom  
## (4473 observations deleted due to missingness)  
## Multiple R-squared: 0.1562, Adjusted R-squared: 0.1559   
## F-statistic: 460 on 1 and 2485 DF, p-value: < 2.2e-16

ggplot(young\_sporophyte, aes(x = Blade, y = Bulb, na.rm = TRUE)) +  
 geom\_point() +  
 geom\_smooth(method = "lm") +  
 theme\_minimal() +  
 labs(y = "Bulb diameter (cm)", x = "Blade width (cm)", title = "Relationship between blade width and bulb diameter (young sporophytes)") +   
 annotate("text", x = 6, y=12, label = "Adj. R-squared = 0.16, p < 0.0001")



## What is the mean diameter of holdfasts?

ggplot(young\_sporophyte, aes(y = `Ho Fa`)) +  
 geom\_boxplot() +  
 theme\_minimal() +  
 labs(y = "Holdfast diameter (cm)", title = "Range of holdfast diameters") +  
 annotate("text", x = 0.2, y = 11, label = "mean = 3.00; \n median = 2.60; \n range = 0.3 - 12.00")

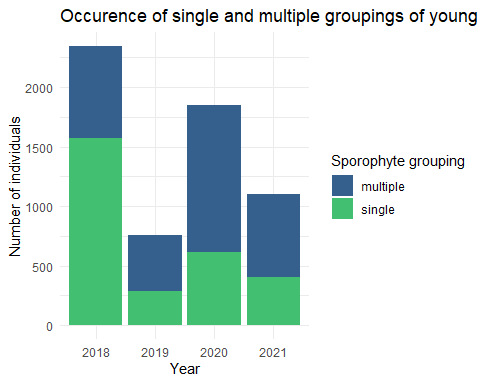


## For each year, how many sporophytes were singular, and how many were clusters?

young\_sporophyte\_q3 <- young\_sporophyte %>%  
 separate(Date, c("year", "month", "day"), sep = "-") %>%  
 unite("month-day", month:day, remove = FALSE) %>%   
 filter(Single %in% c("1", "0")) %>%  
 mutate(Single = case\_when(  
 Single == "1" ~ "single",  
 Single == "0" ~ "multiple")) %>%  
 group\_by(year)  
  
young\_sporophyte\_q3 %>%  
 group\_by(year, Single) %>%  
 summarise(length(year))

## # A tibble: 8 × 3  
## # Groups: year [4]  
## year Single `length(year)`  
## <chr> <chr> <int>  
## 1 2018 multiple 774  
## 2 2018 single 1571  
## 3 2019 multiple 464  
## 4 2019 single 290  
## 5 2020 multiple 1234  
## 6 2020 single 618  
## 7 2021 multiple 699  
## 8 2021 single 401

ggplot(young\_sporophyte\_q3, aes(x = year, label = Single)) +  
 geom\_bar(aes(fill = Single)) +  
 theme\_minimal() +  
 scale\_fill\_viridis(discrete = TRUE, option = "D", begin = 0.3, end = 0.7) +  
 labs(title = "Occurence of single and multiple groupings of young sporophytes", fill = "Sporophyte grouping", x = "Year", y = "Number of individuals")

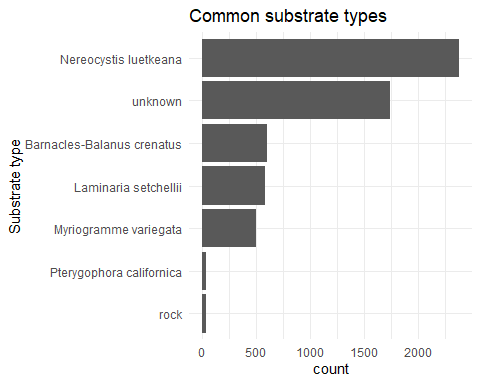


## What is the count and percentage of each substrate type?

# THIS COULD BE A GREAT FIGURE IF ALL OTHER DATA ADDED IN - INFOGRAPHIC STYLE  
  
young\_sporophyte\_q4 <- young\_sporophyte %>%  
 mutate(Subst = na\_if(Subst, "nd")) %>%  
 mutate(Subst = na\_if(Subst, "N A")) %>%  
 mutate(Subst = na\_if(Subst, "Na")) %>%  
 filter(!is.na(Subst)) %>%  
 left\_join(substrate, by = "Subst")  
  
young\_sporophyte\_q4 %>%  
 group\_by(Substrate) %>%  
 summarise('total' = length(Substrate)) %>%  
 arrange(desc(total)) %>%  
 mutate(percent = (total/sum(total))\*100)

## # A tibble: 30 × 3  
## Substrate total percent  
## <chr> <int> <dbl>  
## 1 Nereocystis luetkeana 2380 39.8   
## 2 unknown 1736 29.0   
## 3 Barnacles-Balanus crenatus 605 10.1   
## 4 Laminaria setchellii 587 9.81   
## 5 Myriogramme variegata 502 8.39   
## 6 Pterygophora californica 38 0.635  
## 7 rock 35 0.585  
## 8 shell 20 0.334  
## 9 mussel-Mytilus califorianus 14 0.234  
## 10 Cystoseira osmundacea 9 0.150  
## # ℹ 20 more rows

ggplot(young\_sporophyte\_q4 %>%  
 group\_by(Substrate) %>%  
 summarise(count = length(Substrate)) %>%  
 filter(count > 20) %>%  
 arrange(desc(count)), aes(fct\_reorder(Substrate, count), count)) +  
 geom\_col() +  
 theme\_minimal() +  
 labs(title = "Common substrate types", x = "Substrate type", "Number of occurrences") +  
 coord\_flip()

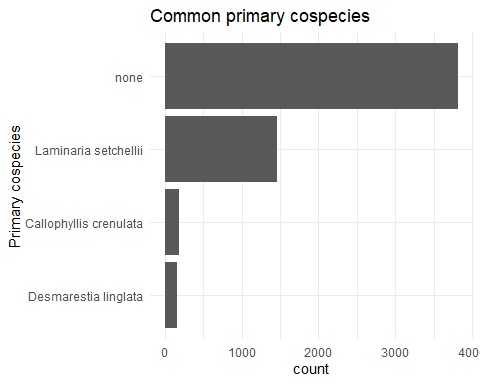


## How common were cospecies? Numbers and percentages.

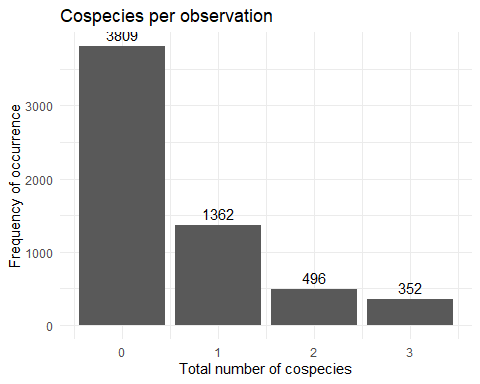
# primary  
  
CoSp1\_data <- cospecies %>%  
 select(CoSp1, Cospecies)  
  
young\_sporophyte\_q8 <- young\_sporophyte %>%  
 mutate(CoSp1 = na\_if(CoSp1, "nd")) %>%  
 mutate(CoSp1 = na\_if(CoSp1, "N A")) %>%  
 mutate(CoSp1 = na\_if(CoSp1, "Na")) %>%  
 filter(!is.na(CoSp1)) %>%  
 left\_join(CoSp1\_data, by = "CoSp1")  
  
young\_sporophyte\_q8 %>%  
 group\_by(Cospecies) %>%  
 summarise('total' = length(Cospecies)) %>%  
 arrange(desc(total)) %>%  
 mutate(percent = (total/sum(total))\*100)

## # A tibble: 33 × 3  
## Cospecies total percent  
## <chr> <int> <dbl>  
## 1 none 3810 63.2   
## 2 Laminaria setchellii 1459 24.2   
## 3 Callophyllis crenulata 192 3.18   
## 4 Desmarestia linglata 162 2.69   
## 5 Alaria marginata 76 1.26   
## 6 Callophyllis pinnata 59 0.979  
## 7 Callophyllis fabellulata 57 0.945  
## 8 hydroid 39 0.647  
## 9 Hymenena flabelligera 35 0.581  
## 10 Polyneura latissima 17 0.282  
## # ℹ 23 more rows

ggplot(young\_sporophyte\_q8 %>%  
 group\_by(Cospecies) %>%  
 summarise(count = length(Cospecies)) %>%  
 filter(count > 100) %>%  
 arrange(desc(count)), aes(fct\_reorder(Cospecies, count), count)) +  
 geom\_col() +  
 theme\_minimal() +  
 labs(title = "Common primary cospecies", x = "Primary cospecies", "Number of occurrences") +  
 coord\_flip()

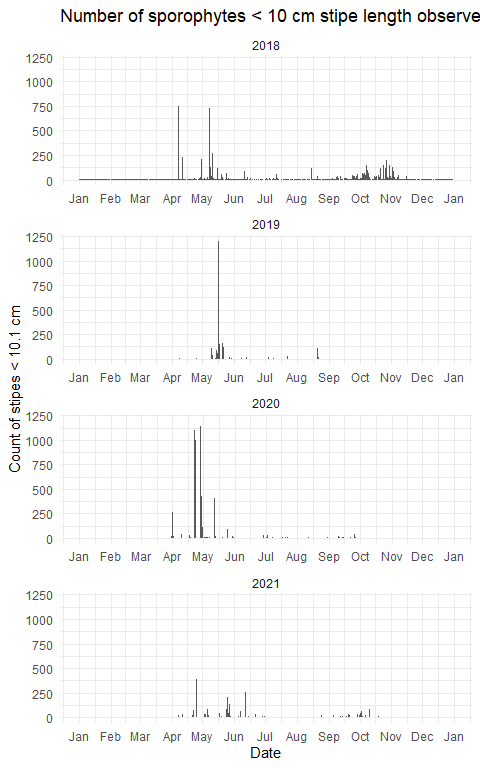


# percent with multiples  
  
young\_sporophyte\_q8\_1 <- young\_sporophyte %>%  
 mutate(CoSp1 = na\_if(CoSp1, "nd")) %>%  
 mutate(CoSp1 = na\_if(CoSp1, "N A")) %>%  
 mutate(CoSp1 = na\_if(CoSp1, "Na")) %>%  
 mutate(CoSp2 = na\_if(CoSp2, "nd")) %>%  
 mutate(CoSp2 = na\_if(CoSp2, "N A")) %>%  
 mutate(CoSp2 = na\_if(CoSp2, "Na")) %>%  
 mutate(CoSp3 = na\_if(CoSp3, "nd")) %>%  
 mutate(CoSp3 = na\_if(CoSp3, "N A")) %>%  
 mutate(CoSp3 = na\_if(CoSp3, "Na")) %>%  
 mutate(CoSp1 = ifelse(CoSp1 == "none", "0", "1")) %>%  
 mutate(CoSp2 = ifelse(CoSp2 == "none", "0", "1")) %>%   
 mutate(CoSp3 = ifelse(CoSp3 == "none", "0", "1")) %>%  
 mutate\_at(c("CoSp1", "CoSp2", "CoSp3"), as.numeric) %>%  
 mutate(Total\_cospecies = CoSp1 + CoSp2 + CoSp3) %>%  
 filter(!is.na(Total\_cospecies))  
  
cospecies\_totals <- young\_sporophyte\_q8\_1 %>%  
 group\_by(Total\_cospecies) %>%  
 summarise('total' = length(Total\_cospecies)) %>%  
 arrange(desc(total)) %>%  
 mutate(percent = (total/sum(total))\*100)  
  
ggplot(cospecies\_totals, aes(x = Total\_cospecies, y = total)) +  
 geom\_col() +  
 theme\_minimal() +  
 labs(title = "Cospecies per observation", x = "Total number of cospecies", y = "Frequency of occurrence") +  
 geom\_text(aes(label = total), vjust = -0.5)



## For each year, when does recruitment occur? (as function of % of 0.1-10cm length stipes throughout the year)

# HOW LONG TILL STIPES REACH 10CM FROM SETTLEMENT?  
  
# THIS IS SUPER SEXY AND POTENTIAL THE CRUX OF THIS DATASET  
  
young\_sporophyte\_q9 <- young\_sporophyte %>%  
 complete(Date = seq.Date(as.Date("2018-01-01"), as.Date("2021-12-31"), by="day")) %>%  
 separate(Date, c("year", "month", "day"), sep = "-", remove = FALSE) %>%  
 unite("month-day", month:day, remove = FALSE) %>%   
 mutate(Stipe = ifelse(is.na(Stipe), 0.5, Stipe)) %>%  
 filter(Stipe < 10.1)  
  
ggplot(young\_sporophyte\_q9, aes(x = Date, y = Stipe)) +  
 geom\_col() +  
 theme\_minimal() +  
 labs(title = "Number of sporophytes < 10 cm stipe length observed by day", x = "Date", y = "Count of stipes < 10.1 cm") +  
 scale\_x\_date(breaks = "1 month", labels = date\_format("%b")) +  
 facet\_wrap(~year, nrow = 4, scales = "free\_x")



# References