Polyploidy in Plants - Mystery Datasets

Thomas Holland (th675)

January 11th, 2023

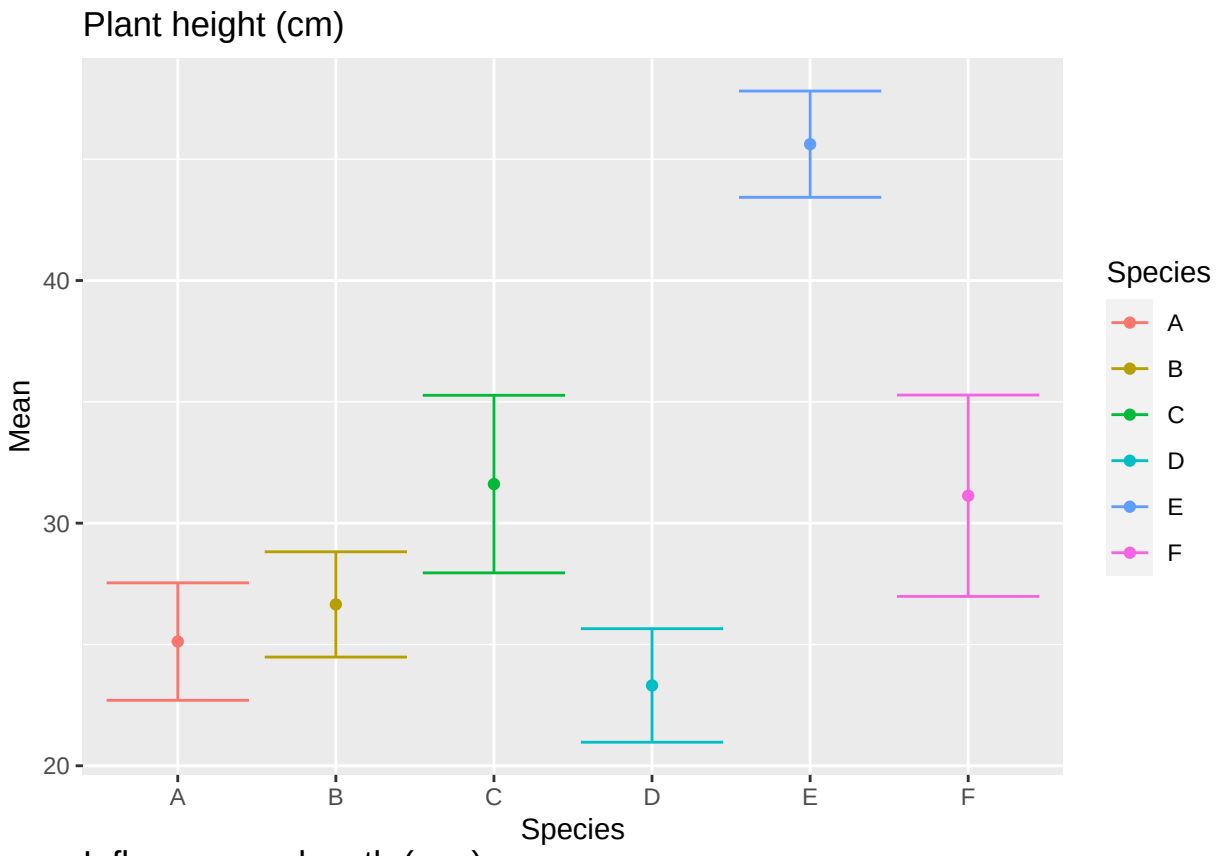
The Rmarkdown file used to generate this document can be found here (https://github.com/th-holland/Mini-Research-Projects/blob/main/Evolution%20Practicals/Polyploidy%20Practical/actual/polyploidy.rmd).

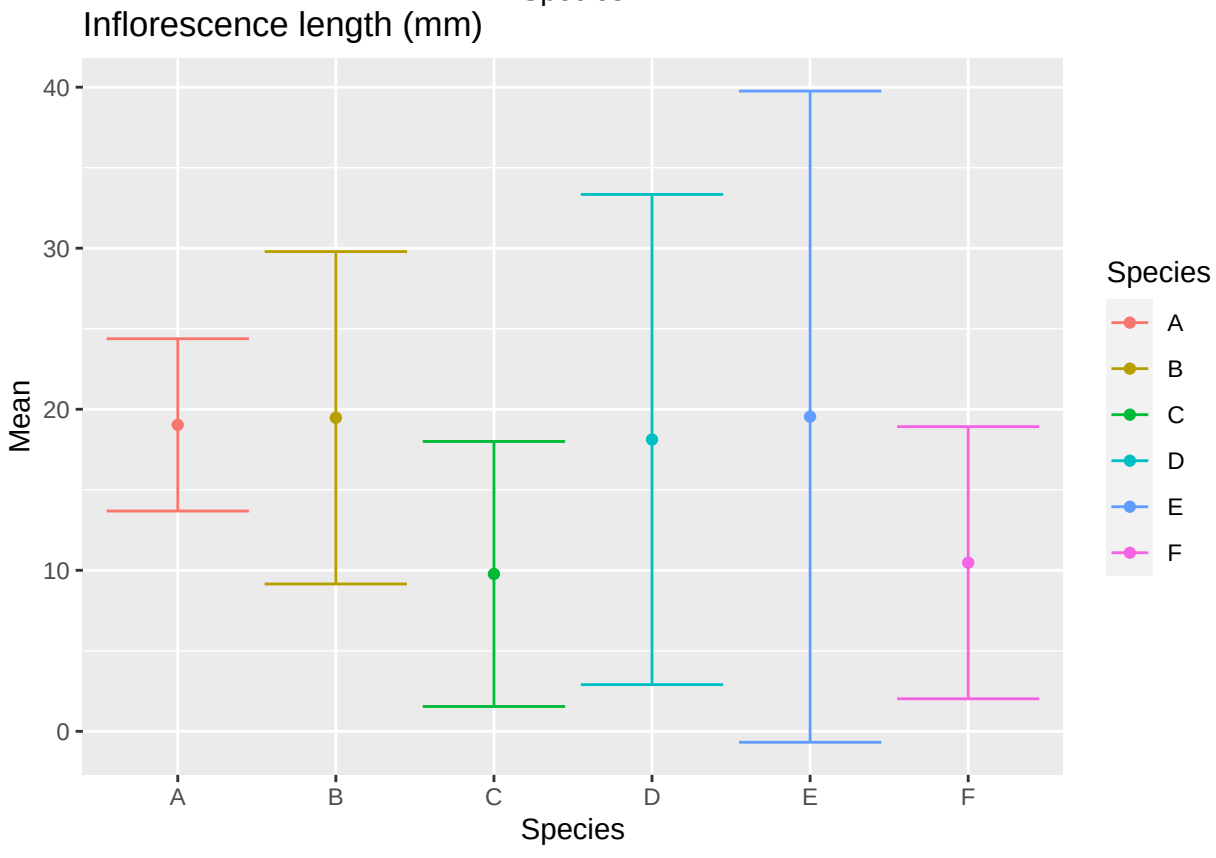
Plots

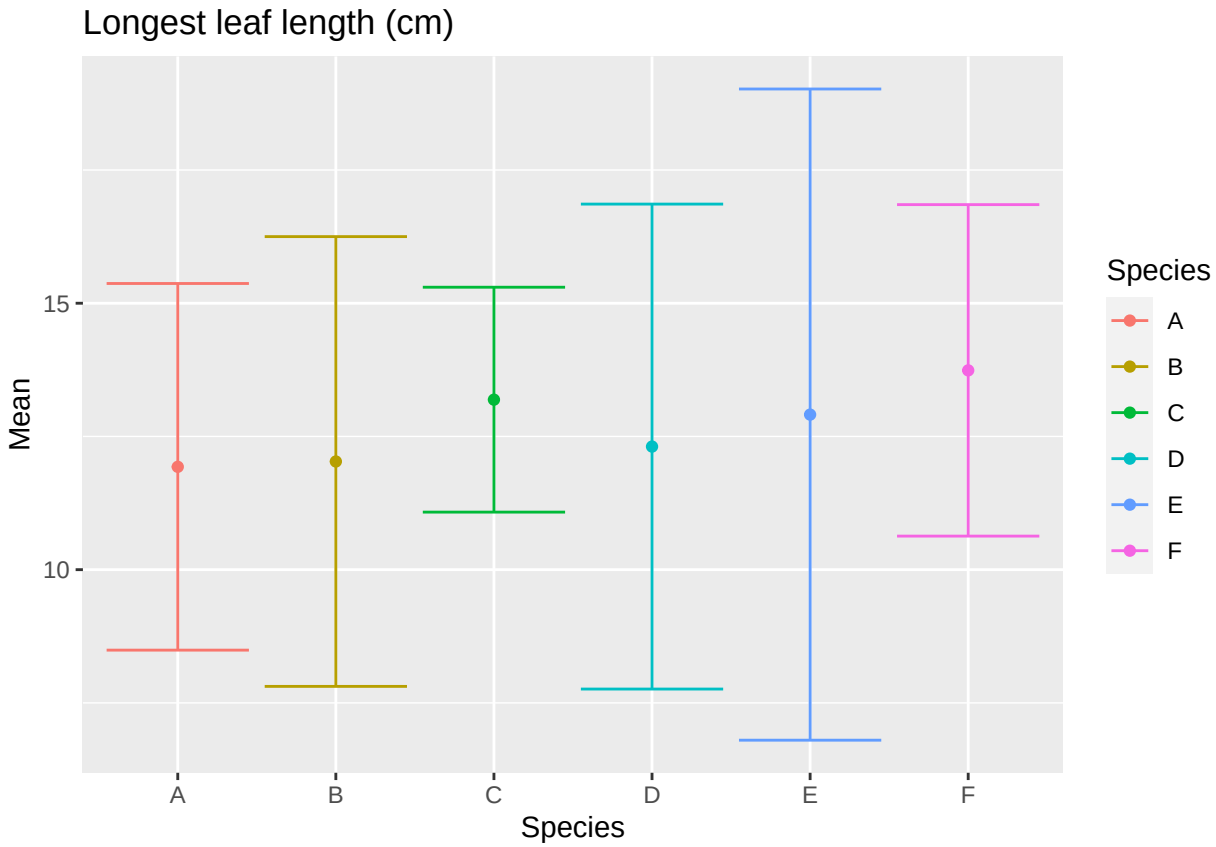
```
morph_data_avg <- read_excel("data1.xlsx", sheet = "mean")
morph_data_sd <- read_excel("data1.xlsx", sheet = "SD")

morph_data_avg <- t(morph_data_avg)
morph_data_sd <- t(morph_data_sd)
colnames(morph_data_avg) <- morph_data_avg[1, ]
colnames(morph_data_sd) <- morph_data_sd[1, ]
morph_data_avg <- morph_data_avg[-1, ]
morph_data_sd <- morph_data_sd[-1, ]
# convert to data frame
morph_data_avg <- as.data.frame(morph_data_avg)
morph_data_sd <- as.data.frame(morph_data_sd)

# append .SD to the column names of morph_data_SD
colnames(morph_data_sd) <- paste(colnames(morph_data_sd), ".sd",
    sep = "")

Species <- c("A", "B", "C", "D", "E", "F")

# combine Species and morph_data_avg
morph_data <- cbind(morph_data_avg, morph_data_sd)
```

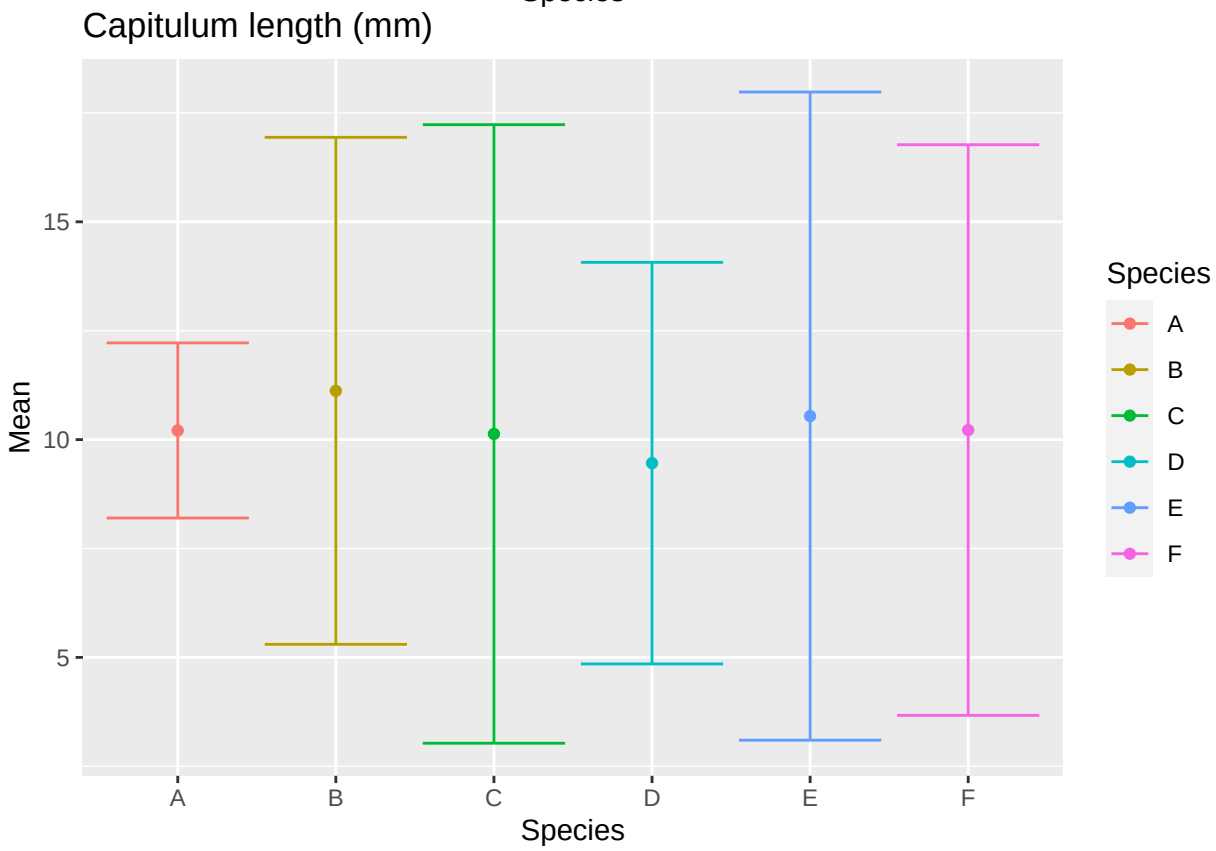
```
# recursively generate a plot of the means and error bars
# for the SD for each trait with each species with a
# different colour (species is first column of
# morph_data_avg)

for (i in 1:8) {
    means <- as.numeric(morph_data[, i])
    high <- as.numeric(means) + as.numeric(morph_data[, i + 7])
    low <- as.numeric(means) - as.numeric(morph_data[, i + 7])
    measurment <- colnames(morph_data)[i]
    data <- data.frame(Species, means, high, low)
    plot <- ggplot(data, aes(x = Species, y = means, colour = Species)) +
```

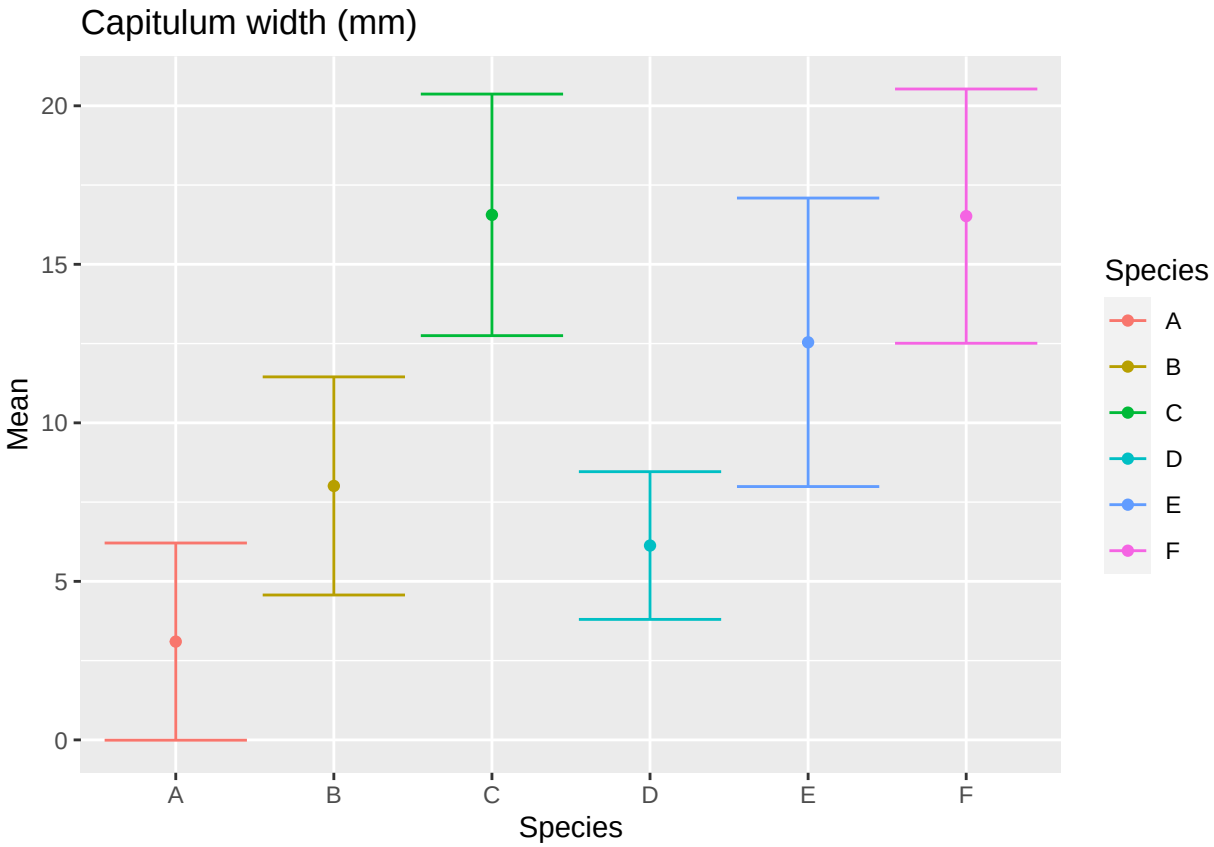
```
        geom_point() + geom_errorbar(aes(ymin = low, ymax = high)) +
        labs(title = measurment, x = "Species", y = "Mean")
    print(plot, width = 7, height = 5)
}
```

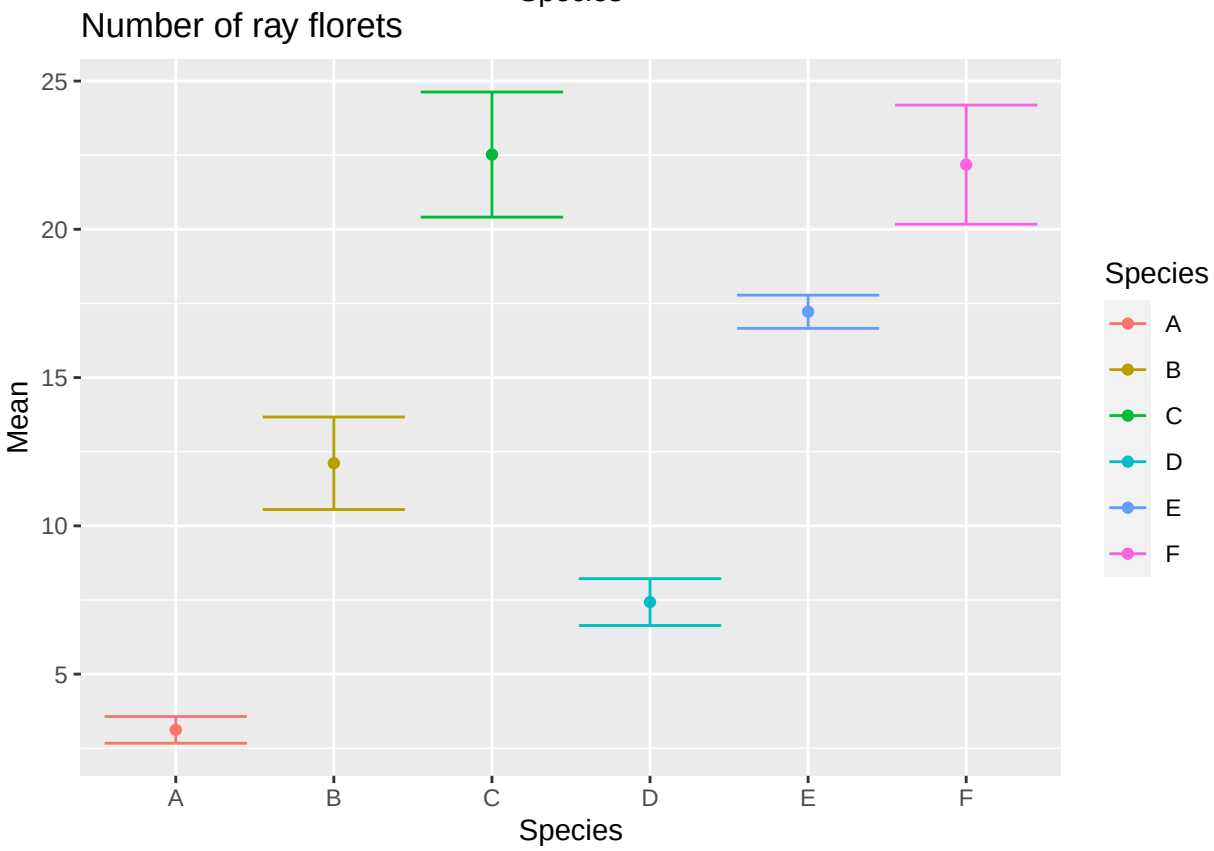


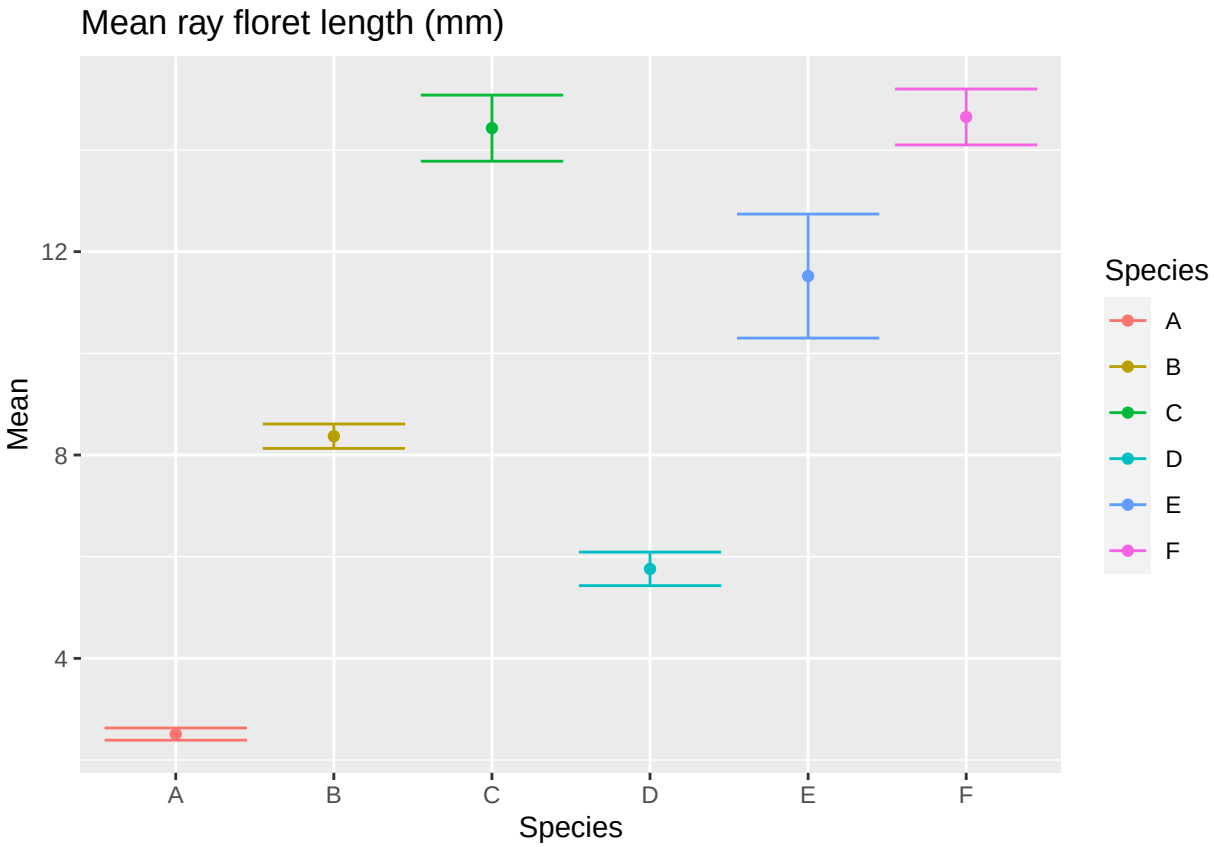


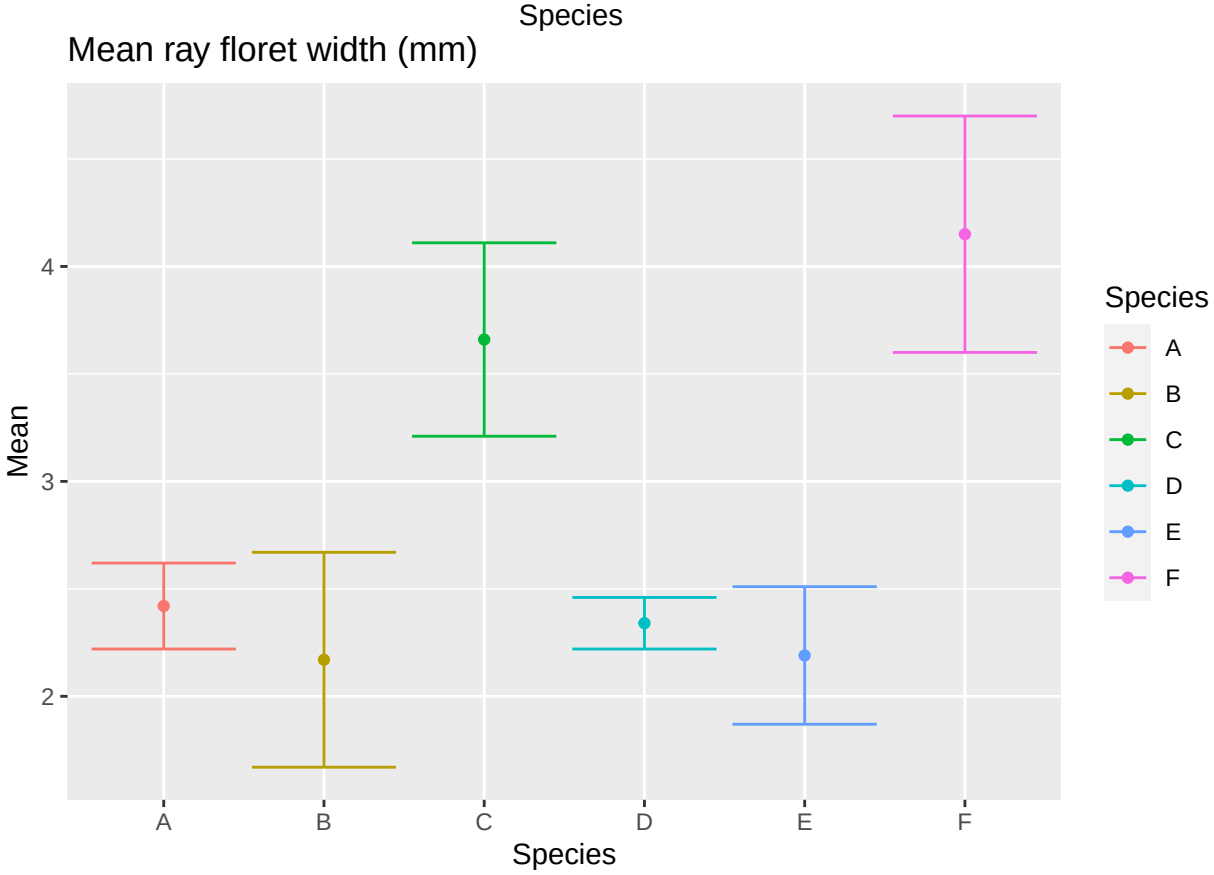

Longest leaf length (cm)
Mean
15
10
A
B
C
D
E
F
Species
Species
A
B
C
D
E
F


Capitulum length (mm)
Mean
15
10
5
A
B
C
D
E
F
Species
Species
A
B
C
D
E
F


Capitulum width (mm)
Mean
20
15
10
5
0
A
B
C
D
E
F
Species
Species
A
B
C
D
E
F


Number of ray florets
Mean
25
20
15
10
5
A
B
C
D
E
F
Species
Species
A
B
C
D
E
F


Mean ray floret length (mm)
Mean
12
8
4
A
B
C
D
E
F
Species
Species
A
B
C
D
E
F


Mean ray floret width (mm)
Mean
4
3
2
A
B
C
D
E
F
Species
Species
A
B
C
D
E
F

Questions

1) For each species A, B, C, D, E, & F, identify which species are diploid and which species are polyploid. Explain your reasoning which respect to the chromosome numbers of diploids and polyploids.

A, B and C all have 8 chromosomes and are likely diploid, D, E and F all have 16 chromosomes and are likely polyploid.

2) You need to advise a field team as to which morphological characters are most diagnostic for each diploid and polyploid species. Make graphs of the mean and standard deviation of each species for each trait. Ensure that graphs are fully and clearly labelled. Indicate which traits should be measured, and what would be expected to be observed for each species.

Please see above for the graphs.

Suggestions for identification of diploid vs polyploid species: - Capitulum length is not very diagnostic, only difference is with A where the variance is much lower than the other species. - Plant height is quite diagnostic for E, as E is much higher (above 40cm) than the other species, with no overlap in the error bars. However this could be due to the environment that E grows in/was sampled and thus due to environmental factors.

3) Explain to the team why some morphological traits are more useful than others.

- Some traits are additive, such that the polyploid phenotype is the sum of the diploid phenotypes.
- However, some traits are not additive, such that the polyploid phenotype is not the sum of the diploid phenotypes, and more intermediate between the diploid phenotypes, and as such less diagnostic of the polyploid species.

4) For each polyploid species, indicate which ALFP bands (using the numbering system running down the side of the AFLP diagram) are informative in discerning which diploid progenitors have given rise to each polyploid.

1 Shows D originated from A

2 D and E from B

3 E and F from C

10, 24, 30, 33 shows the same as 1

5, 13, 18, 25, 31, 37 shows the same as 2

4, 7, 19, 26, 28,32 show the same as 3

Bands 6, 8 9, 11, 12, 14, 15, 16, 17, 20 through 24, 27 29, 24 through 36, 38 and 39 are not informative as the banding is common to all.

The fact that the patterns seen in 1, 2 and 3 are repeated throughout the AFLP data suggests that the AFLP data is accurate and that the lineage can be distinguished by AFLP.

5) For each of the three polyploid species, explain their relationship to the diploid species that gave rise to them. Explain your answer with reference to ALL the evidence you have used to discern the relationships.

It is likely that D originated from a combination of A and B, E from a combination of B and C and F from a doubling of C. This is because the AFLP data shows that the bands that originated in A appear in D only, the bands that originated in B appear in both D and E, and the bands that originated in C appear in both E and F. This suggests that D is a combination of A and B; E is a combination of B and C; and F is a doubling of C.

The morphological data also supports this, as mostly D is intermediate between A and B, E is intermediate between B and C and F is similar to C. The only data-point that doesn't fit this trend is with the plant height where for E the effect seems to be additive but for D it appears to be negative (D is shorter than A or B); although the height of the plant could be predominantly determined by the environment that the plant is growing in.

The karyotype data also supports this, as D has 8 chromosomes from A and 8 chromosomes from B, E has 8 chromosomes from B and 8 chromosomes from C and F has 2 sets of 8 chromosomes from C.

6) Suggest why allopolyploids are more commonly described than autopolyploids.

Allopolyploids are more commonly described than autopolyploids because allopolyploids are more likely to be fertile than autopolyploids. This is because allopolyploids have two sets of chromosomes from two different species, and as such are more likely to have the correct number of chromosomes for each gene. Autopolyploids have two sets of chromosomes from the same species, and as such are more likely to have the incorrect number of chromosomes for each gene.

Also, allopolyploids tend to have more variation in their phenotypes from their diploid progenitors than autopolyploids, and as such are more likely to be identified as a new species.

7) Give three selective advantages of a polyploid over its diploid progenitors? [1]

- More copies of a gene means that deleterious mutations are less likely to be expressed in the phenotype of the polyploid when compared to the diploid progenitor
- The redundancy of the genes means that neofunctionalization or subfunctionalization can occur, which can lead to new traits in the polyploid that are not present in the diploid progenitor
- Also there is a possible benefit where heterosis occurs, where the polyploid has a higher fitness (from an increased biomass, fertility and speed of development) than the diploid progenitor [2]

8) Give three potential evolutionary disadvantages of polyploidy? [3]

- The increased number of chromosomes leads to an complexity of the pairing and segregation interactions in meiosis, which can lead to having an abnormal number of chromosomes in a haploid set (aneuploidy)
- Due to the increased genome, the cell size typically is increased and so the cell architecture is altered, which can lead to a decrease in fitness (e.g. changed SA:V ratio)

[1] Madlung, A. Polyploidy and its effect on evolutionary success: old questions revisited with new tools. Heredity 110, 99–104 (2013). https://doi.org/10.1038/hdy.2012.79

[2] James A. Birchler, Hong Yao, Sivanandan Chudalayandi, Daniel Vaiman, Reiner A. Veitia, Heterosis, The Plant Cell, Volume 22, Issue 7, July 2010, Pages 2105–2112, https://doi.org/10.1105/tpc.110.076133

[3] Madlung, A. Polyploidy and its effect on evolutionary success: old questions revisited with new tools. Heredity 110, 99–104 (2013). https://doi.org/10.1038/hdy.2012.79

- Altered gene expression can be changed due to many factors, predominantly due to an altered transcriptome, genomic architecture and epigenetic landscape, which can lead to excess gene expression, reduced gene expression or altered gene expression.

9) Morphology, AFLP and Karyotyping each give different but complementary insights into polyploidy. Construct a table and for each technique give two advantages and two disadvantages of each method. [4]

Type	Advantages	Disadvantages
Morphology	- Easily to observe in the field	- Might be additive, intermediate or neither, so requires analysis that may be inaccurate
	- Can be done with living samples without damaging them (e.g. with endangered species)	-Plant development is non-deterministic and so some features, such as height or length between features may not be genetically determined
AFLP	- One of the least labour intensive methods	- It is unable to distinguish dominant homozygous from dominant heterozygous individuals
	- High repeatability when compared to other methods such as RAPD	- Require higher levels of DNA than other methods (such as using Micro-satellite DNA markers)
Karyotyping	- Easily determine the number of sets of chromosomes	- Requires a skilled technician to perform the karyotyping as the cells need to put into stasis during prophase
	- Can be used to distinguish the sizes of the chromosomes relative to each other and the other species (can give indication of origins)	- Not indicative of the genetic content of the chromosome so working out the origins from the parents is harder than with AFLP

10) Hypothesise why certain diploid species combinations in this system do not apparently (at least given this dataset) form allopolyploids?

- The tetraploids would be unable to for viable offspring with the diploid species as this would cause a triploid block to occur, as the triploids are (usually) unable to produce viable gametes.
- The chromosome shapes/sizes appear to be a large mismatch between A and C in the karyotype dataset, and their lack of allopolyploids suggests that they are unable to produce viable offspring.
- Similarly, the morphological data shows that, for some features, A and C are very different, and so it is likely that they are unable to produce viable offspring as there would be clashing phenotypes.

[4] Yang, W., Kang, X., Yang, Q. et al. Review on the development of genotyping methods for assessing farm animal diversity. J Animal Sci Biotechnol 4, 2 (2013). https://doi.org/10.1186/2049-1891-4-2

- From the morphological data, it appears that B could be a hybridization of A and C, which is somewhat supported by the mixture of the karyotypes (but not supported particularly well by the AFLP data). This would explain how B is able to produce viable offspring with both A and C, when A and C cannot produce viable offspring with each other.