Lab3_2

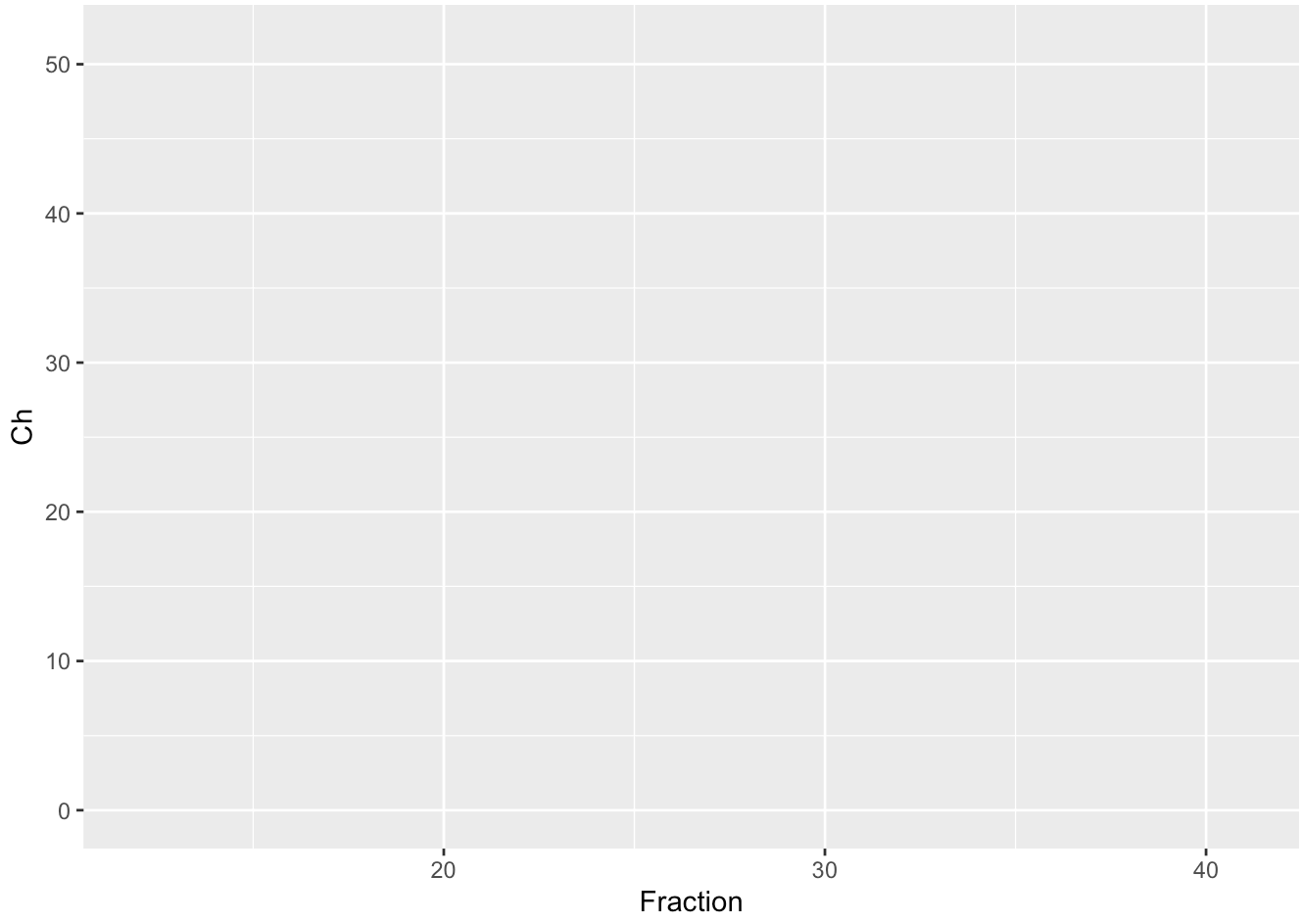
2024-04-18

Section 1

```
# 3
library("ggplot2")
```

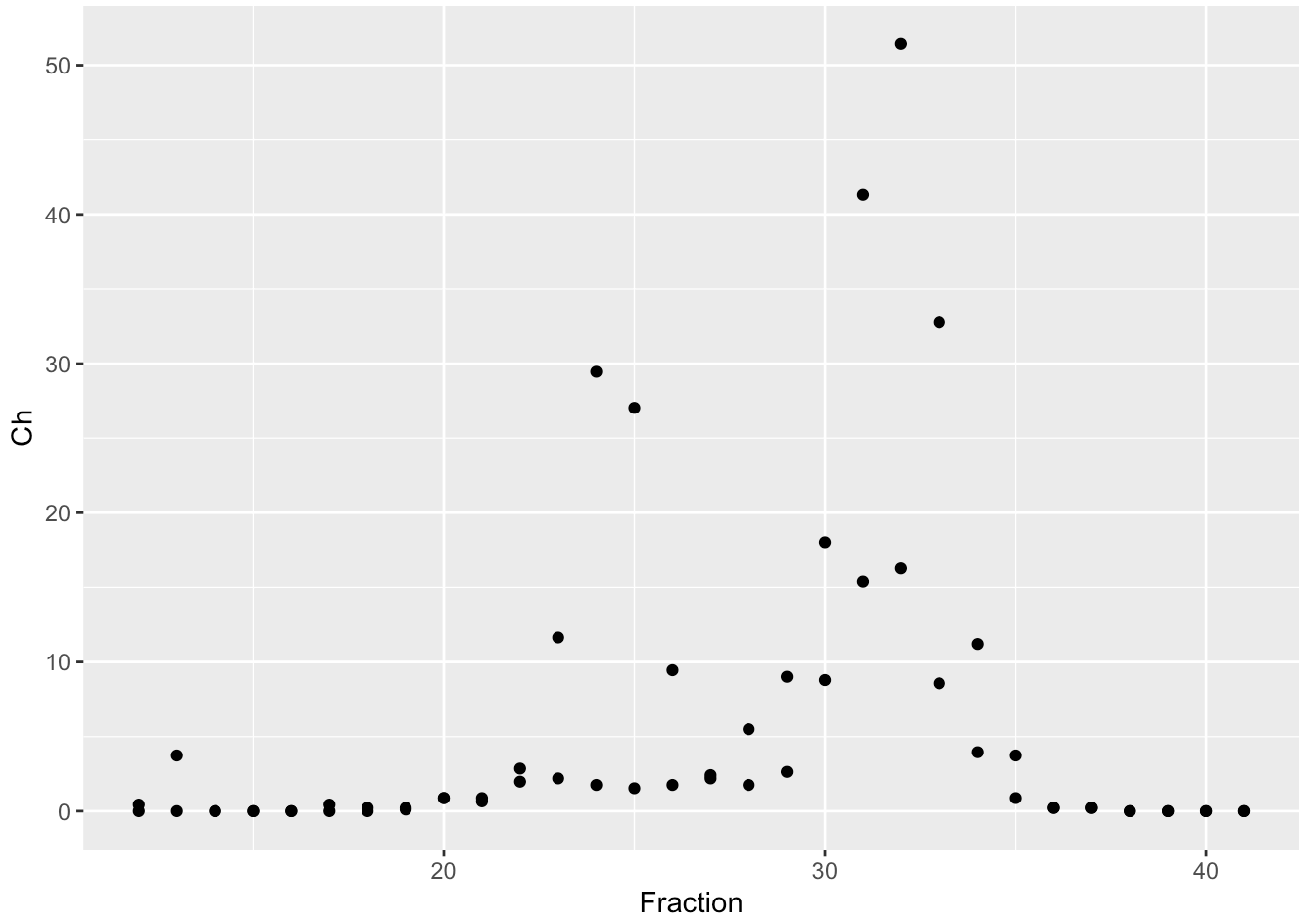
```
## Warning: package 'ggplot2' was built under R version 4.3.2
```

```
# 4
Ch.profiles <- read.csv("cholesterol_profiles.csv")
# 5, 6
ggplot( data=Ch.profiles, aes(x=Fraction, y=Ch) )
```

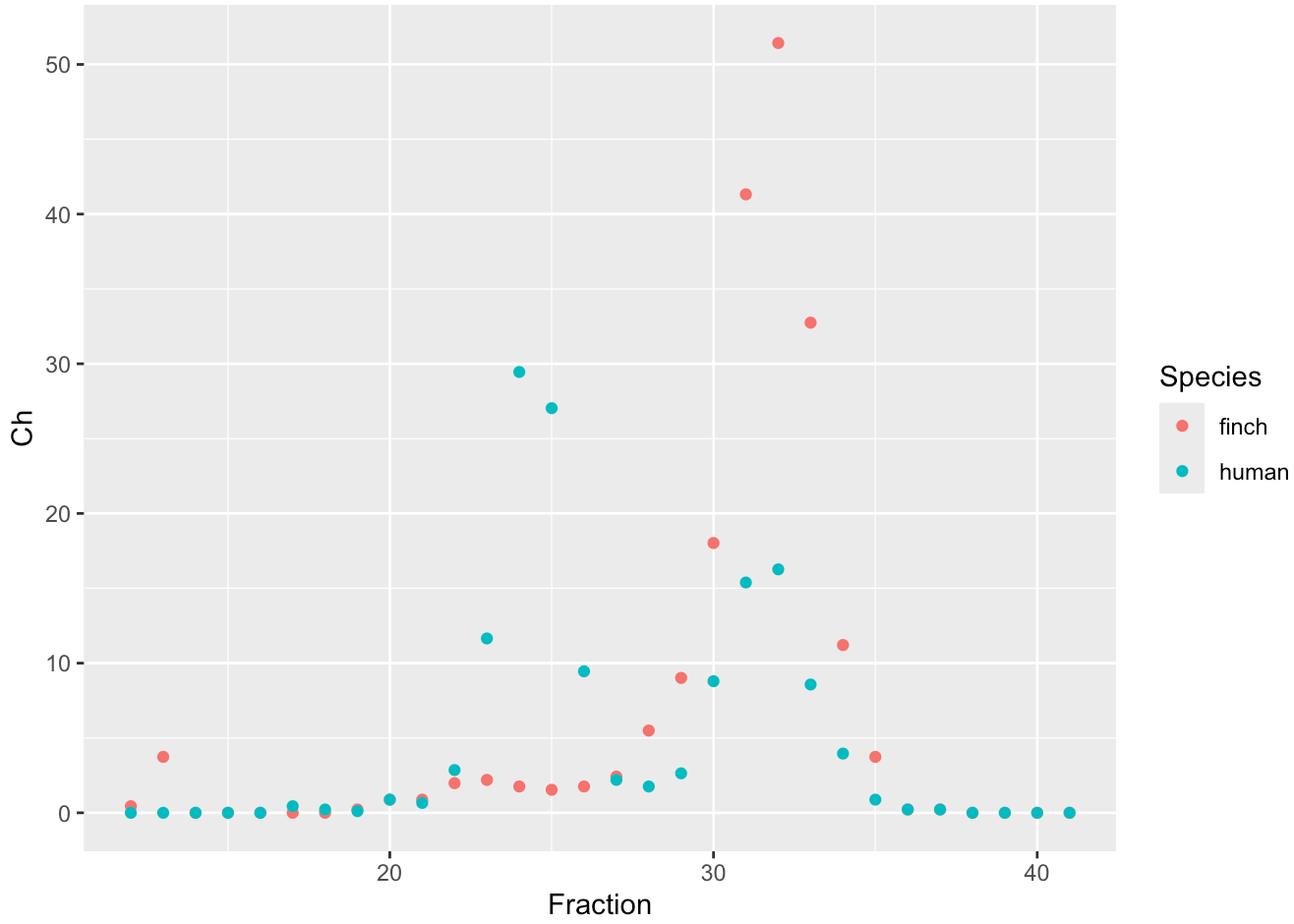


Section 2

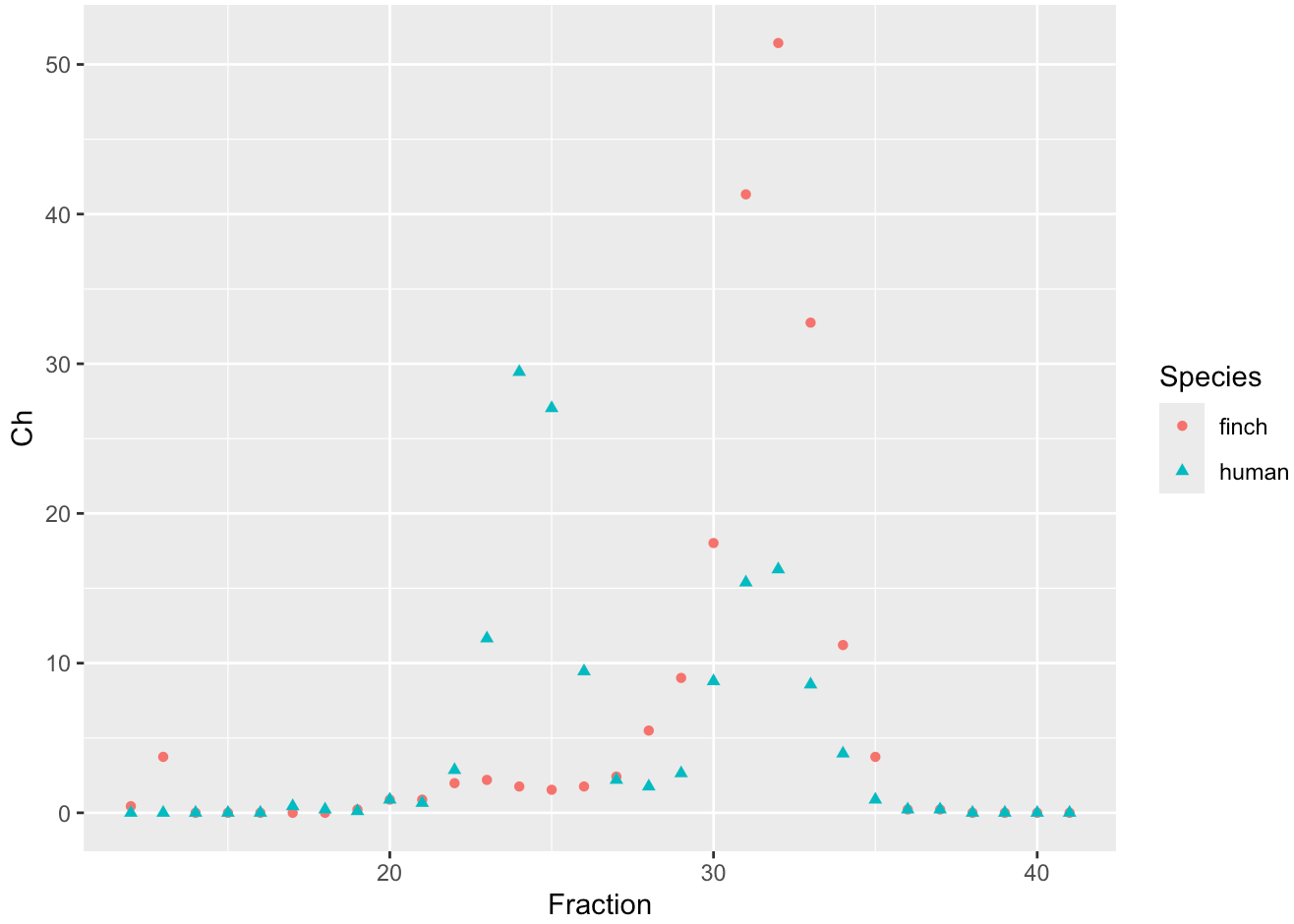
```
# 7
ggplot( data=Ch.profiles, aes(x=Fraction, y=Ch) ) + geom_point( )
```



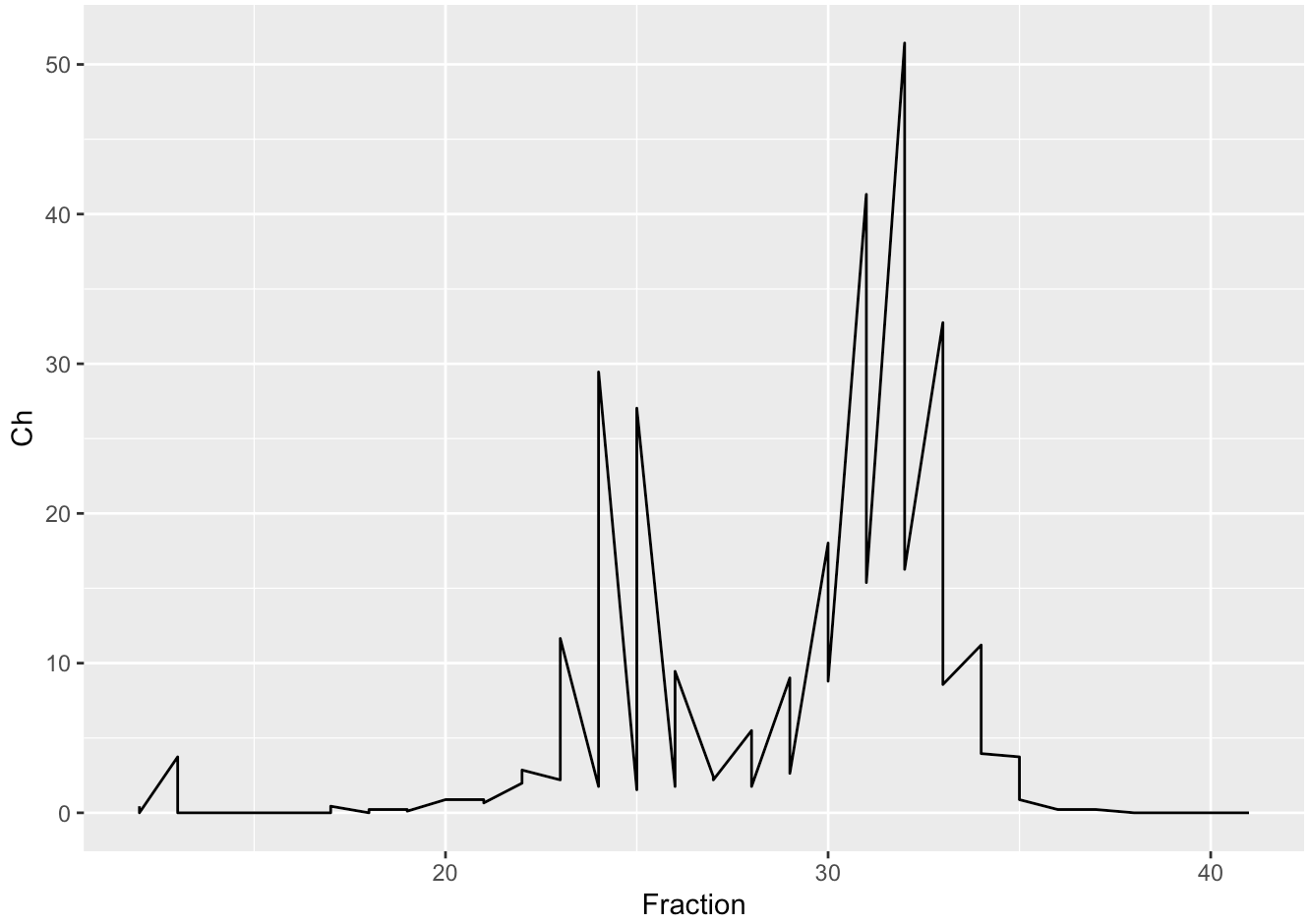

```
# 8
ggplot( data=Ch.profiles, aes(x=Fraction, y=Ch, colour=Species) ) + geom_point( )
```



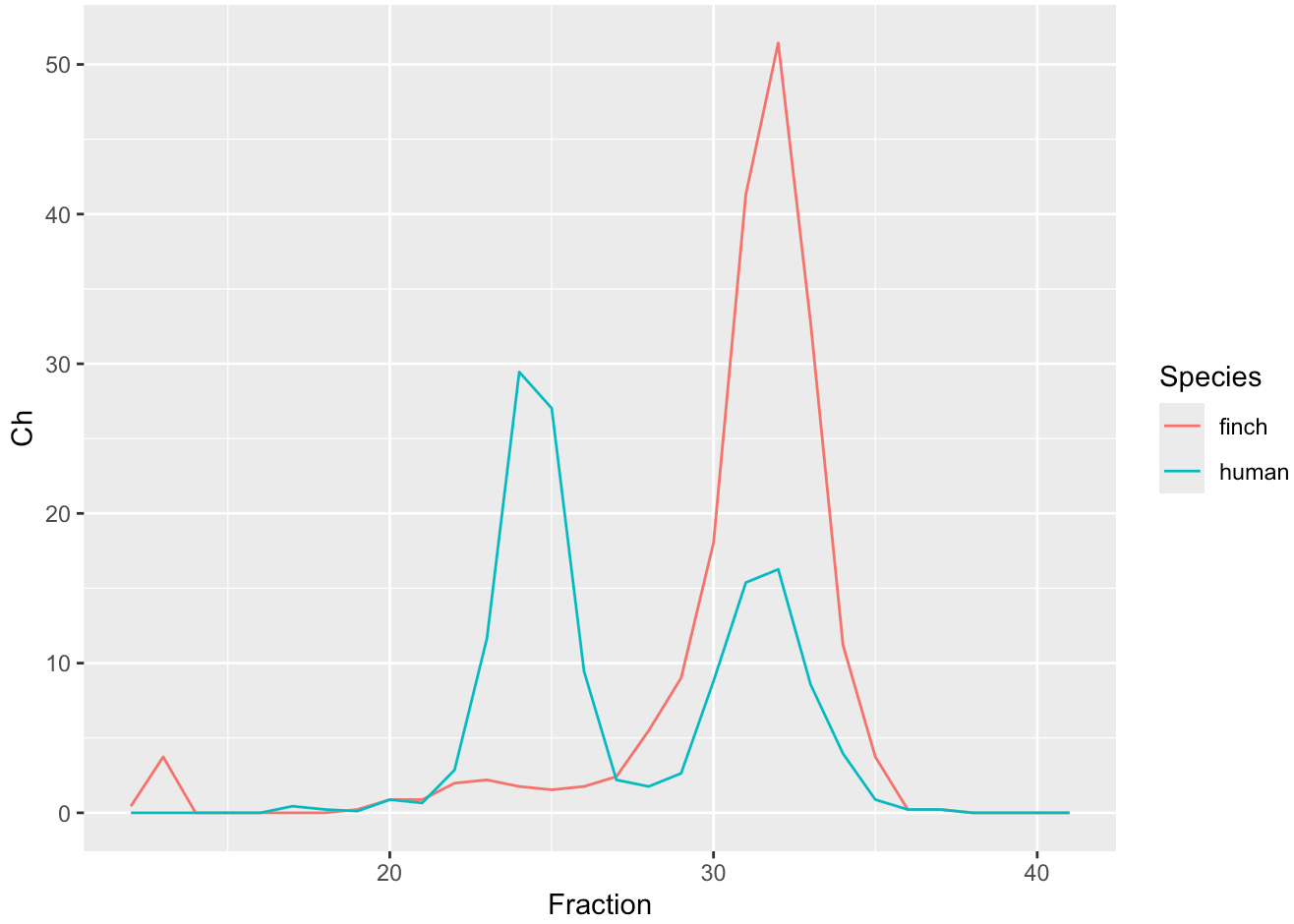

```
# 9
ggplot( data=Ch.profiles, aes(x=Fraction, y=Ch, colour=Species, shape=Species)) + geom_p
oint()
```



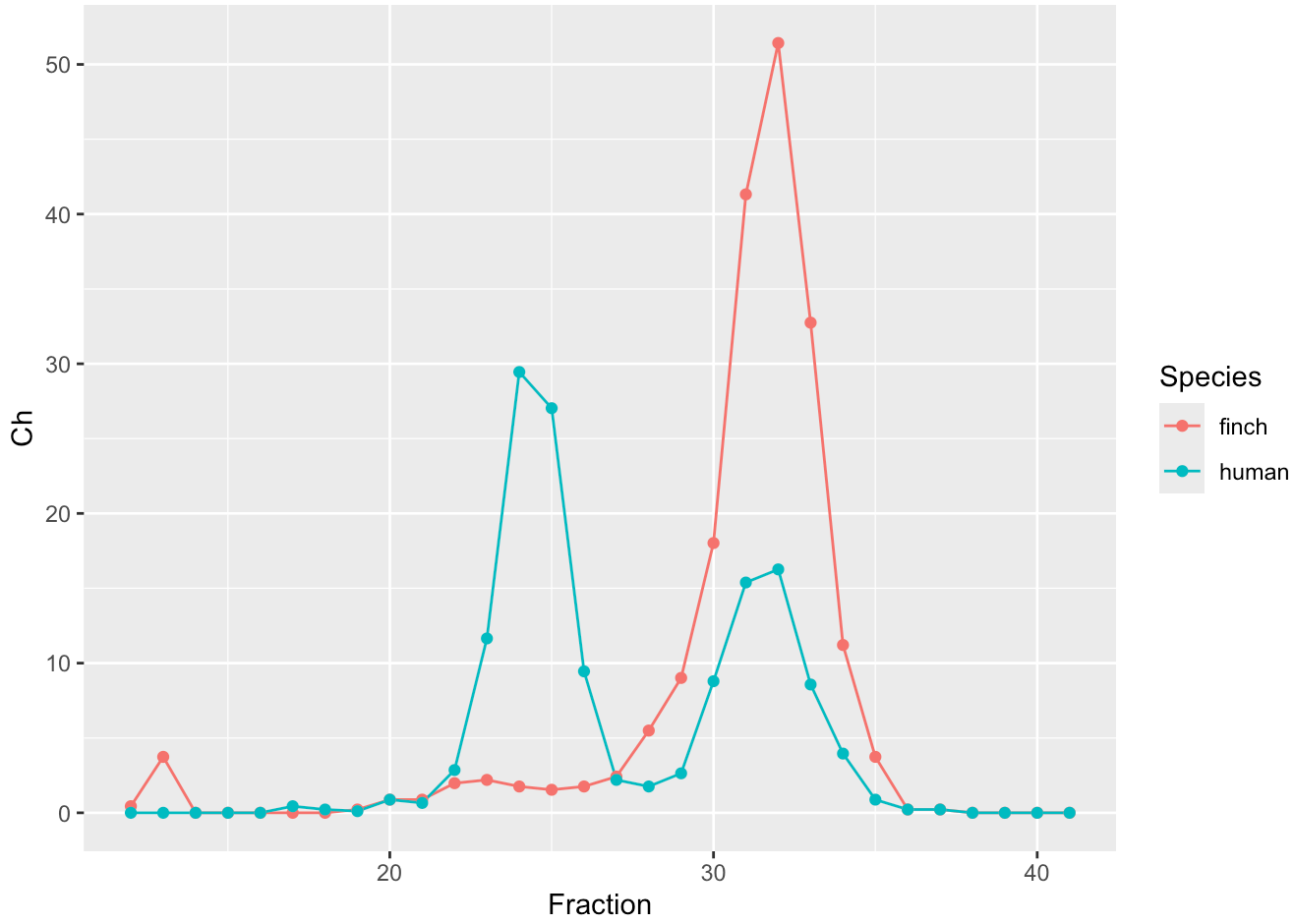

```
# 10
ggplot( data=Ch.profiles, aes(x=Fraction, y= Ch)) + geom_line()
```



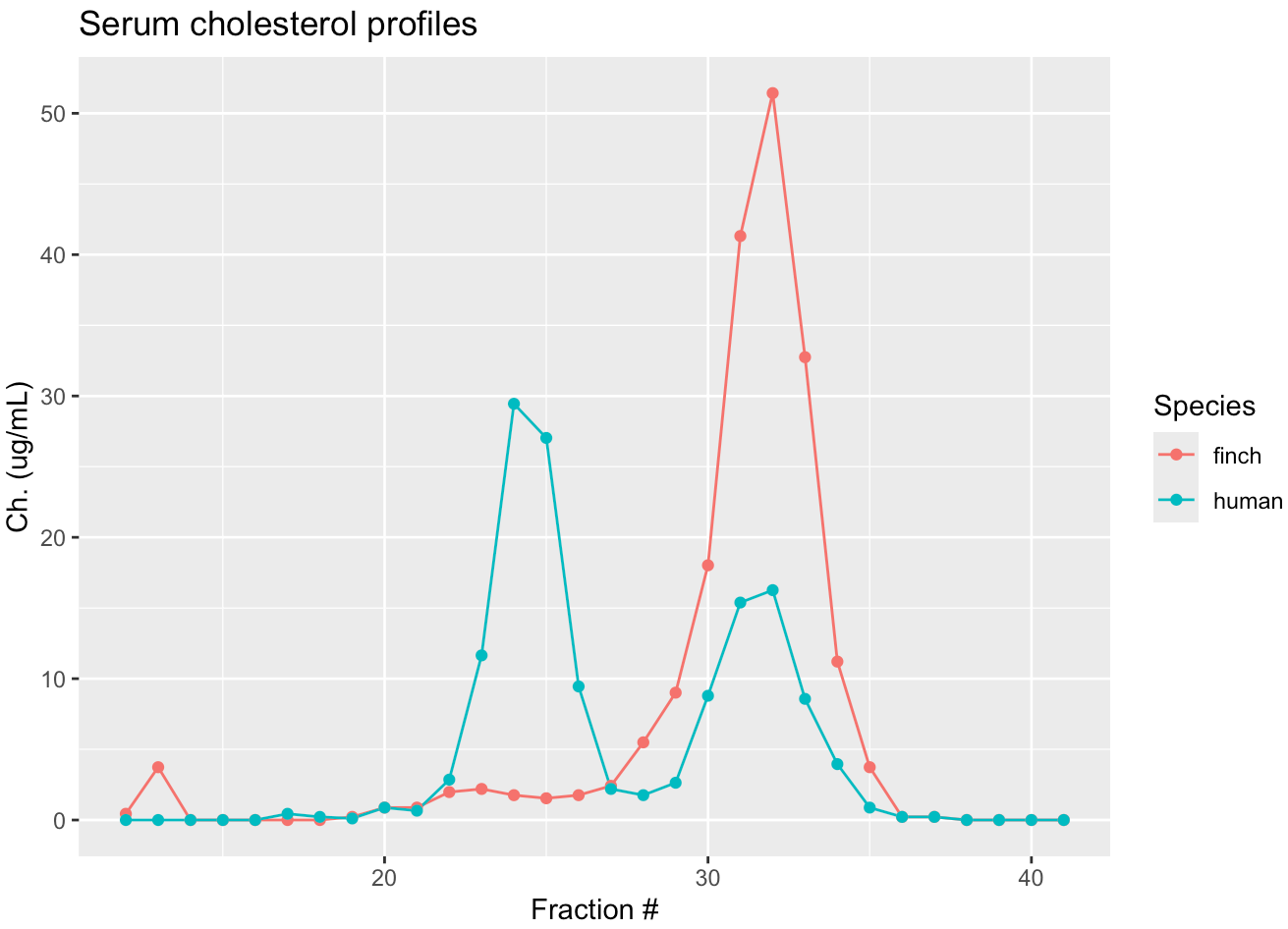

```
# 11
ggplot( data=Ch.profiles, aes(x=Fraction, y= Ch, colour=Species)) + geom_line()
```




```
# 12
Ch.plot<-ggplot( data=Ch.profiles, aes(x=Fraction, y= Ch, colour=Species)) + geom_line()
# 13
Ch.plot+geom_point()
```

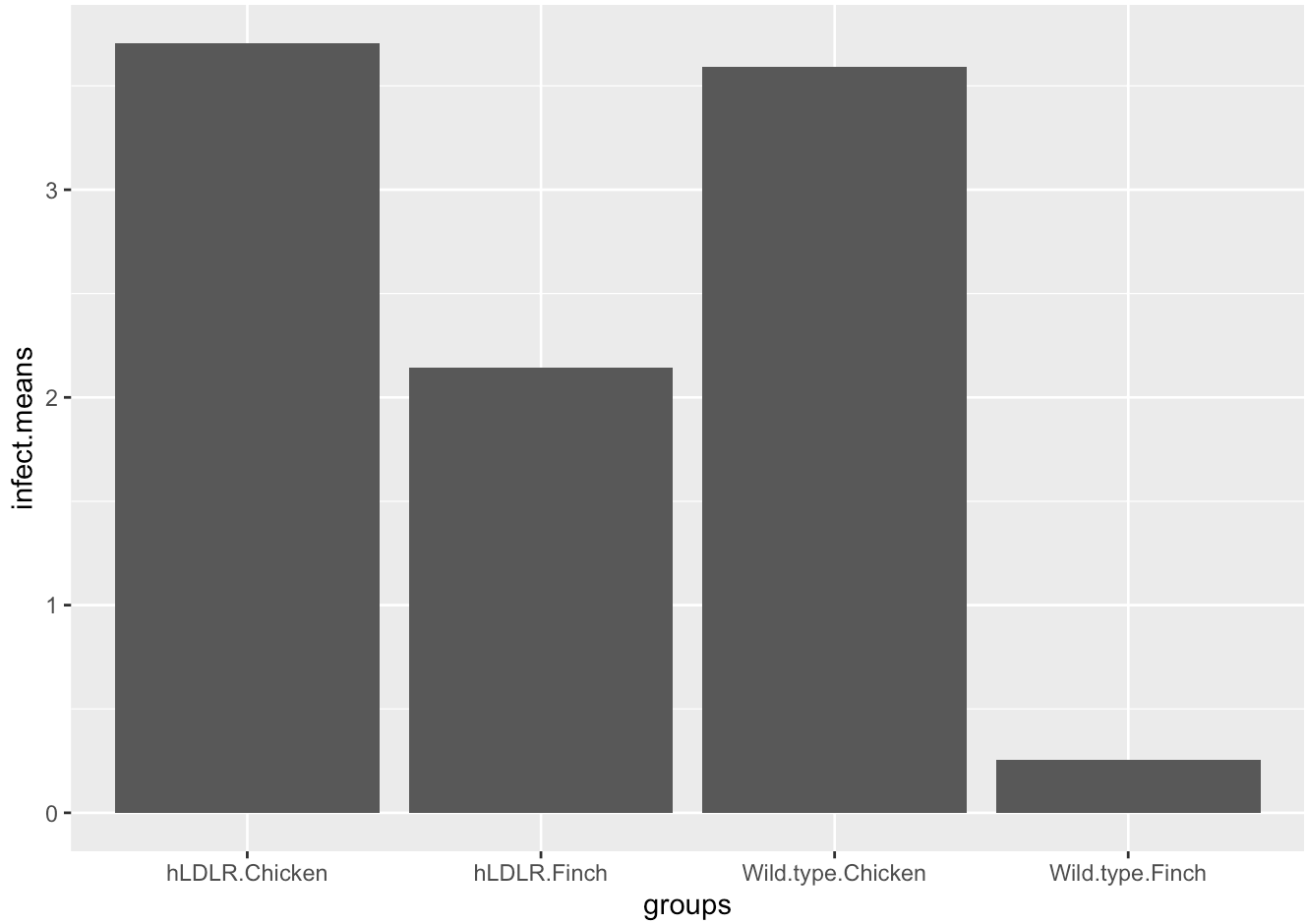



```
# 14
Ch.plot+geom_point()+ggtitle("Serum cholesterol profiles") + # plot title
xlab("Fraction #") + # x-axis label
ylab("Ch. (ug/mL)")
```

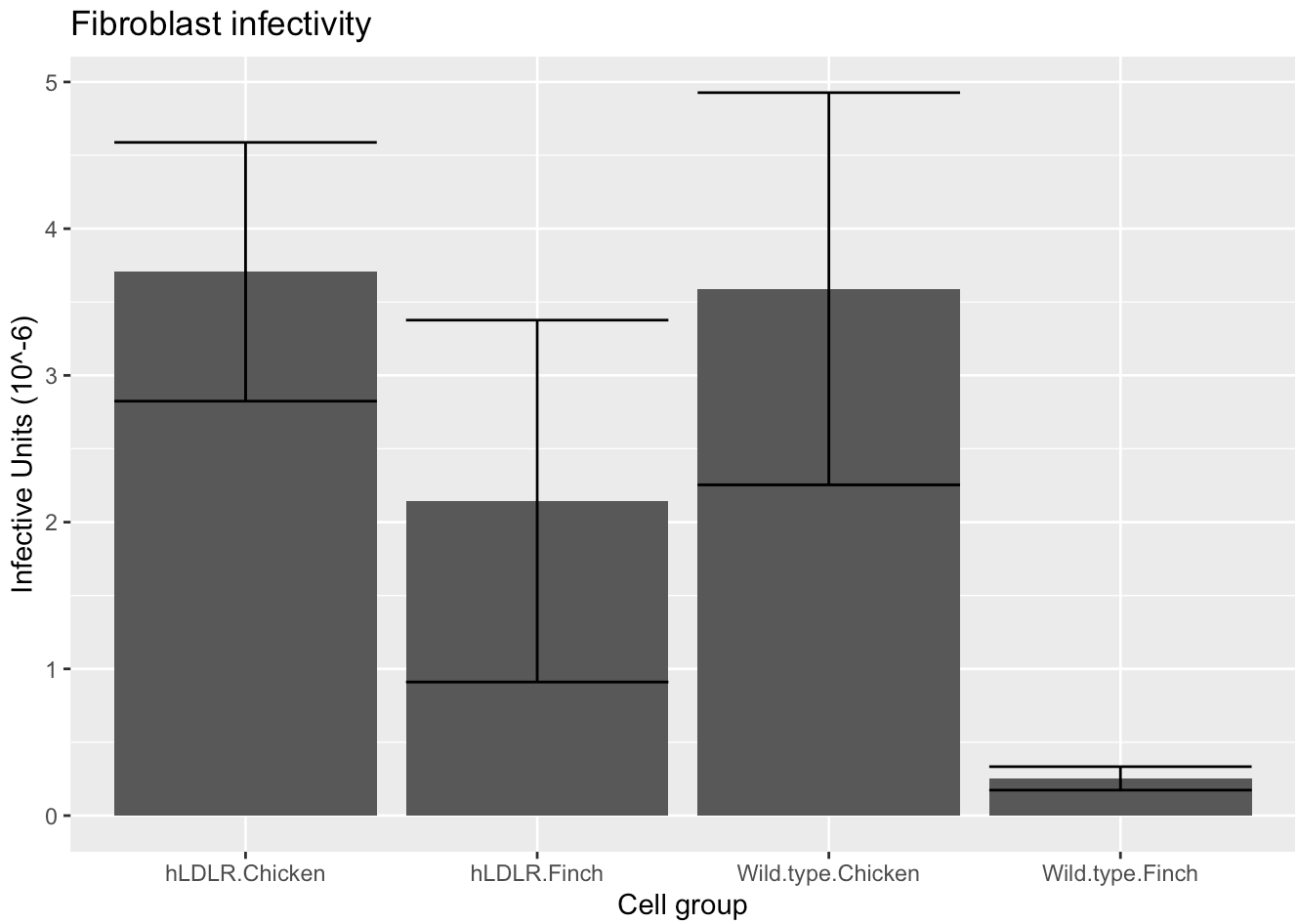


Section 3

```
# 15
infectivity.w <- read.csv("cell_infectivity_wide.csv")
# 16
infect.means <- colMeans(infectivity.w)
# infect.means
# 17
groups <- colnames(infectivity.w)
# groups
# 18
data.bar <- data.frame(groups, infect.means)
# 19
ggplot( data=data.bar, aes(x=groups, y=infect.means))+geom_col()
```




```
# 20
bar.plot <- ggplot( data=data.bar, aes(x=groups, y=infect.means))+ geom_col()
# 21
infect.sd <- apply(infectivity.w, 2, sd)
# 22
data.bar$infect.sd <- infect.sd
# 23
bar.plot +
  geom_errorbar(aes(ymin = infect.means-infect.sd, ymax = infect.means+infect.sd))+
    # 24
    ggtitle("Fibroblast infectivity") + # plot title
    xlab("Cell group") + # x-axis label
    ylab("Infective Units (10^-6)")
```



Section 4

Q1:

It is important as it can increase expression promotes the growth of prostate cancer. It can help us understand how cancer progression can be linked to the evolution of genes.

Q2:

1. They obtained sequence from human testis library (Bai et al. 2005).
2. They obtained the sequence for other primates from the national center fo biotechnology information (NCBI: https://www.ncbi.nlm.nih.gov/ (https://www.ncbi.nlm.nih.gov/)).

Q3:

We get the accession number: NM_005366.4

Homo sapiens MAGE family member A11 (MAGEA11), transcript variant 1, mRNA
https://www.ncbi.nlm.nih.gov/nuccore/NM_005366.4/ (https://www.ncbi.nlm.nih.gov/nuccore/NM_005366.4/)

Q4:

Chimpanzee aligns best with the human as period indicate matches in the alignment to human sequence. And there is only 3 to 4 mismatch or gaps for the alignment with human sequence.

Lemur aligns least with the human as it has the least amount of period compare to other species. And it has a lot of letters and dash, which indicate there are many mismatches and gaps when align to the human's amino acid sequence.

Q5:

Is there additional analysis for other primates on the same genes as it can help understand the potential impact of MAGE-A11 on different species.