

# CM515 Day 2: Plotting with ggplot()

## 1. Putting it all together

---

We are going to take some of my data and combine the principles from day1 of the ggplot section to make an awesome graph!

Context:

TS559-Parent Strain TK0887-deleted for Endonuclease Q TK1898-deleted for Endonuclease MS TK2143-deleted for Uracil DNA Glycosylase

TK0887/TK1898-deleted for both genes TK0887/TK1898/TK2143-deleted for all three genes

Cultures were grown at the biological temperature of the organism-85 degrees celsius.

Growth measured every hour for 26 hours in biological quadruplicate by optical density at 600nm.

```
####  
#01#  
####
```

```
dU_GC<-read.csv("Growth_Curve_dU.csv",fill=TRUE,header=TRUE)
```

```
head(dU_GC)
```

```
##   Hours Blank.1 Blank.2 Blank.3 Blank.4 TS559.1 TS559.2 TS559.3 TS559.4  
## 1     0   0.001  -0.035  -0.047   0.000  -0.003   0.018  -0.017   0.000  
## 2     1  -0.023  -0.030  -0.036  -0.051   0.049   0.020   0.023   0.032  
## 3     2  -0.043  -0.033  -0.041  -0.035   0.093   0.082   0.074   0.055  
## 4     3  -0.045  -0.054  -0.045  -0.034   0.144   0.148   0.133   0.129  
## 5     4  -0.045  -0.044  -0.045  -0.035   0.269   0.245   0.252   0.243  
## 6     5  -0.042  -0.050  -0.046  -0.036   0.393   0.389   0.409   0.397  
##   TK0887.1 TK0887.2 TK0887.3 TK0887.4 TK1898.1 TK1898.2 TK1898.3 TK1898.4  
## 1   -0.011  -0.011  -0.008  -0.018   0.000   0.004  -0.015  -0.013  
## 2    0.050   0.034   0.021   0.033   0.061   0.037   0.017   0.020  
## 3    0.090   0.086   0.096   0.076   0.056   0.061   0.073   0.091  
## 4    0.162   0.158   0.153   0.146   0.168   0.111   0.130   0.145  
## 5    0.253   0.256   0.276   0.259   0.239   0.199   0.237   0.247  
## 6    0.360   0.350   0.382   0.371   0.316   0.314   0.386   0.404  
##   TK2143.1 TK2143.2 TK2143.3 TK2143.4 TK0887.TK1898.1 TK0887.TK1898.2  
## 1   -0.009  -0.015   0.040   0.004                -0.012                -0.007  
## 2    0.007   0.031   0.022   0.037                0.027                0.017  
## 3    0.066   0.072   0.069   0.062                0.068                0.080  
## 4    0.094   0.125   0.121   0.125                0.130                0.140  
## 5    0.179   0.246   0.231   0.226                0.310                0.225
```

```
## 6      0.314      0.398      0.399      0.404              0.310              0.324
##      TK0887.TK1898.3 TK0887.TK1898.4 TK0887.TK2143.1 TK0887.TK2143.2
## 1              0.002              0.012              0.003              -0.025
## 2              0.022              0.043              -0.001              0.006
## 3              0.077              0.082              0.037              0.037
## 4              0.073              0.080              0.070              0.107
## 5              0.245              0.294              0.133              0.143
## 6              0.344              0.468              0.232              0.267
##      TK0887.TK2143.3 TK0887.TK2143.4 TK1898.TK2143.1 TK1898.TK2143.2
## 1              -0.017              -0.009              -0.019              -0.009
## 2              0.006              0.003              0.000              0.006
## 3              0.033              0.035              0.029              0.032
## 4              0.073              0.080              0.063              0.075
## 5              0.144              0.153              0.113              0.124
## 6              0.256              0.274              0.219              0.246
##      TK1898.TK2143.3 TK1898.TK2143.4 TK0887.TK1898.TK2143.1 TK0887.TK1898.TK2143.2
## 1              -0.020              0.047              0.008              -0.025
## 2              0.010              0.073              0.030              -0.003
## 3              0.039              0.043              0.050              0.021
## 4              0.079              0.134              0.072              0.053
## 5              0.137              0.142              0.073              0.088
## 6              0.260              0.308              0.092              0.092
##      TK0887.TK1898.TK2143.3 TK0887.TK1898.TK2143.4
## 1              -0.001              -0.016
## 2              0.007              0.001
## 3              0.035              0.030
## 4              0.079              0.064
## 5              0.068              0.090
## 6              0.086              0.106
```

## 2. What happens when you use not tidy data.

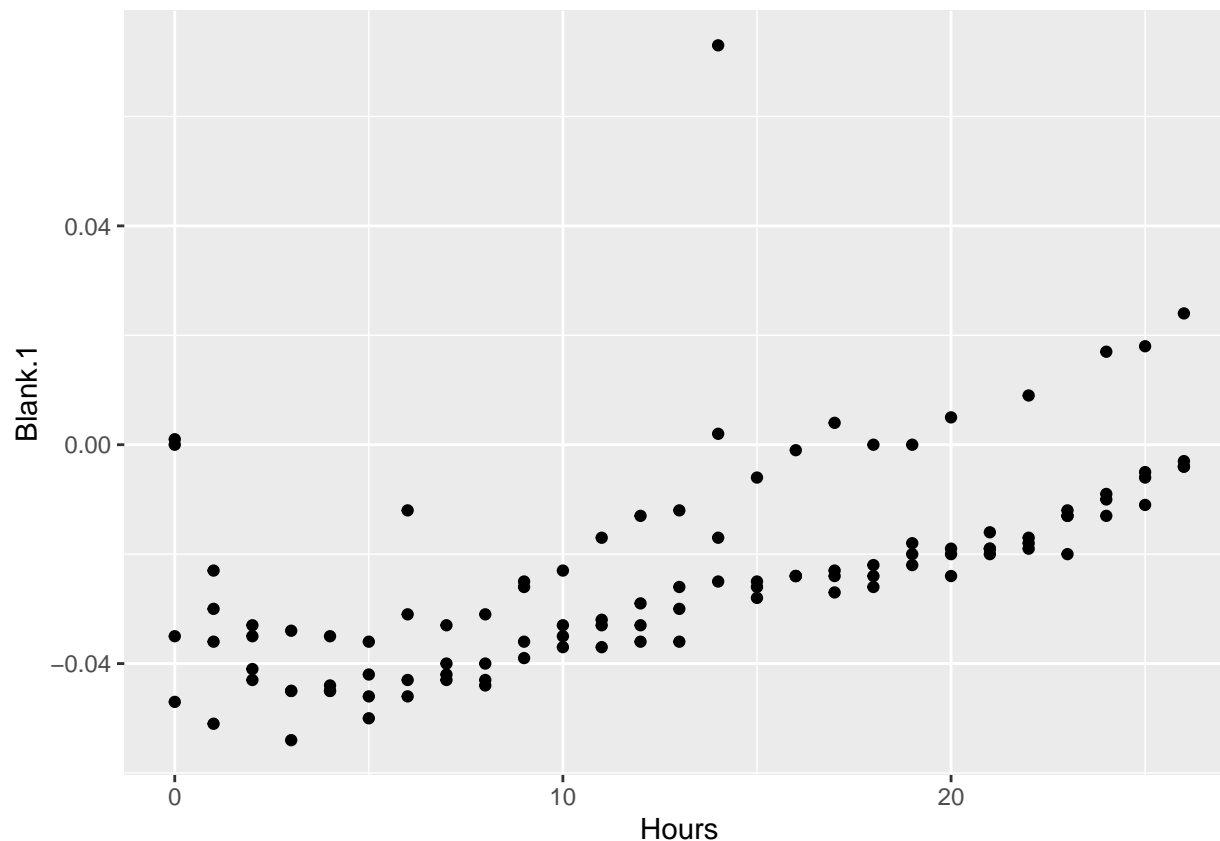
---

This data is not tidy, which means that we will have a lot lot lot of issues.

1. Because the strains and replicates are separated, I can only graph one strain at a time, one replicate at a time.

```
####
#02#
####

dU_GC%>%
  ggplot() +
  geom_point(aes(x=Hours, y=Blank.1)) +
  geom_point(aes(x=Hours, y=Blank.2)) +
  geom_point(aes(x=Hours, y=Blank.3)) +
  geom_point(aes(x=Hours, y=Blank.4))
```



## 2. What happens when you do statistics on not tidy data

I want to graph the average ODs for each strain at every hour.

If I am using the current data in not tidy format, this is the amount of statistical analysis I have to do:

1. Pivot the data longer and make individual tables for each strain that summarize the data. I basically have to make tidy data using R which, while doable, is a lot of unnecessary work.
2. I then have to combine every single table for each individual strain together using rbind.

```
####
#02#
####

GC_85_Blank <-dU_GC %>%
pivot_longer(cols = starts_with("Blank"),
              names_to = "Blank Replicates",
              values_to = "Blank Absorbance",
              values_drop_na = TRUE,
              values_transform = list(Absorbance=as.character)) %>%
select(c(Hours, `Blank Replicates`, `Blank Absorbance`)) %>%
group_by(Hours) %>%
summarize(mean = mean(`Blank Absorbance`), sd= sd(`Blank Absorbance`)) %>%
```

```

add_column(Strain = "Blank")

GC_85_TS559 <- dU_GC %>%
  pivot_longer(cols = starts_with("TS559"),
    names_to = "TS559 Replicates",
    values_to = "TS559 Absorbance",
    values_drop_na = TRUE) %>%
  select(c(Hours, `TS559 Replicates`, `TS559 Absorbance`)) %>%
  group_by(Hours) %>%
  summarize(mean = mean(`TS559 Absorbance`), sd = sd(`TS559 Absorbance`)) %>%
  add_column(Strain = "TS559")

GC_85_0887 <- dU_GC %>%
  pivot_longer(cols = starts_with("TK0887"),
    names_to = "0887 Replicates",
    values_to = "0887 Absorbance",
    values_drop_na = TRUE) %>%
  select(c(Hours, `0887 Replicates`, `0887 Absorbance`)) %>%
  group_by(Hours) %>%
  summarize(mean = mean(`0887 Absorbance`), sd = sd(`0887 Absorbance`)) %>%
  add_column(Strain = "TK0887-Endonuclease Q")

GC_85_1898 <- dU_GC %>%
  pivot_longer(cols = starts_with("TK1898"),
    names_to = "1898 Replicates",
    values_to = "1898 Absorbance",
    values_drop_na = TRUE) %>%
  select(c(Hours, `1898 Replicates`, `1898 Absorbance`)) %>%
  group_by(Hours) %>%
  summarize(mean = mean(`1898 Absorbance`), sd = sd(`1898 Absorbance`)) %>%
  add_column(Strain = "TK1898-Endonuclease MS")

GC_85_2143 <- dU_GC %>%
  pivot_longer(cols = starts_with("TK2143"),
    names_to = "2143 Replicates",
    values_to = "2143 Absorbance",
    values_drop_na = TRUE) %>%
  select(c(Hours, `2143 Replicates`, `2143 Absorbance`)) %>%
  group_by(Hours) %>%
  summarize(mean = mean(`2143 Absorbance`), sd = sd(`2143 Absorbance`)) %>%
  add_column(Strain = "TK2143-Uracil DNA Glycosylase")

GC_85_0887_1898 <- dU_GC %>%
  pivot_longer(cols = starts_with("TK0887.TK1898"),
    names_to = "TK0887/TK1898 Replicates",
    values_to = "TK0887/TK1898 Absorbance",
    values_drop_na = TRUE) %>%
  select(c(Hours, `TK0887/TK1898 Replicates`, `TK0887/TK1898 Absorbance`)) %>%
  group_by(Hours) %>%
  summarize(mean = mean(`TK0887/TK1898 Absorbance`), sd = sd(`TK0887/TK1898 Absorbance`)) %>%
  add_column(Strain = "TK0887/TK1898")

```

```

GC_85_0887_2143 <- dU_GC %>%
  pivot_longer(cols = starts_with("TK0887.TK2143"),
    names_to = "TK0887/TK2143 Replicates",
    values_to = "TK0887/TK2143 Absorbance",
    values_drop_na = TRUE) %>%
  select(c(Hours, `TK0887/TK2143 Replicates`, `TK0887/TK2143 Absorbance`)) %>%
  group_by(Hours) %>%
  summarize(mean = mean(`TK0887/TK2143 Absorbance`), sd = sd(`TK0887/TK2143 Absorbance`)) %>%
  add_column(Strain = "TK0887/TK2143")

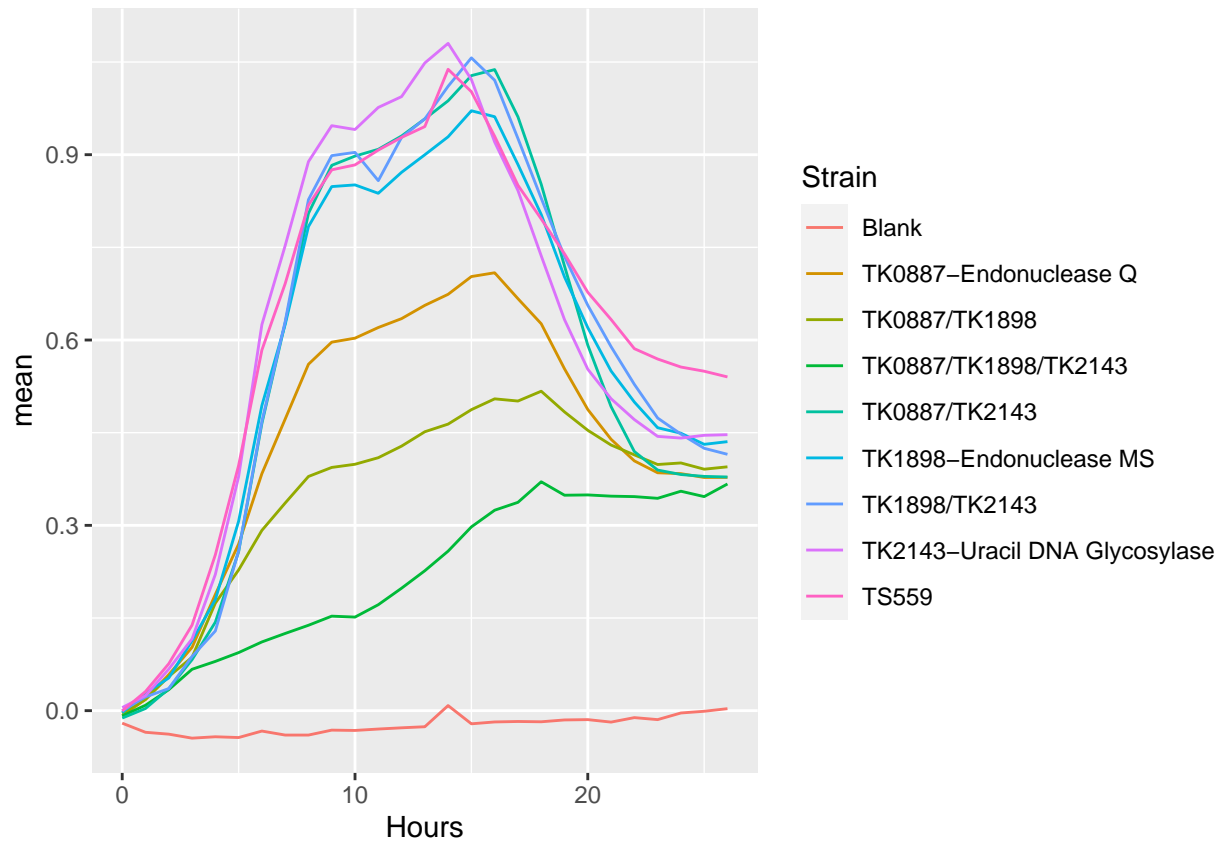
GC_85_1898_2143 <- dU_GC %>%
  pivot_longer(cols = starts_with("TK1898.TK2143"),
    names_to = "TK1898/TK2143 Replicates",
    values_to = "TK1898/TK2143 Absorbance",
    values_drop_na = TRUE) %>%
  select(c(Hours, `TK1898/TK2143 Replicates`, `TK1898/TK2143 Absorbance`)) %>%
  group_by(Hours) %>%
  summarize(mean = mean(`TK1898/TK2143 Absorbance`), sd = sd(`TK1898/TK2143 Absorbance`)) %>%
  add_column(Strain = "TK1898/TK2143")

GC_85_0887_1898_2143 <- dU_GC %>%
  pivot_longer(cols = starts_with("TK0887.TK1898.TK2143"),
    names_to = "TK0887/TK1898/TK2143 Replicates",
    values_to = "TK0887/TK1898/TK2143 Absorbance",
    values_drop_na = TRUE) %>%
  select(c(Hours, `TK0887/TK1898/TK2143 Replicates`, `TK0887/TK1898/TK2143 Absorbance`)) %>%
  group_by(Hours) %>%
  summarize(mean = mean(`TK0887/TK1898/TK2143 Absorbance`), sd = sd(`TK0887/TK1898/TK2143 Absorbance`)) %>%
  add_column(Strain = "TK0887/TK1898/TK2143")

# combine tables
dU_Table <- rbind(GC_85_Blank, GC_85_TS559)
dU_Table <- rbind(dU_Table, GC_85_0887)
dU_Table <- rbind(dU_Table, GC_85_1898)
dU_Table <- rbind(dU_Table, GC_85_2143)
dU_Table <- rbind(dU_Table, GC_85_0887_1898)
dU_Table <- rbind(dU_Table, GC_85_0887_2143)
dU_Table <- rbind(dU_Table, GC_85_1898_2143)
dU_Table_Final <- rbind(dU_Table, GC_85_0887_1898_2143)

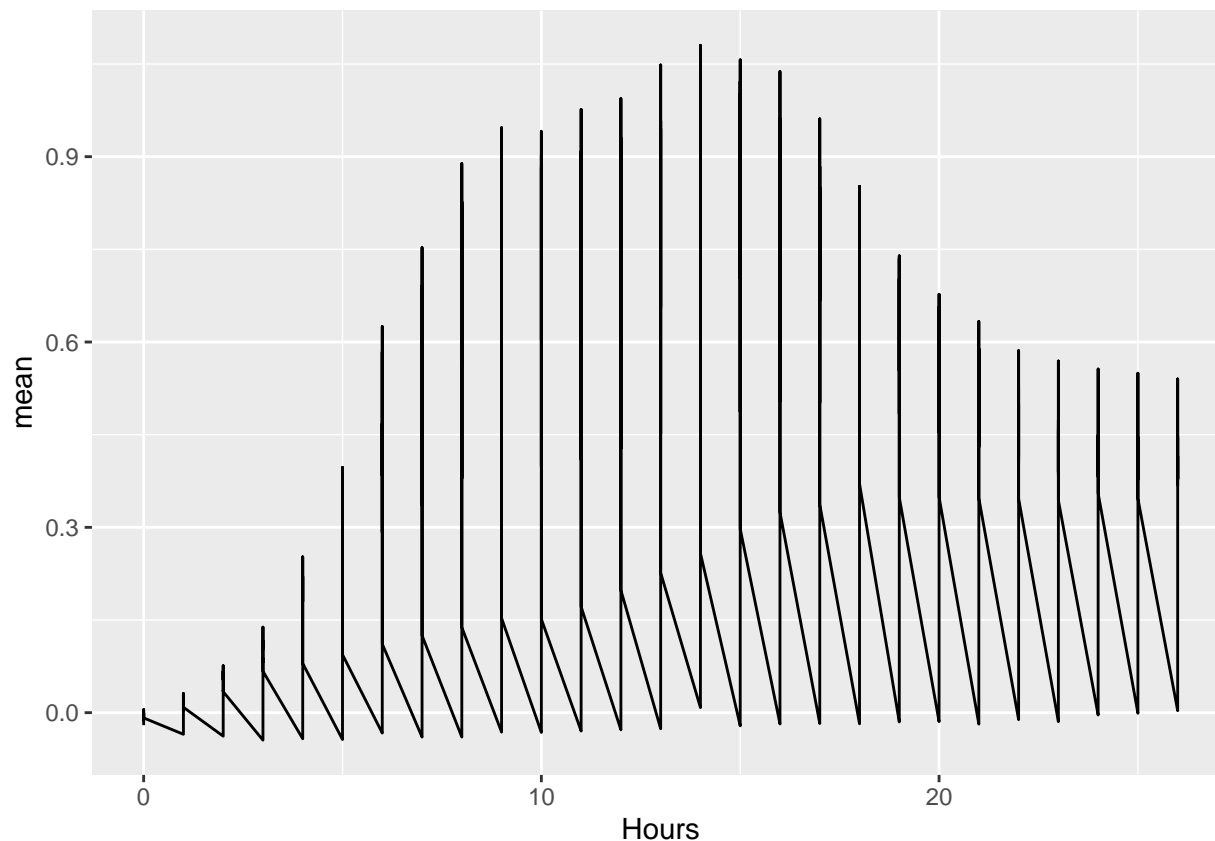
dU_Table_Final %>%
  ggplot() +
  geom_line(aes(x=Hours, y=mean, color=Strain))

```



Looks great right? But if you remove the color factor, what does the raw mapping actually look like?

```
dU_Table_Final %>%
  ggplot() +
  geom_line(aes(x=Hours, y=mean))
```



While this is a great way to learn data manipulation skills in R, this is very unnecessary.

Not only is that an insane amount of work (took me several days to figure out) and a lot of lines of code, the graph is factually incorrect.

Without factoring the data to a strain, the mapping is completely off.

Even when we factor the data, if we go back to the raw data, the lines for each strain are not right.

Because TK0887 is a repeating column, the 'startswith' argument cannot differentiate between each column that has the name TK0887.

This means way way more work, more time being frustrated, and less time making your graph awesome!

Let's make the data tidy, and then start over.

### 3. Working with tidy data. Mapping Aesthetics

We can complete the same statistical analysis in a SIGNIFICANTLY shorter time span.

```
####
#03#
####

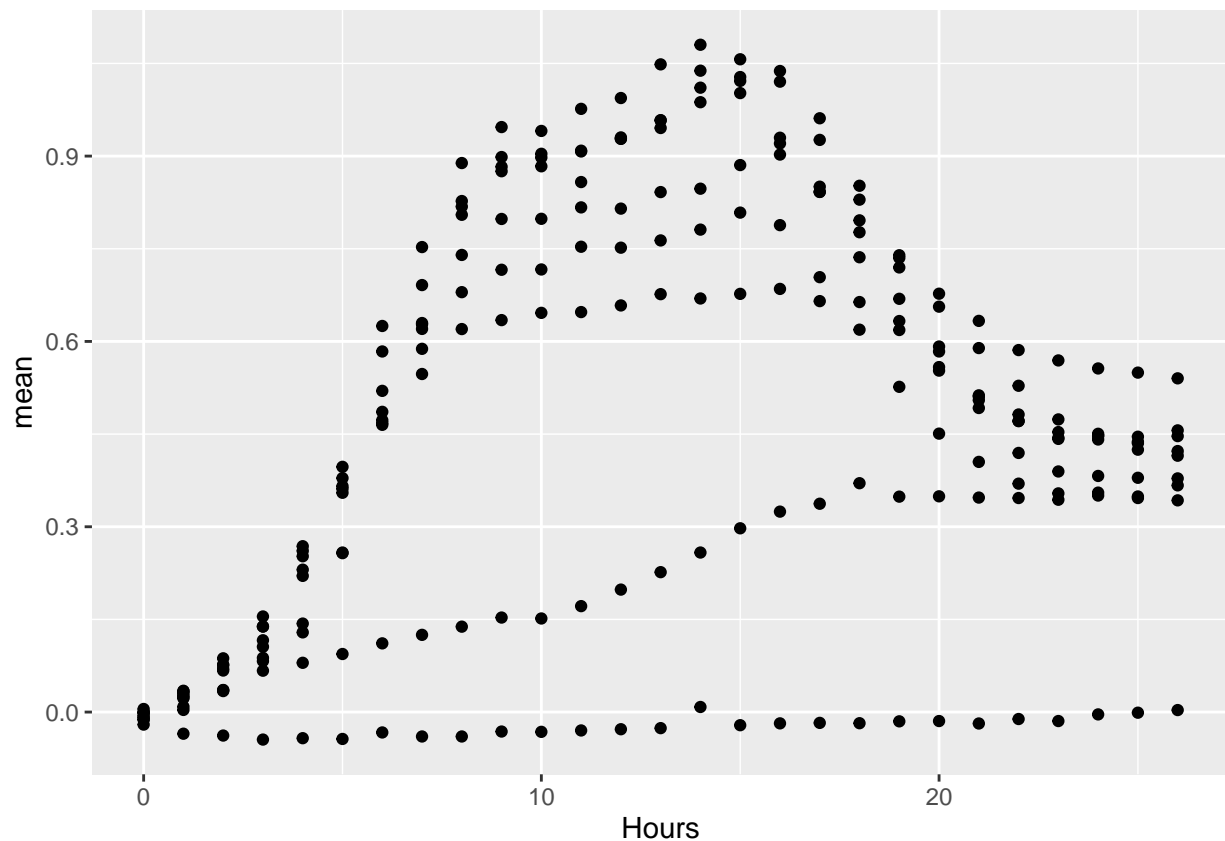
dU_GC_tidy<-read.csv("Growth_Curve_dU_tidy.csv",fill=TRUE,header=TRUE)

head(dU_GC_tidy)
```

```
##   Hours Strain Replicates    OD
## 1     0 Blank           1 0.001
## 2     1 Blank           1 -0.023
## 3     2 Blank           1 -0.043
## 4     3 Blank           1 -0.045
## 5     4 Blank           1 -0.045
## 6     5 Blank           1 -0.042
```

```
dU_GC_tidy_Stat <- dU_GC_tidy %>%
  group_by(Hours, Strain) %>%
  summarize(mean = mean(OD), sd = sd(OD)) %>%
  group_by(Strain)

dU_GC_tidy_Stat%>%
  ggplot() +
  geom_point(aes(x=Hours, y=mean))
```



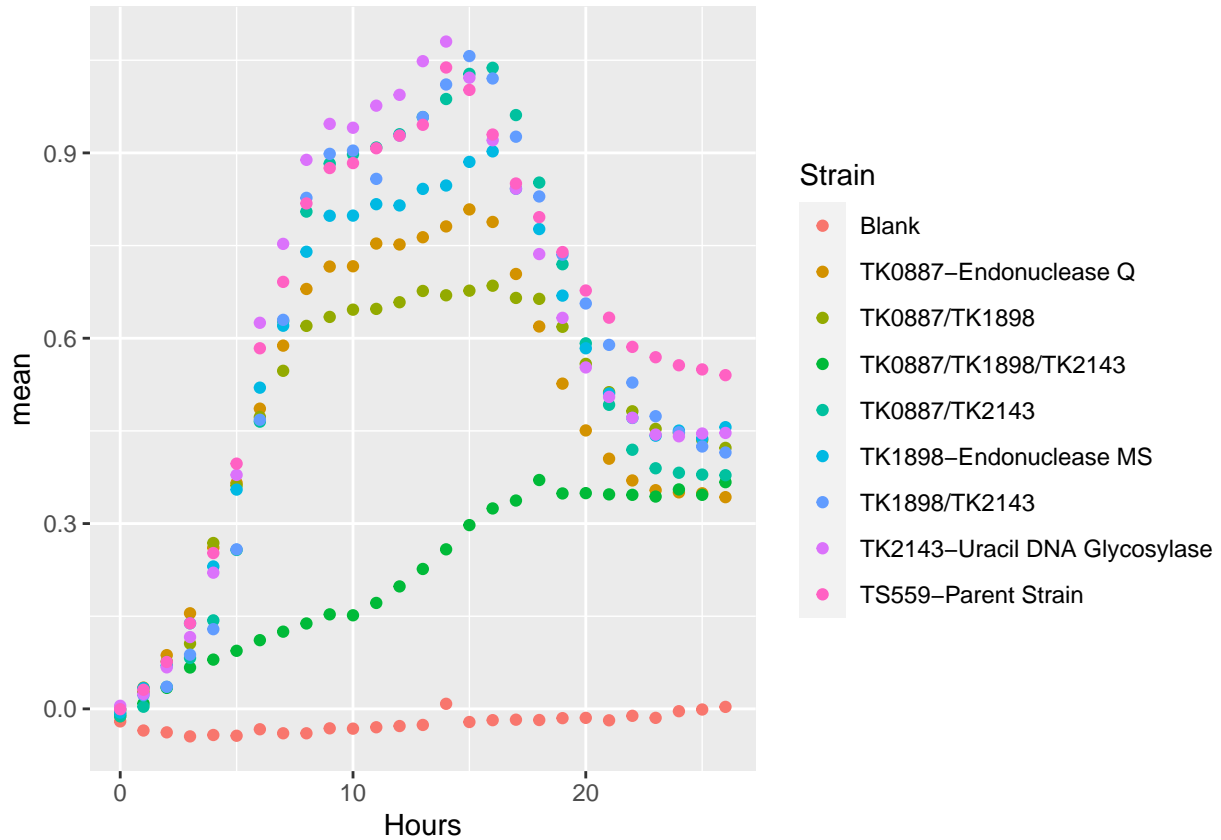
## 4. Factoring the Mapping

Now the mapping looks great! We have given R readable data that it can properly interpret into a graph.

However, we have no idea which set of points are related to which strain, so let's add to our mapping some kind of factor to differentiate the Strains.



```
####  
#O4#  
####  
  
dU_GC_tidy_Stat%>%  
  ggplot() +  
  geom_point(aes(x=Hours, y=mean, color=Strain))
```



## # 5. Factoring the Mapping

Our legend labels do not say in any way that these are gene deletions. I did not include deltas because R would not interpret this from a csv file.

I recommend always fixing the names of columns, Rows, and categorical variables before graphing because afterwards it is a lot of code to fix those issues.

We should put a delta sign in front of each gene number, but how?

Fortunately, R can interpret unicode.

Using labels, we can rename our legend to include deltas in front of each gene.

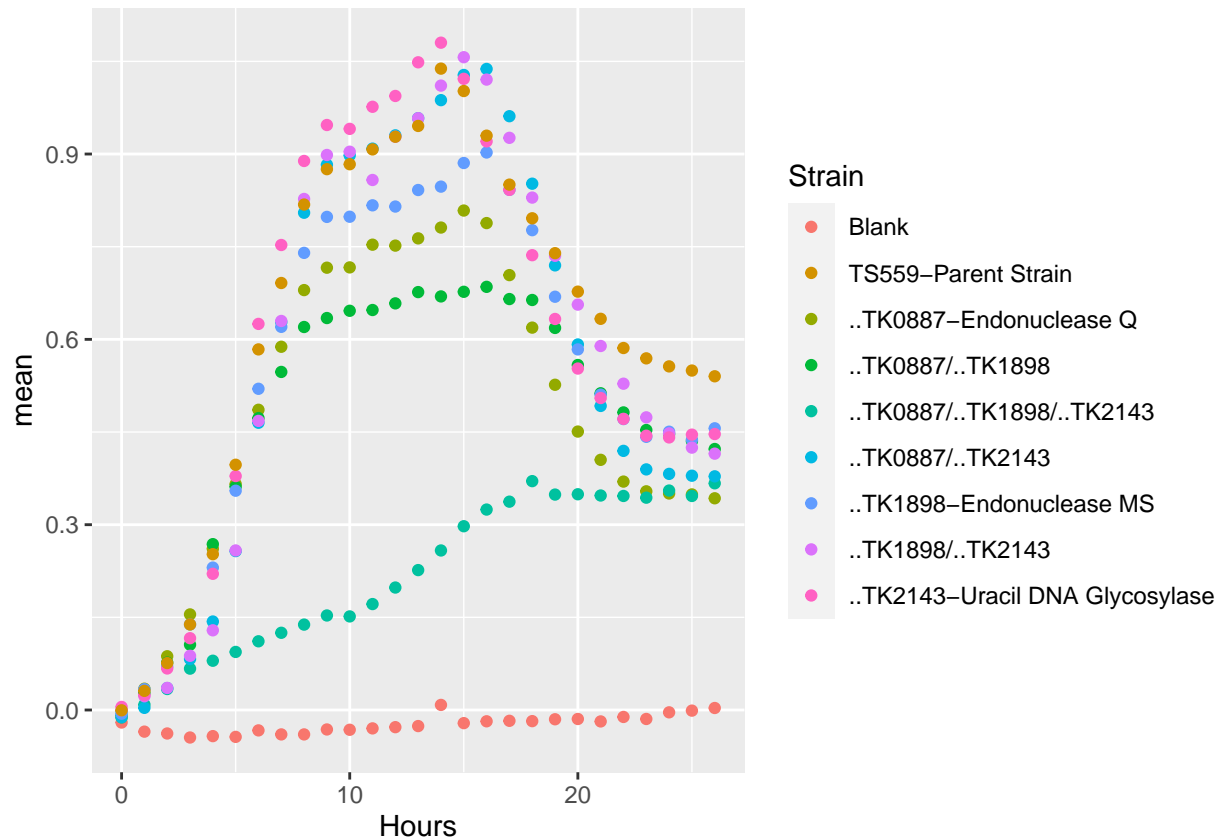
[https://en.wikipedia.org/wiki/List\\_of\\_Unicode\\_characters](https://en.wikipedia.org/wiki/List_of_Unicode_characters)

You can find all greek letters under Greek and Coptic. We put the unicode right into the string (unicode for delta is 0394)

####  
#05#  
####

```
dU_GC_tidy_Stat <- dU_GC_tidy_Stat %>%
  mutate(Strain = recode(Strain, "TK0887-Endonuclease Q" = "\u0394TK0887-Endonuclease Q", "TK1898-Endonuclease Q" = "\u0394TK1898-Endonuclease Q"))

dU_GC_tidy_Stat %>%
  ggplot() +
  geom_point(aes(x=Hours, y=mean, color=Strain))
```



## 6. Picking the right geometry

Now the data has received proper statistics and been mapped.

We can now tell the dots apart, but, maybe a scatter plot is not the best way to represent this data.

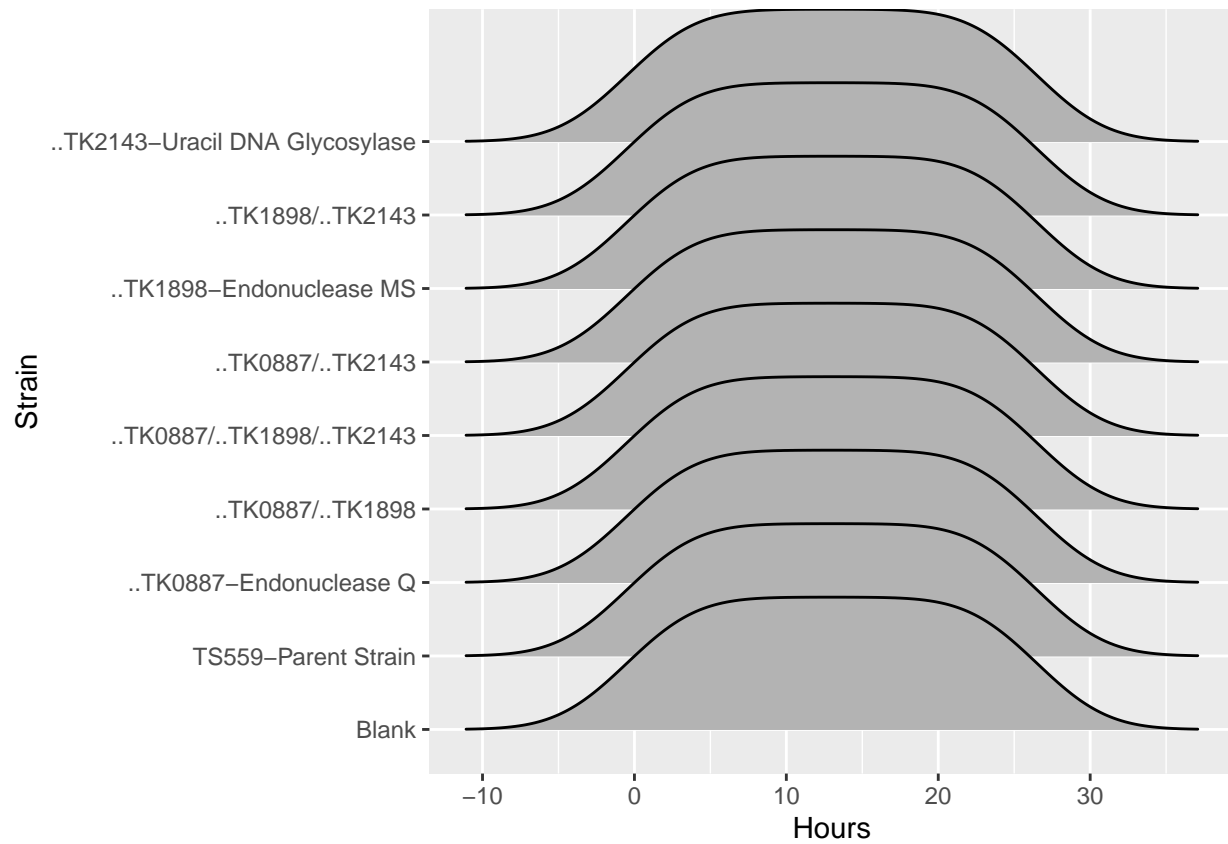
Lets look at a few different geometries.

```
####
#06#
####

#install.packages("ggribes")
library(ggribes)

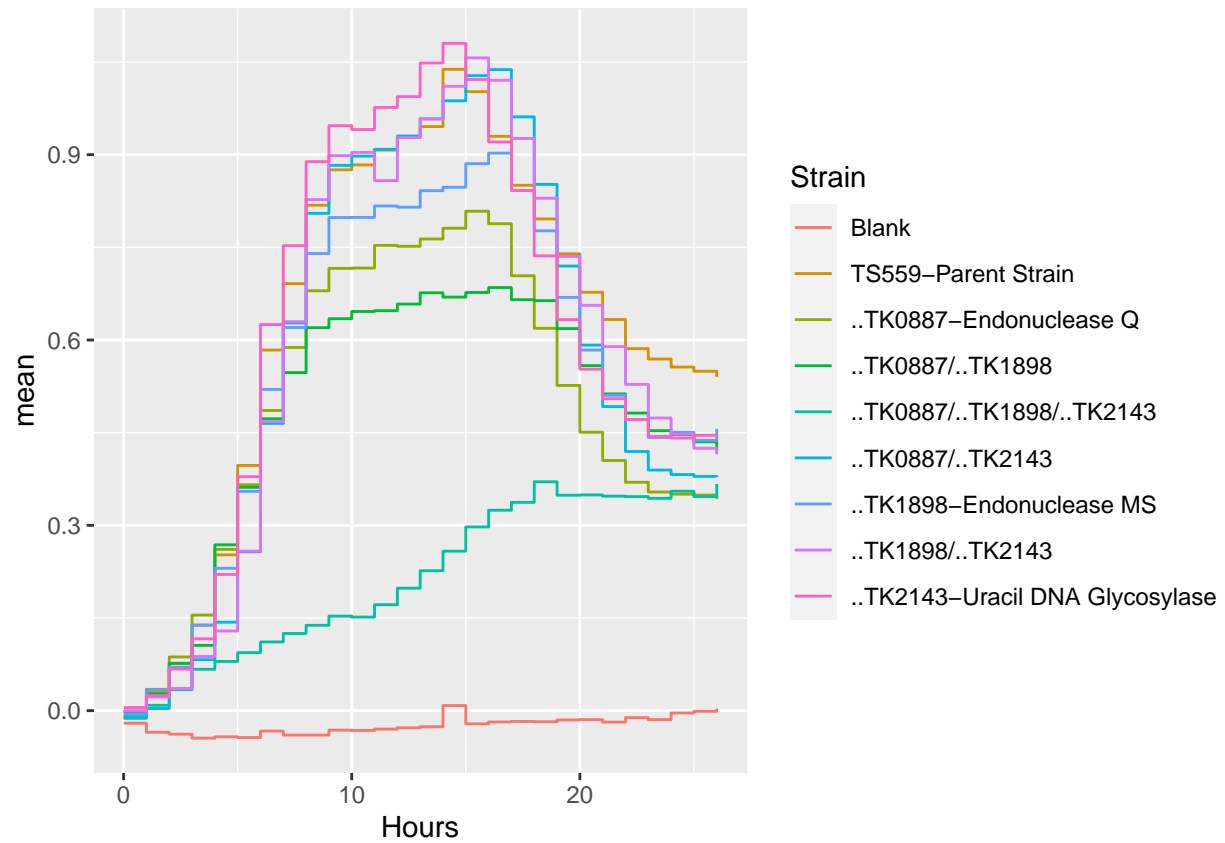
dU_GC_tidy_Stat %>%
```

```
ggplot() +  
  geom_density_ridges_gradient(aes(x=Hours, y= Strain))
```



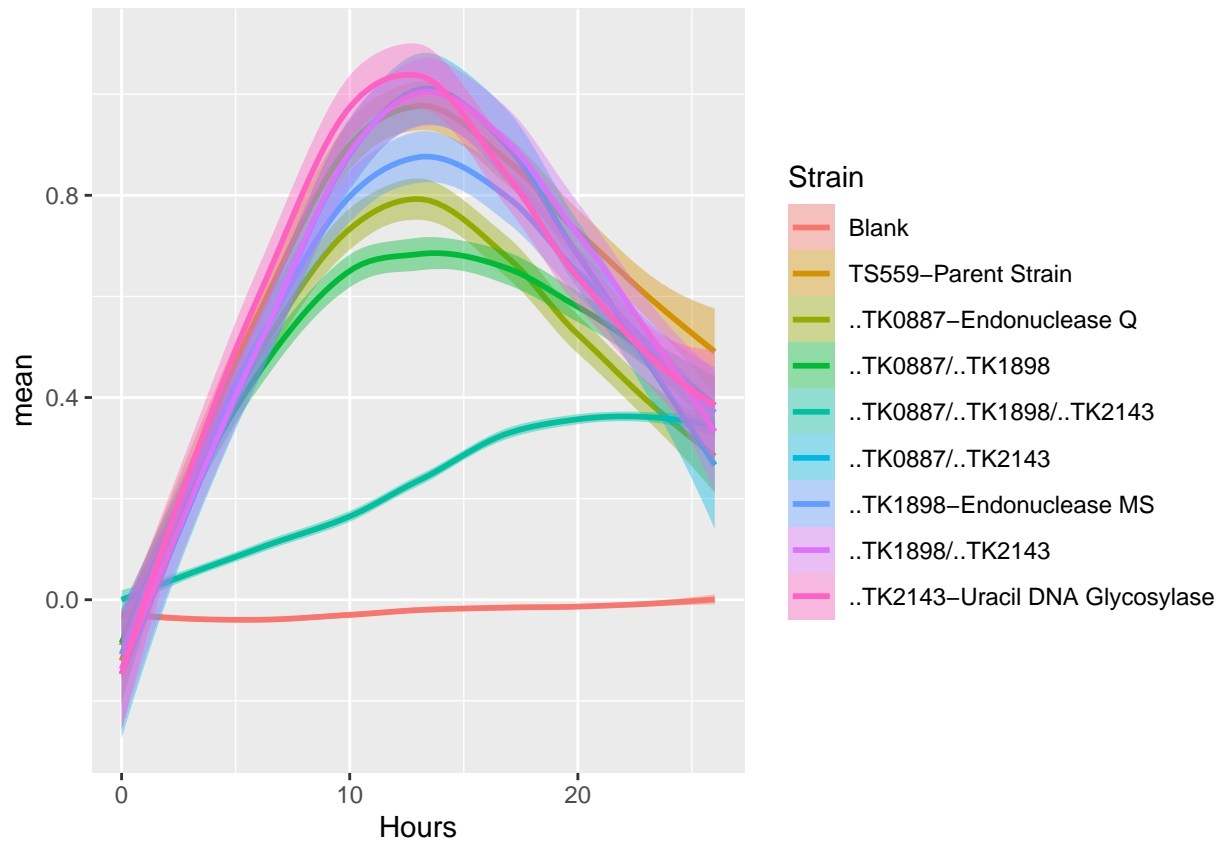
The {ggridges} package draws density curves which are by definition scaled to an area of 1. The height of the ridges therefore cannot be mapped to the mean OD so we must use a different graph.

```
dU_GC_tidy_Stat%>%  
  ggplot() +  
  geom_step(aes(x=Hours, y=mean, color=Strain, fill=Strain))
```



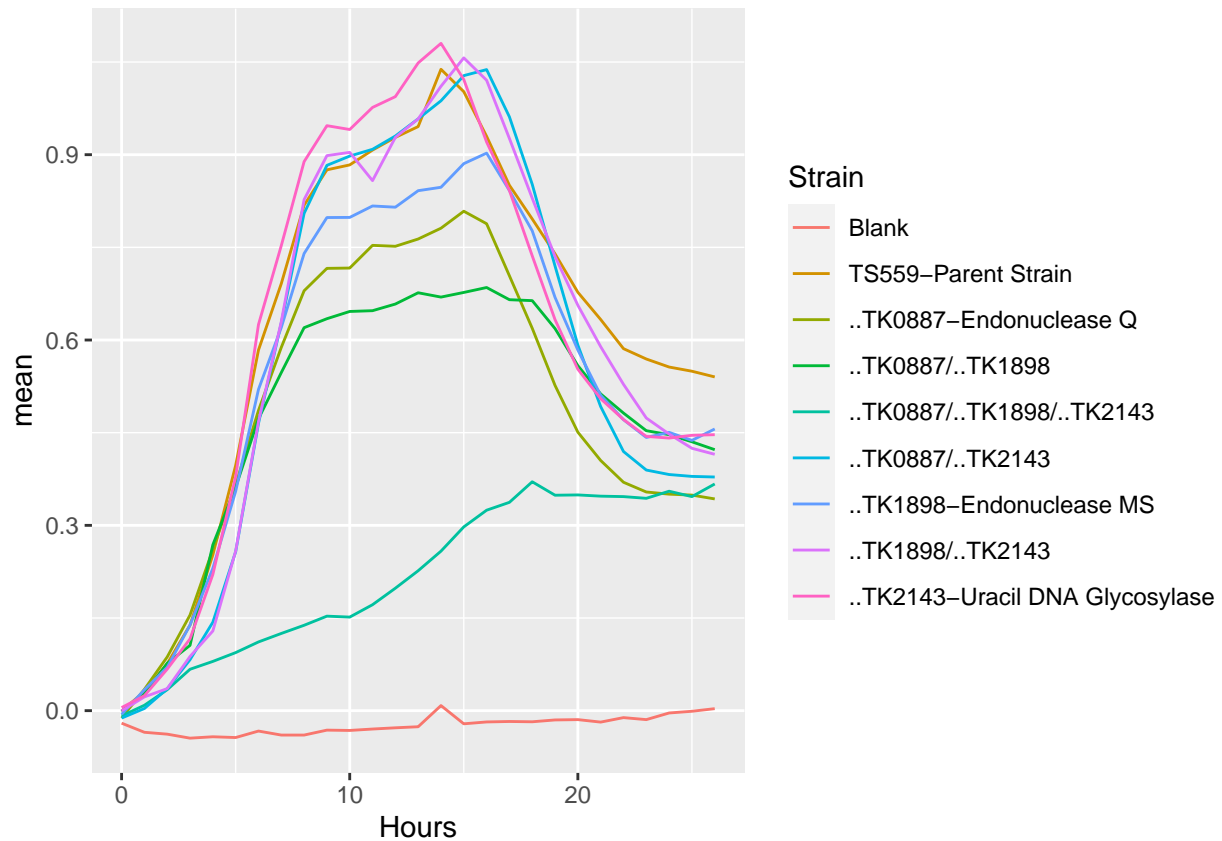
Makes a step-wise line go through each point, but in this context it is quite ugly.

```
dU_GC_tidy_Stat%>%
  ggplot() +
  geom_smooth(aes(x=Hours, y=mean, color=Strain, fill=Strain))
```



This one is pretty great when you are only measuring growth for a few strains, but with nine it looks overwhelming.

```
dU_GC_tidy_Stat%>%
  ggplot() +
  geom_line(aes(x=Hours, y=mean, color=Strain))
```



*#Perfect!*

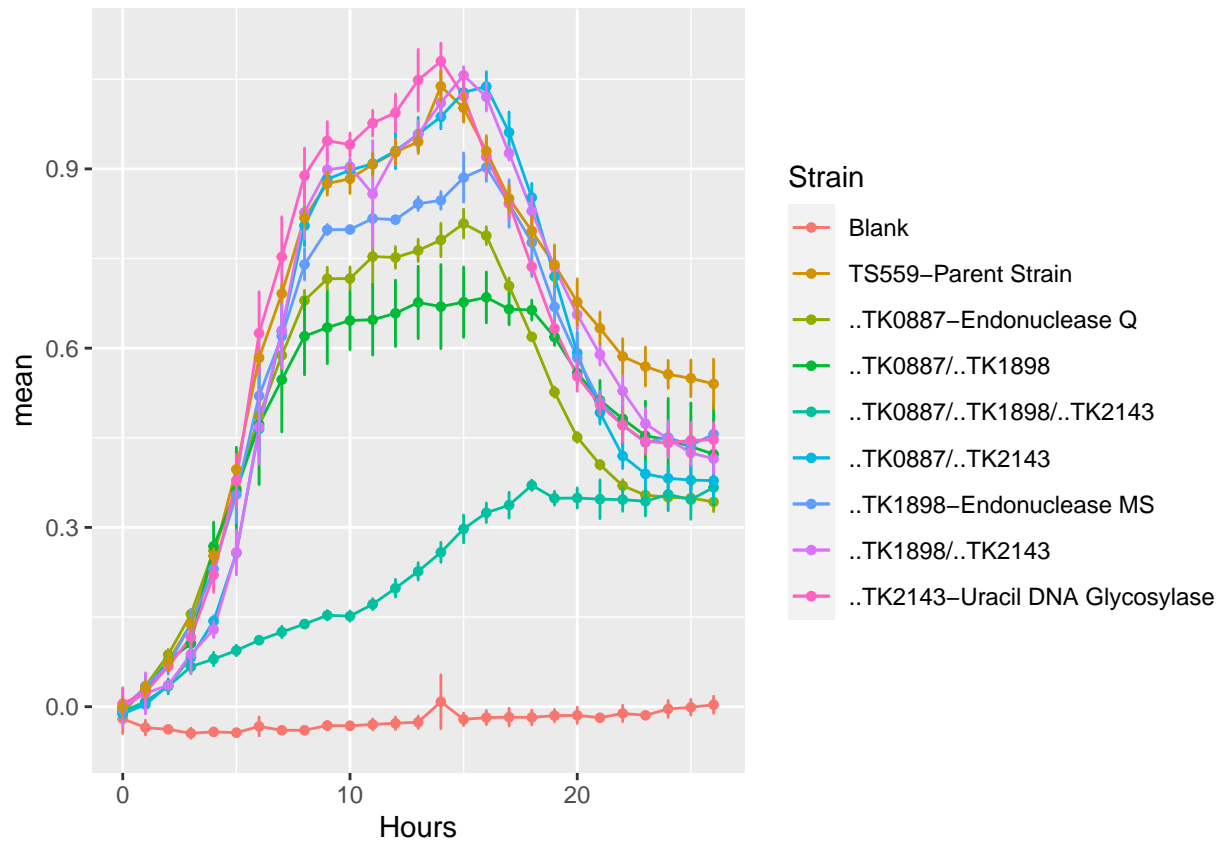
## 6b. Overlapping Geometries

While the line follows the trend, it is not a bad idea to include the points where the data are.

We can also include error bars.

```
####
#06b#
####

dU_GC_tidy_Stat%>%
  ggplot() +
  geom_line(aes(x=Hours, y=mean, color=Strain), size=0.5) +
  geom_point(aes(x=Hours, y=mean, color=Strain), size=1.2) +
  geom_errorbar(width=0.1, aes(x=Hours, ymax = mean+sd, ymin = mean-sd, color=Strain))
```



## 7. Labels

Now this graph is starting to display our data appropriately.

Let's add changing the axis labels

We want to specify that the y axis is optical density at 600nm. We also want to specify that these strains are deletions with a delta.

Bquote allows us to customize our text to be bold, italicized and underlined. We can also add superscripts, subscripts, and many other components.

For more on bquote-use `help("plotmath")`

```
####
#07#
####

help("plotmath")

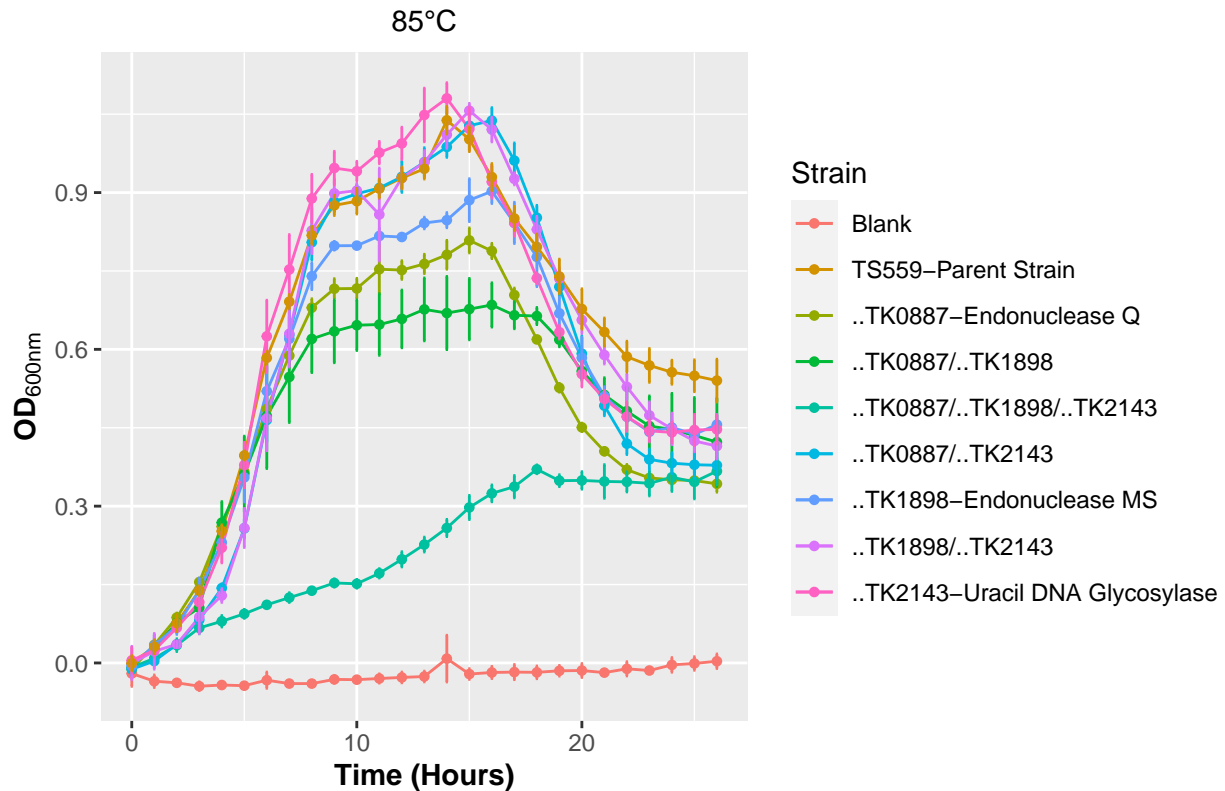
dU_GC_tidy_Stat%>%
  ggplot() +
  geom_line(aes(x=Hours, y=mean, color=Strain), size=0.5) +
  geom_point(aes(x=Hours, y=mean, color=Strain), size=1.2) +
  geom_errorbar(width=0.1, aes(x=Hours, ymax = mean+sd, ymin = mean-sd, color=Strain)) +
  labs(title = "Only deletion of all dU repair components impacts growth", subtitle="85\u00B0C",
```

```

y = bquote(bold("OD") [600] [nm]),
x = bquote(bold('Time (Hours)')) +
theme(plot.title = element_text(hjust = 0.5), plot.subtitle= element_text(hjust = 0.5))

```

Only deletion of all dU repair components impacts growth



## 8. Scales

The scales are currently not very informative.

Instead of every 10 hours, lets look at every 4 hours.

```

####
#08#
####

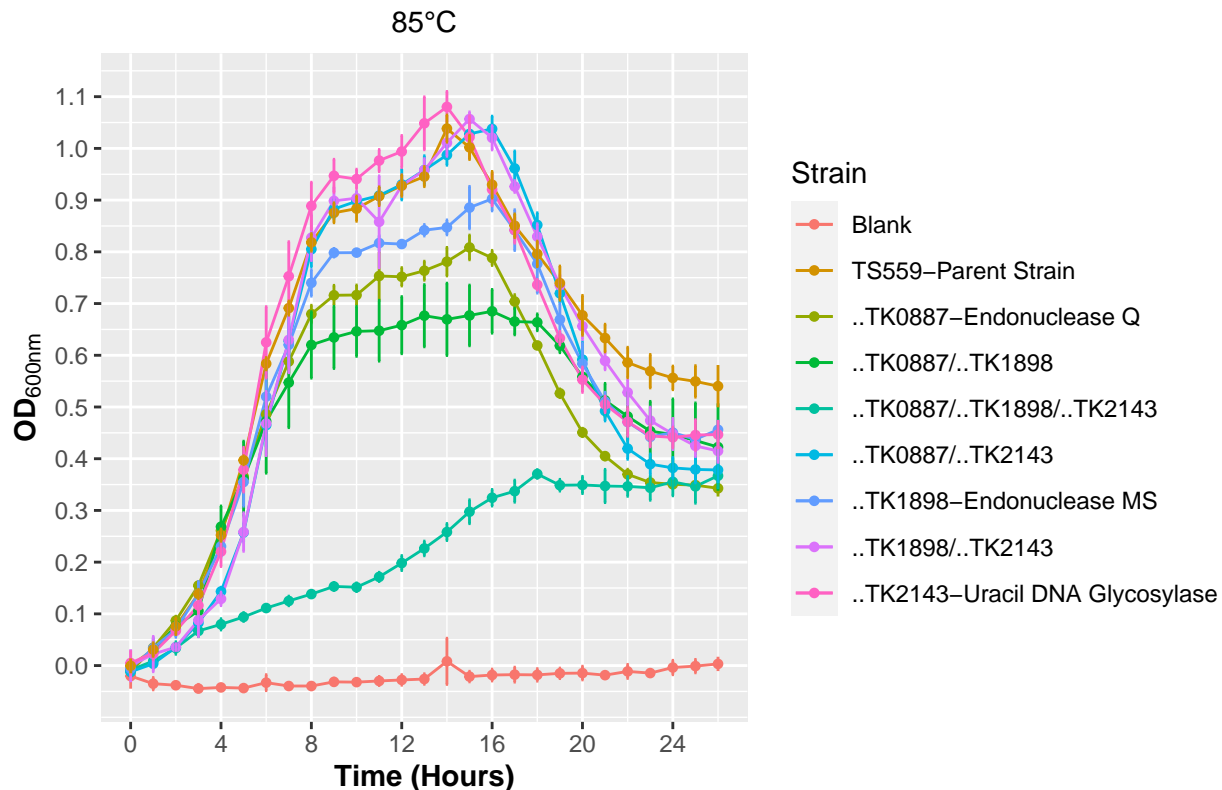
dU_GC_tidy_Stat%>%
  ggplot() +
  geom_line(aes(x=Hours, y=mean, color=Strain), size=0.5) +
  geom_point(aes(x=Hours, y=mean, color=Strain), size=1.2) +
  geom_errorbar(width=0.1, aes(x=Hours, ymax = mean+sd, ymin = mean-sd, color=Strain)) +
  labs(title = "Only deletion of all dU repair components impacts growth", subtitle="85\u00B0C",
    y = bquote(bold("OD") [600] [nm]),
    x = bquote(bold('Time (Hours)')) +
  theme(plot.title = element_text(hjust = 0.5), plot.subtitle= element_text(hjust = 0.5)) +

```



```
scale_x_continuous(breaks=c(0, 4, 8, 12, 16, 20, 24), limits=c(0,26)) +
scale_y_continuous(breaks=c(0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0, 1.1), limits=c(-0.05
```

Only deletion of all dU repair components impacts growth



## 8b. Fitting the Data

The scales are better... but the data does not start at 0,0. It also goes out past 26 to empty space.

We can mess around with the limits, or we can use coordinates.

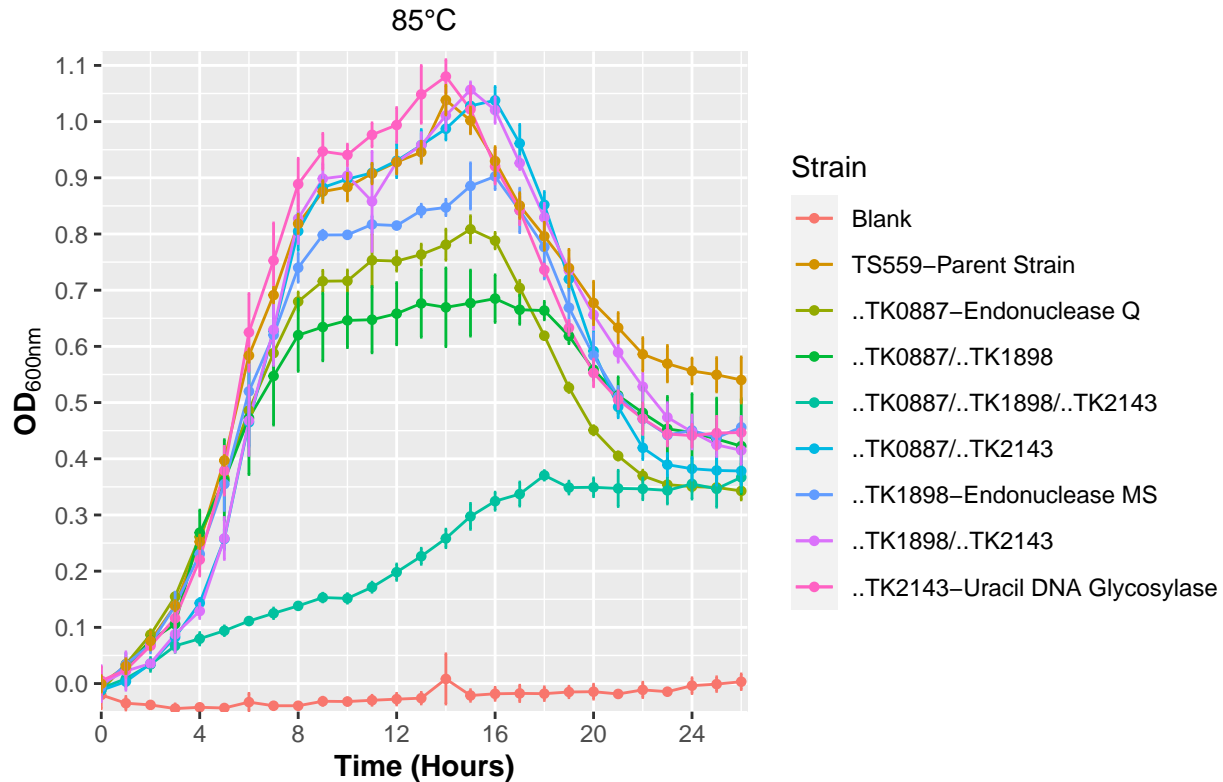
We can also add an expand argument to our continuous scales to set the data at the origin.

```
####
#08b#
####

dU_GC_tidy_Stat%>%
  ggplot() +
  geom_line(aes(x=Hours, y=mean, color=Strain), size=0.5) +
  geom_point(aes(x=Hours, y=mean, color=Strain), size=1.2) +
  geom_errorbar(width=0.1, aes(x=Hours, ymax = mean+sd, ymin = mean-sd, color=Strain)) +
  labs(title = bquote(bold("Only deletion of all dU repair components impacts growth")), subtitle="85\u00b0C",
       y = bquote(bold("OD") [600] [nm]),
       x = bquote(bold('Time (Hours)')))) +
  theme(plot.title = element_text(hjust = 0.5), plot.subtitle= element_text(hjust = 0.5)) +
```

```
scale_x_continuous(breaks=c(0, 4, 8, 12, 16, 20, 24), expand=c(0,0)) +
scale_y_continuous(breaks=c(0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0, 1.1), expand=c(0,0))
coord_cartesian(xlim=c(0,26.25), ylim=c(-0.05, 1.125))
```

## ly deletion of all dU repair components impacts growth



## 8c. Transforming to log scale

That looks good!

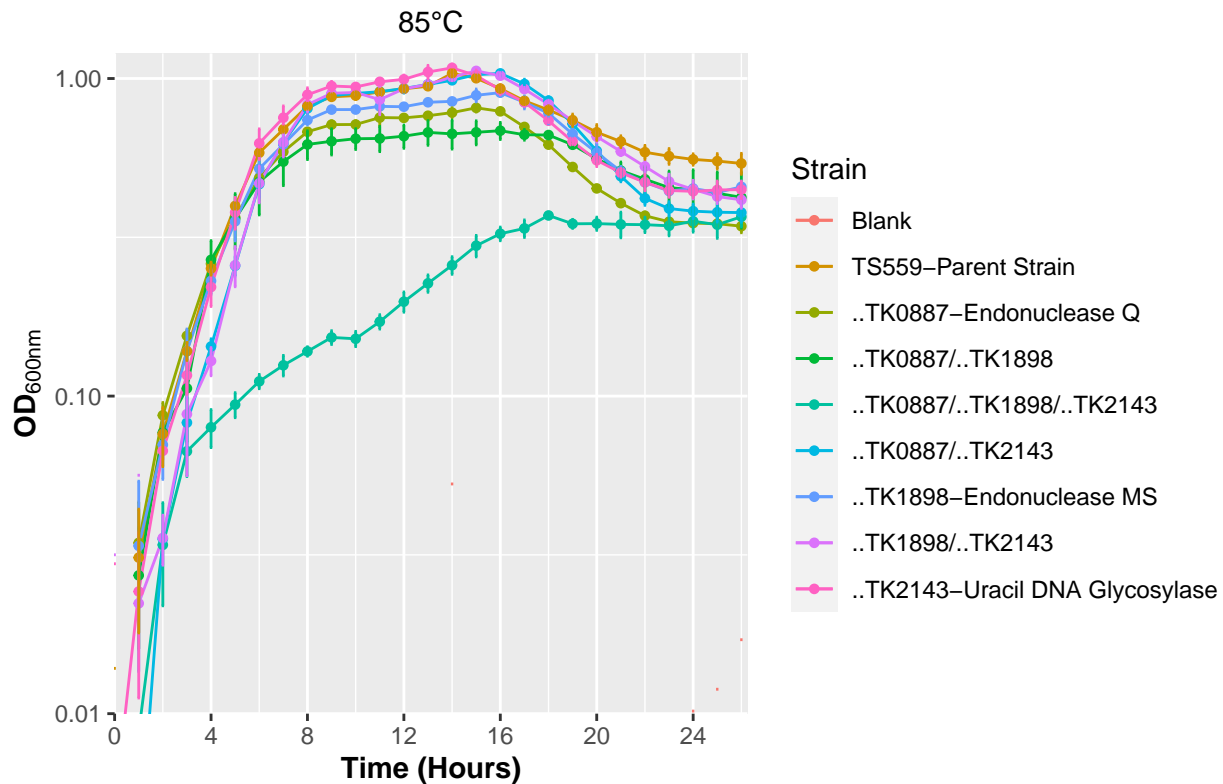
We could absolutely present the data this way. This will straighten out exponential growth patterns and reduce heteroscedasticity (i.e., stabilizes variance). Publishers will definitely prefer the log scale.

```
####
#08c#
####

dU_GC_tidy_Stat%>%
  ggplot() +
  geom_line(aes(x=Hours, y=mean, color=Strain), size=0.5) +
  geom_point(aes(x=Hours, y=mean, color=Strain), size=1.2) +
  geom_errorbar(width=0.1, aes(x=Hours, ymax = mean+sd, ymin = mean-sd, color=Strain)) +
  labs(title = bquote(bold("Only deletion of all dU repair components impacts growth")), subtitle="85\u00b0C",
    y = bquote(bold("OD" [600] [nm])),
    x = bquote(bold("Time (Hours)")) +
  theme(plot.title = element_text(hjust = 0.5), plot.subtitle= element_text(hjust = 0.5)) +
```

```
scale_x_continuous(breaks=c(0, 4, 8, 12, 16, 20, 24), expand=c(0,0)) +
scale_y_continuous(breaks=c(0.01, 0.1, 1.0), expand=c(0,0), trans='log10') +
coord_cartesian(xlim=c(0,26.25), ylim=c(0.01, 1.2))
```

## Only deletion of all dU repair components impacts growth



## 9. Scales-Colors

Right now the legend is pretty out of order and the rainbow of colors is hard to follow.

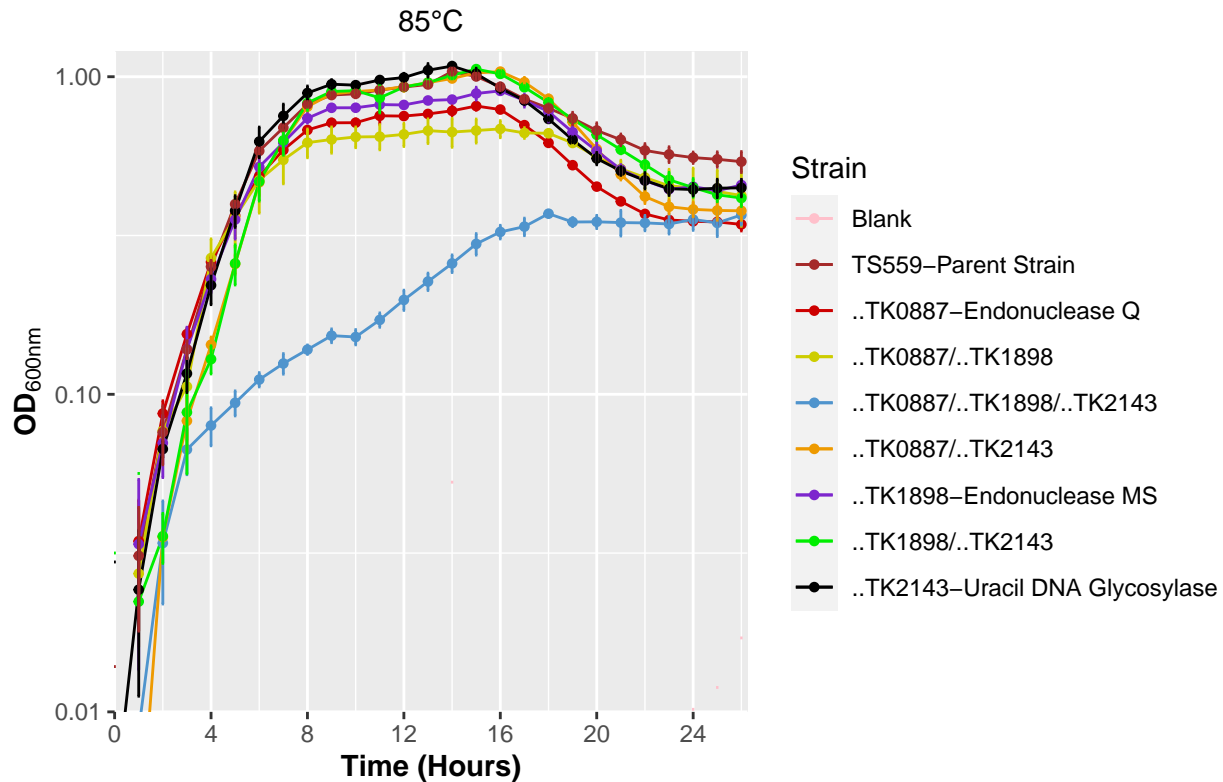
Let's make a color palette we can work with!

```
####
#09#
####

dU_GC_tidy_Stat%>%
  ggplot() +
  geom_line(aes(x=Hours, y=mean, color=Strain), size=0.5) +
  geom_point(aes(x=Hours, y=mean, color=Strain), size=1.2) +
  geom_errorbar(width=0.1, aes(x=Hours, ymax = mean+sd, ymin = mean-sd, color=Strain)) +
  labs(title = bquote(bold("Only deletion of all dU repair components impacts growth")), subtitle="85°C",
       y = bquote(bold("OD") [600] [nm]),
       x = bquote(bold("Time (Hours)")) +
  theme(plot.title = element_text(hjust = 0.5), plot.subtitle= element_text(hjust = 0.5)) +
  scale_x_continuous(breaks=c(0, 4, 8, 12, 16, 20, 24), expand=c(0,0)) +
```

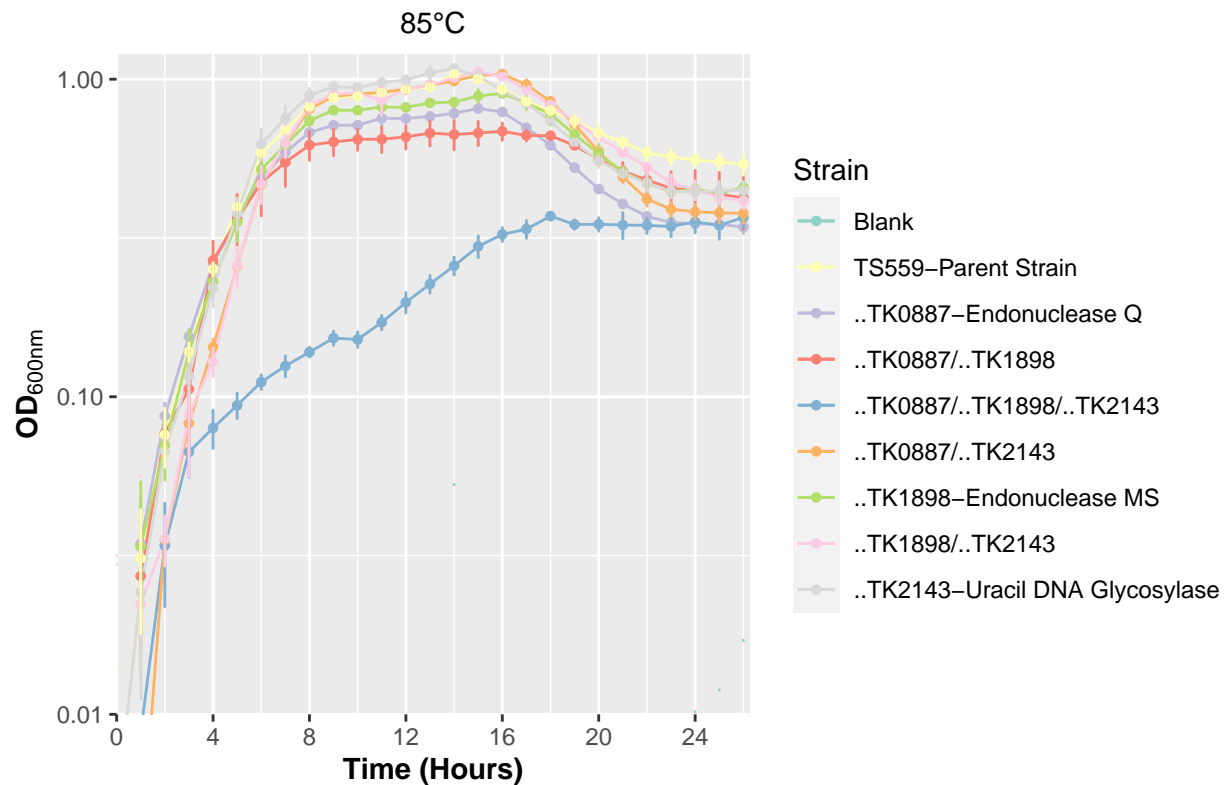
```
scale_y_continuous(breaks=c(0.01, 0.1, 1.0), expand=c(0,0), trans='log10') +
coord_cartesian(xlim=c(0,26.25), ylim=c(0.01, 1.2)) +
scale_color_manual(values=c("pink", "brown", "red3", "yellow3", "steelblue3", "orange2", "purple3", "black"))
```

## Only deletion of all dU repair components impacts growth



```
dU_GC_tidy_Stat%>%
ggplot() +
geom_line(aes(x=Hours, y=mean, color=Strain), size=0.5) +
geom_point(aes(x=Hours, y=mean, color=Strain), size=1.2) +
geom_errorbar(width=0.1, aes(x=Hours, ymax = mean+sd, ymin = mean-sd, color=Strain)) +
labs(title = bquote(bold("Only deletion of all dU repair components impacts growth")), subtitle="85°C",
      y = bquote(bold("OD") [600] [nm]),
      x = bquote(bold('Time (Hours)')))) +
theme(plot.title = element_text(hjust = 0.5), plot.subtitle= element_text(hjust = 0.5)) +
scale_x_continuous(breaks=c(0, 4, 8, 12, 16, 20, 24), expand=c(0,0)) +
scale_y_continuous(breaks=c(0.01, 0.1, 1.0), expand=c(0,0), trans='log10') +
coord_cartesian(xlim=c(0,26.25), ylim=c(0.01, 1.2)) +
scale_color_brewer(palette="Set3")
```

### ly deletion of all dU repair components impacts growth



# 9b. Set your colors to your legend

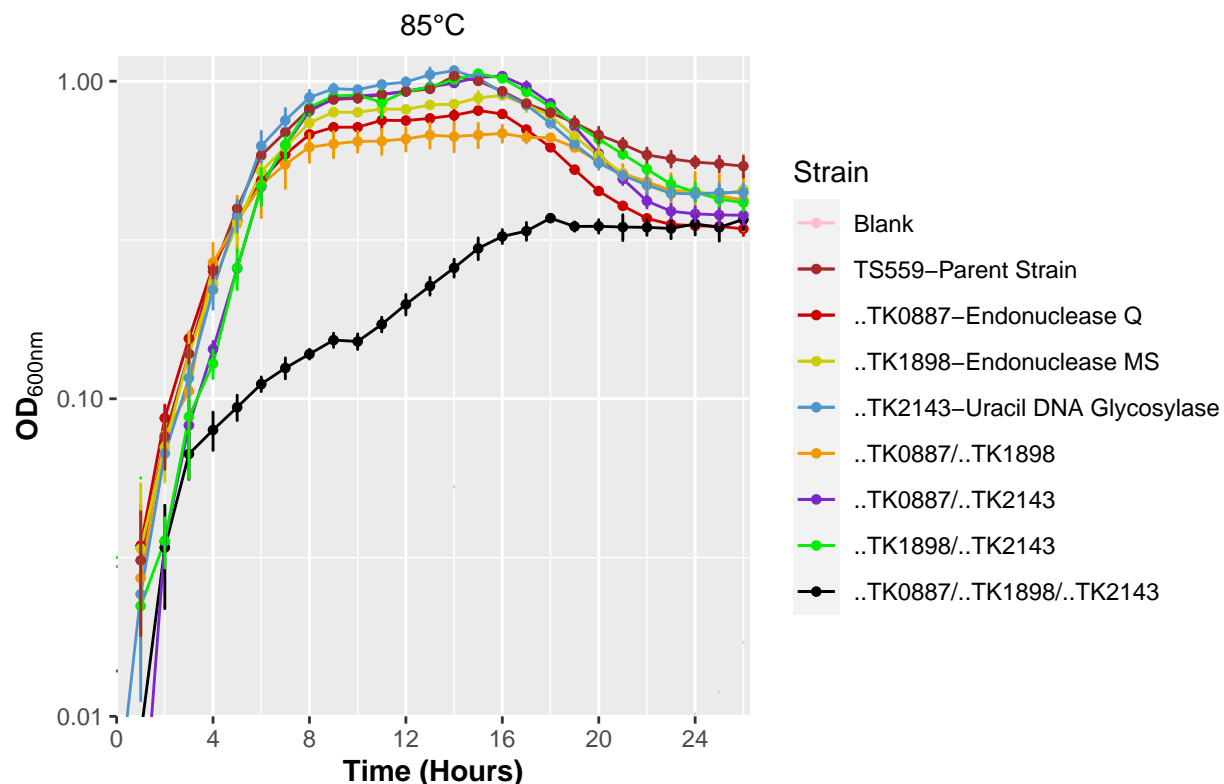
Now we have our own colors, but the legend is not in any logical order!

As with `scale_x_continuous`, we can also add breaks to `scale_color_manual` to reorder our legend names to our colors.

```
#####
#09b#
#####

dU_GC_tidy_Stat%>%
  ggplot() +
  geom_line(aes(x=Hours, y=mean, color=Strain), size=0.5) +
  geom_point(aes(x=Hours, y=mean, color=Strain), size=1.2) +
  labs(title = bquote(bold("Only deletion of all dU repair components impacts.")), subtitle="85\u00B0C"
        y = bquote(bold("OD")[600][nm]),
        x = bquote(bold('Time (Hours)')))) +
  theme(plot.title = element_text(hjust = 0.5), plot.subtitle= element_text(hjust = 0.5)) +
  scale_x_continuous(breaks=c(0, 4, 8, 12, 16, 20, 24), expand=c(0,0)) +
  scale_y_continuous(breaks=c(0.01, 0.1, 1.0), expand=c(0,0), trans='log10') +
  coord_cartesian(xlim=c(0,26.25), ylim=c(0.01, 1.2)) +
  scale_color_manual(values=c("pink", "brown", "red3", "yellow3", "steelblue3", "orange2", "purple3", "black")) +
  geom_errorbar(width=0.1, aes(x=Hours, ymax = mean+sd, ymin = mean-sd, color=Strain)) +
  labs(title = bquote(bold("Only deletion of all dU repair components impacts growth")), subtitle="85\u00B0C")
```

## ly deletion of all dU repair components impacts growth



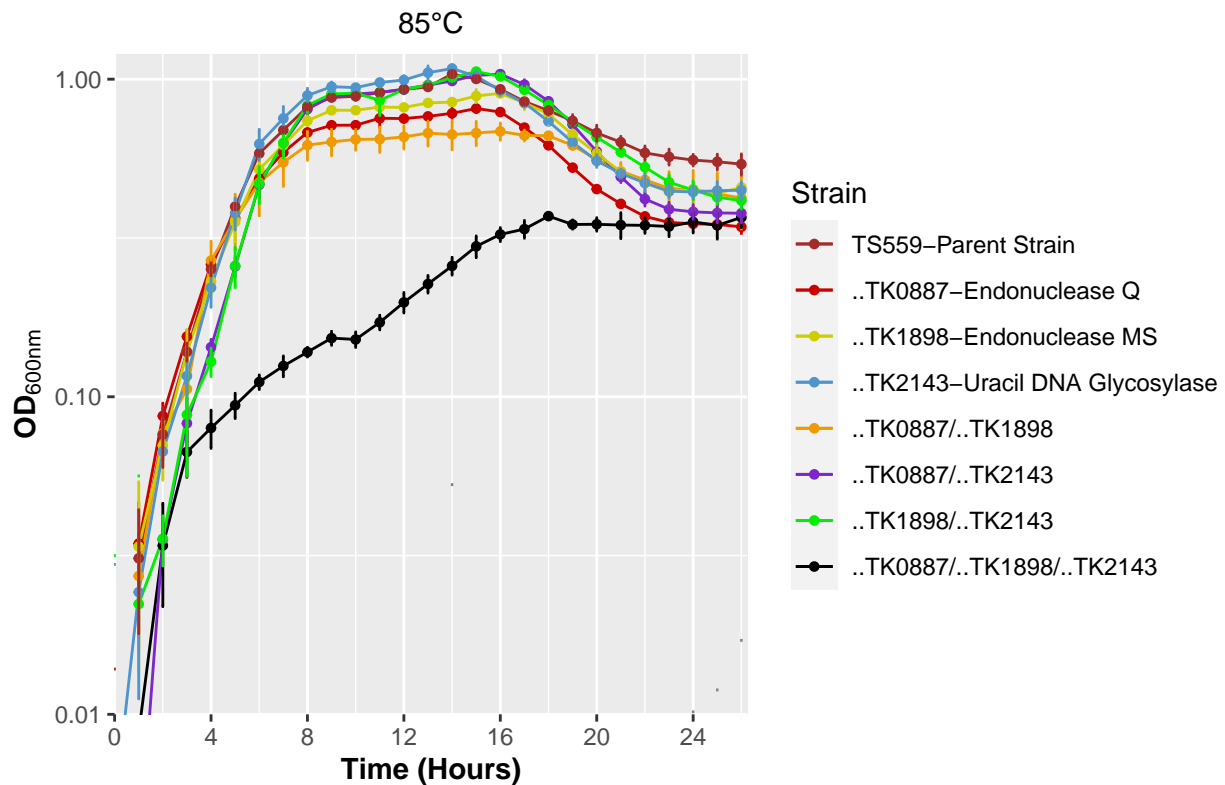
# 9c. Setting limits to your legend.

It is a bit strange having the Blank labeled when we do not see it. In a log ten scale we do not expect to see our Blank at all, so we can just remove it from the legend entirely by adding a limit to our color scaling.

```
####
#09c#
####

dU_GC_tidy_Stat%>%
  ggplot() +
  geom_line(aes(x=Hours, y=mean, color=Strain), size=0.5) +
  geom_point(aes(x=Hours, y=mean, color=Strain), size=1.2) +
  geom_errorbar(width=0.1, aes(x=Hours, ymax = mean+sd, ymin = mean-sd, color=Strain)) +
  labs(title = bquote(bold("Only deletion of all dU repair components impacts growth")), subtitle="85°C",
       y = bquote(bold("OD") [600] [nm]),
       x = bquote(bold('Time (Hours)')))) +
  theme(plot.title = element_text(hjust = 0.5), plot.subtitle= element_text(hjust = 0.5)) +
  scale_x_continuous(breaks=c(0, 4, 8, 12, 16, 20, 24), expand=c(0,0)) +
  scale_y_continuous(breaks=c(0.01, 0.1, 1.0), expand=c(0,0), trans='log10') +
  coord_cartesian(xlim=c(0,26.25), ylim=c(0.01, 1.2)) +
  scale_color_manual(values=c("white", "brown", "red3", "yellow3", "steelblue3", "orange2", "purple3",
```

## Only deletion of all dU repair components impacts growth



## 10. Theme

Now we are done graphing, choose your favorite theme.

I prefer minimalism in my axis.

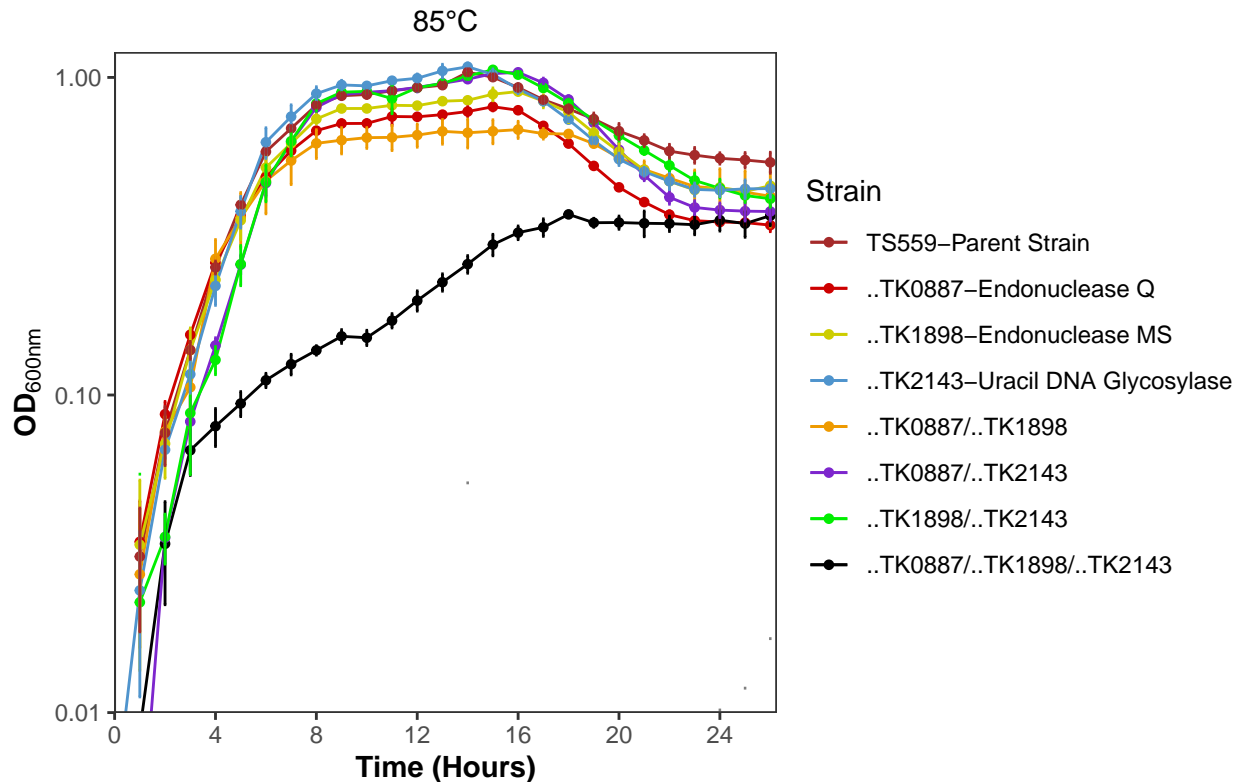
If you put a theme, you will want to put it early in your ggplot code. Otherwise the theme is building off of every previous component and this can cause errors.

```
####
#10#
####

dU_GC_tidy_Stat%>%
  ggplot() +
  theme_test() +
  geom_line(aes(x=Hours, y=mean, color=Strain), size=0.5) +
  geom_point(aes(x=Hours, y=mean, color=Strain), size=1.2) +
  geom_errorbar(width=0.1, aes(x=Hours, ymax = mean+sd, ymin = mean-sd, color=Strain)) +
  labs(title = bquote(bold("Only deletion of all dU repair components impacts growth")), subtitle="85°C",
       y = bquote(bold("OD") [600] [nm]),
       x = bquote(bold('Time (Hours)')))) +
  theme(plot.title = element_text(hjust = 0.5), plot.subtitle= element_text(hjust = 0.5)) +
  scale_x_continuous(breaks=c(0, 4, 8, 12, 16, 20, 24), expand=c(0,0)) +
  scale_y_continuous(breaks=c(0.01, 0.1, 1.0), expand=c(0,0), trans='log10') +
```

```
coord_cartesian(xlim=c(0,26.25), ylim=c(0.01, 1.2)) +
scale_color_manual(values=c("white", "brown", "red3", "yellow3", "steelblue3", "orange2", "purple3", "black"))
```

## Only deletion of all dU repair components impacts growth



# 10b. Adjusting Themes

But what if I wanted a different theme and wanted no axis?

Watch out-different themes will resize your axis, title, and legend, so you will have to fix this in post.

```
####
#10b#
####

#using a theme

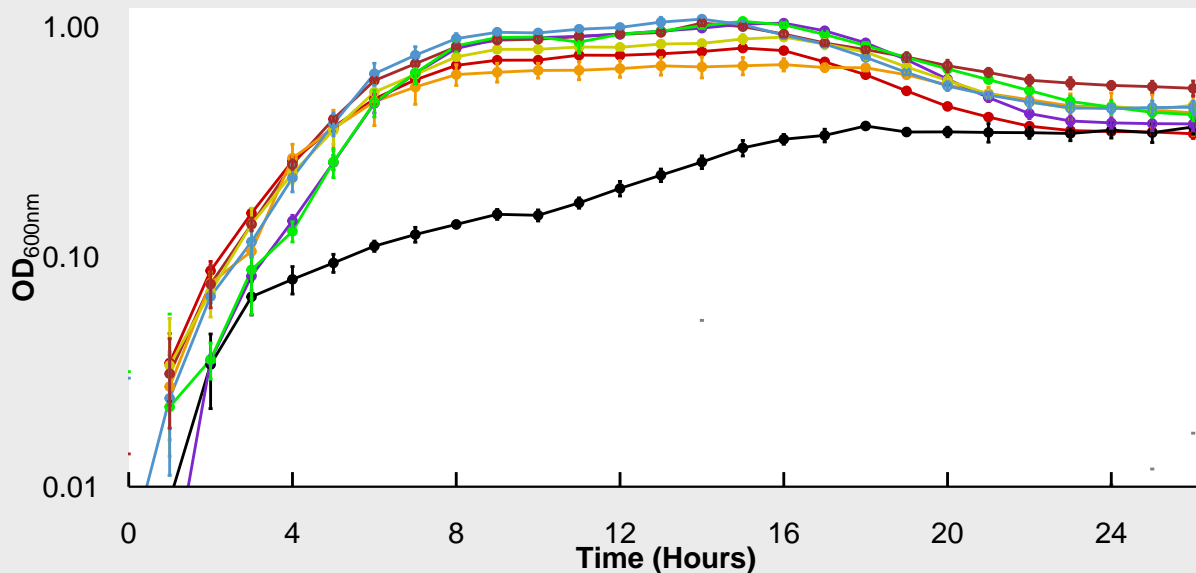
dU_GC_tidy_Stat%>%
  ggplot() +
  theme_economist_white() +
  geom_line(aes(x=Hours, y=mean, color=Strain), size=0.5) +
  geom_point(aes(x=Hours, y=mean, color=Strain), size=1.2) +
  geom_errorbar(width=0.1, aes(x=Hours, ymax = mean+sd, ymin = mean-sd, color=Strain)) +
  labs(title = bquote(bold("Only deletion of all dU repair components impacts growth")), subtitle="85°C",
       y = bquote(bold("OD") [600] [nm]),
       x = bquote(bold('Time (Hours)')))) +
  theme(plot.title = element_text(hjust = 0.5), plot.subtitle= element_text(hjust = 0.5)) +
  scale_x_continuous(breaks=c(0, 4, 8, 12, 16, 20, 24), expand=c(0,0)) +
```



```
scale_y_continuous(breaks=c(0.01, 0.1, 1.0), expand=c(0,0), trans='log10') +
coord_cartesian(xlim=c(0,26.25), ylim=c(0.01, 1.2)) +
scale_color_manual(values=c("white", "brown", "red3", "yellow3", "steelblue3", "orange2", "purple3",
theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank())
```

## Only deletion of all dU repair components impacts growth 85°C

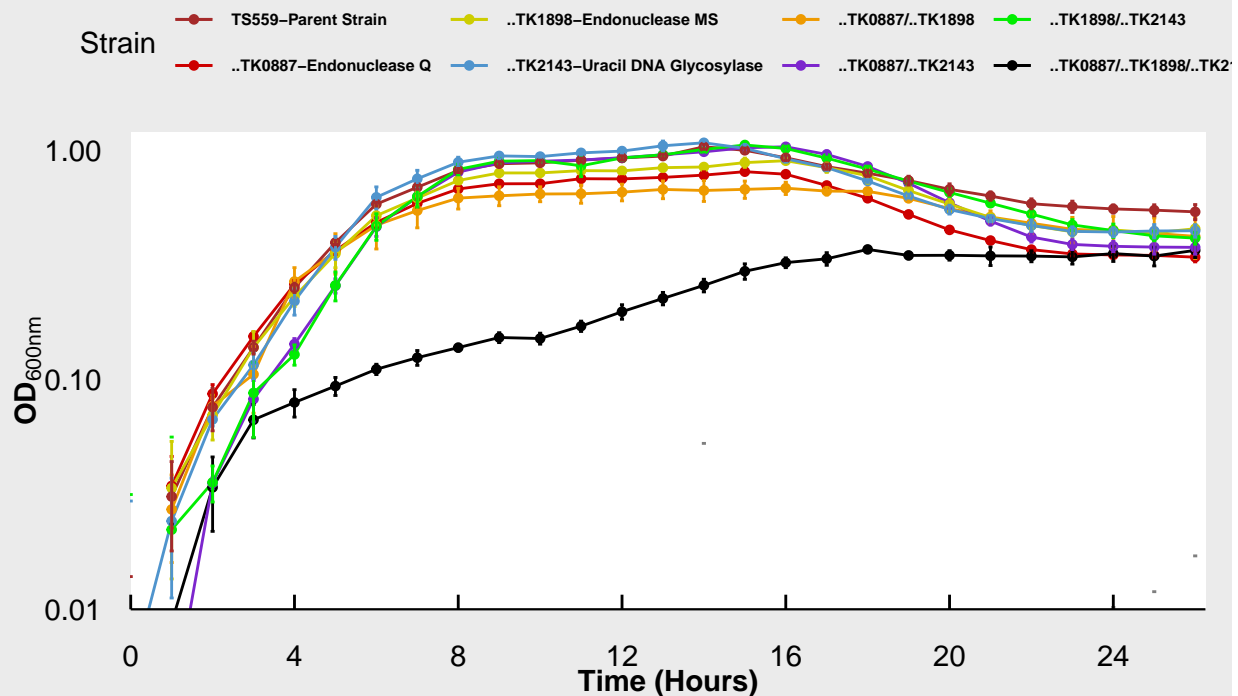
Strain      —●— ..TK1898–Endonuclease MS      —●— ..TK0887/..TK1898  
nuclease Q    —●— ..TK2143–Uracil DNA Glycosylase    —●— ..TK0887/..TK2143



*#using and adjusting the theme*

```
dU_GC_tidy_Stat%>%
ggplot() +
  theme_economist_white() +
  geom_line(aes(x=Hours, y=mean, color=Strain), size=0.5) +
  geom_point(aes(x=Hours, y=mean, color=Strain), size=1.2) +
  geom_errorbar(width=0.1, aes(x=Hours, ymax = mean+sd, ymin = mean-sd, color=Strain)) +
  labs(title = bquote(bold("Only deletion of all dU repair components impacts growth")), subtitle=bquote(
    y = bquote(bold("OD") [600] [nm]),
    x = bquote(bold('Time (Hours)')))) +
  theme(plot.title = element_text(hjust = 0.5), plot.subtitle= element_text(hjust = 0.5)) +
  scale_x_continuous(breaks=c(0, 4, 8, 12, 16, 20, 24), expand=c(0,0)) +
  scale_y_continuous(breaks=c(0.01, 0.1, 1.0), expand=c(0,0), trans='log10') +
  coord_cartesian(xlim=c(0,26.25), ylim=c(0.01, 1.2)) +
  scale_color_manual(values=c("white", "brown", "red3", "yellow3", "steelblue3", "orange2", "purple3",
  theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank()) +
  theme(legend.text = element_text(family = '', face = 'bold', size = 6))
```

## Only deletion of all dU repair components impacts growth 85°C



# 11. ggSave

It looks like we are ready to save the graph!

ggplot has ggsave which will save the file.

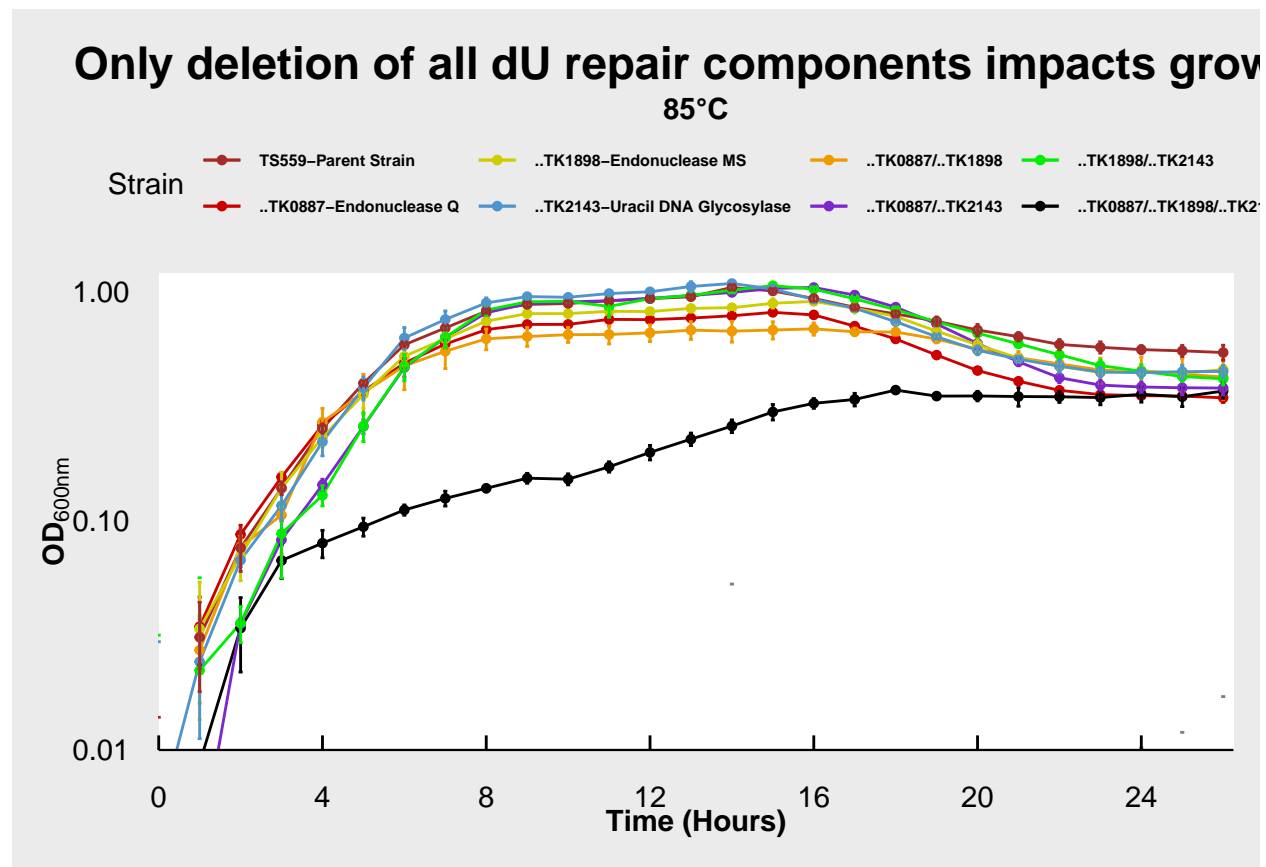
If you think the graph we have looks very... lame... go ahead and open the saved file.

Looks pretty crisp right?

```
####
#10#
####
```

```
dU_GC_tidy_Stat%>%
  ggplot() +
  theme_economist_white() +
  geom_line(aes(x=Hours, y=mean, color=Strain), size=0.5) +
  geom_point(aes(x=Hours, y=mean, color=Strain), size=1.2) +
  geom_errorbar(width=0.1, aes(x=Hours, ymax = mean+sd, ymin = mean-sd, color=Strain)) +
  labs(title = bquote(bold("Only deletion of all dU repair components impacts growth")), subtitle=bquote(
    y = bquote(bold("OD") [600] [nm]),
    x = bquote(bold('Time (Hours)')))) +
  theme(plot.title = element_text(hjust = 0.5), plot.subtitle= element_text(hjust = 0.5)) +
  scale_x_continuous(breaks=c(0, 4, 8, 12, 16, 20, 24), expand=c(0,0)) +
  scale_y_continuous(breaks=c(0.01, 0.1, 1.0), expand=c(0,0), trans='log10') +
  coord_cartesian(xlim=c(0,26.25), ylim=c(0.01, 1.2)) +
  scale_color_manual(values=c("white", "brown", "red3", "yellow3", "steelblue3", "orange2", "purple3",
```

```
theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank()) +
theme(legend.text = element_text(face = 'bold', size = 6))
```



```
ggsave("dU Growth Curve.jpg")
```

## 11. Facet Wrapping

BUT WAIT! THERE IS MORE!

We never did figure out the issue of overlapping graphs after all.

We can separate the graphs into smaller individual graphs for each strain.

```
####
#11#
####

dU_GC_tidy_Stat%>%
  ggplot() +
  theme_test() +
  geom_line(aes(x=Hours, y=mean, color=Strain), size=0.5) +
  geom_point(aes(x=Hours, y=mean, color=Strain), size=1.2) +
  geom_errorbar(width=0.1, aes(x=Hours, ymax = mean+sd, ymin = mean-sd, color=Strain)) +
  labs(title = bquote(bold("Only deletion of all dU repair components impacts growth")), subtitle=bquote
```

### only deletion of all dU repair components impacts growth



The only way to remove data you are not using is to make independent graphs for each strain... or more easily create a new data set lacking Blank.

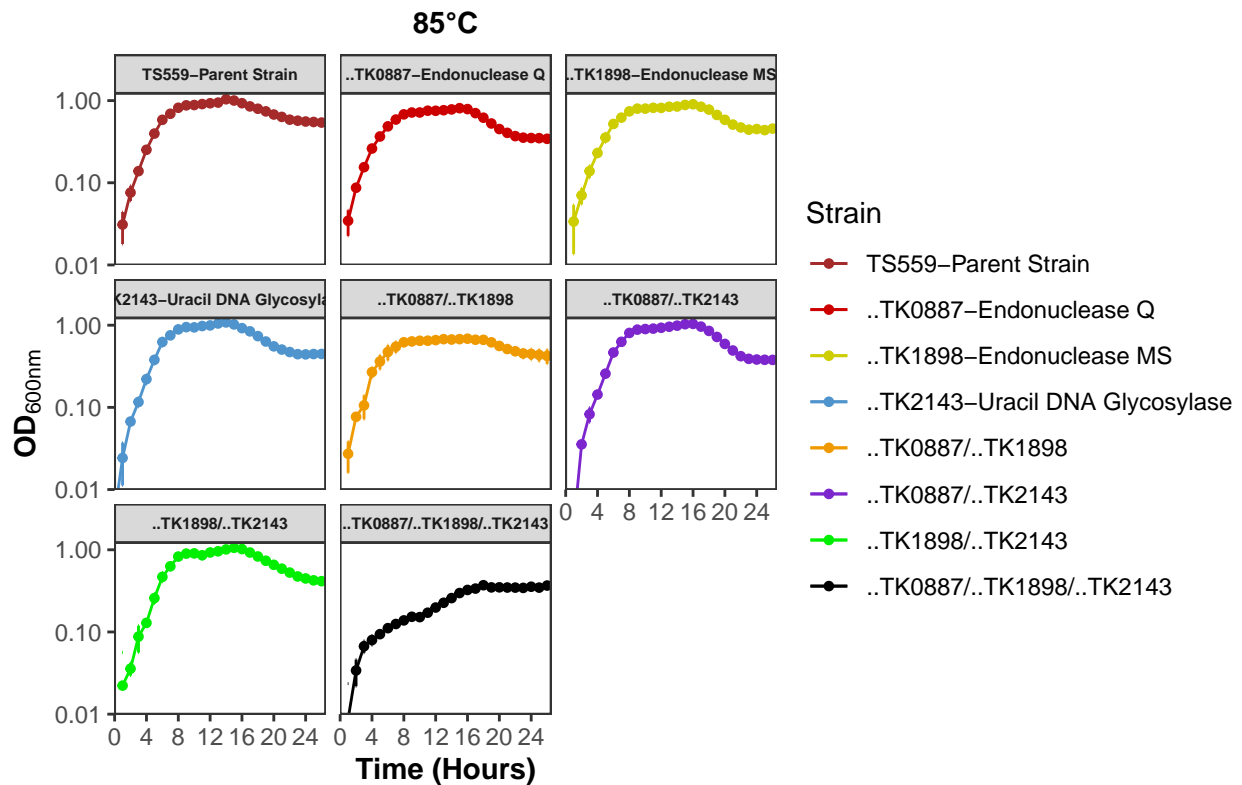
We can drop the legend or drop the panels.

28

```
dU_GC_tidy_Stat_Clean <- dU_GC_tidy_Stat[-(which(dU_GC_tidy_Stat$Strain %in% "Blank" )),]

dU_GC_tidy_Stat_Clean%>%
  ggplot() +
  theme_test() +
  geom_line(aes(x=Hours, y=mean, color=Strain), size=0.5) +
  geom_point(aes(x=Hours, y=mean, color=Strain), size=1.2) +
  geom_errorbar(width=0.1, aes(x=Hours, ymax = mean+sd, ymin = mean-sd, color=Strain)) +
  labs(title = bquote(bold("Only deletion of all dU repair components impacts growth")), subtitle=bquote(
    y = bquote(bold("OD") [600] [nm]),
    x = bquote(bold('Time (Hours)')))) +
  theme(plot.title = element_text(hjust = 0.5), plot.subtitle= element_text(hjust = 0.5)) +
  scale_x_continuous(breaks=c(0, 4, 8, 12, 16, 20, 24), expand=c(0,0)) +
  scale_y_continuous(breaks=c(0.01, 0.1, 1.0), expand=c(0,0), trans='log10') +
  coord_cartesian(xlim=c(0,26.5), ylim=c(0.01, 1.225)) +
  scale_color_manual(values=c("brown", "red3", "yellow3", "steelblue3", "orange2", "purple3", "green2",
  theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank()) +
  facet_wrap(~Strain) +
  facet_wrap(~factor(Strain, levels=c( "TS559-Parent Strain", "\u0394TK0887-Endonuclease Q", "\u0394TK1898-Endonuclease MS",
  theme(strip.text = element_text(face = "bold", size=6))
```

## Only deletion of all dU repair components impacts growth



## 11c. Facet-Wrap - Removing Legends

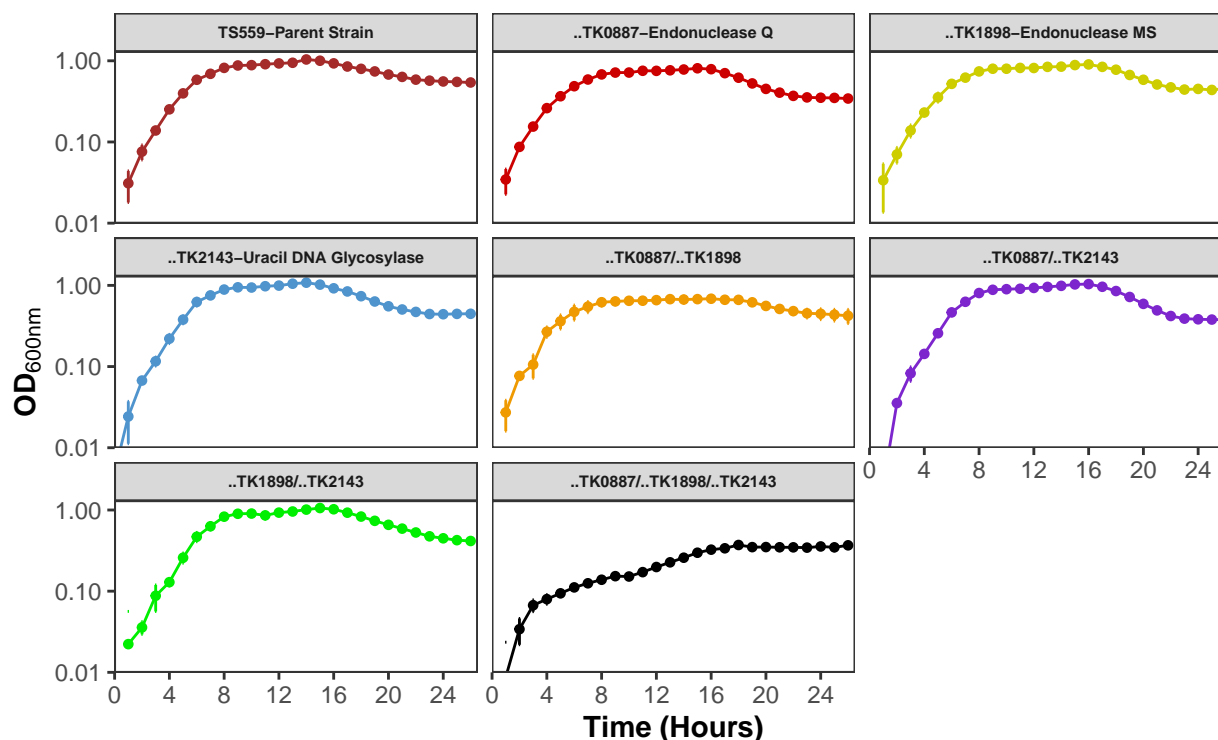
It is a bit redundant to have the panel labels and the strain labels.

We can remove one or the other.

```
####  
#11c#  
####  
  
#without the legend  
  
dU_GC_tidy_Stat_Clean%>%  
  ggplot() +  
  theme_test() +  
  geom_line(aes(x=Hours, y=mean, color=Strain), size=0.5) +  
  geom_point(aes(x=Hours, y=mean, color=Strain), size=1.2) +  
  geom_errorbar(width=0.1, aes(x=Hours, ymax = mean+sd, ymin = mean-sd, color=Strain)) +  
  labs(title = bquote(bold("Only deletion of all dU repair components impacts growth")), subtitle=bquote(  
    y = bquote(bold("OD") [600] [nm]),  
    x = bquote(bold('Time (Hours)')))) +  
  theme(plot.title = element_text(hjust = 0.5), plot.subtitle= element_text(hjust = 0.5)) +  
  scale_x_continuous(breaks=c(0, 4, 8, 12, 16, 20, 24), expand=c(0,0)) +  
  scale_y_continuous(breaks=c(0.01, 0.1, 1.0), expand=c(0,0), trans='log10') +  
  coord_cartesian(xlim=c(0,26.5), ylim=c(0.01, 1.3)) +  
  scale_color_manual(values=c("brown", "red3", "yellow3", "steelblue3", "orange2", "purple3", "green2",  
  theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank()) +  
  facet_wrap(~Strain) +  
  facet_wrap(~factor(Strain, levels=c( "TS559-Parent Strain", "\u0394TK0887-Endonuclease Q", "\u0394TK1  
    theme(strip.text = element_text(face = "bold", size=6)) +  
  theme(legend.position = "none")
```

## Only deletion of all dU repair components impacts growth

85°C



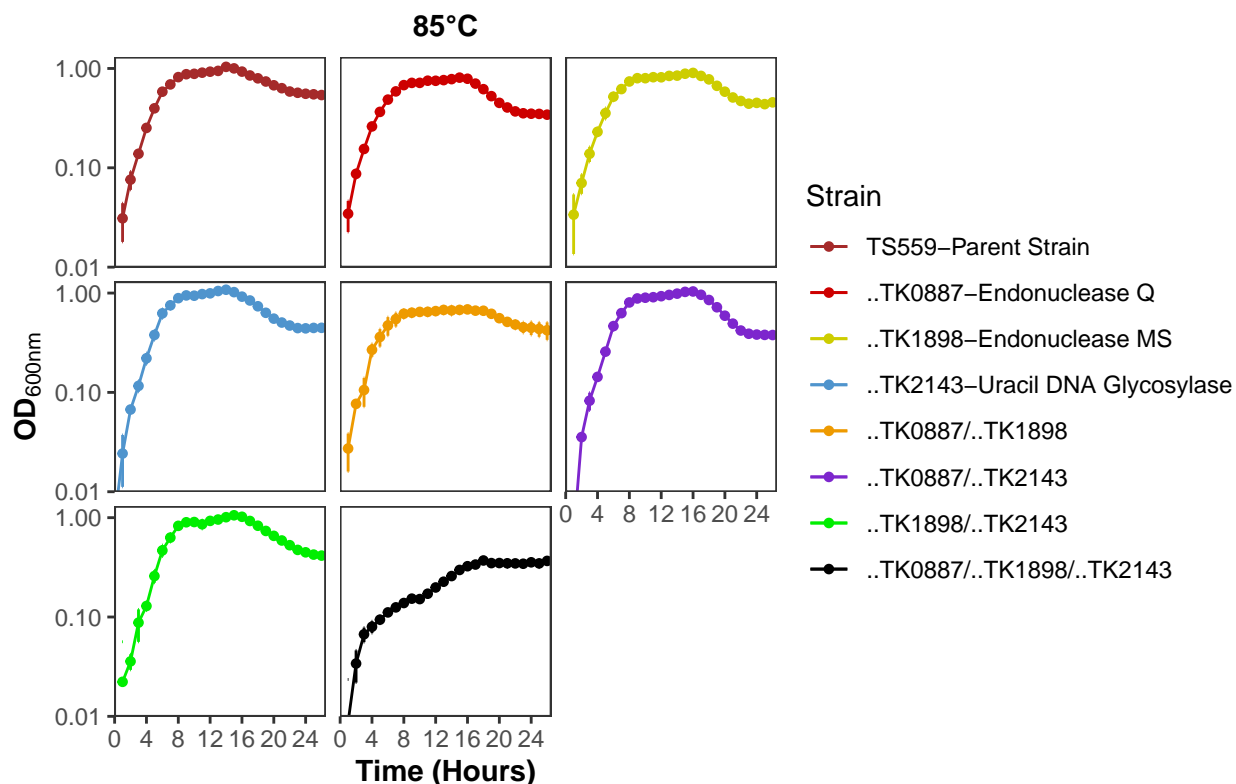
```
ggsave("dU Growth Curve_Facet_Wrapped.jpg")
```

```
#without the panel labels
```

```
dU_GC_tidy_Stat_Clean%>%
```

```
  ggplot() +
  theme_test() +
  geom_line(aes(x=Hours, y=mean, color=Strain), size=0.5) +
  geom_point(aes(x=Hours, y=mean, color=Strain), size=1.2) +
  geom_errorbar(width=0.1, aes(x=Hours, ymax = mean+sd, ymin = mean-sd, color=Strain)) +
  labs(title = bquote(bold("Only deletion of all dU repair components impacts growth")), subtitle=bquote(
    y = bquote(bold("OD") [600] [nm]),
    x = bquote(bold('Time (Hours)')))) +
  theme(plot.title = element_text(hjust = 0.5), plot.subtitle= element_text(hjust = 0.5)) +
  scale_x_continuous(breaks=c(0, 4, 8, 12, 16, 20, 24), expand=c(0,0)) +
  scale_y_continuous(breaks=c(0.01, 0.1, 1.0), expand=c(0,0), trans='log10') +
  coord_cartesian(xlim=c(0,26.5), ylim=c(0.01, 1.3)) +
  scale_color_manual(values=c("brown", "red3", "yellow3", "steelblue3", "orange2", "purple3", "green2",
  theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank()) +
  facet_wrap(~Strain) +
  facet_wrap(~factor(Strain, levels=c( "TS559-Parent Strain", "\u0394TK0887-Endonuclease Q", "\u0394TK1898-Endonuclease MS",
    theme(strip.text = element_text(face = "bold", size=6)) +
  theme( strip.text.x = element_blank() )
```

## nly deletion of all dU repair components impacts growth



```
ggsave("dU Growth Curve_Facet_Wrapped_2.jpg")
```

## 12. Bonus- ggpubr

There are a TON of packages that can take ggplot to the next level.

I actually collected data for both 85 degrees and 95 degrees C.

If I want, I can use the ggpubr package to get an image with two graphs.

```
####
#11#
####

#install.packages("ggpubr")
library(ggpubr)

#make 85 degree data into an item

dU_85 <- dU_GC_tidy_Stat_Clean%>%
  ggplot() +
  theme_test() +
  geom_line(aes(x=Hours, y=mean, color=Strain), size=0.5) +
  geom_point(aes(x=Hours, y=mean, color=Strain), size=1.2) +
  geom_errorbar(width=0.1, aes(x=Hours, ymax = mean+sd, ymin = mean-sd, color=Strain)) +
```





```

dU_Plot <- ggarrange(dU_85, dU_95, ncol=2, common.legend = TRUE, legend="bottom")

dU_Plot <- annotate_figure(dU_Plot, top = text_grob("Only deletion of all dU repair components impacts g
              color = "black", face = "bold", size = 14))

ggsave("dU_GC_Final.jpg", dU_Plot)

```

```

# Appendix

```

```

'''r
library(datasets)
library(tidyverse)
library(knitr)
library(ggplot2)
library(ggthemes)

```

```

####
#01#
####

```

```

dU_GC<-read.csv("Growth_Curve_dU.csv",fill=TRUE,header=TRUE)

head(dU_GC)

```

```

####
#02#
####

```

```

dU_GC%>%
  ggplot() +
    geom_point(aes(x=Hours, y=Blank.1)) +
    geom_point(aes(x=Hours, y=Blank.2)) +
    geom_point(aes(x=Hours, y=Blank.3)) +
    geom_point(aes(x=Hours, y=Blank.4))

```

```

####
#02#
####

```

```

GC_85_Blank <-dU_GC %>%
pivot_longer(cols = starts_with("Blank"),
              names_to = "Blank Replicates",
              values_to = "Blank Absorbance",
              values_drop_na = TRUE,
              values_transform = list(Absorbance=as.character)) %>%

```

```

select(c(Hours, 'Blank Replicates', 'Blank Absorbance')) %>%
group_by(Hours) %>%
summarize(mean = mean('Blank Absorbance'), sd= sd('Blank Absorbance')) %>%
add_column(Strain = "Blank")

GC_85_TS559 <- dU_GC %>%
  pivot_longer(cols = starts_with("TS559"),
               names_to = "TS559 Replicates",
               values_to = "TS559 Absorbance",
               values_drop_na = TRUE) %>%
  select(c(Hours, 'TS559 Replicates', 'TS559 Absorbance')) %>%
  group_by(Hours) %>%
  summarize(mean = mean('TS559 Absorbance'), sd = sd('TS559 Absorbance')) %>%
  add_column(Strain = "TS559")

GC_85_0887 <- dU_GC %>%
  pivot_longer(cols = starts_with("TK0887"),
               names_to = "0887 Replicates",
               values_to = "0887 Absorbance",
               values_drop_na = TRUE) %>%
  select(c(Hours, '0887 Replicates', '0887 Absorbance')) %>%
  group_by(Hours) %>%
  summarize(mean = mean('0887 Absorbance'), sd = sd('0887 Absorbance')) %>%
  add_column(Strain = "TK0887-Endonuclease Q")

GC_85_1898 <- dU_GC %>%
  pivot_longer(cols = starts_with("TK1898"),
               names_to = "1898 Replicates",
               values_to = "1898 Absorbance",
               values_drop_na = TRUE) %>%
  select(c(Hours, '1898 Replicates', '1898 Absorbance')) %>%
  group_by(Hours) %>%
  summarize(mean = mean('1898 Absorbance'), sd = sd('1898 Absorbance')) %>%
  add_column(Strain = "TK1898-Endonuclease MS")

GC_85_2143 <- dU_GC %>%
  pivot_longer(cols = starts_with("TK2143"),
               names_to = "2143 Replicates",
               values_to = "2143 Absorbance",
               values_drop_na = TRUE) %>%
  select(c(Hours, '2143 Replicates', '2143 Absorbance')) %>%
  group_by(Hours) %>%
  summarize(mean = mean('2143 Absorbance'), sd = sd('2143 Absorbance')) %>%
  add_column(Strain = "TK2143-Uracil DNA Glycosylase")

GC_85_0887_1898 <- dU_GC %>%
  pivot_longer(cols = starts_with("TK0887.TK1898"),
               names_to = "TK0887/TK1898 Replicates",
               values_to = "TK0887/TK1898 Absorbance",
               values_drop_na = TRUE) %>%
  select(c(Hours, 'TK0887/TK1898 Replicates', 'TK0887/TK1898 Absorbance')) %>%
  group_by(Hours) %>%
  summarize(mean = mean('TK0887/TK1898 Absorbance'), sd = sd('TK0887/TK1898 Absorbance')) %>%

```

```

add_column(Strain = "TK0887/TK1898")

GC_85_0887_2143 <- dU_GC %>%
  pivot_longer(cols = starts_with("TK0887.TK2143"),
    names_to = "TK0887/TK2143 Replicates",
    values_to = "TK0887/TK2143 Absorbance",
    values_drop_na = TRUE) %>%
  select(c(Hours, 'TK0887/TK2143 Replicates', 'TK0887/TK2143 Absorbance')) %>%
  group_by(Hours) %>%
  summarize(mean = mean('TK0887/TK2143 Absorbance'), sd = sd('TK0887/TK2143 Absorbance')) %>%
  add_column(Strain = "TK0887/TK2143")

GC_85_1898_2143 <- dU_GC %>%
  pivot_longer(cols = starts_with("TK1898.TK2143"),
    names_to = "TK1898/TK2143 Replicates",
    values_to = "TK1898/TK2143 Absorbance",
    values_drop_na = TRUE) %>%
  select(c(Hours, 'TK1898/TK2143 Replicates', 'TK1898/TK2143 Absorbance')) %>%
  group_by(Hours) %>%
  summarize(mean = mean('TK1898/TK2143 Absorbance'), sd = sd('TK1898/TK2143 Absorbance')) %>%
  add_column(Strain = "TK1898/TK2143")

GC_85_0887_1898_2143 <- dU_GC %>%
  pivot_longer(cols = starts_with("TK0887.TK1898.TK2143"),
    names_to = "TK0887/TK1898/TK2143 Replicates",
    values_to = "TK0887/TK1898/TK2143 Absorbance",
    values_drop_na = TRUE) %>%
  select(c(Hours, 'TK0887/TK1898/TK2143 Replicates', 'TK0887/TK1898/TK2143 Absorbance')) %>%
  group_by(Hours) %>%
  summarize(mean = mean('TK0887/TK1898/TK2143 Absorbance'), sd = sd('TK0887/TK1898/TK2143 Absorbance')) %>%
  add_column(Strain = "TK0887/TK1898/TK2143")

# combine tables
dU_Table <- rbind(GC_85_Blank, GC_85_TS559)
dU_Table <- rbind(dU_Table, GC_85_0887)
dU_Table <- rbind(dU_Table, GC_85_1898)
dU_Table <- rbind(dU_Table, GC_85_2143)
dU_Table <- rbind(dU_Table, GC_85_0887_1898)
dU_Table <- rbind(dU_Table, GC_85_0887_2143)
dU_Table <- rbind(dU_Table, GC_85_1898_2143)
dU_Table_Final <- rbind(dU_Table, GC_85_0887_1898_2143)

dU_Table_Final %>%
  ggplot() +
  geom_line(aes(x=Hours, y=mean, color=Strain))
dU_Table_Final %>%
  ggplot() +
  geom_line(aes(x=Hours, y=mean))

####
#03#
####

```

```

dU_GC_tidy<-read.csv("Growth_Curve_dU_tidy.csv",fill=TRUE,header=TRUE)

head(dU_GC_tidy)

dU_GC_tidy_Stat <- dU_GC_tidy %>%
  group_by(Hours, Strain) %>%
  summarize(mean = mean(OD), sd = sd(OD)) %>%
  group_by(Strain)

dU_GC_tidy_Stat%>%
  ggplot() +
  geom_point(aes(x=Hours, y=mean))

####
#04#
####

dU_GC_tidy_Stat%>%
  ggplot() +
  geom_point(aes(x=Hours, y=mean, color=Strain))

####
#05#
####

dU_GC_tidy_Stat <- dU_GC_tidy_Stat %>%
  mutate(Strain = recode(Strain, "TK0887-Endonuclease Q" = "\u0394TK0887-Endonuclease Q", "TK1898-Endonuclease Q" = "\u0394TK1898-Endonuclease Q"))

dU_GC_tidy_Stat%>%
  ggplot() +
  geom_point(aes(x=Hours, y=mean, color=Strain))

####
#06#
####

#install.packages("ggribes")
library(ggribes)

dU_GC_tidy_Stat%>%
  ggplot() +
  geom_density_ridges_gradient(aes(x=Hours, y= Strain))
dU_GC_tidy_Stat%>%
  ggplot() +
  geom_step(aes(x=Hours, y=mean, color=Strain, fill=Strain))
dU_GC_tidy_Stat%>%
  ggplot() +
  geom_smooth(aes(x=Hours, y=mean, color=Strain, fill=Strain))
dU_GC_tidy_Stat%>%

```

```

ggplot() +
  geom_line(aes(x=Hours, y=mean, color=Strain))

#Perfect!

####
#06b#
####

dU_GC_tidy_Stat%>%
  ggplot() +
  geom_line(aes(x=Hours, y=mean, color=Strain), size=0.5) +
  geom_point(aes(x=Hours, y=mean, color=Strain), size=1.2) +
  geom_errorbar(width=0.1, aes(x=Hours, ymax = mean+sd, ymin = mean-sd, color=Strain))

####
#07#
####

help("plotmath")

dU_GC_tidy_Stat%>%
  ggplot() +
  geom_line(aes(x=Hours, y=mean, color=Strain), size=0.5) +
  geom_point(aes(x=Hours, y=mean, color=Strain), size=1.2) +
  geom_errorbar(width=0.1, aes(x=Hours, ymax = mean+sd, ymin = mean-sd, color=Strain)) +
  labs(title = "Only deletion of all dU repair components impacts growth", subtitle="85\u00B0C",
       y = bquote(bold("OD") [600] [nm]),
       x = bquote(bold('Time (Hours)')))) +
  theme(plot.title = element_text(hjust = 0.5), plot.subtitle= element_text(hjust = 0.5))

####
#08#
####

dU_GC_tidy_Stat%>%
  ggplot() +
  geom_line(aes(x=Hours, y=mean, color=Strain), size=0.5) +
  geom_point(aes(x=Hours, y=mean, color=Strain), size=1.2) +
  geom_errorbar(width=0.1, aes(x=Hours, ymax = mean+sd, ymin = mean-sd, color=Strain)) +
  labs(title = "Only deletion of all dU repair components impacts growth", subtitle="85\u00B0C",
       y = bquote(bold("OD") [600] [nm]),
       x = bquote(bold('Time (Hours)')))) +
  theme(plot.title = element_text(hjust = 0.5), plot.subtitle= element_text(hjust = 0.5)) +
  scale_x_continuous(breaks=c(0, 4, 8, 12, 16, 20, 24), limits=c(0,26)) +
  scale_y_continuous(breaks=c(0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0, 1.1), limits=c(-0.05

```

```
####
#08b#
####
```

```
dU_GC_tidy_Stat%>%
  ggplot() +
  geom_line(aes(x=Hours, y=mean, color=Strain), size=0.5) +
  geom_point(aes(x=Hours, y=mean, color=Strain), size=1.2) +
  geom_errorbar(width=0.1, aes(x=Hours, ymax = mean+sd, ymin = mean-sd, color=Strain)) +
  labs(title = bquote(bold("Only deletion of all dU repair components impacts growth")), subtitle="85\u00b5M",
        y = bquote(bold("OD") [600] [nm]),
        x = bquote(bold('Time (Hours)')))) +
  theme(plot.title = element_text(hjust = 0.5), plot.subtitle= element_text(hjust = 0.5)) +
  scale_x_continuous(breaks=c(0, 4, 8, 12, 16, 20, 24), expand=c(0,0)) +
  scale_y_continuous(breaks=c(0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0, 1.1), expand=c(0,0)) +
  coord_cartesian(xlim=c(0,26.25), ylim=c(-0.05, 1.125))
```

```
####
#08c#
####
```

```
dU_GC_tidy_Stat%>%
  ggplot() +
  geom_line(aes(x=Hours, y=mean, color=Strain), size=0.5) +
  geom_point(aes(x=Hours, y=mean, color=Strain), size=1.2) +
  geom_errorbar(width=0.1, aes(x=Hours, ymax = mean+sd, ymin = mean-sd, color=Strain)) +
  labs(title = bquote(bold("Only deletion of all dU repair components impacts growth")), subtitle="85\u00b5M",
        y = bquote(bold("OD") [600] [nm]),
        x = bquote(bold('Time (Hours)')))) +
  theme(plot.title = element_text(hjust = 0.5), plot.subtitle= element_text(hjust = 0.5)) +
  scale_x_continuous(breaks=c(0, 4, 8, 12, 16, 20, 24), expand=c(0,0)) +
  scale_y_continuous(breaks=c(0.01, 0.1, 1.0), expand=c(0,0), trans='log10') +
  coord_cartesian(xlim=c(0,26.25), ylim=c(0.01, 1.2))
```

```
####
#09#
####
```

```
dU_GC_tidy_Stat%>%
  ggplot() +
  geom_line(aes(x=Hours, y=mean, color=Strain), size=0.5) +
  geom_point(aes(x=Hours, y=mean, color=Strain), size=1.2) +
  geom_errorbar(width=0.1, aes(x=Hours, ymax = mean+sd, ymin = mean-sd, color=Strain)) +
  labs(title = bquote(bold("Only deletion of all dU repair components impacts growth")), subtitle="85\u00b5M",
        y = bquote(bold("OD") [600] [nm]),
        x = bquote(bold('Time (Hours)')))) +
  theme(plot.title = element_text(hjust = 0.5), plot.subtitle= element_text(hjust = 0.5)) +
  scale_x_continuous(breaks=c(0, 4, 8, 12, 16, 20, 24), expand=c(0,0)) +
  scale_y_continuous(breaks=c(0.01, 0.1, 1.0), expand=c(0,0), trans='log10') +
  coord_cartesian(xlim=c(0,26.25), ylim=c(0.01, 1.2)) +
```





```
#10#  
####
```

```
dU_GC_tidy_Stat%>%  
  ggplot() +  
  theme_test() +  
  geom_line(aes(x=Hours, y=mean, color=Strain), size=0.5) +  
  geom_point(aes(x=Hours, y=mean, color=Strain), size=1.2) +  
  geom_errorbar(width=0.1, aes(x=Hours, ymax = mean+sd, ymin = mean-sd, color=Strain)) +  
  labs(title = bquote(bold("Only deletion of all dU repair components impacts growth")), subtitle="85\ud  
    y = bquote(bold("OD") [600] [nm]),  
    x = bquote(bold('Time (Hours)')))) +  
  theme(plot.title = element_text(hjust = 0.5), plot.subtitle= element_text(hjust = 0.5)) +  
  scale_x_continuous(breaks=c(0, 4, 8, 12, 16, 20, 24), expand=c(0,0)) +  
  scale_y_continuous(breaks=c(0.01, 0.1, 1.0), expand=c(0,0), trans='log10') +  
  coord_cartesian(xlim=c(0,26.25), ylim=c(0.01, 1.2)) +  
  scale_color_manual(values=c("white", "brown", "red3", "yellow3", "steelblue3", "orange2", "purple3", "
```

```
####  
#10b#  
####
```

```
#using a theme
```

```
dU_GC_tidy_Stat%>%  
  ggplot() +  
  theme_economist_white() +  
  geom_line(aes(x=Hours, y=mean, color=Strain), size=0.5) +  
  geom_point(aes(x=Hours, y=mean, color=Strain), size=1.2) +  
  geom_errorbar(width=0.1, aes(x=Hours, ymax = mean+sd, ymin = mean-sd, color=Strain)) +  
  labs(title = bquote(bold("Only deletion of all dU repair components impacts growth")), subtitle="85\ud  
    y = bquote(bold("OD") [600] [nm]),  
    x = bquote(bold('Time (Hours)')))) +  
  theme(plot.title = element_text(hjust = 0.5), plot.subtitle= element_text(hjust = 0.5)) +  
  scale_x_continuous(breaks=c(0, 4, 8, 12, 16, 20, 24), expand=c(0,0)) +  
  scale_y_continuous(breaks=c(0.01, 0.1, 1.0), expand=c(0,0), trans='log10') +  
  coord_cartesian(xlim=c(0,26.25), ylim=c(0.01, 1.2)) +  
  scale_color_manual(values=c("white", "brown", "red3", "yellow3", "steelblue3", "orange2", "purple3", "  
  theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank())
```

```
#using and adjusting the theme
```

```
dU_GC_tidy_Stat%>%  
  ggplot() +  
  theme_economist_white() +  
  geom_line(aes(x=Hours, y=mean, color=Strain), size=0.5) +  
  geom_point(aes(x=Hours, y=mean, color=Strain), size=1.2) +  
  geom_errorbar(width=0.1, aes(x=Hours, ymax = mean+sd, ymin = mean-sd, color=Strain)) +  
  labs(title = bquote(bold("Only deletion of all dU repair components impacts growth")), subtitle=bquote
```

```

      y = bquote(bold("OD") [600] [nm]),
      x = bquote(bold('Time (Hours)')))) +
  theme(plot.title = element_text(hjust = 0.5), plot.subtitle= element_text(hjust = 0.5)) +
  scale_x_continuous(breaks=c(0, 4, 8, 12, 16, 20, 24), expand=c(0,0)) +
  scale_y_continuous(breaks=c(0.01, 0.1, 1.0), expand=c(0,0), trans='log10') +
  coord_cartesian(xlim=c(0,26.25), ylim=c(0.01, 1.2)) +
  scale_color_manual(values=c("white", "brown", "red3", "yellow3", "steelblue3", "orange2", "purple3",
  theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank()) +
theme(legend.text = element_text(family = '', face = 'bold', size = 6))
####
#10#
####

dU_GC_tidy_Stat%>%
  ggplot() +
  theme_economist_white() +
  geom_line(aes(x=Hours, y=mean, color=Strain), size=0.5) +
  geom_point(aes(x=Hours, y=mean, color=Strain), size=1.2) +
  geom_errorbar(width=0.1, aes(x=Hours, ymax = mean+sd, ymin = mean-sd, color=Strain)) +
  labs(title = bquote(bold("Only deletion of all dU repair components impacts growth")), subtitle=bquote(
      y = bquote(bold("OD") [600] [nm]),
      x = bquote(bold('Time (Hours)')))) +
  theme(plot.title = element_text(hjust = 0.5), plot.subtitle= element_text(hjust = 0.5)) +
  scale_x_continuous(breaks=c(0, 4, 8, 12, 16, 20, 24), expand=c(0,0)) +
  scale_y_continuous(breaks=c(0.01, 0.1, 1.0), expand=c(0,0), trans='log10') +
  coord_cartesian(xlim=c(0,26.25), ylim=c(0.01, 1.2)) +
  scale_color_manual(values=c("white", "brown", "red3", "yellow3", "steelblue3", "orange2", "purple3",
  theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank()) +
theme(legend.text = element_text(face = 'bold', size = 6))

ggsave("dU Growth Curve.jpg")

####
#11#
####

dU_GC_tidy_Stat%>%
  ggplot() +
  theme_test() +
  geom_line(aes(x=Hours, y=mean, color=Strain), size=0.5) +
  geom_point(aes(x=Hours, y=mean, color=Strain), size=1.2) +
  geom_errorbar(width=0.1, aes(x=Hours, ymax = mean+sd, ymin = mean-sd, color=Strain)) +
  labs(title = bquote(bold("Only deletion of all dU repair components impacts growth")), subtitle=bquote(
      y = bquote(bold("OD") [600] [nm]),
      x = bquote(bold('Time (Hours)')))) +
  theme(plot.title = element_text(hjust = 0.5), plot.subtitle= element_text(hjust = 0.5)) +
  scale_x_continuous(breaks=c(0, 4, 8, 12, 16, 20, 24), expand=c(0,0)) +
  scale_y_continuous(breaks=c(0.01, 0.1, 1.0), expand=c(0,0), trans='log10') +
  coord_cartesian(xlim=c(0,26.5), ylim=c(0.01, 1.2)) +
  scale_color_manual(values=c("white", "brown", "red3", "yellow3", "steelblue3", "orange2", "purple3",
  theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank()) +
  facet_wrap(~Strain) +
  theme(strip.text = element_text(face = "bold", size=6))

```

```
####
#11b#
####
```

```
dU_GC_tidy_Stat_Clean <- dU_GC_tidy_Stat[-(which(dU_GC_tidy_Stat$Strain %in% "Blank" )),]
```

```
dU_GC_tidy_Stat_Clean%>%
  ggplot() +
  theme_test() +
  geom_line(aes(x=Hours, y=mean, color=Strain), size=0.5) +
  geom_point(aes(x=Hours, y=mean, color=Strain), size=1.2) +
  geom_errorbar(width=0.1, aes(x=Hours, ymax = mean+sd, ymin = mean-sd, color=Strain)) +
  labs(title = bquote(bold("Only deletion of all dU repair components impacts growth")), subtitle=bquote(
    y = bquote(bold("OD") [600] [nm])),
    x = bquote(bold('Time (Hours)')))) +
  theme(plot.title = element_text(hjust = 0.5), plot.subtitle= element_text(hjust = 0.5)) +
  scale_x_continuous(breaks=c(0, 4, 8, 12, 16, 20, 24), expand=c(0,0)) +
  scale_y_continuous(breaks=c(0.01, 0.1, 1.0), expand=c(0,0), trans='log10') +
  coord_cartesian(xlim=c(0,26.5), ylim=c(0.01, 1.225)) +
  scale_color_manual(values=c("brown", "red3", "yellow3", "steelblue3", "orange2", "purple3", "green2",
  theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank()) +
  facet_wrap(~Strain) +
  facet_wrap(~factor(Strain, levels=c( "TS559-Parent Strain", "\u0394TK0887-Endonuclease Q", "\u0394TK1
    theme(strip.text = element_text(face = "bold", size=6))
```

```
####
#11c#
####
```

```
#without the legend
```

```
dU_GC_tidy_Stat_Clean%>%
  ggplot() +
  theme_test() +
  geom_line(aes(x=Hours, y=mean, color=Strain), size=0.5) +
  geom_point(aes(x=Hours, y=mean, color=Strain), size=1.2) +
  geom_errorbar(width=0.1, aes(x=Hours, ymax = mean+sd, ymin = mean-sd, color=Strain)) +
  labs(title = bquote(bold("Only deletion of all dU repair components impacts growth")), subtitle=bquote(
    y = bquote(bold("OD") [600] [nm])),
    x = bquote(bold('Time (Hours)')))) +
  theme(plot.title = element_text(hjust = 0.5), plot.subtitle= element_text(hjust = 0.5)) +
  scale_x_continuous(breaks=c(0, 4, 8, 12, 16, 20, 24), expand=c(0,0)) +
  scale_y_continuous(breaks=c(0.01, 0.1, 1.0), expand=c(0,0), trans='log10') +
  coord_cartesian(xlim=c(0,26.5), ylim=c(0.01, 1.3)) +
  scale_color_manual(values=c("brown", "red3", "yellow3", "steelblue3", "orange2", "purple3", "green2",
  theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank()) +
  facet_wrap(~Strain) +
  facet_wrap(~factor(Strain, levels=c( "TS559-Parent Strain", "\u0394TK0887-Endonuclease Q", "\u0394TK1
    theme(strip.text = element_text(face = "bold", size=6)) +
  theme(legend.position = "none")
```

```
ggsave("dU Growth Curve_Facet_Wrapped.jpg")
```

```
#without the panel labels
```

```
dU_GC_tidy_Stat_Clean%>%
  ggplot() +
  theme_test() +
  geom_line(aes(x=Hours, y=mean, color=Strain), size=0.5) +
  geom_point(aes(x=Hours, y=mean, color=Strain), size=1.2) +
  geom_errorbar(width=0.1, aes(x=Hours, ymax = mean+sd, ymin = mean-sd, color=Strain)) +
  labs(title = bquote(bold("Only deletion of all dU repair components impacts growth")), subtitle=bquote(
    y = bquote(bold("OD") [600] [nm])),
    x = bquote(bold('Time (Hours)')))) +
  theme(plot.title = element_text(hjust = 0.5), plot.subtitle= element_text(hjust = 0.5)) +
  scale_x_continuous(breaks=c(0, 4, 8, 12, 16, 20, 24), expand=c(0,0)) +
  scale_y_continuous(breaks=c(0.01, 0.1, 1.0), expand=c(0,0), trans='log10') +
  coord_cartesian(xlim=c(0,26.5), ylim=c(0.01, 1.3)) +
  scale_color_manual(values=c("brown", "red3", "yellow3", "steelblue3", "orange2", "purple3", "green2",
  theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank()) +
  facet_wrap(~Strain) +
  facet_wrap(~factor(Strain, levels=c( "TS559-Parent Strain", "\u0394TK0887-Endonuclease Q", "\u0394TK1
    theme(strip.text = element_text(face = "bold", size=6)) +
  theme( strip.text.x = element_blank() )
```

```
ggsave("dU Growth Curve_Facet_Wrapped_2.jpg")
```

```
####
#11#
####
```

```
#install.packages("ggpubr")
library(ggpubr)
```

```
#make 85 degree data into an item
```

```
dU_85 <- dU_GC_tidy_Stat_Clean%>%
  ggplot() +
  theme_test() +
  geom_line(aes(x=Hours, y=mean, color=Strain), size=0.5) +
  geom_point(aes(x=Hours, y=mean, color=Strain), size=1.2) +
  geom_errorbar(width=0.1, aes(x=Hours, ymax = mean+sd, ymin = mean-sd, color=Strain)) +
  labs(subtitle=bquote(bold("85\u00B0C")),
    y = bquote(bold("OD") [600] [nm])),
    x = bquote(bold('Time (Hours)')))) +
  theme(plot.subtitle= element_text(hjust = 0.5)) +
  scale_x_continuous(breaks=c(0, 4, 8, 12, 16, 20, 24), expand=c(0,0)) +
  scale_y_continuous(breaks=c(0.01, 0.1, 1.0), expand=c(0,0), trans='log10') +
  coord_cartesian(xlim=c(0,26.5), ylim=c(0.01, 1.3)) +
  scale_color_manual(values=c("brown", "red3", "yellow3", "steelblue3", "orange2", "purple3", "green2",
  theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank()) +
  theme(legend.position="bottom", legend.text = element_text(size=5))
```

```
#95 degree data
```

```
dU_GC_tidy_95<-read.csv("Growth_Curve_dU_tidy_95.csv",fill=TRUE,header=TRUE)
```

```
head(dU_GC_tidy)
```

```
dU_GC_tidy_Stat_95 <- dU_GC_tidy_95 %>%
```

```
  group_by(Hours, Strain) %>%
```

```
  summarize(mean = mean(OD), sd = sd(OD)) %>%
```

```
  group_by(Strain) %>%
```

```
  mutate(Strain = recode(Strain, "TK0887-Endonuclease Q" = "\u0394TK0887-Endonuclease Q", "TK1898-Endonuclease Q" = "\u0394TK1898-Endonuclease Q"))
```

```
dU_GC_tidy_Stat_Clean_95 <- dU_GC_tidy_Stat_95[-(which(dU_GC_tidy_Stat_95$Strain %in% "Blank" )),]
```

```
dU_95 <- dU_GC_tidy_Stat_Clean_95 %>%
```

```
  ggplot() +
```

```
  theme_test() +
```

```
  geom_line(aes(x=Hours, y=mean, color=Strain), size=0.5) +
```

```
  geom_point(aes(x=Hours, y=mean, color=Strain), size=1.2) +
```

```
  geom_errorbar(width=0.1, aes(x=Hours, ymax = mean+sd, ymin = mean-sd, color=Strain)) +
```

```
  labs(subtitle=bquote(bold("95\u00B0C")),
```

```
        y = bquote(bold("OD") [600] [nm]),
```

```
        x = bquote(bold('Time (Hours)')))) +
```

```
  theme(plot.subtitle= element_text(hjust = 0.5)) +
```

```
  scale_x_continuous(breaks=c(0, 4, 8, 12, 16, 20, 24), expand=c(0,0)) +
```

```
  scale_y_continuous(breaks=c(0.01, 0.1, 1.0), expand=c(0,0), trans='log10') +
```

```
  coord_cartesian(xlim=c(0,26.5), ylim=c(0.01, 1.3)) +
```

```
  scale_color_manual(values=c("brown", "red3", "yellow3", "steelblue3", "orange2", "purple3", "green2",
```

```
  theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank()) +
```

```
  theme(legend.position="none")
```

```
#ggarrange
```

```
dU_Plot <- ggarrange(dU_85, dU_95, ncol=2, common.legend = TRUE, legend="bottom")
```

```
dU_Plot <-annotate_figure(dU_Plot, top = text_grob("Only deletion of all dU repair components impacts growth",  
  color = "black", face = "bold", size = 14))
```

```
ggsave("dU_GC_Final.jpg", dU_Plot)
```