BIO331L - Molecular Biology Lab

ArunK

Notes

Lab 1

- 3 major experiments and 3 lab reports; read lab manual prior to lab section
- review of common units, dilution and solution calculations
- SI units:
  - combine common prefixes with base units

```
library(knitr)
kable(data.frame("prefix" = c("deci","centi","milli","micro","nano","pico","femto","atto"), "Factor bas
```

prefix	Factor.base.10	symbol	Prefix	Factor.base.10.1	symbol.1
deci	-1	d	deca	1	da
centi	-2	c	hecto	2	h
milli	-3	m	kilo	3	k
micro	-6	mu	mega	6	M
nano	-9	n	giga	9	G
pico	-12	p	tera	12	T
femto	-15	f	peta	15	P
atto	-18	a	exa	18	E

Lab 2

- goal of understanding water stress tolerance mechanisms in plants
- abscisic acid (ABA) is involved in stress-response signalling pathways
  - ABA is a terpene-based regulator, synthesized similarly to other steroids and lipids

9/2/2020 Lecture

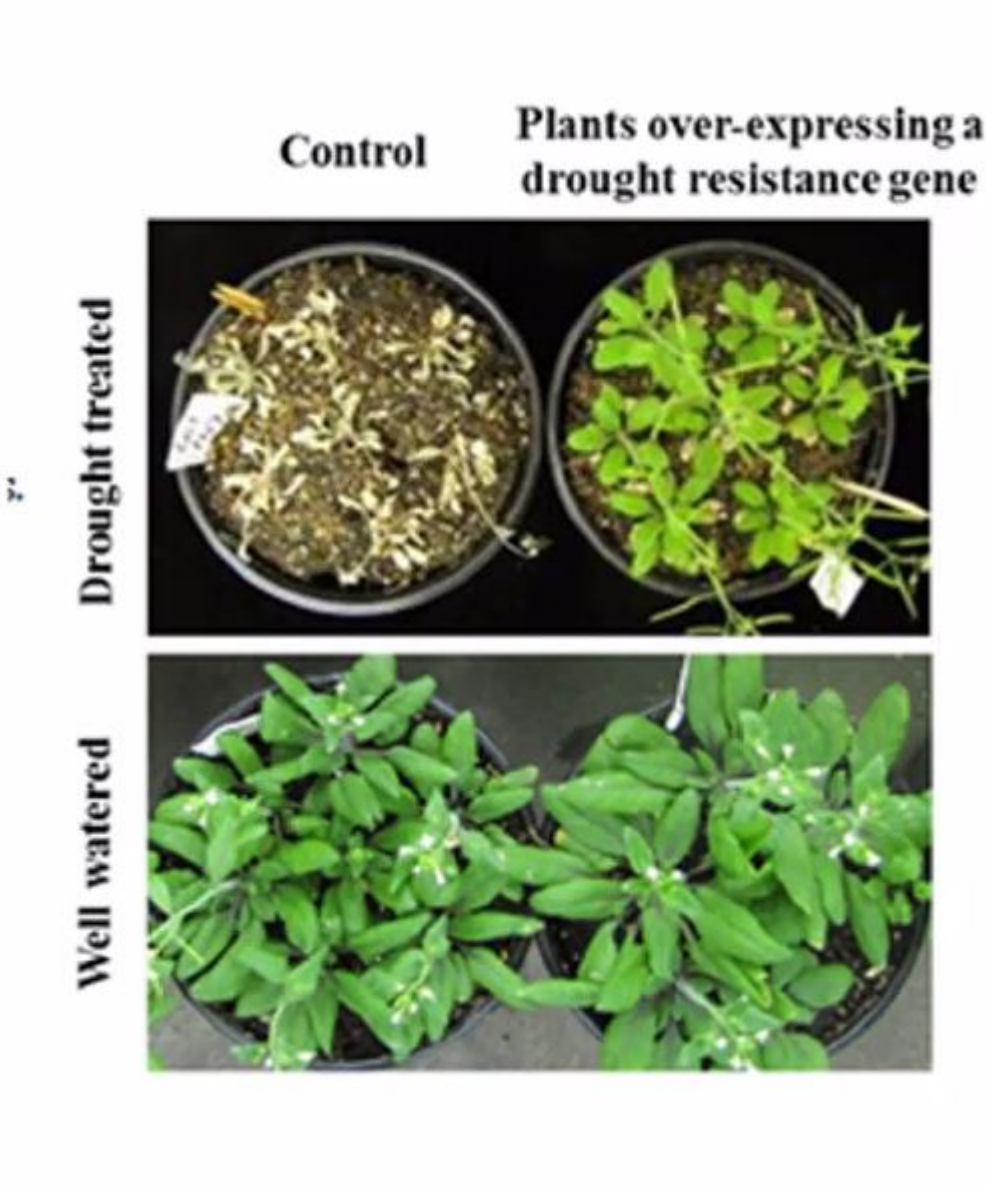
- Following a 2010 earthquake in Haiti, which caused massive amounts of destruction, several cases of Cholera popped up. Since Haiti had not had any lingering Cholera infections in the country, there was concern as cases seemed to crop up among a major waterway. Rapid sequence analysis revealed that the disease agent had come from Nepal with peacekeeping troops. Following accidental sewage system damage, Cholera was introduced to the watershed and spread further.
- Isle Royale, small island somewhat near mainland; low number of wolves present with natural prey of moose. Tend to come on and off the island using ice bridges; as a result of general warming, ice bridges became less frequent. Wolf genetic diversity plummeted due to inbreeding, causing severe population decline; wolves brought in from distinct populations to improve fitness.
- Papaya ringspot virus was completely eradicated in Hawaii by making papaya highly resistant to the virus

9/2/2020 Lab

- Abscisic acid (ABA): major plant hormone, used in gene expression, signalling pathways. Can circumvent gene expression to directly interact with biomolecule for faster cellular response: rapid interaction with guard cells to close them in response to water loss.
- Arabinogalactan (AGP31 in Arabidopsis): glycoprotein in the cell wall. Use of Wild Type, AGP31 mutant, and AGP31 rescued Adabidopsis to see how plant response to Abscisic acid varied
- RNA Purification:
  - RNA experiment seeks to compare gene expression in wild type and two AGP31 altered transgenic lines when exposed to ABA in media
  - Use of mechanical force to grind seedling samples
  - Tissue lyser is used to rapidly grind sample material, kept at low temperature to avoid RNA degradation
  - AGP31 rescued line occurs when wild gene copies are transferred to the knockout -Spectroscopic analysis of RNA samples
  - MS grow playes used to germinate Arabidopsis seeds
  - Nanodrop Spectrophotometer

Lecture 9/9/2020

- ABA is almost ubiquitous in plants, serves several signalling functions
  - has also been found to play a role in signalling in some animal lineages
- most ecosystems (terrestrial, freshwater, and marine) are changing genetically and shifting in response to anthropogenic climate change
  - experiment conducted on Arabidopsis shows promise of dought modified plant lines



  - first ABA-pathway deficient Arabidopsis grows poorly even when properly watered
- ABA pathway:
  - ABAR (PYR/PYL) receptor located in cytoplasm; under cellular dehydration, there will be a higher concentration of ABA in the cell (can be synthesized within the cell or transported from elsewhere)
    - PP2C: protein phosphatase (removed phospates from proteins), becomes bound to ABA ABAR complex
    - SnRK2: usually without phosphate due to PP2C activity; active in the phosphorylated form
      - adds phosphates to target proteins
    - AB15: activated by SnRK2 phosphorylation
    - SLAC1/anion transport group is also activated by SnRK2 phosphorylation

Lab 9/9/2020

- What did we do last week?
  - purified RNA from seedlings treated with ABA or control, then determine RNA quality using spectroscopic analysis and electrophoresis

- cDNA synthesis: use purified RNA in reverse transcription to produce complementary DNA
  - thermostable DNA pol requires a DNA template to synthesize DNA; we use RNA to produce cDNA to make a template for synthesis
    * we prepare samples that are both +RT and -RT (control samples)
    * incubate with DNAse to reduce any contaminating genomic DNA still present; DNase is then inactivated
      · DNase will not be 100% comoplete, some residual genomic DNA will be present and be used as a template for DNA; we need -RT samples to quantify this effect
    *
- we need to determine with primers for Arabidopsis *OLE1* gene expression show acceptable properties for use in qPCR

Lab 9/16/2020

- 3.3 is the number of cycles for 2 starting molecules to equate with 10 starting molecules

Lecture 9/30/2020

- use of ABA and ethanol treatments during incubation
  - can isolate any handling stress related, or media derived effects
- we know RD29B and RAB18 (well known ABA marker genes) show less increase in 6134 than WT Ler-0; interested in seeing if oleosin genes show lower increase in mRNA levels

Lab 9/30/2020

Lecture 10/07/2020

- AGP shows up in variety of food products, including gum; occurs as gooey sap from gum arabic trees
- plant cell wall proteins have specific effects in development and in response to environmental conditions
- AGP's consist of large carbohydrate chains in glycoproteins; many different classes of proteins can be bound, all consist of signal peptide + proline rich domain
- CRU3 and other mRNAs are increased in the 6134 compared to Ler-0 WT

Lab 10/07/2020

- 6134 AGP31 mutant line experiences a large spice in CRU3 compared to wild type at the 13-15 day seedling mark, as well as other proteins like OLE1

Lab 10/14/2020

- we know that 6134 mutant has higher gene expression for CRU3, interested in seeing if protein concentration is the same
  - protein purification, concentration by Bradford assay
  - protein gel (SDS-PAGE) and blotting
  - antibody incubation and antigen detection
- Bradford reagent dye experiences change in peak absorbance when bound to protein (465 to 595 nm)

Lecture 10/21/2020

- cruciferins are produced as precursors in seeds; can be in processed or unprocessed form in vegetative tissue
  - seed storage protein happen relatively late in Arabidopsis seed development
  - precursor cruciferins are cleaved during maturation
    * α and β subunits are 30-35 and 21-25 kD
  - seed storage proteins are degradd during seedling growth; is the 6134 mutatnt Cru mRNA increase accompanied by a Cru protein subunit increase

  - high number of storage proteins can provide useful storage reserves in cases of high environmental stress

Lab 10/21/2020

- after transfer to membrane, blocking with protein solution should occur to prevent unwanted binding
  - primary antibody and secondary antibody are bound to target protein and detected with fluorescence or other method
- treatment of Ler wild-type and 6134, separated into root and aerial parts
  - calculate BSA absorbance values with standard curve (6x difference to BSA protein amount)
  - use protein concentration to load a consistent amount of protein for each well
- SDS PAGE is used to separate proteins based on molecular weight
  - SDS denatures protein hydrogen, ionic bonds, and hydrophobic interactions; beta-mercatoethanol breaks disulfide bonds; glycine and chloride ions are used to push proteins in a group through the stacking gel, and then spread the proteins into bands in the separating gel

Lecture 10/28/2020

- SSP and oleosins are primarily found in cotyledons (become aerial green part)
  - upregulation in 6134 may be easier to see in root portions, low SSP presence
  - similarity of legumin and cruciferins in Arabidopsis, somewhat homologous

Lab 10/28/2020

- after transferring to membrane, add blocking buffer to prevent non-specific binding; incubation with primary and secondary antibodies (signal detection and amplification)
  - PVDF membrane has a high capacity to bind to proteins, need blocking buffer
- experimental improvements: CRU specific antibody (instead of legumin), include CRU null line

Lecture 11/4/2020

- replication of Fall 2019 CRISPR experiment
- CRISPR systems protect prokaryotic cells from invading DNA (just look at microbial)

Lab 11/4/2020

- anneal constructed primers to target sequence, then allow ligation to the cut plasmid
- use chemical transformation to insert plasmid to the E. Coli, select for KanR

Lecture 11/11/2020

- Core ABA response pathway: snRK2s activate ABI4 and ABI5 (possibly ABI3) by phosphorylation, which then bind to increase transcription
  - ABI5/ABFs-ABI3: formation of heterodimer, transcription factor is active during germination and seedling development
- Mehdy lab results: AGP31 null mutatnt 6134 shows up-regulation of certain SSP and OLE mRNAs during seedling development
  - these genes are known to be transcriptionally controlled (ABA pathway)
  - hypothesis: ABI5 and ABI3 are active in transcriptional activation of select SSP and OLE genes in 6134 seedlings
    - AGP31 ABI3 homozygous mutant was constructed, and failed to exhibit SSP/OLE mRNA upregulation (ABI3 required for upregulation); ABA regulation pathway via ABI3/5 have primary effects during germination
  - Fall 2019 work: WRKY41 regulates ABI3 transcript levels, directly binds to ABI3 promoter to increase gene transcription
  - plasmid used includes sgRNA, Cas9 sequence (splicing factor as promoter and terminator)

    * restriction enzyme Bsa1 excises sgRNA (should be gel purified, linearized plasmid should not ligate to reform circular genome)
    * use KanR to select for plasmid presence
    * Bsa1 Type IIs cleavage occurs outside recognition region, creation of 5' sticky overhangs; plasmid is engineered so overhangs are not complementary after excision

Lab 11/11/2020

- molecular weight standard curve:
  - made to determine size of bands of interest, make standard curve with distances and known sizes (15 to 75 kD, middle of band, x dist y molecular weight)
- recombinant plasmids are made by bidirectional annealing of WRKY41 sense and antisense segments, ligation and transformation
  - KanR colonies are selected and counted (separate samples based on oligio): uncut pHEE results in high colony number, random oligos result in some background colonies (review justification for why the cut pHEE results in lower background transformation than the random controls)
    * include analysis of group and control colony counts in report 3
- plasmid purification and analysis
  - colonies are grown, inoculation with oligos of interest; purify plasmids and measure purity (Nanodrop)
    * concurrently perform colony PCR, detection of bands of interest; if validated, then send samples to sequencing
  - nanodrop results have same interpretation as the RNA experiment
- we design forward and reverse primer flanking the target sgRNA sequence, can use gels to determine if correct cassette is in the plasmid (large size difference)
- typical PCR cycle: 1 cycle at high heat, 25 cycles of high-low-medium heat, then medium and cooling
  - high heat steps create single strands, low heat anneals to oligo primers, medium allows polymerization and ligation
  - PCR verifies that all sets produced correct plasmid, verified against positive control
- gRNA design tutorial:
  - find gene of interest in Arabidopsis site, download CDS sequence, BLAST
  - search for PAM (5'N20NGG3'), use BLAST to verify gene only matches other genes
- RNA data for later classes; gene expression for 6134 line is too variable
  - limited BRs with high variation, only 1 time treatment point
  - new RNA experiment,

Lab 2

- drought historically has limited ecosystems, agriculture
  - drying trends have increased due to anthropogenic climate change
  - understanding water stress tolerance mechanisms in plants helps mitigate damage
- stresses elicit tolerance responses in plants, often involve abscisic acid; ABA regulates normal developmental processes
- ABA is a terpene based regulator, serves diverse functions in plants and higher sponges and animals
  - in plants, ABA can originate intra or extra cellularly; synthesized in chloroplast, in a carotene pathway
  - continuing research on the primary ABA receptor in higher mammals; thought to be the C-like 2 cellular membrane protein
- RNA experiment wants to compare gene expression in wild type and two AGP31 altered lines
  - RNA is less stable than DNA and more prone to hydrolysis
    * mechanical disruption: grind plant tissue with liquid N_2

```
#primer standardization curve should be aggregate
m <- matrix(c(200,100,10,23.9517084757487, 24.5212860107421, 27.2934220631917, 23.481273651123, 24.1533
m
```

```
##      [,1]     [,2]     [,3]     [,4]     [,5]
## [1,]  200 23.95171 23.48127 27.49259 24.01634
## [2,]  100 24.52129 24.15337 28.80624 24.65372
## [3,]   10 27.29342 27.21988 30.23062 28.01442
```

```
m[,1] <- log(m[,1], base = 10)
m
```

```
##         [,1]     [,2]     [,3]     [,4]     [,5]
## [1,] 2.30103 23.95171 23.48127 27.49259 24.01634
## [2,] 2.00000 24.52129 24.15337 28.80624 24.65372
## [3,] 1.00000 27.29342 27.21988 30.23062 28.01442
```

```
colnames(m) <- c("x","y1","y2","y3","y4")
m <- data.frame(m)
m
```

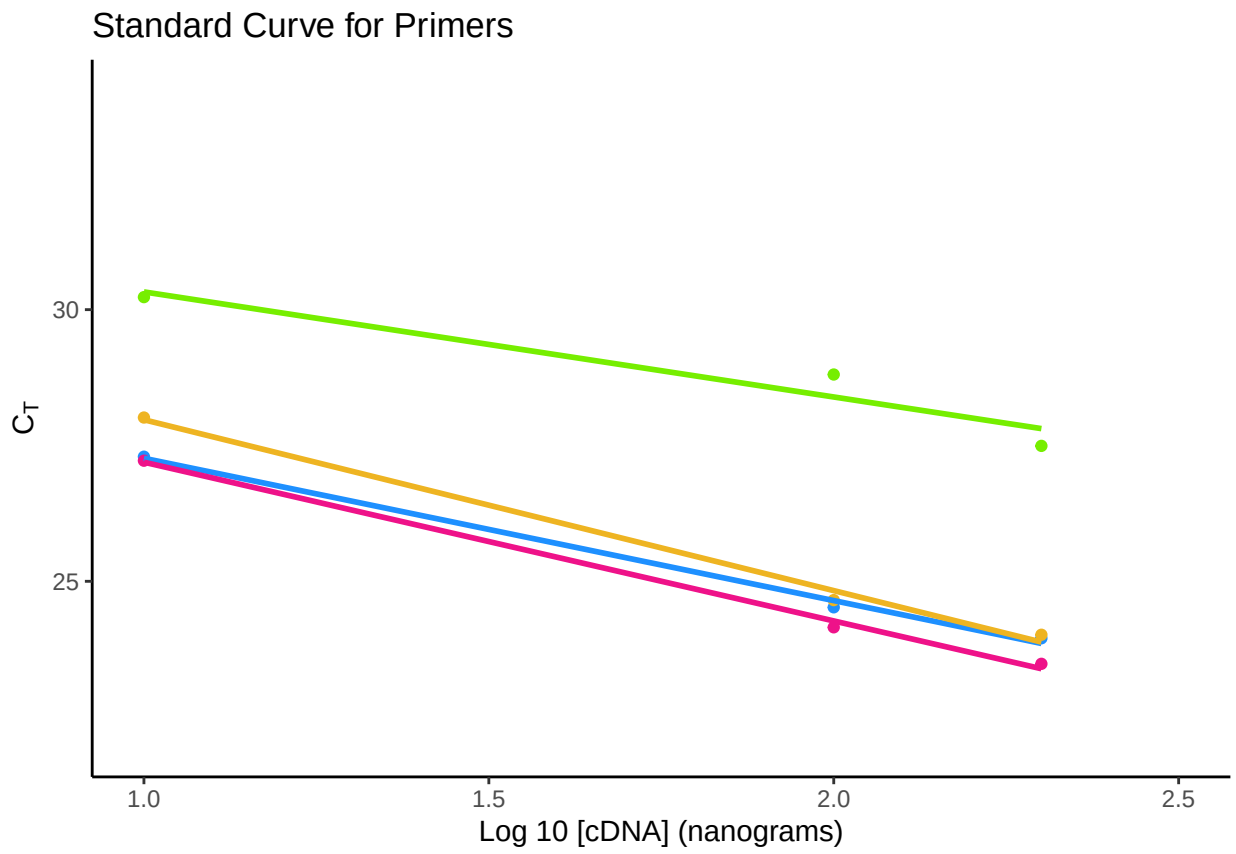
```
##         x       y1       y2       y3       y4
## 1 2.30103 23.95171 23.48127 27.49259 24.01634
## 2 2.00000 24.52129 24.15337 28.80624 24.65372
## 3 1.00000 27.29342 27.21988 30.23062 28.01442
```

```
lm(y1~x, data  = m)
```

```
##
## Call:
## lm(formula = y1 ~ x, data = m)
##
## Coefficients:
## (Intercept)            x
##       29.88        -2.62
```

```
library(ggplot2)

ggplot(m, aes(x=x)) +
  geom_point(aes(y=y1), col = "dodgerblue") + geom_point(aes(y=y2), col = "deeppink2") + geom_point(aes
  geom_smooth(aes(y=y1), method = "lm",se= F,col = "dodgerblue") +  geom_smooth(aes(y=y2), method = "lm
  theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank(),
        panel.background = element_blank(), axis.line = element_line(colour = "black"))
```

```
## `geom_smooth()` using formula 'y ~ x'
## `geom_smooth()` using formula 'y ~ x'
## `geom_smooth()` using formula 'y ~ x'
## `geom_smooth()` using formula 'y ~ x'
```




```
one <- lm(y1~x, data = m)
two <- lm(y2~x, data = m)
three <- lm(y3~x, data = m)
four <- lm(y4~x, data = m)

summary(one)
```

```
##
## Call:
## lm(formula = y1 ~ x, data = m)
##
## Residuals:
##        1        2        3
##  0.09518 -0.12383  0.02865
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  29.8844     0.3054   97.86  0.00651 **
## x            -2.6196     0.1649  -15.89  0.04001 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1588 on 1 degrees of freedom
## Multiple R-squared:  0.9961, Adjusted R-squared:  0.9921
## F-statistic: 252.5 on 1 and 1 DF,  p-value: 0.04001
```

```
summary(two)
```

```
##
## Call:
## lm(formula = y2 ~ x, data = m)
##
## Residuals:
##        1        2        3
##  0.09018 -0.11733  0.02715
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  30.1148     0.2894  104.07  0.00612 **
## x            -2.9220     0.1562  -18.71  0.03400 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1505 on 1 degrees of freedom
## Multiple R-squared:  0.9972, Adjusted R-squared:  0.9943
## F-statistic: 349.9 on 1 and 1 DF,  p-value: 0.034
```

```
summary(three)
```

```
##
## Call:
## lm(formula = y3 ~ x, data = m)
##
## Residuals:
##       1       2       3
## -0.3179  0.4136 -0.0957
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  32.2600     1.0201  31.625   0.0201 *
## x            -1.9337     0.5507  -3.512   0.1766
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.5304 on 1 degrees of freedom
## Multiple R-squared:  0.925,  Adjusted R-squared:   0.85
## F-statistic: 12.33 on 1 and 1 DF,  p-value: 0.1766
```

```
summary(four)
```

```
##
## Call:
## lm(formula = y4 ~ x, data = m)
##
## Residuals:
##        1        2        3
##  0.13447 -0.17496  0.04048
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  31.1192     0.4315   72.12  0.00883 **
```

```
## x            -3.1453     0.2329  -13.50  0.04706 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.2243 on 1 degrees of freedom
## Multiple R-squared:  0.9945, Adjusted R-squared:  0.9891
## F-statistic: 182.4 on 1 and 1 DF,  p-value: 0.04706
```

Primer Set 1

$$y = -2.6196x + 29.8844$$
$$\mathrm{R}^2 = 0.9911$$

Primer Set 2

$$y = -2.9220x + 30.1148$$
$$\mathrm{R}^2 = 0.9943$$

Primer Set 3

$$y = -1.9337x + 32.26$$
$$\mathrm{R}^2 = 0.85$$

Primer Set 4

$$y = -3.1453x + 31.1192$$
$$\mathrm{R}^2 = 0.98$$

```
### relative expression plot
### LER A AND LER ETH
library(tidyverse)
```

```
## -- Attaching packages --------------------------------- tidyverse 1.3.0 --
```

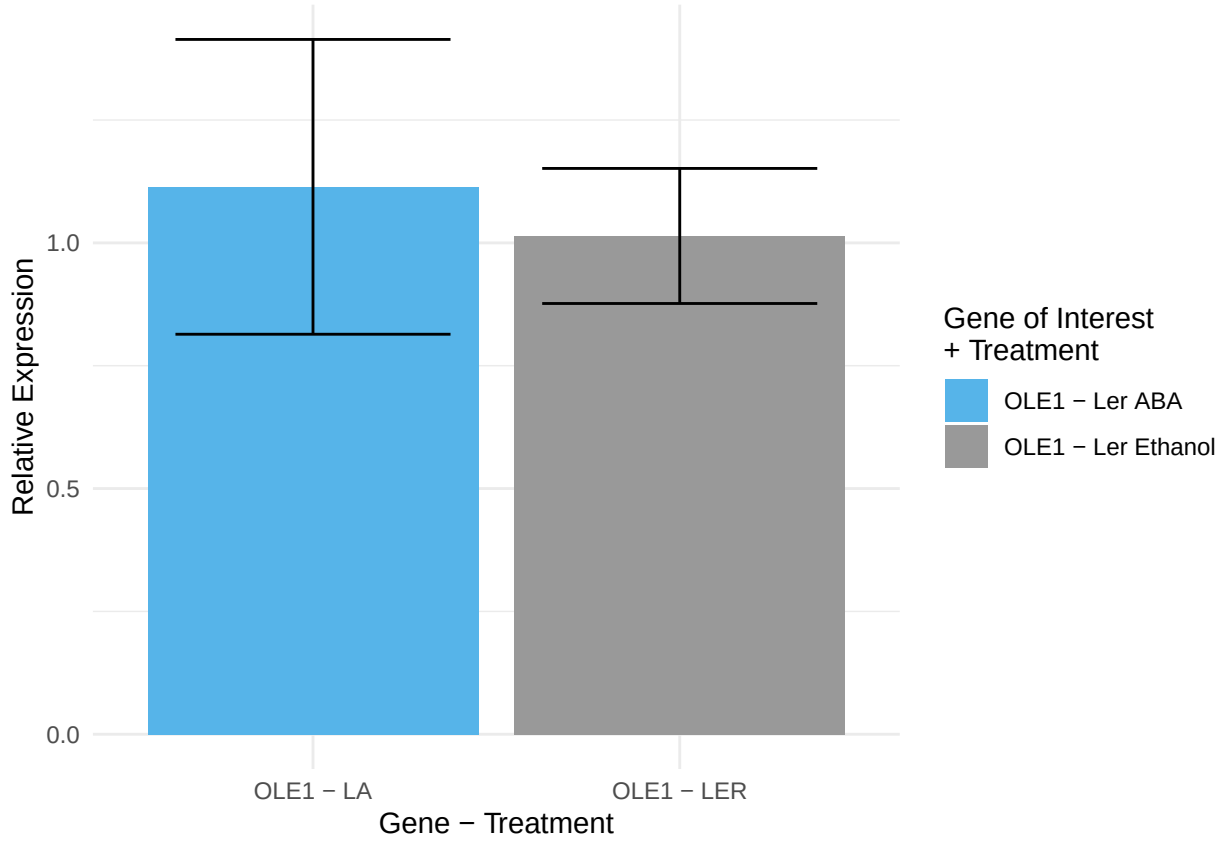
```
## v tibble  3.0.3     v dplyr   1.0.2
## v tidyr   1.1.2     v stringr 1.4.0
## v readr   1.3.1     v forcats 0.5.0
## v purrr   0.3.4
```

```
## -- Conflicts ------------------------------------ tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

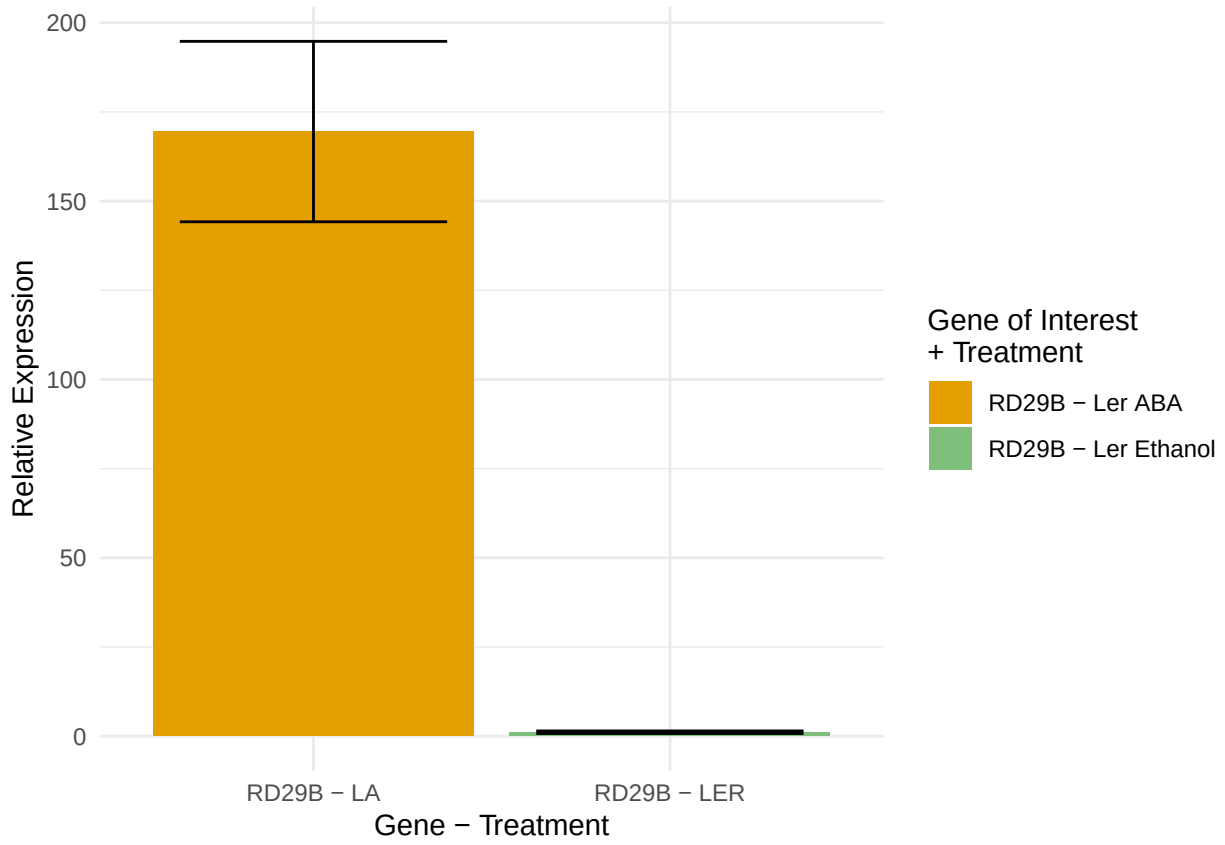
```
library(ggpubr)
t <- matrix(c(1.11403189024852, 1.0140646013784, 169.490832162453, 1.10480500761417, 0.814039999061835,
t
```

```
##            [,1]        [,2]       [,3]
## [1,]   1.114032   0.8140400   1.414024
## [2,]   1.014065   0.8766431   1.151486
## [3,] 169.490832 144.1992607 194.782404
## [4,]   1.104805   0.7429001   1.466710
```

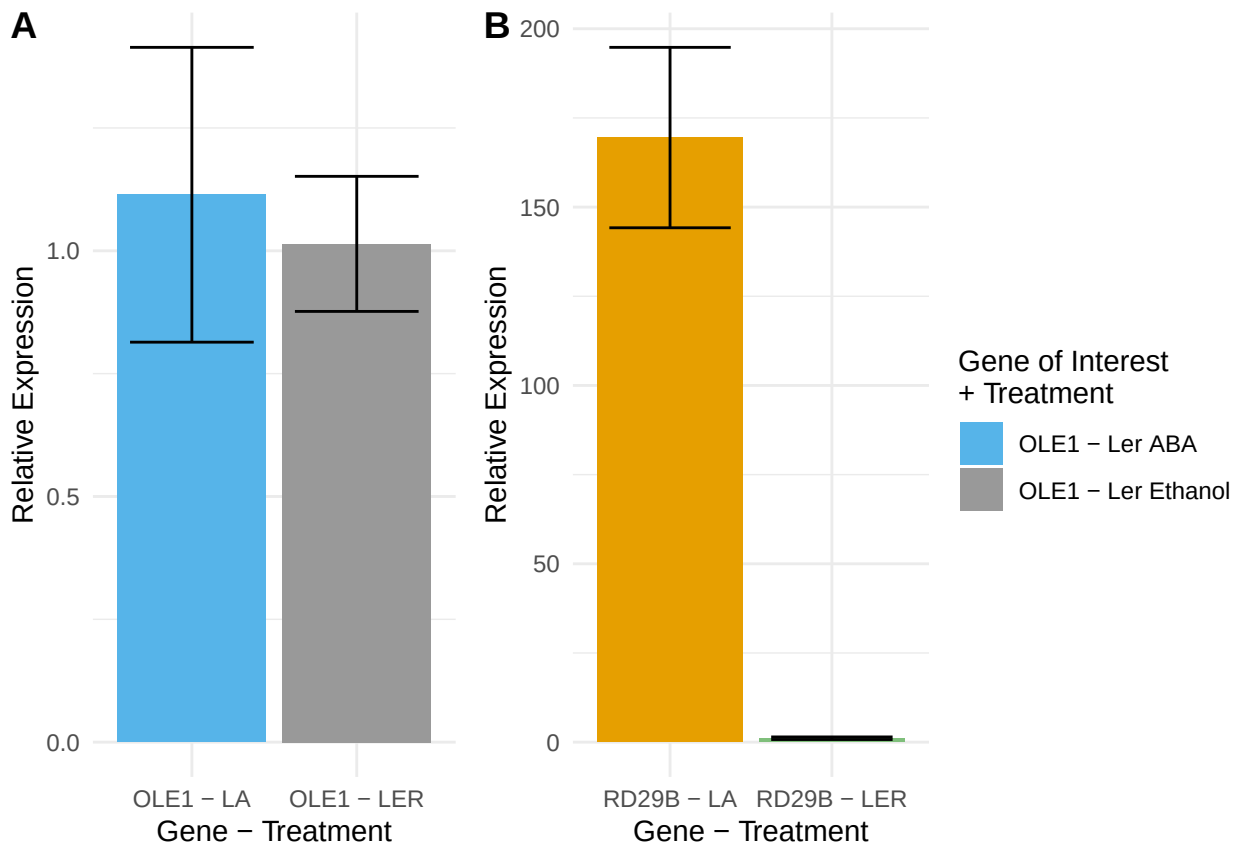
```r
t <- data.frame(t)

rownames(t) <- c("OLE1 - LA", "OLE1 - LER", "RD29B - LA", "RD29B - LER")
t <- rownames_to_column(t)

colnames(t) <- c("Gene","expr1", "Lower", "Upper")
t_ole <- t[1:2,]
t_rd <- t[3:4,]

olegg1 <- ggplot(t_ole, aes(Gene, expr1, fill = Gene)) + geom_col()  + geom_errorbar(aes(ymin = Lower, y
                        breaks=c("OLE1 - LA", "OLE1 - LER"),
                        labels=c("OLE1 - Ler ABA", "OLE1 - Ler Ethanol")) +theme_minimal()
olegg1
```



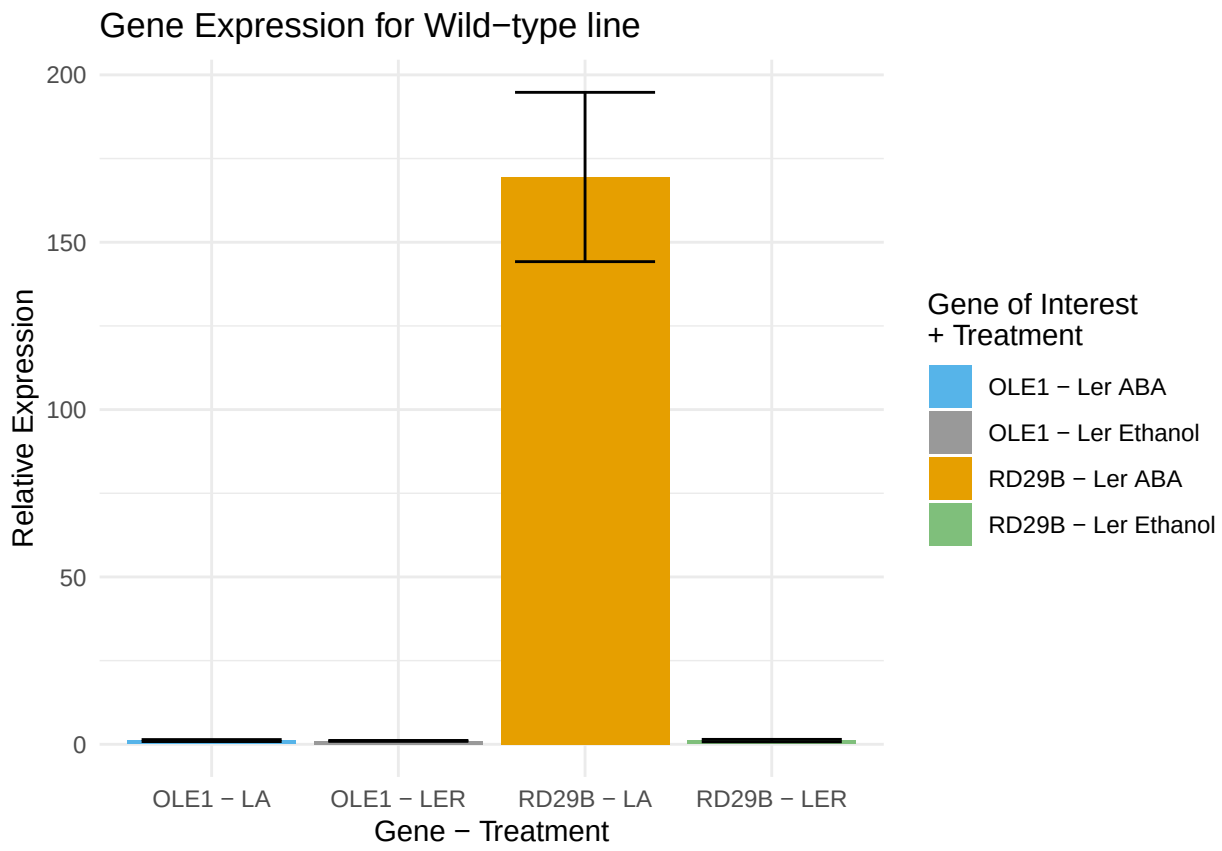

```r
rdgg1 <- ggplot(t_rd, aes(Gene, expr1,fill = Gene)) + geom_col()  + geom_errorbar(aes(ymin = Lower, yma
                        breaks=c("RD29B - LA", "RD29B - LER"),
                        labels=c("RD29B - Ler ABA", "RD29B - Ler Ethanol")) +theme_minimal()
rdgg1
```




```
fig <- ggarrange(olegg1,rdgg1, labels = c("A","B"), nrow=1, ncol = 2,
                 common.legend = TRUE, legend = "right")

fig
```



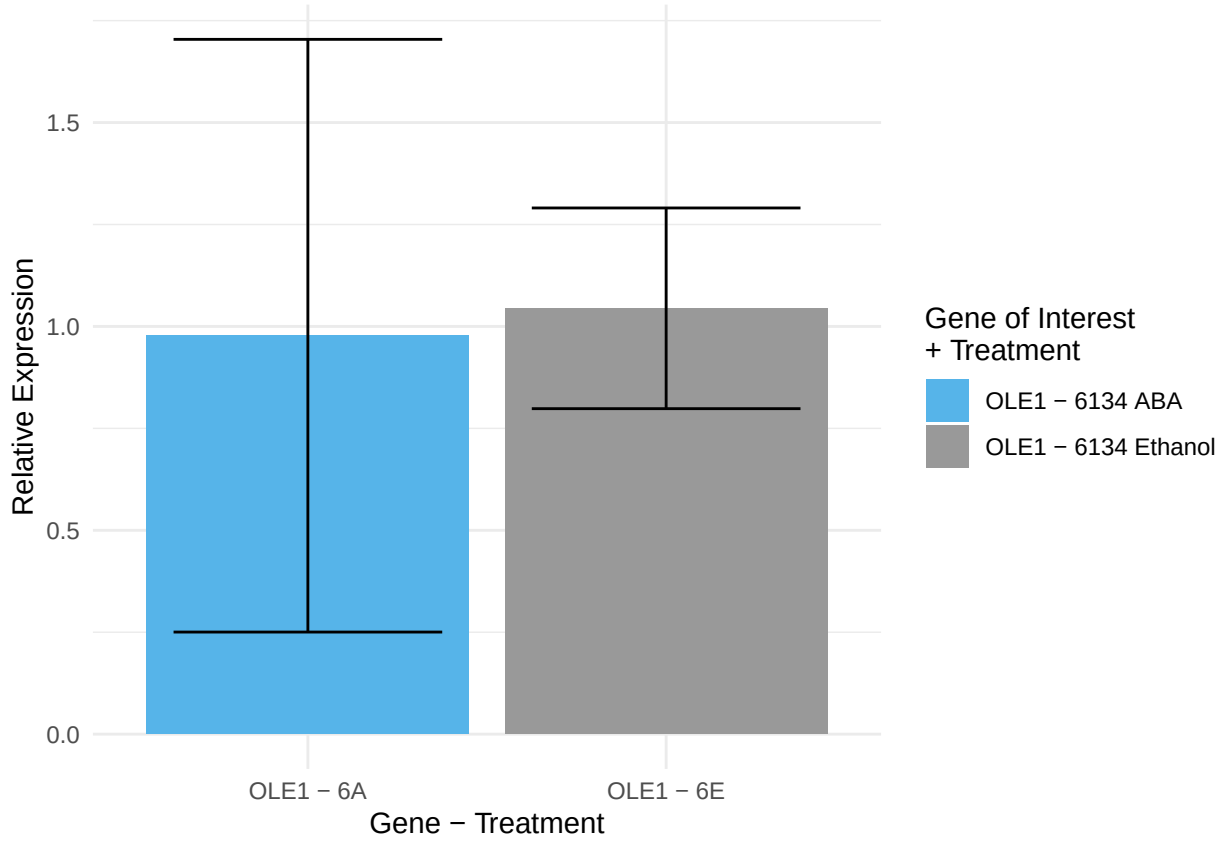

```
ggplot(t, aes(Gene, expr1, fill = Gene)) + geom_col()  + geom_errorbar(aes(ymin = Lower, ymax = Upper),
                        breaks=c("OLE1 - LA", "OLE1 - LER","RD29B - LA", "RD29B - LER"),
                        labels=c("OLE1 - Ler ABA", "OLE1 - Ler Ethanol","RD29B - Ler ABA", "RD29B - Le
```



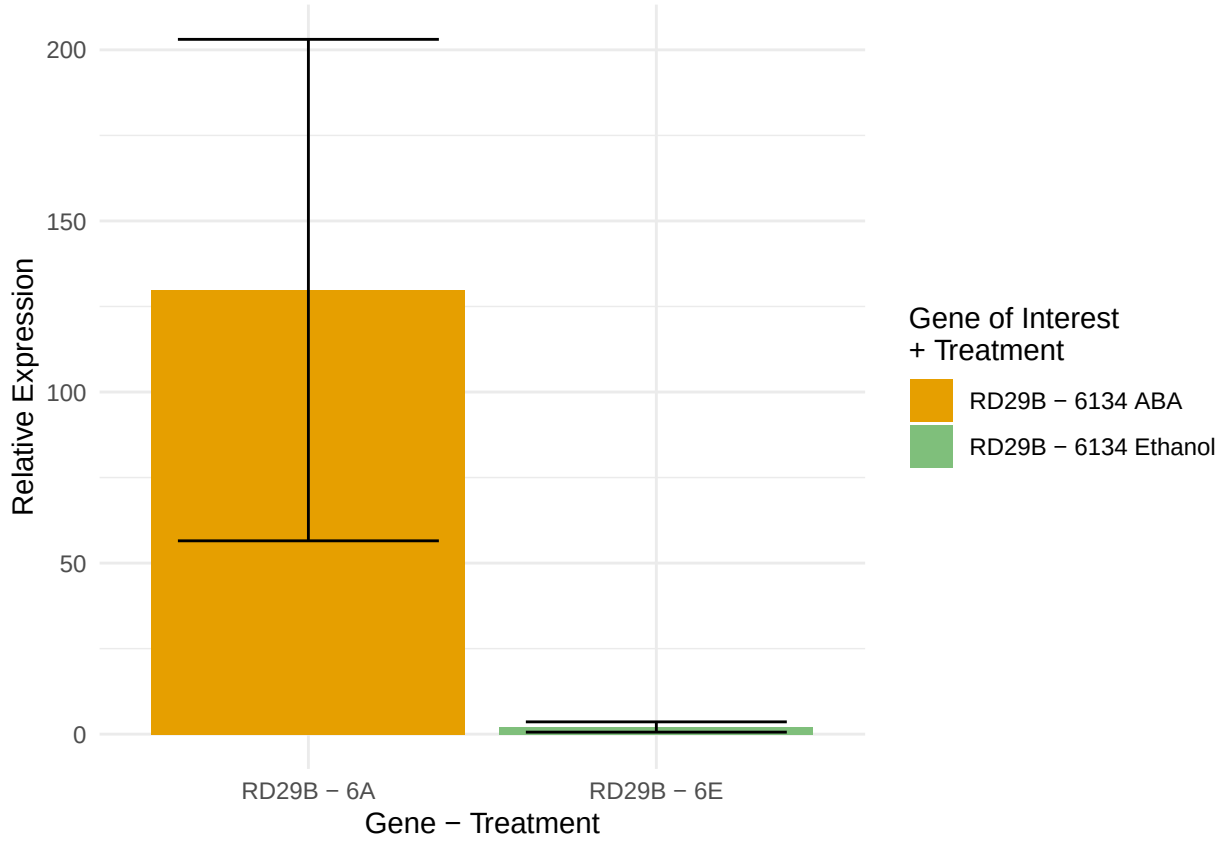

```
## relative expression
## 6134 ABA AND 6134 ETH
t <- matrix(c(0.977269186280336, 1.0444301621343, 129.789739009814, 2.09451616953587, 0.250437616778727
t
```

```
##             [,1]       [,2]       [,3]
## [1,]   0.9772692  0.2504376   1.704101
## [2,]   1.0444302  0.7983484   1.290512
## [3,] 129.7897390 56.5040597 203.075418
## [4,]   2.0945162  0.5918523   3.597180
```

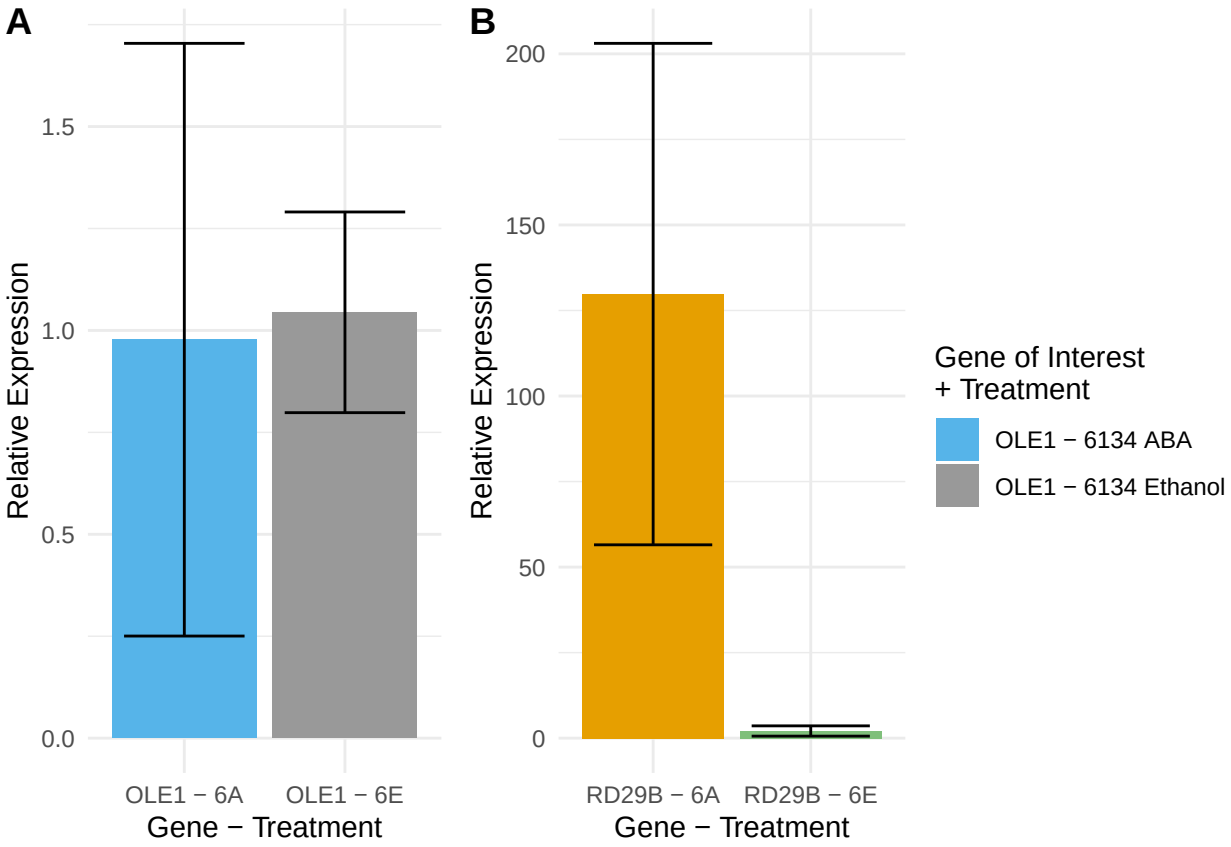
```
t <- data.frame(t)

rownames(t) <- c("OLE1 - 6A", "OLE1 - 6E", "RD29B - 6A", "RD29B - 6E")
t <- rownames_to_column(t)

colnames(t) <- c("Gene","expr1", "Lower", "Upper")
t_ole <- t[1:2,]
t_rd <- t[3:4,]

olegg1 <- ggplot(t_ole, aes(Gene, expr1, fill = Gene)) + geom_col()  + geom_errorbar(aes(ymin = Lower, y
                        breaks=c("OLE1 - 6A", "OLE1 - 6E"),
                        labels=c("OLE1 - 6134 ABA", "OLE1 - 6134 Ethanol")) +theme_minimal()
olegg1
```



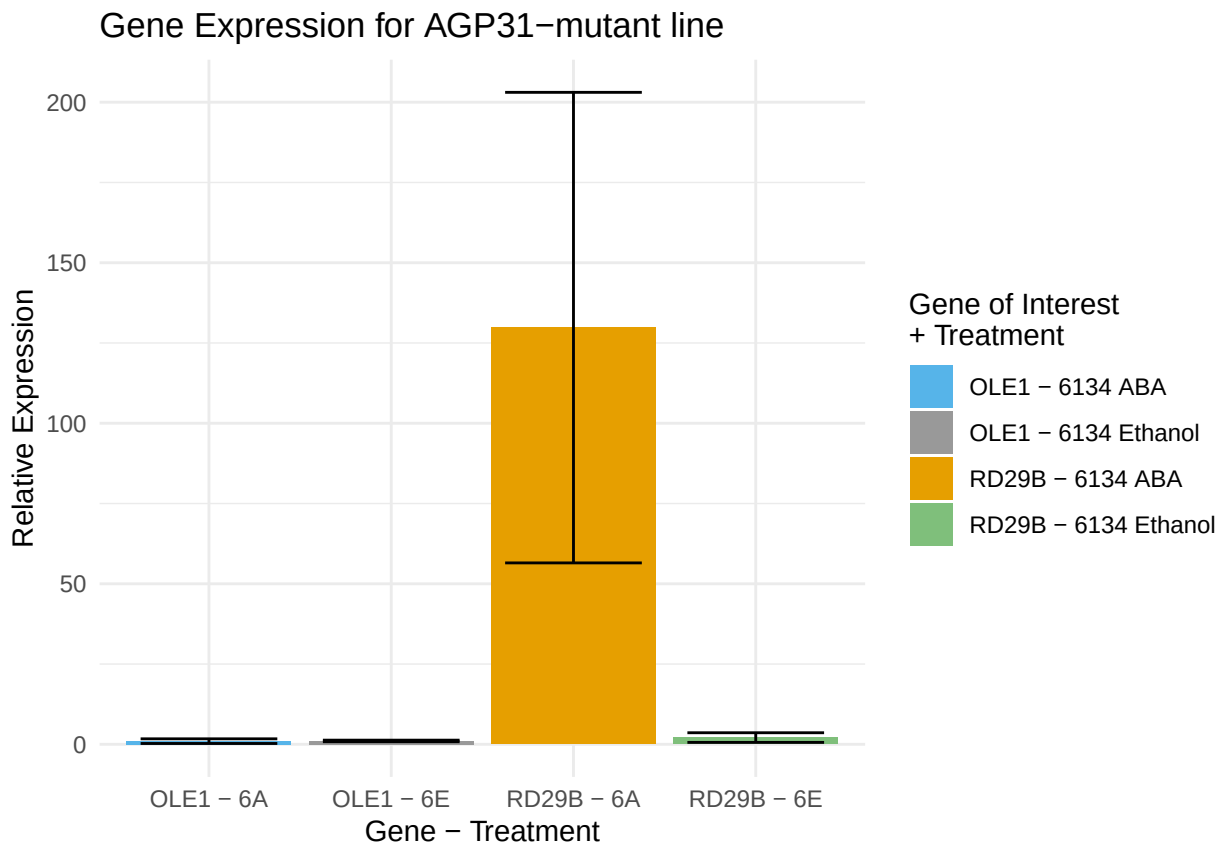

```
rdgg1 <- ggplot(t_rd, aes(Gene, expr1,fill = Gene)) + geom_col()  + geom_errorbar(aes(ymin = Lower, yma
                          breaks=c("RD29B - 6A", "RD29B - 6E"),
                          labels=c("RD29B - 6134 ABA", "RD29B - 6134 Ethanol")) +theme_minimal()
rdgg1
```



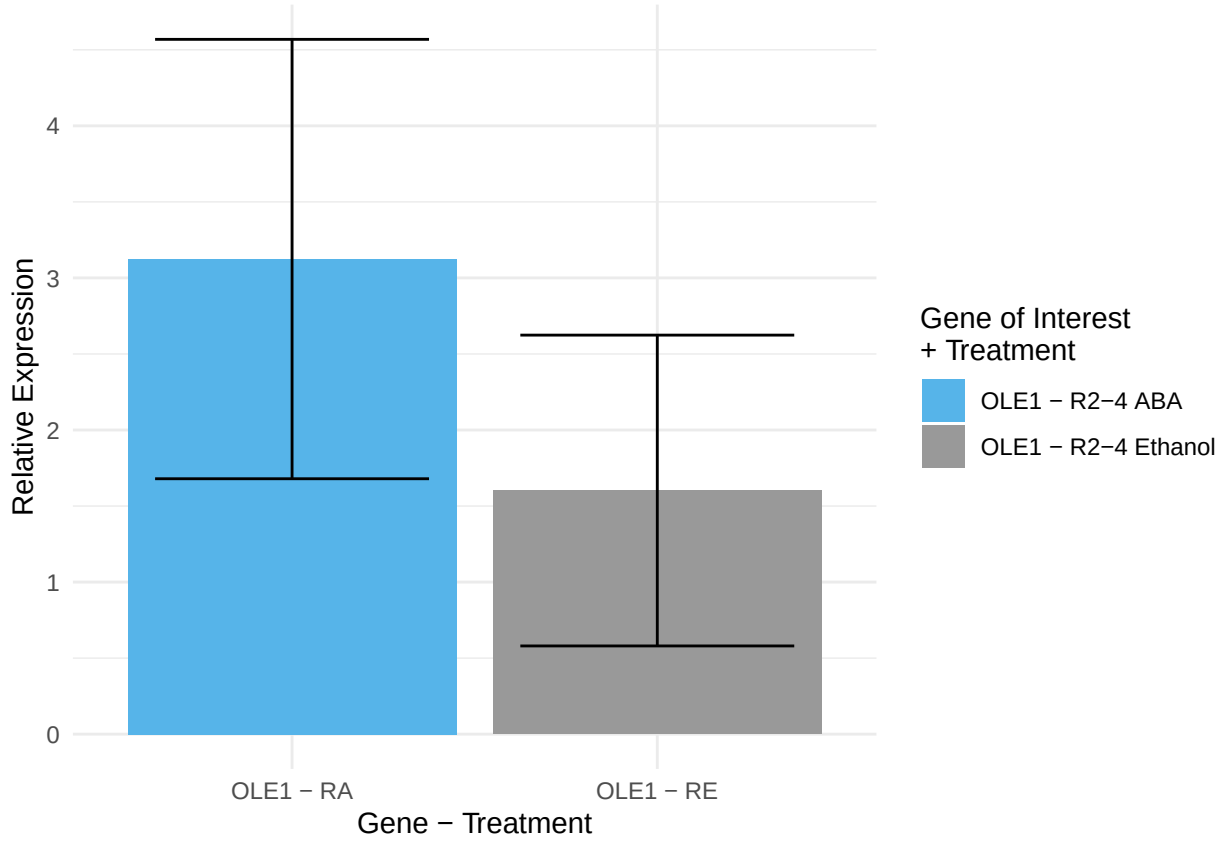

```
fig <- ggarrange(olegg1,rdgg1, labels = c("A","B"), nrow=1, ncol = 2,
                 common.legend = TRUE, legend = "right")

fig
```



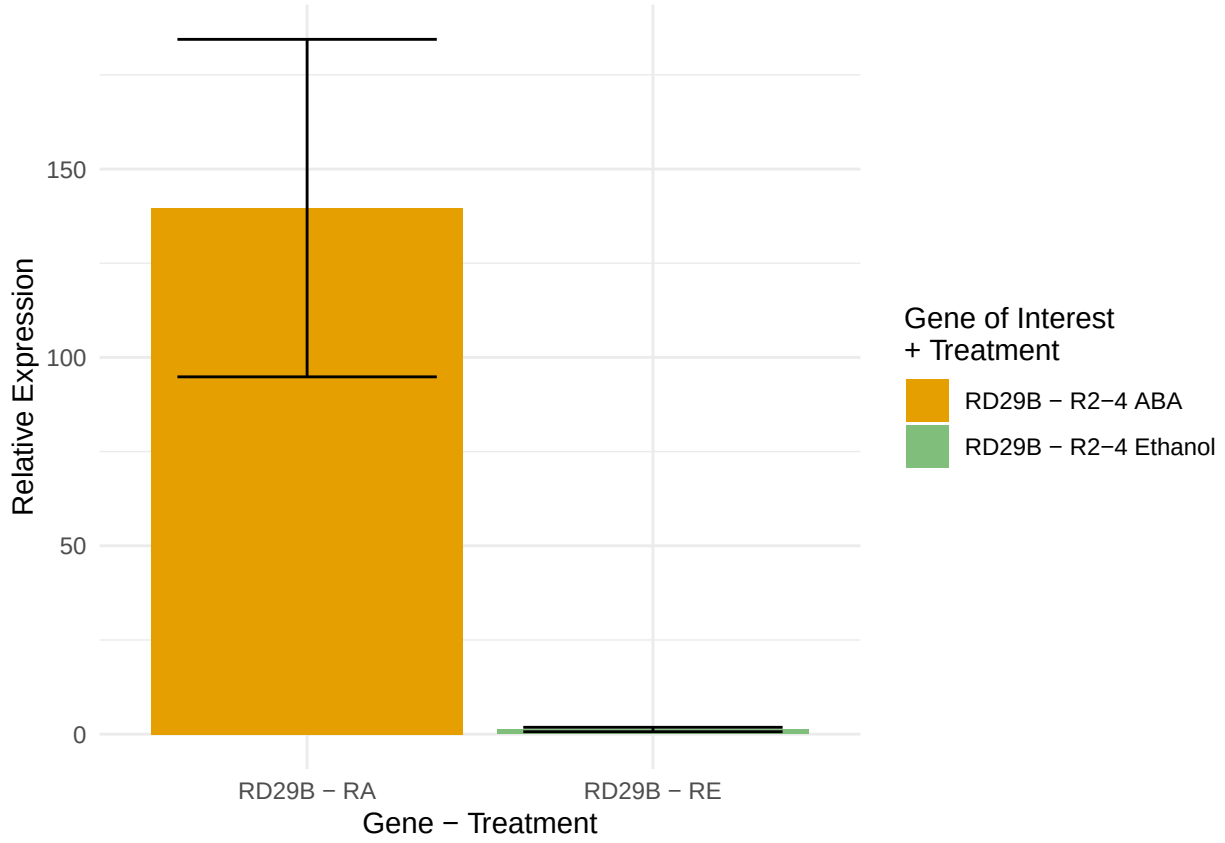

```
ggplot(t, aes(Gene, expr1, fill = Gene)) + geom_col()  + geom_errorbar(aes(ymin = Lower, ymax = Upper),
                        breaks=c("OLE1 - 6A", "OLE1 - 6E", "RD29B - 6A", "RD29B - 6E"),
                        labels=c("OLE1 - 6134 ABA", "OLE1 - 6134 Ethanol","RD29B - 6134 ABA", "RD29B -
```



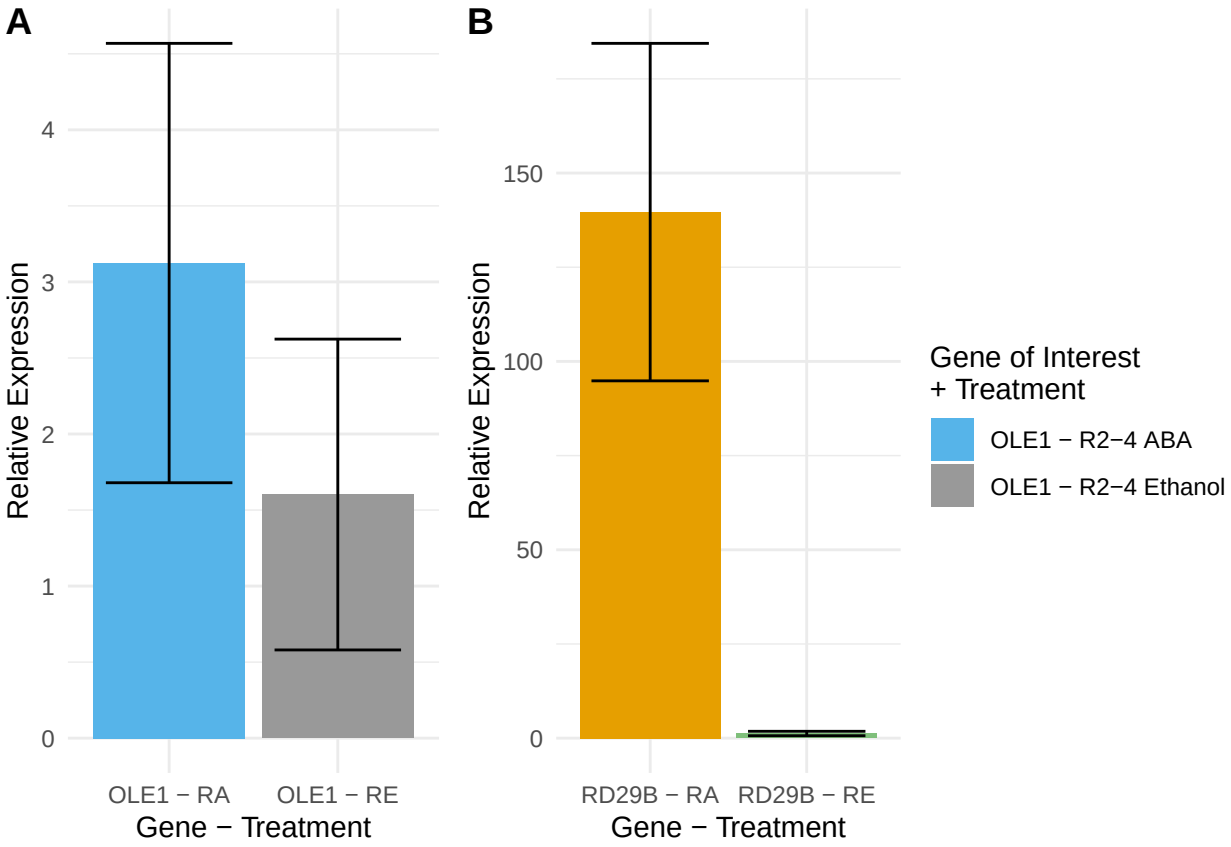

```
## REMAKING RESCUE LINE PLOTS
t <- matrix(c(3.1241226911872, 1.60205112996825, 139.63573981855, 1.23105532903857, 1.67950402187815, 0
t <- data.frame(t)

rownames(t) <- c("OLE1 - RA", "OLE1 - RE", "RD29B - RA", "RD29B - RE")
t <- rownames_to_column(t)
colnames(t) <- c("Gene","expr1","Lower","Upper")

t_ole <- t[1:2,]
t_rd <- t[3:4,]

olegg1 <- ggplot(t_ole, aes(Gene, expr1, fill = Gene)) + geom_col() + geom_errorbar(aes(ymin = Lower, yr
                          breaks=c("OLE1 - RA", "OLE1 - RE"),
                          labels=c("OLE1 - R2-4 ABA", "OLE1 - R2-4 Ethanol"))+theme_minimal()
olegg1
```



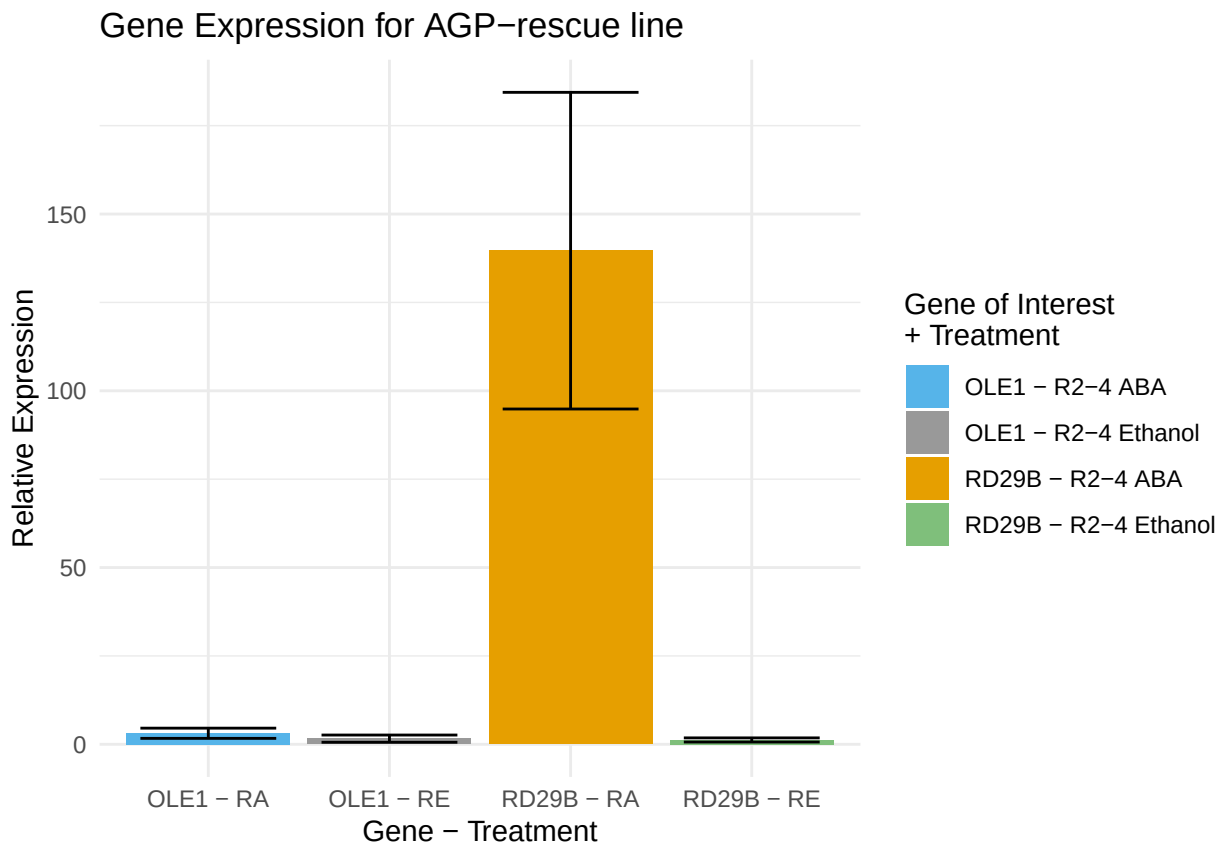

```
rdgg1 <- ggplot(t_rd, aes(Gene, expr1,fill = Gene)) + geom_col()  + geom_errorbar(aes(ymin = Lower, yma
                          breaks=c("RD29B - RA", "RD29B - RE"),
                          labels=c("RD29B - R2-4 ABA", "RD29B - R2-4 Ethanol")) +theme_minimal()
rdgg1
```




```
fig <- ggarrange(olegg1,rdgg1, labels = c("A","B"), nrow=1, ncol = 2,
                 common.legend = TRUE, legend = "right")

fig
```




```
ggplot(t, aes(Gene, expr1, fill = Gene)) + geom_col() + geom_errorbar(aes(ymin = Lower, ymax = Upper), w
                        breaks=c("OLE1 - RA", "OLE1 - RE", "RD29B - RA", "RD29B - RE"),
                        labels=c("OLE1 - R2-4 ABA", "OLE1 - R2-4 Ethanol","RD29B - R2-4 ABA", "RD29B -
```



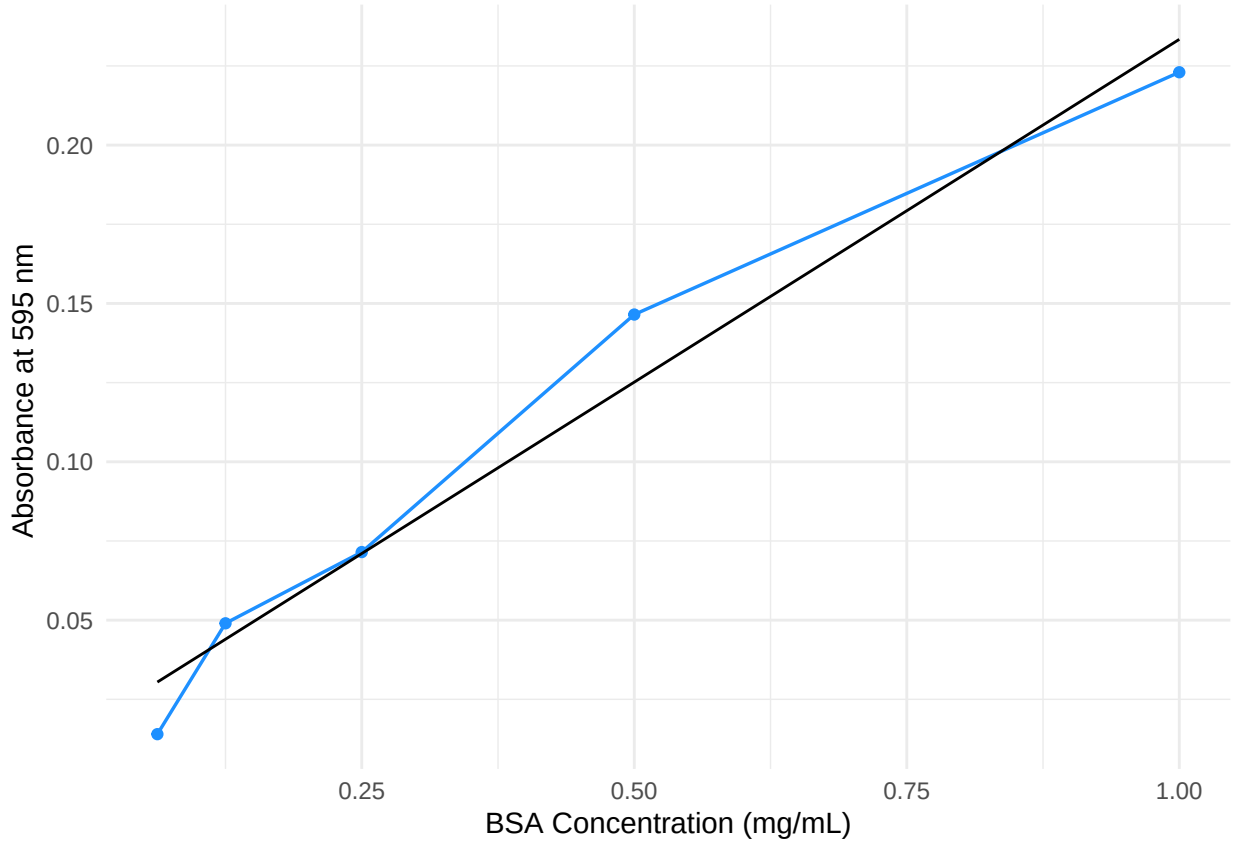

```
library(knitr)
kable(data.frame("Sample"= c("LG1","LG2","LR1","LR2","6G1","6G2","6R1","6R2"), "Total Weight (g)" = c(0
```

report 2 figure generation

Sample	Total Weight (g)
LG1	0.208
LG2	0.210
LR1	0.164
LR2	0.159
6G1	0.224
6G2	0.220
6R1	0.187
6R2	0.208

```
library(tidyverse)
bsa <- data.frame("conc" = c(0.0625,0.125,0.25,0.5,1), "abs" = c(0.014,0.049,0.0715,0.1465,0.223))

bsa %>% ggplot(aes(x = conc, y = abs)) + geom_line(col = "dodgerblue", size = .6) + geom_point(col = "d
```

```
## `geom_smooth()` using formula 'y ~ x'
```




```
lm(bsa$abs ~ bsa$conc) %>% summary()
```

```
##
## Call:
## lm(formula = bsa$abs ~ bsa$conc)
##
## Residuals:
##          1          2          3          4          5
## -0.0164462  0.0050242  0.0004651  0.0213468 -0.0103898
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.01692    0.01146   1.476  0.23638
## bsa$conc     0.21647    0.02220   9.750  0.00229 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.01693 on 3 degrees of freedom
## Multiple R-squared:  0.9694, Adjusted R-squared:  0.9592
## F-statistic: 95.06 on 1 and 3 DF,  p-value: 0.002292
```

$$y = 0.21647x + 0.01692$$

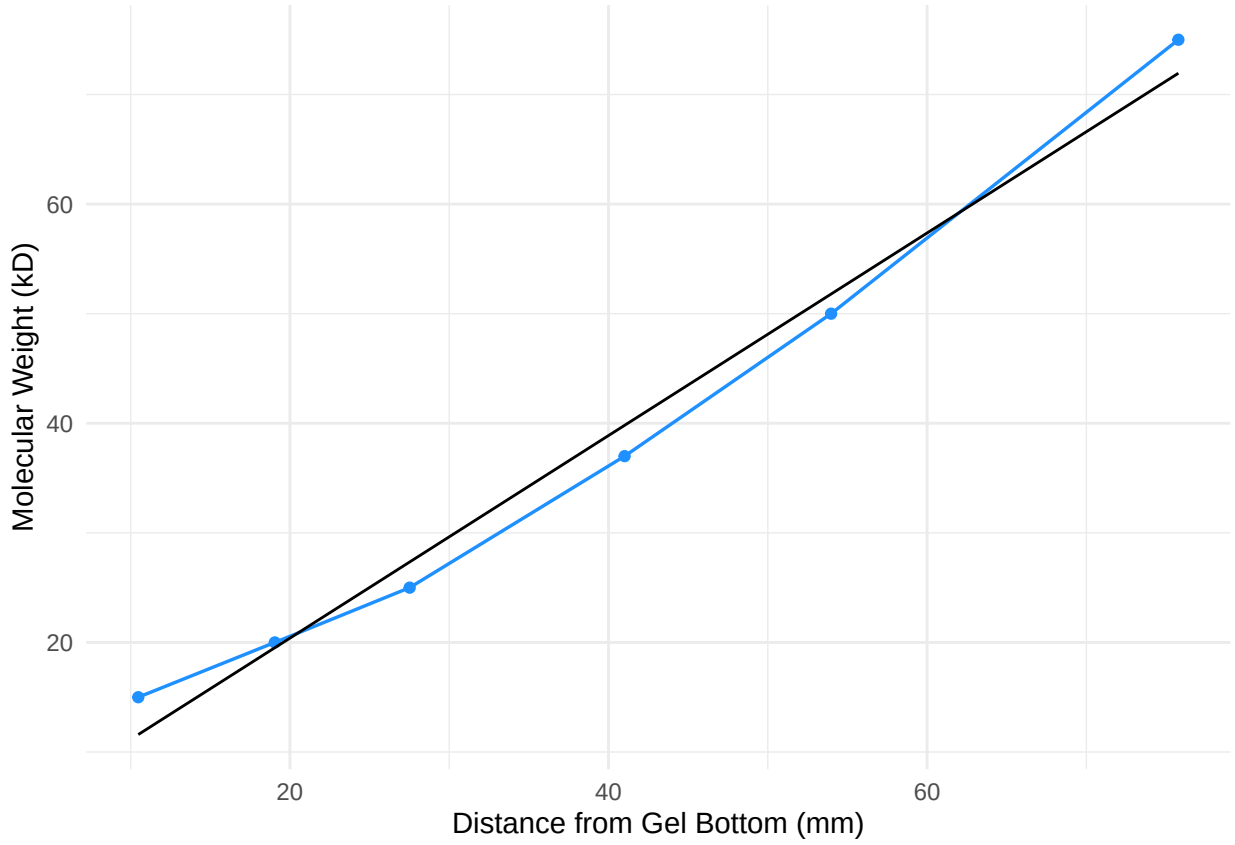
$$\mathrm{R}^2 = 0.9592$$

```
data.frame(a = c('LG1','LG2','LR1','LR2','6G1','6G2','6R1','6R2'), t = c(117,95,88,88,98,108,88,85), b =
```

Sample Name	Protein Sample Stock Volume (uL)	Absorbance at 595 nm	Sample Protein Concentration (mg/mL)	Total Sample Protein (g)	Yield (mg/g)
LG1	117	0.141	3.829960	0.4481053	2.154352
LG2	95	0.150	4.079413	0.3875443	1.845449
LR1	88	0.096	2.582692	0.2272769	1.385835
LR2	88	0.098	2.638126	0.2321551	1.460095
6G1	98	0.166	4.522886	0.4432429	1.978763
6G2	108	0.142	3.857677	0.4166291	1.893769
6R1	88	0.144	3.913111	0.3443538	1.841464
6R2	85	0.117	3.164750	0.2690038	1.293287

```
data.frame(a = c(15,20,25,37,50,75), b = c(10.477499868,19.04999976,27.51666632,41.01041615,53.97499932
pp %>% ggplot(aes(x = b, y = a)) + geom_point(col = "dodgerblue") + geom_line(col = "dodgerblue", size =
```

```
## `geom_smooth()` using formula 'y ~ x'
```




```
lm(pp$a~pp$b) %>% summary()
```

```
##
## Call:
## lm(formula = pp$a ~ pp$b)
##
```

```
## Residuals:
##       1       2       3       4       5       6
##  3.4063  0.4836 -2.3412 -2.8121 -1.7938  3.0572
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   1.9105     2.4984   0.765    0.487
## pp$b          0.9242     0.0569  16.242 8.41e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3.074 on 4 degrees of freedom
## Multiple R-squared:  0.9851, Adjusted R-squared:  0.9813
## F-statistic: 263.8 on 1 and 4 DF,  p-value: 8.408e-05
```

$$y = 0.9242x + 1.9105$$

$$\mathrm{R}^2 = 0.9813$$

```
data.frame(a = c('A','B','C','D','E','F','G','H'), b = c(7.38021785499206,20.8356474960431,22.8396476553
```

Protein Marker	Molecular Weight (kD)
A	7.380218
B	20.835648
C	22.839648
D	54.331079
E	19.690504
F	5.662503
G	17.113933
H	35.149934