Cerebellum Hypoxia QPCR

Mari Williams

Data Prep

```
library(dplyr)
```

```
Warning: package 'dplyr' was built under R version 4.3.3
```

```

Attaching package: 'dplyr'

The following objects are masked from 'package:stats':

    filter, lag

The following objects are masked from 'package:base':

    intersect, setdiff, setequal, union
```

```
library(tidyr)
```

```
Warning: package 'tidyr' was built under R version 4.3.3
```

```
library(ggplot2)
```

```
Warning: package 'ggplot2' was built under R version 4.3.3
```

```r
df <- read.csv("qpcr practice - Sheet2.csv",row.names=1)
df$treatment <- c(rep("fent", 6), rep("saline", 5))
```

```r
gene_names <- c("GAPDH", "PGK1", "ENO1", "BNIP3", "Pdk1", "EPO", "SLC2a1", "HIF1a")
rep_names  <- rep(1:3, times = length(gene_names))
gene_labels <- rep(gene_names, each = 3)
new_names <- paste0(gene_labels, "_", rep_names)

colnames(df)[1:24] <- new_names
colnames(df)[25] <- "treatment"   # last column is treatment

df$sample_id <- seq_len(nrow(df))
head(df)
```

```
    GAPDH_1  GAPDH_2  GAPDH_3   PGK1_1   PGK1_2   PGK1_3   ENO1_1   ENO1_2
M1 18.84041 18.60793 18.40817 21.57485 21.76817 21.59148 21.22606 21.10789
M2 18.82198 18.62519 18.41288 21.50123 21.72028 21.54622 21.26596 21.17082
M3 18.50397 18.07368 18.09683 21.24134 21.37162 21.22058 20.97619 20.85441
M4 18.29610 18.13960 18.08584 21.41430 21.26513 21.24518 20.79418 21.02328
M5 18.35345 18.25200 18.42662 21.32344 21.50715 21.35677 21.09079 20.81473
M6 18.36421 18.36075 18.28182 21.02036 21.30590 20.94598 20.73579 20.65836
     ENO1_3  BNIP3_1  BNIP3_2  BNIP3_3   Pdk1_1   Pdk1_2   Pdk1_3    EPO_1
M1 21.10416 23.50314 23.54776 23.65755 27.73711 27.80666 27.59428 30.46519
M2 21.28153 23.54336 23.58366 23.83811 27.78936 27.75134 27.93373 31.20307
M3 20.85714 23.11276 23.08525 23.38539 27.27982 27.32600 27.30245 31.08594
M4 20.82464 23.06449 23.09985 23.15351 27.18481 27.07314 27.23758 31.08182
M5 21.05201 23.00901 23.10994 23.32766 27.45444 27.39637 27.47734 31.01667
M6 20.74387 23.00564 23.08596 23.24686 26.89393 26.95865 27.41794 30.66494
      EPO_2    EPO_3 SLC2a1_1 SLC2a1_2 SLC2a1_3  HIF1a_1  HIF1a_2  HIF1a_3
M1 30.39515 30.74418 29.27820 29.11687 29.31546 24.58430 24.65959 24.52542
M2 31.22270 31.37660 29.81025 29.87586 30.03294 25.04854 24.96117 25.24032
M3 31.26204 31.27904 28.95826 28.82185 29.12280 24.45475 24.43622 25.17597
M4 30.57889 31.03318 28.96546 29.23028 29.36372 24.58650 24.59768 24.55592
M5 31.28418 31.65812 28.92071 29.23221 29.16674 24.65250 24.49878 24.54183
M6 30.23802 30.83582 29.72502 29.24234 29.29040 24.66163 25.12429 25.18002
   treatment sample_id
M1      fent         1
M2      fent         2
M3      fent         3
M4      fent         4
M5      fent         5
M6      fent         6
```

```r
df_long <- df %>%
  pivot_longer(
    cols = all_of(new_names),
    names_to = "gene_rep",
    values_to = "CT"
  ) %>%
  separate(gene_rep, into = c("gene", "rep"), sep = "_", remove = TRUE)
head(df_long)
```

```
# A tibble: 6 x 5
  treatment sample_id gene  rep      CT
  <chr>         <int> <chr> <chr> <dbl>
1 fent              1 GAPDH 1      18.8
2 fent              1 GAPDH 2      18.6
3 fent              1 GAPDH 3      18.4
4 fent              1 PGK1  1      21.6
5 fent              1 PGK1  2      21.8
6 fent              1 PGK1  3      21.6
```

Averaging and Computing delta CT

```r
# average the triplicates
gene_means <- df_long %>%
  group_by(sample_id, treatment, gene) %>%
  summarise(mean_CT = mean(CT, na.rm = TRUE), .groups = "drop")

# make a df with the GAPDH means
gapdh_means <- gene_means %>%
  filter(gene == "GAPDH") %>%
  select(sample_id, treatment, gapdh_mean = mean_CT)
gene_means
```

```
# A tibble: 88 x 4
   sample_id treatment gene   mean_CT
       <int> <chr>     <chr>    <dbl>
 1         1 fent      BNIP3     23.6
 2         1 fent      ENO1      21.1
 3         1 fent      EPO       30.5
 4         1 fent      GAPDH     18.6
```

```
 5         1 fent      HIF1a     24.6
 6         1 fent      PGK1      21.6
 7         1 fent      Pdk1      27.7
 8         1 fent      SLC2a1    29.2
 9         2 fent      BNIP3     23.7
10         2 fent      ENO1      21.2
# i 78 more rows
```

```
gapdh_means
```

```
# A tibble: 11 x 3
   sample_id treatment gapdh_mean
       <int> <chr>          <dbl>
 1         1 fent            18.6
 2         2 fent            18.6
 3         3 fent            18.2
 4         4 fent            18.2
 5         5 fent            18.3
 6         6 fent            18.3
 7         7 saline          18.3
 8         8 saline          18.7
 9         9 saline          18.3
10        10 saline          18.3
11        11 saline          18.1
```

```
#add the gapdh mean for that mouse to the column, and then subtract it from the CT of each o
dCT <- gene_means %>%
  left_join(gapdh_means, by = c("sample_id", "treatment")) %>%
  mutate(dCT = mean_CT - gapdh_mean) %>%
  filter(gene != "GAPDH")

dCT_summary <- dCT %>%
  group_by(treatment, gene) %>%
  summarise(mean_dCT = mean(dCT, na.rm = TRUE),
            sd_dCT   = sd(dCT, na.rm = TRUE),
            .groups = "drop")
dCT_summary
```

```
# A tibble: 14 x 4
   treatment gene   mean_dCT sd_dCT
   <chr>     <chr>     <dbl>  <dbl>
```

```
 1 fent      BNIP3      4.91 0.100
 2 fent      ENO1       2.59 0.121
 3 fent      EPO       12.6  0.424
 4 fent      HIF1a      6.36 0.237
 5 fent      PGK1       3.00 0.130
 6 fent      Pdk1       9.04 0.154
 7 fent      SLC2a1    10.9  0.252
 8 saline    BNIP3      4.85 0.0997
 9 saline    ENO1       2.62 0.132
10 saline    EPO       12.1  0.446
11 saline    HIF1a      6.30 0.236
12 saline    PGK1       3.07 0.157
13 saline    Pdk1       8.85 0.0899
14 saline    SLC2a1    10.8  0.195
```

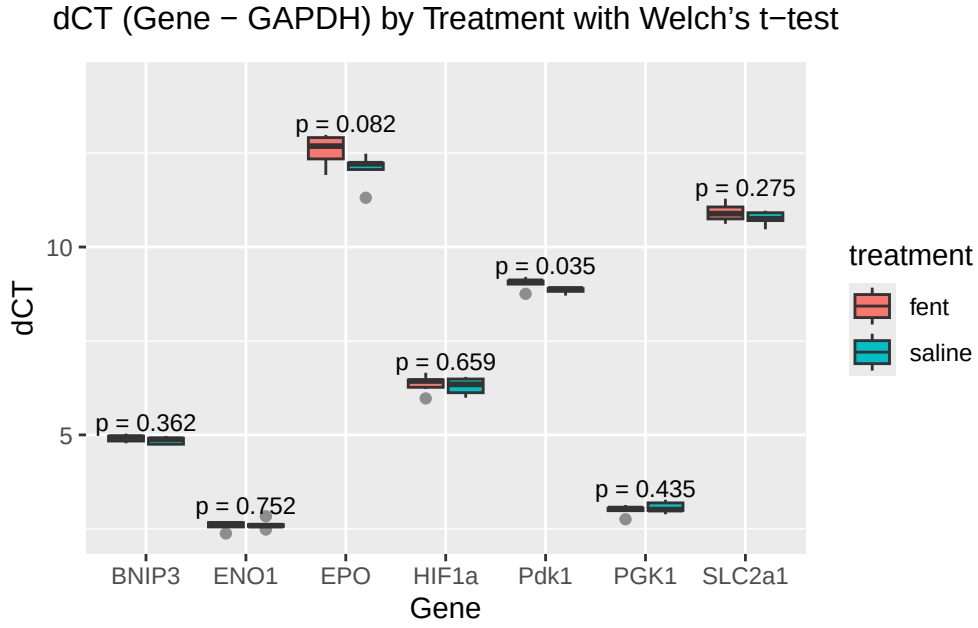
Welch test

```
t_results <- dCT %>%
  group_by(gene) %>%
  summarise(
    t_pvalue   = t.test(dCT ~ treatment, var.equal = FALSE)$p.value,
    t_stat     = unname(t.test(dCT ~ treatment, var.equal = FALSE)$statistic),
    mean_fent   = mean(dCT[treatment == "fent"], na.rm = TRUE),
    mean_saline = mean(dCT[treatment == "saline"], na.rm = TRUE),
    diff = mean_fent - mean_saline,
    .groups = "drop"
  ) %>%
  mutate(p_label = paste0("p = ", formatC(t_pvalue, format = "f", digits = 3)))

# print results sorted by p-value
t_results <- t_results %>% arrange(t_pvalue)
t_results
```

```
# A tibble: 7 x 7
  gene   t_pvalue t_stat mean_fent mean_saline    diff p_label
  <chr>     <dbl>  <dbl>     <dbl>       <dbl>   <dbl> <chr>
1 Pdk1     0.0353  2.52       9.04        8.85  0.188  p = 0.035
2 EPO      0.0819  1.97      12.6         12.1   0.521  p = 0.082
3 SLC2a1   0.275   1.16      10.9         10.8   0.157  p = 0.275
4 BNIP3    0.362   0.963      4.91         4.85  0.0582 p = 0.362
```

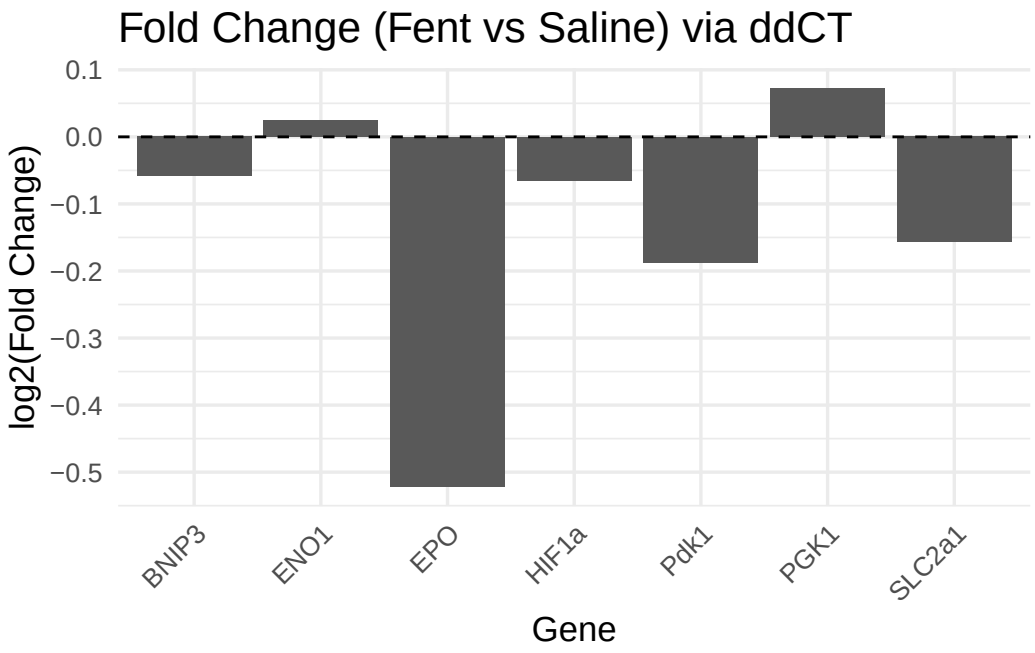
```
5 PGK1     0.435  -0.822      3.00        3.07 -0.0725 p = 0.435
6 HIF1a    0.659   0.456      6.36        6.30  0.0653 p = 0.659
7 ENO1     0.752  -0.326      2.59        2.62 -0.0251 p = 0.752
```

```r
y_pos <- dCT %>%
  group_by(gene) %>%
  summarise(y = max(dCT, na.rm = TRUE) + 0.3, .groups = "drop")

sig_df <- t_results %>% left_join(y_pos, by = "gene")

# 3) Boxplot + significance text
pd <- position_dodge(width = 0.8)

ggplot(dCT, aes(x = gene, y = dCT, fill = treatment)) +
  geom_boxplot(width = 0.7, position = pd, outlier.alpha = 0.5) +

  geom_text(
    data = sig_df,
    inherit.aes = FALSE,
    aes(x = gene, y = y, label = p_label),
    size = 3
  ) +
  scale_y_continuous(expand = expansion(mult = c(0.05, 0.15))) +  # extra headroom
  labs(
    title = "dCT (Gene - GAPDH) by Treatment with Welch's t-test",
    x = "Gene",
    y = "dCT"
  ) +
   theme(plot.margin = margin(10, 15, 10, 15), plot.title = element_text(size = 12, hjust = 
coord_cartesian(clip = "off")
```



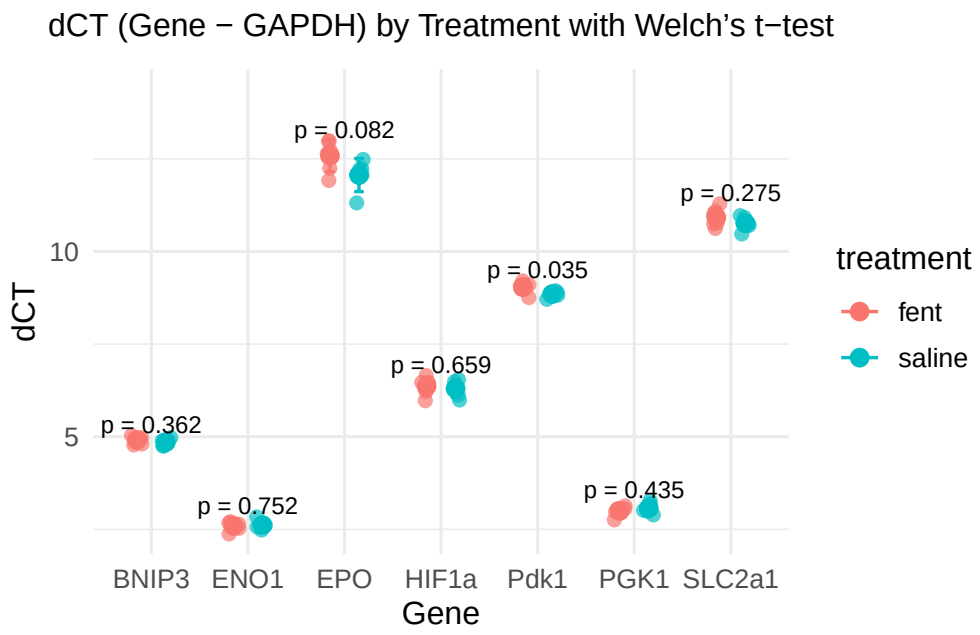

```r
dct_mean <- dCT %>%
  group_by(gene, treatment) %>%
  summarise(mean_dCT = mean(dCT, na.rm = TRUE), .groups = "drop")


ddct <- dct_mean %>%
  pivot_wider(names_from = treatment, values_from = mean_dCT) %>%
  mutate(
    ddCT = fent - saline,                 # ΔΔCT
    fold_change = 2^(-ddCT),              # relative expression
    log2_fc = -ddCT                       # log2 FC for plotting
  )


## 5) Plot log2 fold change (fent vs saline)
ggplot(ddct, aes(x = gene, y = log2_fc)) +
  geom_col() +
  geom_hline(yintercept = 0, linetype = "dashed") +
  labs(
    title = "Fold Change (Fent vs Saline) via ddCT",
    x = "Gene",
    y = "log2(Fold Change)"
  ) +
```

```
  theme_minimal(base_size = 13) +
  theme(
    axis.text.x = element_text(angle = 45, hjust = 1)  # avoid guide position warnings
  )
```




```
dCT_summary <- dCT %>%
  dplyr::group_by(gene, treatment) %>%
  dplyr::summarise(
    mean_dCT = mean(dCT, na.rm = TRUE),
    sd_dCT   = sd(dCT, na.rm = TRUE),
    .groups = "drop"
  )

# 2) Positioning
pd  <- position_dodge(width = 0.6)
pjd <- position_jitterdodge(jitter.width = 0.12, dodge.width = 0.6)

# 3) Replicate dots + mean ± SD + p-value labels
ggplot(dCT, aes(x = gene, y = dCT, color = treatment)) +
  # all replicates as dots
  geom_point(position = pjd, alpha = 0.7, size = 2) +
```

```r
  # mean ± SD per gene × treatment
  geom_errorbar(
    data = dCT_summary,
    inherit.aes = FALSE,
    aes(x = gene,
        ymin = mean_dCT - sd_dCT,
        ymax = mean_dCT + sd_dCT,
        color = treatment),
    position = pd,
    width = 0.2,
    linewidth = 0.6
  ) +
  geom_point(
    data = dCT_summary,
    inherit.aes = FALSE,
    aes(x = gene, y = mean_dCT, color = treatment),
    position = pd,
    size = 2.8
  ) +

  # p-value labels you already computed
  geom_text(
    data = sig_df,
    inherit.aes = FALSE,
    aes(x = gene, y = y, label = p_label),
    size = 3
  ) +

  scale_y_continuous(expand = expansion(mult = c(0.05, 0.15))) +
  labs(
    title = "dCT (Gene - GAPDH) by Treatment with Welch's t-test",
    x = "Gene",
    y = "dCT"
  ) +
  theme_minimal(base_size = 12) +
  theme(
    plot.margin = margin(10, 15, 10, 15),
    plot.title  = element_text(size = 12, hjust = 0.5, margin = margin(b = 10))
  ) +
  coord_cartesian(clip = "off")
```




```
library(knitr)
```

```
Warning: package 'knitr' was built under R version 4.3.3
```

```
gene_names <- c("GAPDH","PGK1","ENO1","BNIP3","Pdk1","EPO","SLC2a1","HIF1a")
new_names  <- paste0(rep(gene_names, each = 3), "_", rep(1:3, times = length(gene_names)))

# assume df already loaded; rename first 24 CT columns:
ct_idx <- 1:24
df <- df %>% rename_with(~ new_names, all_of(ct_idx))

# ensure Sample id exists (for grouping)
if (!"Sample" %in% names(df)) df$Sample <- paste0("S", seq_len(nrow(df)))

df_long <- df %>%
  pivot_longer(cols = all_of(new_names), names_to = "gene_rep", values_to = "CT") %>%
  separate(gene_rep, into = c("gene","rep"), sep = "_", remove = TRUE) %>%
  group_by(Sample, treatment, gene) %>%
  summarise(mean_CT = mean(CT, na.rm = TRUE), .groups = "drop")

gapdh <- df_long %>%
  filter(gene == "GAPDH") %>%
```

```r
  select(Sample, treatment, gapdh_CT = mean_CT)

dCT <- df_long %>%
  left_join(gapdh, by = c("Sample","treatment")) %>%
  mutate(dCT = mean_CT - gapdh_CT) %>%
  filter(gene != "GAPDH") %>%
  mutate(expr = 2^(-dCT))  # linearized per-sample

by_trt <- dCT %>%
  group_by(gene, treatment) %>%
  summarise(
    n         = dplyr::n(),
    mean_dCT  = mean(dCT, na.rm = TRUE),
    sd_dCT    = sd(dCT, na.rm = TRUE),
    mean_expr = mean(expr, na.rm = TRUE),
    .groups = "drop"
  )

# --- 4) ΔΔCT (fent - saline) and fold change 2^(-ΔΔCT) ---
wide <- by_trt %>%
  select(gene, treatment, mean_dCT, sd_dCT, mean_expr) %>%
  pivot_wider(names_from = treatment, values_from = c(mean_dCT, sd_dCT, mean_expr), names_sep

summary_tbl <- wide %>%
  mutate(
    ddCT        = mean_dCT.fent - mean_dCT.saline,
    fold_change = 2^(-ddCT)
  )

# --- 5) OPTIONAL: Welch p-value on ΔCT per gene ---
pvals <- dCT %>%
  group_by(gene) %>%
  summarise(p_value = tryCatch(t.test(dCT ~ treatment, var.equal = FALSE)$p.value, error = fu
            .groups = "drop")

summary_tbl <- summary_tbl %>% left_join(pvals, by = "gene")

# --- 6) Make a neat display table (no gt) ---
display_tbl <- summary_tbl %>%
  transmute(
    Gene = gene,
    `ΔCT fent (mean±SD)`   = sprintf("%.3f ± %.3f", mean_dCT.fent,   sd_dCT.fent),
```

```
    `ΔCT saline (mean±SD)` = sprintf("%.3f ± %.3f", mean_dCT.saline, sd_dCT.saline),
    `ΔΔCT (fent - saline)` = sprintf("%.3f", ddCT),
    `Fold change 2^-ΔΔCT`  = sprintf("%.3f", fold_change),
    `2^-ΔCT fent (mean)`   = sprintf("%.3f", mean_expr.fent),
    `2^-ΔCT saline (mean)` = sprintf("%.3f", mean_expr.saline),
    `Welch p`              = ifelse(is.na(p_value), "-", formatC(p_value, format = "f", digi
  ) %>%
  arrange(Gene)

# Render with knitr::kable (works great in Quarto PDF)
kable(display_tbl, align = "lcccccc", caption = "CB Hypoxia qPCR Summary: ΔCT, ΔΔCT, linear
```

Table 1: CB Hypoxia qPCR Summary: ΔCT, ΔΔCT, linearized (2^-ΔCT), and fold change

Gene	ΔCT fent (mean±SD)	ΔCT saline (mean±SD)	ΔΔCT (fent - saline)	Fold change 2^-ΔΔCT	2^-ΔCT fent (mean)	2^-ΔCT saline (mean)	Welch p
BNIP3	4.912 ± 0.100	4.853 ± 0.100	0.058	0.960	0.033	0.035	0.362
ENO1	2.591 ± 0.121	2.616 ± 0.132	-0.025	1.018	0.167	0.164	0.752
EPO	12.582 ± 0.424	12.061 ± 0.446	0.521	0.697	0.000	0.000	0.082
HIF1a	6.363 ± 0.237	6.298 ± 0.236	0.065	0.956	0.012	0.013	0.659
PGK1	2.998 ± 0.130	3.071 ± 0.157	-0.072	1.052	0.126	0.120	0.435
Pdk1	9.037 ± 0.154	8.849 ± 0.090	0.188	0.878	0.002	0.002	0.035
SLC2a1	10.918 ± 0.252	10.761 ± 0.195	0.157	0.897	0.001	0.001	0.275