Evaluation of Energy Intake from BXD Datasets

Dave Bridges

January 19, 2022

Contents

The goal is to identify genetic determinants of energy expenditure and of adaptive thermogenesis from BXD mice. To start we searched gene network for energy expenditure data, ignoring those involved in exercise physiology.

- **BXD_17621** Oxygen intake over 24h on NCD at 16 w age. Also included light/dark. Not adjusted for lean mass. Has mean +/- SE in mL/kg/h. From Prinen 2014 (https://doi.org/10.1016/j.cmet.2014.04.002)
- **BXD_17618** Oxygen intake over 24h at 16 w age on HFD?, males. Not adjusted for lean mass. Has mean +/- SE in mL/kg/h. From Williams (2016). Body weight in BXD_17559, lean mass in BXD_17573
- **BXD_17622** Oxygen intake over 24h at 16 w age on HFD?, males. Not adjusted for lean mass. Has mean +/- SE in mL/kg/h. From Williams (2016). Body weight in BXD_17560, lean mass in BXD_17574

```
library(readr)
ncd.pirinen <- read_csv("BXD_17621.csv", skip=9) %>%
  mutate(Diet="NCD",Age=16,Dataset="Prinen")


williams.ncd.ee <- read_csv("BXD_17618.csv", skip=9)%>% #may be mislabelled on genenetwork, assigned ba
  mutate(Diet="NCD",Age=16,Dataset="Williams")
williams.ncd.bw <- read_csv("BXD_17559.csv" , skip=9)%>%
  mutate(Diet="NCD",Age=16,Dataset="Williams")
williams.ncd.lm <- read_csv("BXD_17573.csv" , skip=9)%>%
  mutate(Diet="NCD",Age=16,Dataset="Williams")

williams.ncd <- full_join(williams.ncd.ee,williams.ncd.bw, suffix=c('_ee','_bw'), by=c("Name","Dataset"
  full_join(williams.ncd.lm) %>%
  mutate(Value_lm = Value,
         SE_lm = SE)

williams.hfd.ee <- read_csv("BXD_17622.csv" , skip=9)%>% #may be mislabelled on genenetwork
  mutate(Diet="HFD",Age=16,Dataset="Williams")
williams.hfd.bw <- read_csv("BXD_17560.csv", skip=9)%>%
  mutate(Diet="HFD",Age=16,Dataset="Williams")
```

```r
williams.hfd.lm <- read_csv("BXD_17574.csv", skip=9)%>%
  mutate(Diet="HFD",Age=16,Dataset="Williams")

williams.hfd <- full_join(williams.hfd.ee,williams.hfd.bw, suffix=c('_ee','_bw'), by=c("Name","Dataset"
  full_join(williams.hfd.lm) %>%
  mutate(Value_lm = Value,
         SE_lm = SE)

data <- bind_rows(#ncd.pirinen,
                  williams.ncd,
                  williams.hfd) %>% # in mL/kg/h
  mutate(VO2_g_min = Value_ee/1000) %>% #in mL/g/h
  mutate(VO2_min = VO2_g_min*Value_bw/60) %>% # in mL/min #this seems reasonable
  mutate(MR_KJ_d = VO2_min * 60 * 24 / 1000 * 4.84 * 4.184,
         MR_KJ_d_SE = SE_ee/1000*Value_bw/60* 60 * 24 / 1000 * 4.84 * 4.184) %>% # 60min/h x 24h/day /
  mutate(MR_W = MR_KJ_d * 0.0115740741,
         MR_W_SE = MR_KJ_d_SE* 0.0115740741) %>% # in Watts
  mutate(Diet = relevel(factor(Diet), ref="NCD"))
```

These data can be found in /Users/davebrid/Documents/GitHub/TissueSpecificTscKnockouts/Other Published Data/Systems Biology. This script was most recently updated on Sun Feb 13 10:20:35 2022.
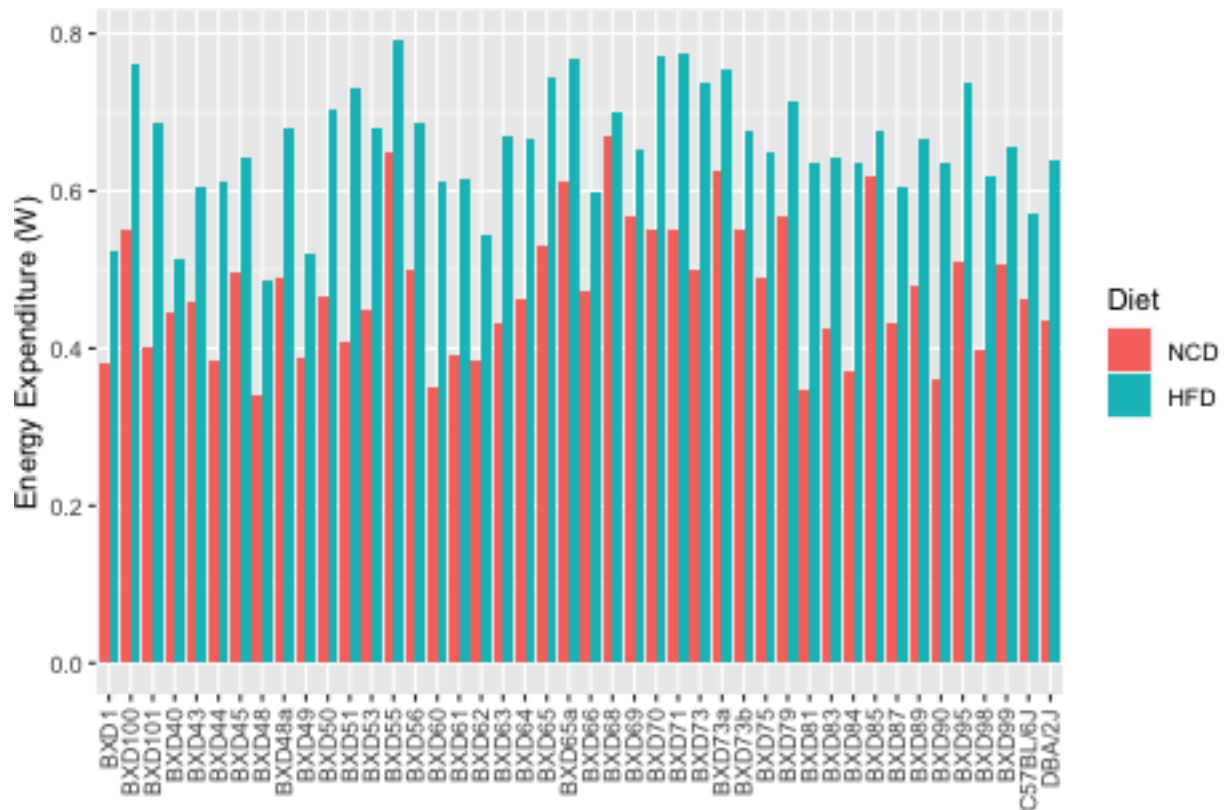
Analysis

Comparason of Datasets

```r
library(ggplot2)
data %>%
  filter(!(is.na(MR_W))) %>% # complete cases only
  ggplot(aes(y=MR_W,
         x=Name,
         ymin=MR_W-MR_W_SE,
         ymax=MR_W-MR_W_SE,
         fill=Diet)) +
  geom_bar(stat='identity',position='dodge') +
  labs(y="Energy Expenditure (W)",
       x="") +
  theme(axis.text.x = element_text(angle = 90, vjust = 0.5, hjust=1))
```

```
#lm(Value~Name+Diet,data=data) %>% summary

mr.order <-
  data %>%
  filter(Diet == "NCD") %>%
  arrange(desc(MR_W)) %>%
  mutate(Name = factor(Name))


data %>%
  filter(!is.na(MR_W)) %>% # complete cases only
  mutate(Name = factor(Name, levels = mr.order$Name, ordered = TRUE)) %>%
  ggplot(aes(y=MR_W,
        x=Name,
        ymin=MR_W-MR_W_SE,
        ymax=MR_W+MR_W_SE,
        fill=Diet)) +
  geom_bar(stat='identity',position='dodge', width=0.75) +
  geom_errorbar(position=position_dodge(width=0.75), width=0.5) +
  labs(y="Energy Expenditure (W)",x="") +
  theme(axis.text.x = element_text(angle = 90, vjust = 0.5, hjust=1))
```
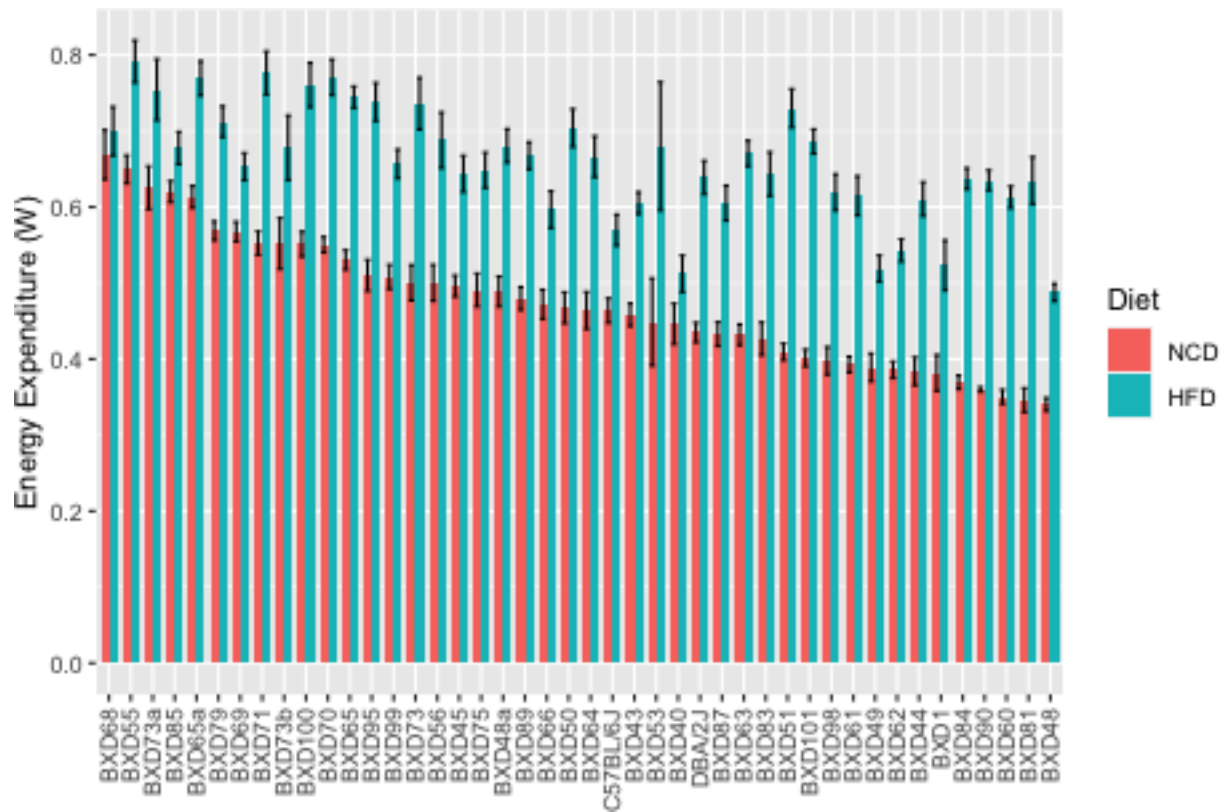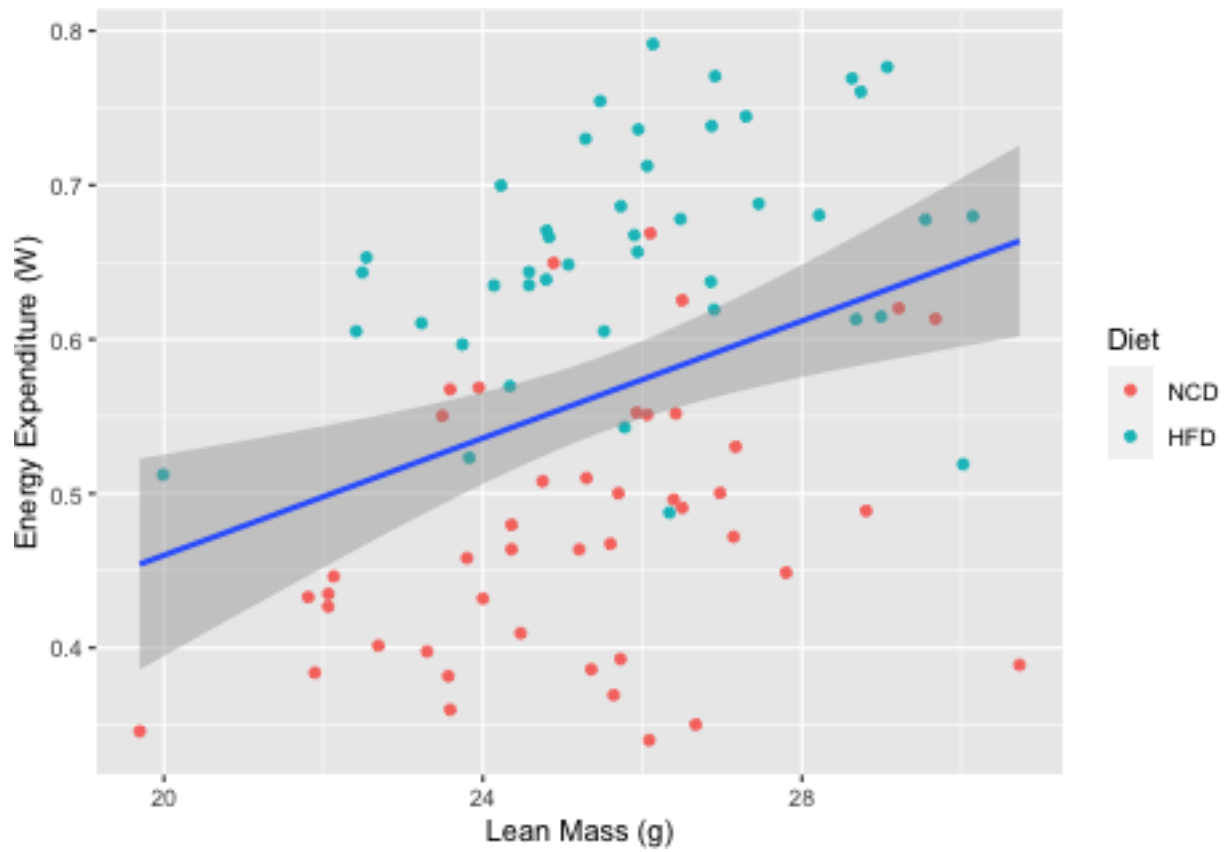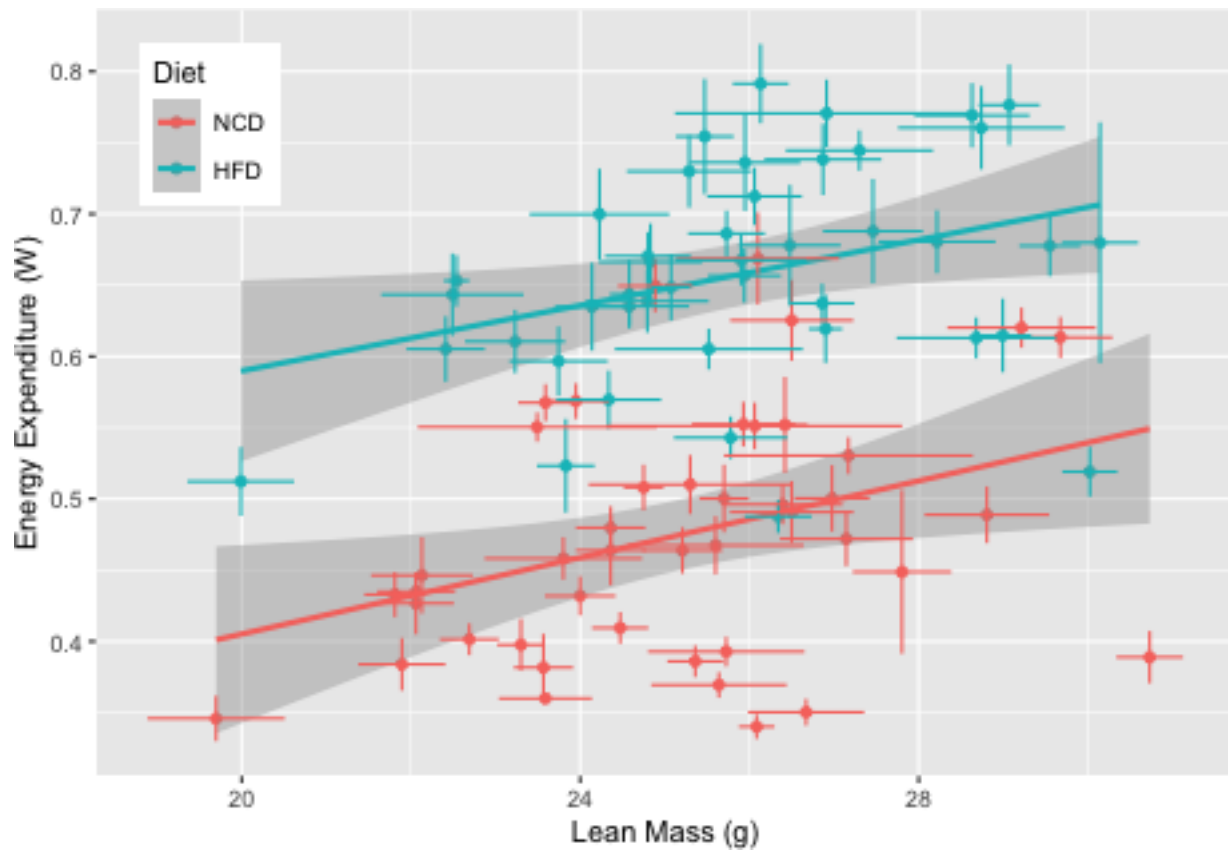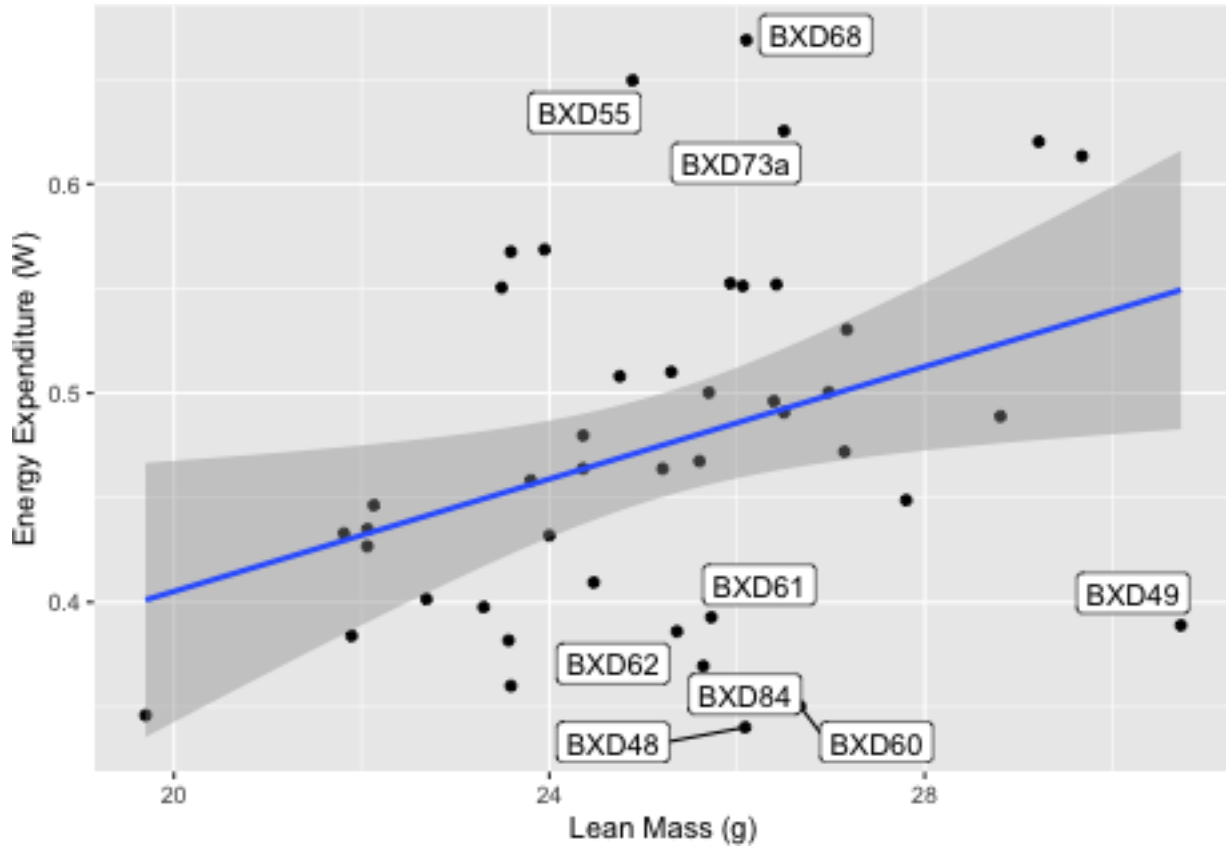
Adjusting for Lean Mass

```r
library(ggrepel)
ggplot(data, aes(y=MR_W,
          x=Value_lm)) +
  geom_point(aes(col=Diet)) +
  geom_smooth(method="lm") +
  #geom_label_repel(data = subset(data, (MR_W < 0.45&Value_lm>25.5)|MR_W>0.65&Value_lm<27), aes(label=N
  labs(y="Energy Expenditure (W)",
      x="Lean Mass (g)")
```

```
ggplot(data, aes(y=MR_W,
            x=Value_lm,
            col=Diet)) +
  geom_point() +
  geom_smooth(method="lm") +
  geom_errorbar(aes(ymin=MR_W-MR_W_SE,
                ymax=MR_W+MR_W_SE)) +
  geom_errorbarh(aes(xmin=Value_lm-SE_lm,
                xmax=Value_lm+SE_lm))+
  #geom_label_repel(data = subset(data, (MR_W < 0.45&Value_lm>25.5)|MR_W>0.65&Value_lm<27),aes(label=Na
  labs(y="Energy Expenditure (W)",
      x="Lean Mass (g)") +
  theme(legend.position = c(0.1,0.85))
```

```
#chow only
ggplot(data %>% filter(Diet=="NCD"), aes(y=MR_W,
        x=Value_lm)) +
  geom_point() +
  geom_smooth(method="lm") +
  geom_label_repel(data = subset(data %>% filter(Diet=="NCD"),
                                (MR_W < 0.43&Value_lm>24.5)|MR_W>0.60&Value_lm<27),
                   aes(label=Name)) +
  guides(fill = guide_legend(override.aes = aes(color = NA))) +
  labs(y="Energy Expenditure (W)",
       x="Lean Mass (g)")
```

```
lm.model.1 <- lm(MR_W~Value_lm,data=data %>% filter(Diet=="NCD")) #model built on only NCD
lm.model.2 <- lm(MR_W~Value_lm+Diet,data=data) #model built on NCD and AT
library(broom)
aov(lm.model.1) %>% tidy %>% kable(caption="Model 1 summary for adjusting for lean mass")
```

Table 1: Model 1 summary for adjusting for lean mass

term	df	sumsq	meansq	statistic	p.value
Value_lm	1	0.040	0.040	5.96	0.019
Residuals	42	0.279	0.007	NA	NA

```
summary(lm.model.1) %>% tidy %>% kable(caption="Model 1 coefficients for adjusting for lean mass")
```

Table 2: Model 1 coefficients for adjusting for lean mass

term	estimate	std.error	statistic	p.value
(Intercept)	0.136	0.139	0.981	0.332
Value_lm	0.013	0.006	2.442	0.019

```
aov(lm.model.2) %>% tidy %>% kable(caption="Model 2 summary for adjusting for lean mass")
```

Table 3: Model 2 summary for adjusting for lean mass

term	df	sumsq	meansq	statistic	p.value
Value_lm	1	0.159	0.159	27.2	0
Diet	1	0.639	0.639	109.5	0
Residuals	84	0.490	0.006	NA	NA

```r
summary(lm.model.2) %>% tidy %>% kable(caption="Model 2 coefficients for adjusting for lean mass")
```

Table 4: Model 2 coefficients for adjusting for lean mass

term	estimate	std.error	statistic	p.value
(Intercept)	0.160	0.094	1.71	0.091
Value_lm	0.012	0.004	3.38	0.001
DietHFD	0.174	0.017	10.47	0.000

```r
data <- data %>%
  mutate(MR_predicted = predict(lm.model.1, newdata = list(Value_lm=Value_lm))) %>%
  mutate(MR_resid = MR_W-MR_predicted) %>%
  mutate(MR_adj = MR_resid + coef(lm.model.1)['(Intercept)'] + coef(lm.model.2)['Value_lm']*mean(data$Va

data %>%
  filter(!is.na(MR_W)) %>% # complete cases only
  ggplot(aes(y=MR_resid,
        x=reorder(Name,-MR_W),
        ymin=MR_resid-MR_W_SE,
        ymax=MR_resid-MR_W_SE,
        fill=Diet)) +
  #geom_label_repel(label=Name) +
  geom_bar(stat='identity',position='dodge') +
  labs(y="EE Observed - EE Predicted (W)") +
  theme(axis.text.x = element_text(angle = 90, vjust = 0.5, hjust=1))
```
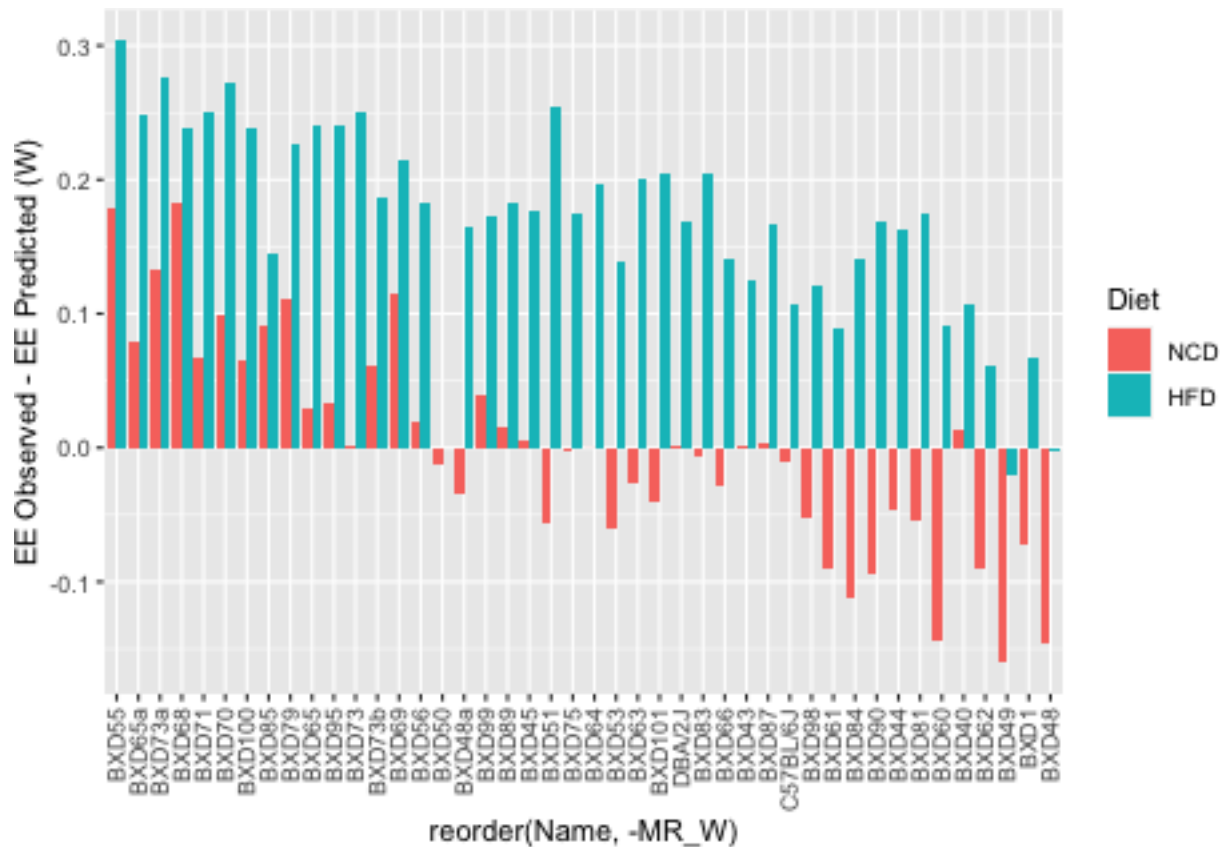
```r
mr.adj.order <-
  data %>%
  filter(Diet == "NCD") %>%
  arrange(desc(MR_adj)) %>%
  mutate(Name = factor(Name))


data %>%
  filter(!is.na(MR_W)) %>% # complete cases only
  mutate(Name = factor(Name, levels = mr.adj.order$Name, ordered = TRUE)) %>%
  ggplot(aes(y=MR_adj,
        x=Name,
        ymin=MR_adj-MR_W_SE,
        ymax=MR_adj+MR_W_SE,
        fill=Diet)) +
  geom_bar(stat='identity',position='dodge', width=0.75) +
  geom_errorbar(position=position_dodge(width=0.75), width=0.5) +
  labs(y="Lean Mass Adjusted Energy Expenditure (W)",x="") +
  theme(axis.text.x = element_text(angle = 90, vjust = 0.5, hjust=1))
```
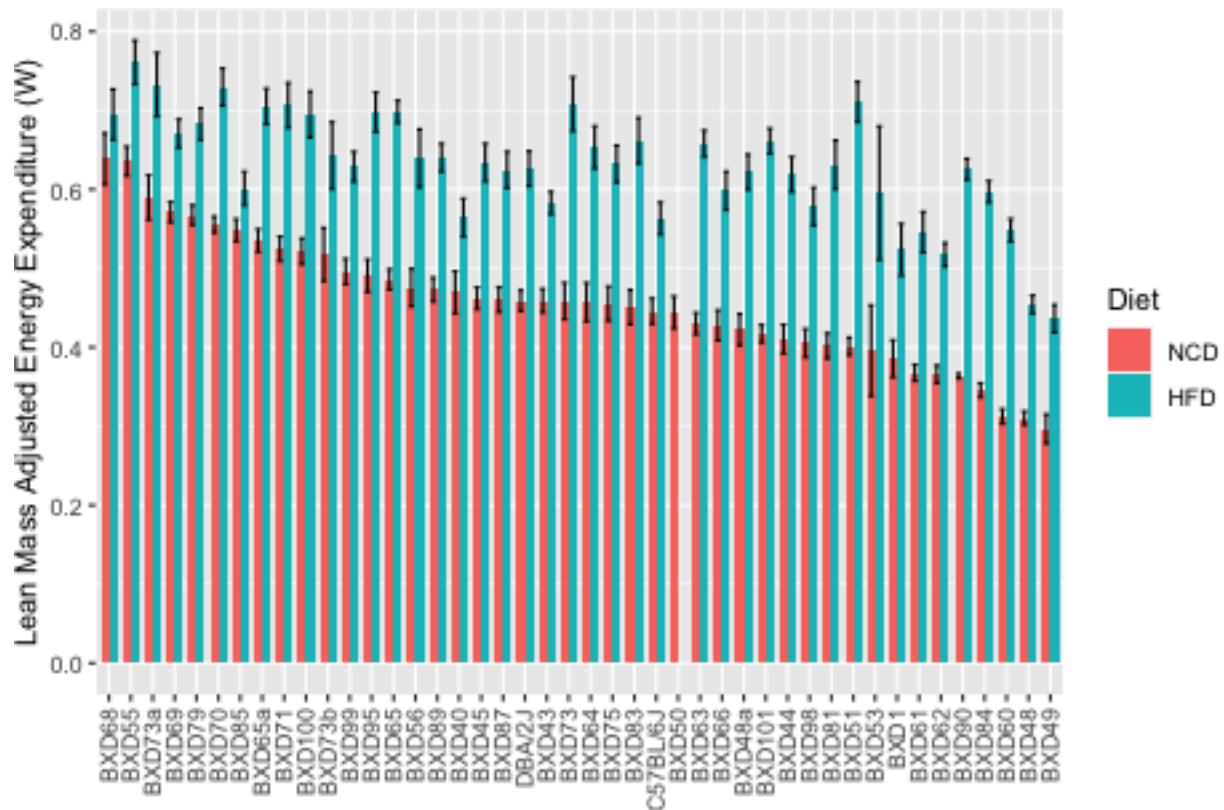
based on this modelling after adjusting for lean mass, HFD increases thermogenesis by `(coef(lm.model.2)["(Intercept)"]-c`

Adaptive Thermogenesis

Defined as lean mass adjusted VO2 from HFD - NCD

```
data.wide <-
  data %>%
  select(Value_lm,Value_bw, MR_W, MR_W_SE, MR_adj,Name,Diet) %>%
  pivot_wider(names_from=Diet,id_cols=Name,values_from=c(Value_lm,Value_bw, MR_W,MR_W_SE,MR_adj)) %>%
  mutate(AT = MR_W_HFD - MR_W_NCD,
         AT_SE = sqrt((MR_W_SE_NCD/MR_W_NCD)^2+(MR_W_SE_HFD/MR_W_HFD)^2)*AT,
         Wt.Gain = Value_bw_HFD-Value_bw_NCD)

data.wide %>%
  filter(!is.na(AT)) %>% # complete cases only
  ggplot(aes(y=AT,
         x=reorder(Name,-AT),
         ymin=AT-AT_SE,
         ymax=AT+AT_SE)) +
  geom_bar(stat='identity',position='dodge') +
    geom_errorbar() +
  labs(y="HFD-Induced Adaptive Thermogenesis (W)",
       x="") +
  theme(axis.text.x = element_text(angle = 90, vjust = 0.5, hjust=1))
```
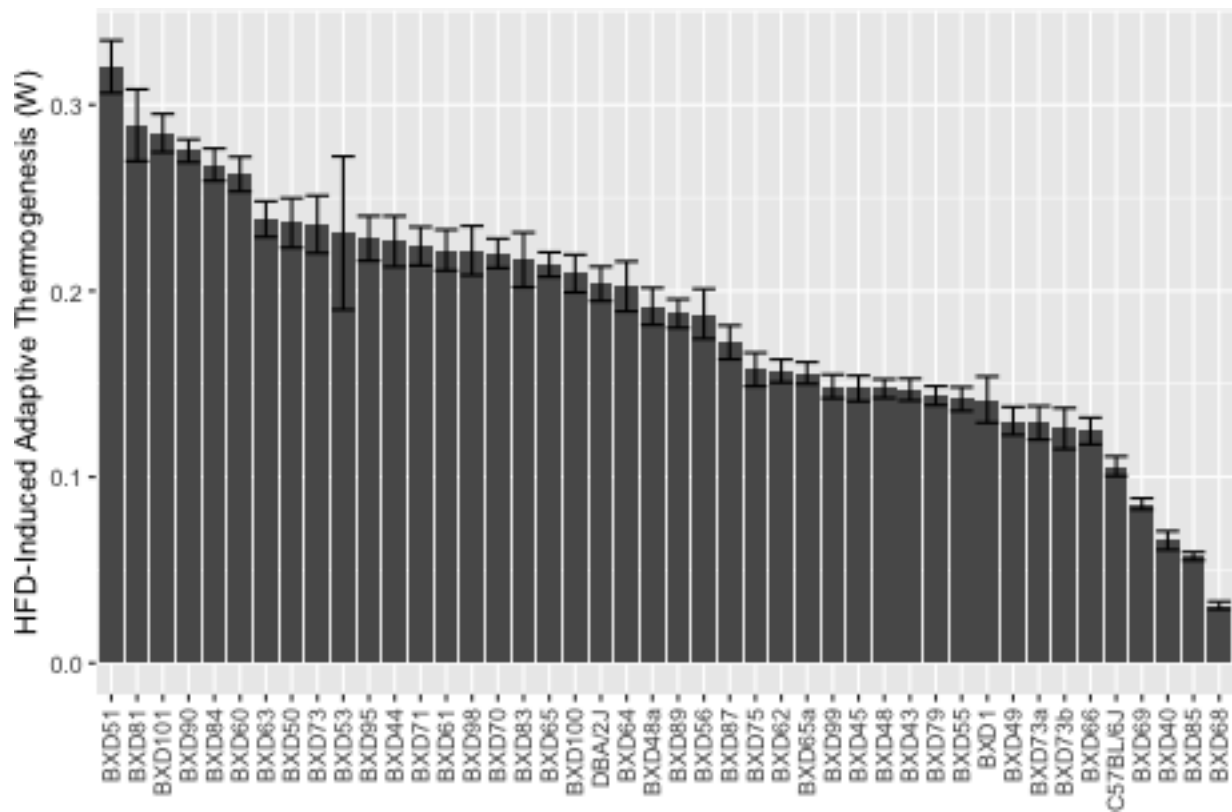
Thermogenesis on NCD as a Predictor of Weight Gain

```
data.wide %>%
  ggplot(aes(y=Wt.Gain,
             x=MR_W_NCD)) +
  labs(y="Weight Gained on HFD",
       x="Energy Expenditure on NCD (W)") +
  geom_point() +
  geom_smooth(method="lm") +
  theme_classic() +
  theme(text=element_text(size=18))
```
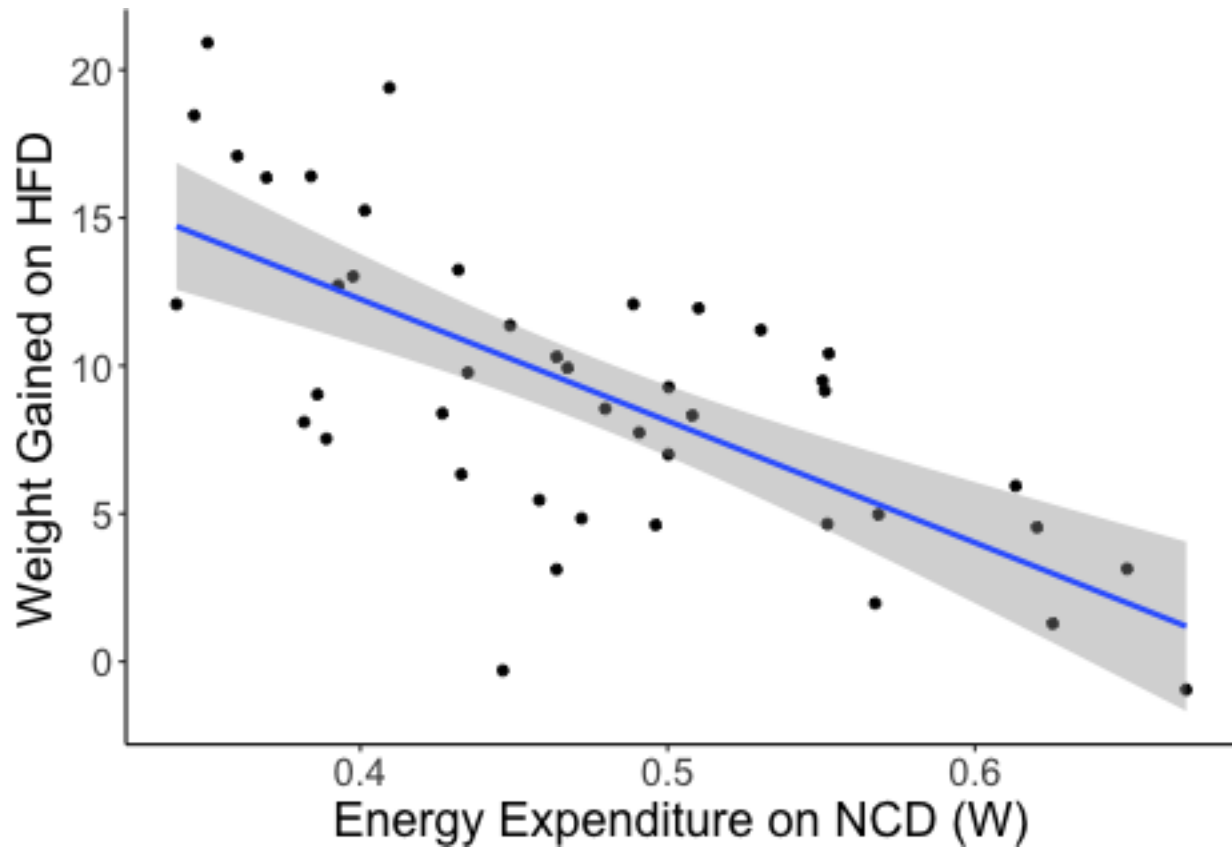
```
lm(Wt.Gain~MR_W_NCD, data=data.wide) %>% glance %>% kable(caption="Summary of relationship between energ
```

Table 5: Summary of relationship between energy expenditure and diet-induced weight gain

r.squared	adj.r.squared	sigma	statistic	p.value	df	logLik	AIC	BIC	deviance	df.residual	nobs
0.474	0.462	3.77	37.9	2.4e-07	1	-120	246	251	598	42	44

```
data.wide %>%
  ggplot(aes(y=Wt.Gain,
             x=MR_adj_NCD)) +
  labs(y="Weight Gained on HFD",
       x="Adjusted Energy Expenditure on NCD (W)") +
  geom_point() +
  geom_smooth(method="lm") +
  theme_classic() +
  theme(text=element_text(size=18))
```
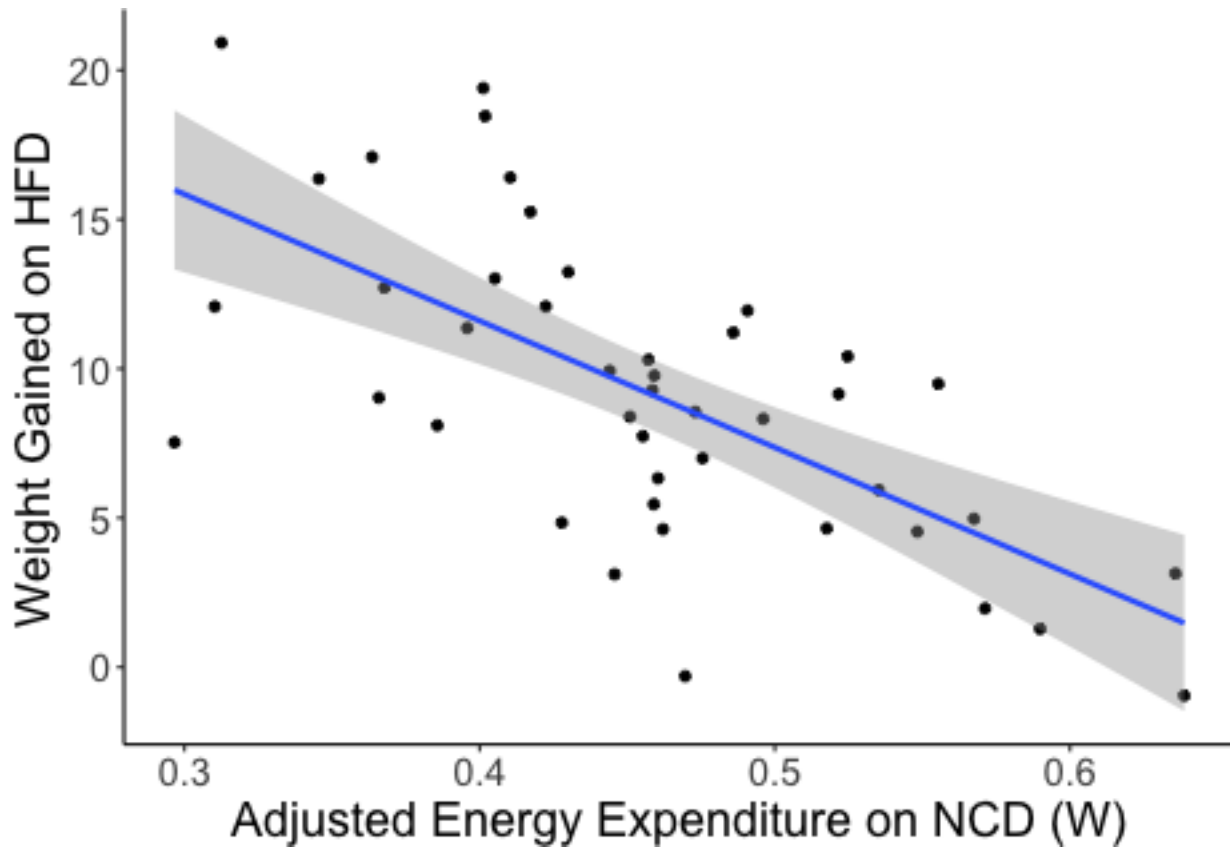
```r
lm(Wt.Gain~MR_adj_NCD, data=data.wide) %>% glance %>% kable(caption="Summary of relationship between lea
```

Table 6: Summary of relationship between lean mass adjusted
energy expenditure and diet-induced weight gain

r.squared	adj.r.squared	sigma	statistic	p.value	df	logLik	AIC	BIC	deviance	df.residual	nobs
0.441	0.427	3.89	33.1	9e-07	1	-121	248	254	636	42	44

```r
gemma.phenotype.export <- 'Strain Level Energy Expenditure Data.csv'
data %>%
  filter(Diet=="NCD") %>%
  select(Name, MR_W, MR_adj) %>%
  write_csv(gemma.phenotype.export)
```

The data on lean mass adjusted thermogenesis was exported to Strain Level Energy Expenditure Data.csv

Heritability of NCD Thermogenesis

Since we dont have individual mouse data we will make fake data based on the mean and se of MR

```r
new.sim.data <- data.frame(Name=NA, Diet=NA,EE=NA)

for (row in 1:dim(data)[1]) {
  strain.data <- data[row,]
  if(!(is.na(strain.data$MR_W))){
  sim.data <- with(strain.data,
```

```r
                  rnorm(mean=MR_W,
                        sd=MR_W_SE * sqrt(N),
                        n=N_ee
                  ))
  sim.lean.data <- with(strain.data,
                  rnorm(mean=Value_lm,
                        sd=SE_lm * sqrt(N),
                        n=N_ee
                  ))
  sim.dataset <- data.frame(Name=strain.data$Name,
                        Diet=strain.data$Diet,
                        EE=sim.data,
                        Lean=sim.lean.data)
  new.sim.data <- bind_rows(new.sim.data,sim.dataset)
  }
else{
    sim.dataset <- data.frame(Name=strain.data$Name,
                        Diet=strain.data$Diet,
                        EE=NA,
                        Lean=NA)
    new.sim.data <- bind_rows(new.sim.data,sim.dataset)
}

  }


aov(EE ~ Name, data=new.sim.data %>% filter(Diet=='NCD')) %>%
  tidy %>%
  mutate(Total.Var=sum(meansq),
         Pct.Var = meansq/Total.Var*100) %>%
  kable(caption="Overall heritability of energy expenditure on NCD mice")
```

Table 7: Overall heritability of energy expenditure on NCD mice

term	df	sumsq	meansq	statistic	p.value	Total.Var	Pct.Var
Name	43	1.431	0.033	16.9	0	0.035	94.43
Residuals	148	0.291	0.002	NA	NA	0.035	5.57

```r
aov(EE ~ Lean + Name, data=new.sim.data %>% filter(Diet=='NCD')) %>%
  tidy %>%
  mutate(Total.Var=sum(meansq),
         Pct.Var = meansq/Total.Var*100) %>%
  kable(caption="Overall heritability of energy expenditure on NCD including lean mass")
```

Table 8: Overall heritability of energy expenditure on NCD including lean mass

term	df	sumsq	meansq	statistic	p.value	Total.Var	Pct.Var
Lean	1	0.102	0.102	53.1	0	0.135	75.64
Name	43	1.335	0.031	16.1	0	0.135	22.93
Residuals	147	0.283	0.002	NA	NA	0.135	1.42

```r
aov(EE ~ Lean + Name, data=new.sim.data %>% filter(Diet=='NCD')) %>%
  tidy %>%
  mutate(Total.Var=sum(meansq[2:3]),
         Pct.Var = meansq/Total.Var*100) -> lean.adj.ee.lean

aov(EE ~ Lean + Name, data=new.sim.data %>% filter(Diet=='NCD')) %>%
  tidy %>%
  mutate(Total.Var=sum(meansq),
         Pct.Var = meansq/Total.Var*100) -> lean.adj.ee.all

lean.adj.ee.lean %>% kable(caption="Overall heritability of energy expenditure on NCD adjusting for lea
```

Table 9: Overall heritability of energy expenditure on NCD adjusting for lean mass

term	df	sumsq	meansq	statistic	p.value	Total.Var	Pct.Var
Lean	1	0.102	0.102	53.1	0	0.033	310.59
Name	43	1.335	0.031	16.1	0	0.033	94.16
Residuals	147	0.283	0.002	NA	NA	0.033	5.84

```r
aov(EE ~ Lean + Name + Diet + Name:Diet, data=new.sim.data) %>%
  tidy %>%
  mutate(Total.Var=sum(meansq),
         Pct.Var = meansq/Total.Var*100) -> hfd.incl.ee

hfd.incl.ee  %>% kable(caption="Overall heritability of energy expenditure including diet and lean mass
```

Table 10: Overall heritability of energy expenditure including diet and lean mass

term	df	sumsq	meansq	statistic	p.value	Total.Var	Pct.Var
Lean	1	0.718	0.718	236.47	0	3.48	20.620
Name	43	2.201	0.051	16.87	0	3.48	1.471
Diet	1	2.700	2.700	889.77	0	3.48	77.589
Name:Diet	42	0.341	0.008	2.68	0	3.48	0.234
Residuals	291	0.883	0.003	NA	NA	3.48	0.087

```r
aov(EE ~ Lean + Name + Diet + Name:Diet, data=new.sim.data) %>%
  tidy %>%
  mutate(Total.Var=sum(meansq[c(2,4,5)]),
         Pct.Var = meansq/Total.Var*100) -> hfd.adj.ee.adj

aov(EE ~ Lean + Name + Diet + Name:Diet, data=new.sim.data) %>%
  tidy %>%
  mutate(Total.Var=sum(meansq[c(2,3,4,5)]),
         Pct.Var = meansq/Total.Var*100) -> hfd.adj.ee.all

hfd.adj.ee.adj %>% kable(caption="Overall heritability of energy expenditure adjusted for diet and lean
```

Table 11: Overall heritability of energy expenditure adjusted for diet and lean mass

term	df	sumsq	meansq	statistic	p.value	Total.Var	Pct.Var
Lean	1	0.718	0.718	236.47	0	0.062	1150.91
Name	43	2.201	0.051	16.87	0	0.062	82.09
Diet	1	2.700	2.700	889.77	0	0.062	4330.66
Name:Diet	42	0.341	0.008	2.68	0	0.062	13.04
Residuals	291	0.883	0.003	NA	NA	0.062	4.87

```
ee.var.data <- bind_rows(lean.adj.ee.lean %>% mutate(Diet="NCD"),hfd.adj.ee.adj %>% mutate(Diet="HFD"))

ggplot(ee.var.data %>% filter(term %in% c('Name','Name:Diet','Residuals')),
       aes(x=reorder(Diet,-Pct.Var),
           y=Pct.Var,
           fill=term)) +
  geom_bar(position="stack",stat='identity') +
  scale_fill_manual(labels = c("Strain", "Strain x Diet", "Other"), values = c("red", "pink","blue"),
                    name="Factor") +
  labs(y="Percent of Variance",
       x="") +
  theme_classic() +
  theme(legend.position="top")+
  theme(text=element_text(size=18))
```
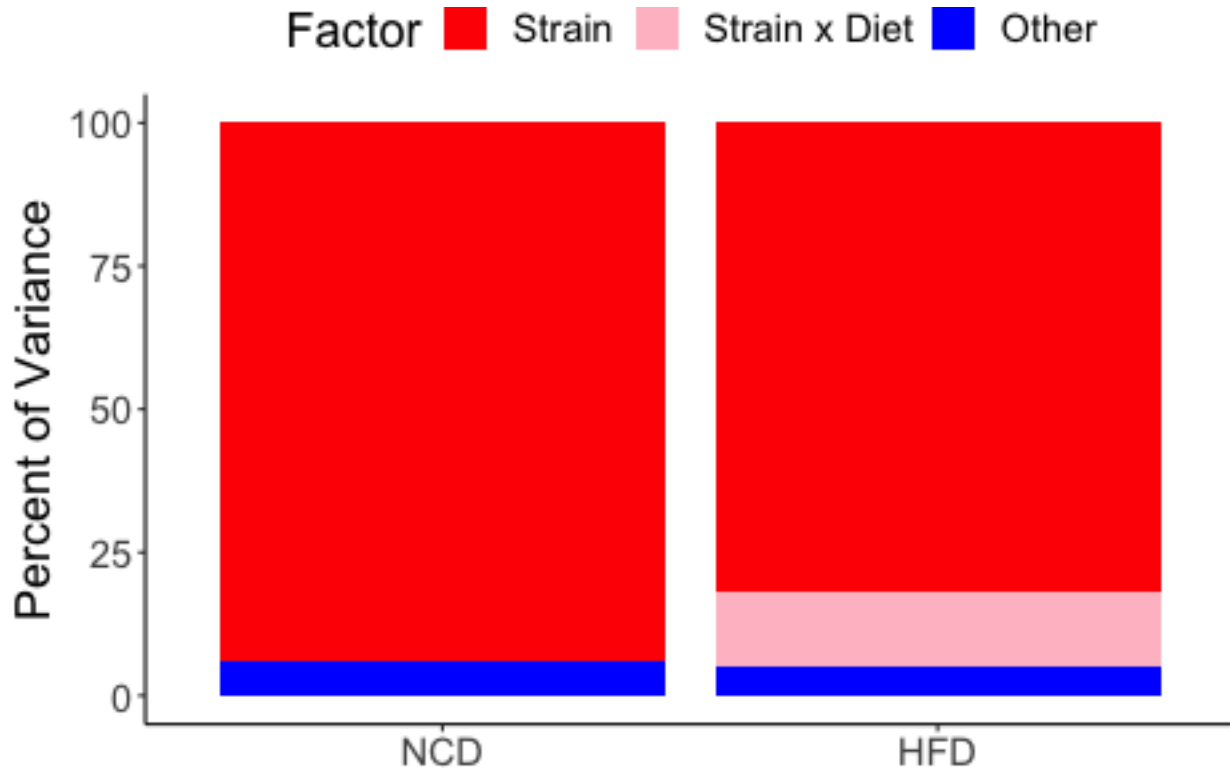


```
ee.var.all.data <- bind_rows(lean.adj.ee.all %>% mutate(Diet="NCD",Corr="None"),
                             lean.adj.ee.lean %>% mutate(Diet="NCD", Corr="Lean Mass"),
                             hfd.adj.ee.all %>% mutate(Diet="HFD", Corr="Lean Mass"),
                             hfd.adj.ee.adj %>% mutate(Diet="HFD", Corr="Lean Mass,Diet")) %>%
```

```r
  mutate(Group=paste(Diet,Corr,sep="-"))

library(RColorBrewer)
ee.var.all.data %>%
        filter(term %in% c('Lean','Name','Name:Diet','Diet','Residuals')) %>%
        filter(!(term=='Lean'&Corr=='Lean Mass')) %>%
        filter(!(term=='Lean'&Corr=="Lean Mass,Diet")) %>%
        filter(!(term=='Diet'&Corr=="Lean Mass,Diet")) %>%
ggplot(aes(x=ordered(Group, levels=c("NCD-None","NCD-Lean Mass", "HFD-Lean Mass", "HFD-Lean Mass,Diet"))
           y=Pct.Var,
           fill=term)) +
  geom_bar(position="stack",stat='identity') +
  scale_fill_manual(name="Factor",
                    labels = c("Diet", "Lean Mass", "Strain", "Strain x Diet", "Residuals"),
                    values=brewer.pal(5, "Set1")) +
  labs(y="Percent of Variance",
       x="Group - Adjustments") +
  theme_classic() +
  theme(text=element_text(size=12))
```
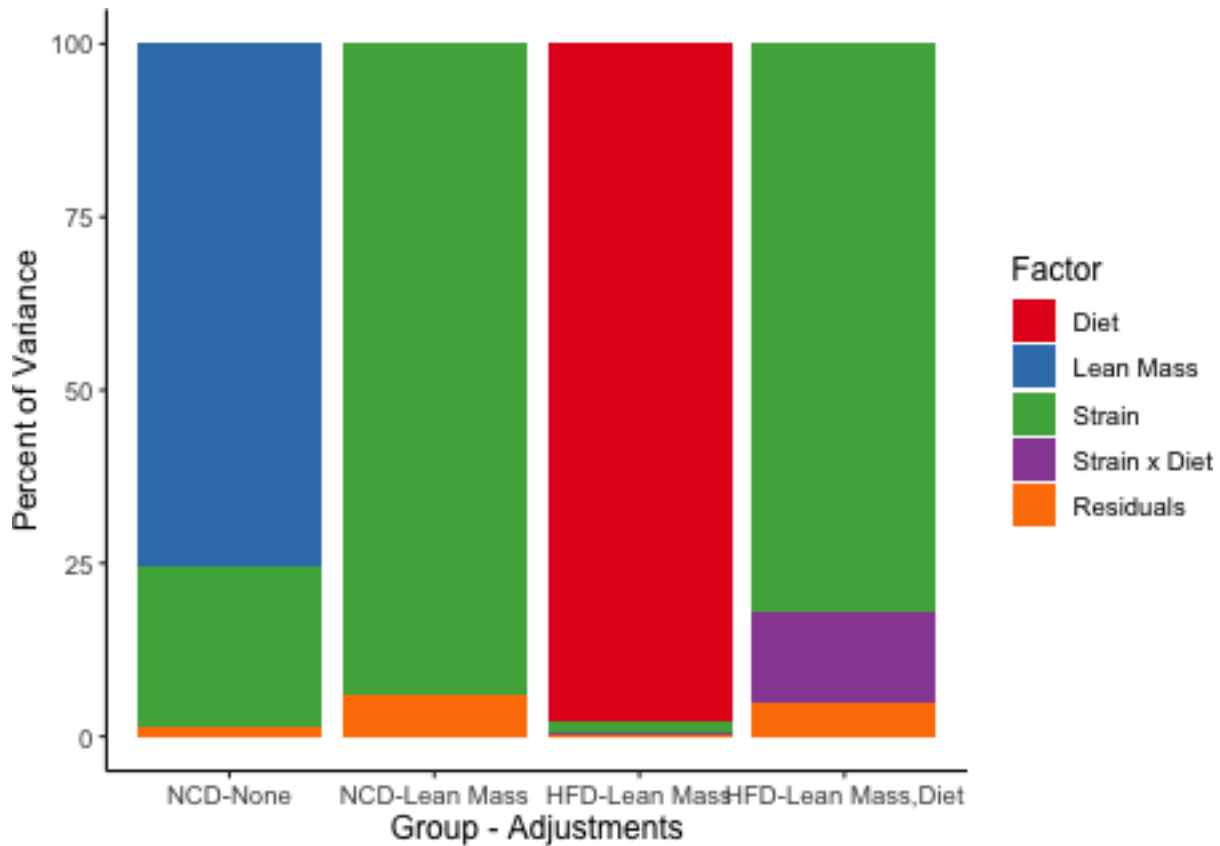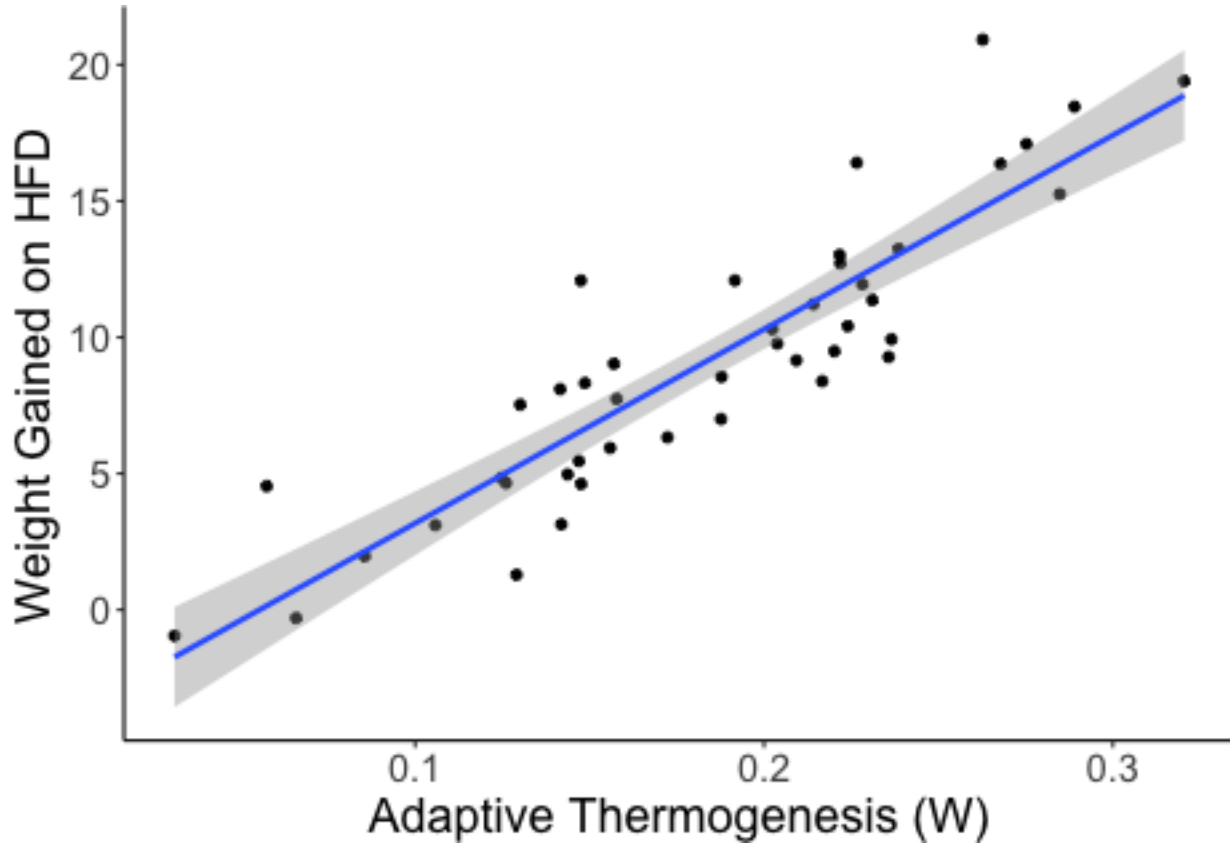

Adaptive Thermogenesis vs Weight Gain

```r
data.wide %>%
  ggplot(aes(y=Wt.Gain,
             x=AT)) +
```

```r
labs(y="Weight Gained on HFD",
     x="Adaptive Thermogenesis (W)") +
geom_point() +
geom_smooth(method="lm") +  theme_classic() +
theme(text=element_text(size=18))
```



```r
lm(Wt.Gain~AT, data=data.wide) %>% glance %>% kable(caption="Summary of relationship between energy expe
```

Table 12: Summary of relationship between energy expenditure and diet-induced weight gain

r.squared	adj.r.squared	sigma	statistic	p.value	df	logLik	AIC	BIC	deviance	df.residual	nobs
0.801	0.796	2.32	169	0	1	-98.5	203	208	226	42	44

Session Information

```r
sessionInfo()
```

```
## R version 4.0.2 (2020-06-22)
## Platform: x86_64-apple-darwin17.0 (64-bit)
## Running under: macOS  10.16
##
## Matrix products: default
## BLAS:   /Library/Frameworks/R.framework/Versions/4.0/Resources/lib/libRblas.dylib
```

```
## LAPACK: /Library/Frameworks/R.framework/Versions/4.0/Resources/lib/libRlapack.dylib
##
## locale:
## [1] en_US.UTF-8/en_US.UTF-8/en_US.UTF-8/C/en_US.UTF-8/en_US.UTF-8
##
## attached base packages:
## [1] stats      graphics  grDevices utils     datasets  methods   base
##
## other attached packages:
## [1] RColorBrewer_1.1-2 broom_0.7.11      ggrepel_0.9.1      ggplot2_3.3.5
## [5] readr_2.1.1        dplyr_1.0.7       tidyr_1.1.4        knitr_1.37
##
## loaded via a namespace (and not attached):
##  [1] tidyselect_1.1.1 xfun_0.29        purrr_0.3.4      splines_4.0.2
##  [5] lattice_0.20-45  colorspace_2.0-2 vctrs_0.3.8      generics_0.1.1
##  [9] htmltools_0.5.2  yaml_2.2.1       mgcv_1.8-38      utf8_1.2.2
## [13] rlang_0.4.12     pillar_1.6.4     glue_1.6.0       withr_2.4.3
## [17] DBI_1.1.2        bit64_4.0.5      lifecycle_1.0.1  stringr_1.4.0
## [21] munsell_0.5.0    gtable_0.3.0     evaluate_0.14    labeling_0.4.2
## [25] tzdb_0.2.0       fastmap_1.1.0    parallel_4.0.2   fansi_1.0.0
## [29] highr_0.9        Rcpp_1.0.7       backports_1.4.1  scales_1.1.1
## [33] vroom_1.5.7      magick_2.7.3     farver_2.1.0     bit_4.0.4
## [37] hms_1.1.1        digest_0.6.29    stringi_1.7.6    grid_4.0.2
## [41] cli_3.1.0        tools_4.0.2      magrittr_2.0.1   tibble_3.1.6
## [45] crayon_1.4.2     pkgconfig_2.0.3  ellipsis_0.3.2   Matrix_1.4-0
## [49] assertthat_0.2.1 rmarkdown_2.11   rstudioapi_0.13  R6_2.5.1
## [53] nlme_3.1-153     compiler_4.0.2
```