

Preliminary Investigation

Ethan Allavarpu (UID: 405287603)

10/27/2020

```
# sample <- read.csv("sample.csv", stringsAsFactors = TRUE)
# sample
```

```
training <- read.csv("training.csv", stringsAsFactors = TRUE)
training$class <- factor(training$class)
levels(training$class) <- c("NG", "OG", "TSG")
dim(training)
```

```
[1] 3177 99
```

```
names(training)[c(1, 99)]
```

```
[1] "id"      "class"
```

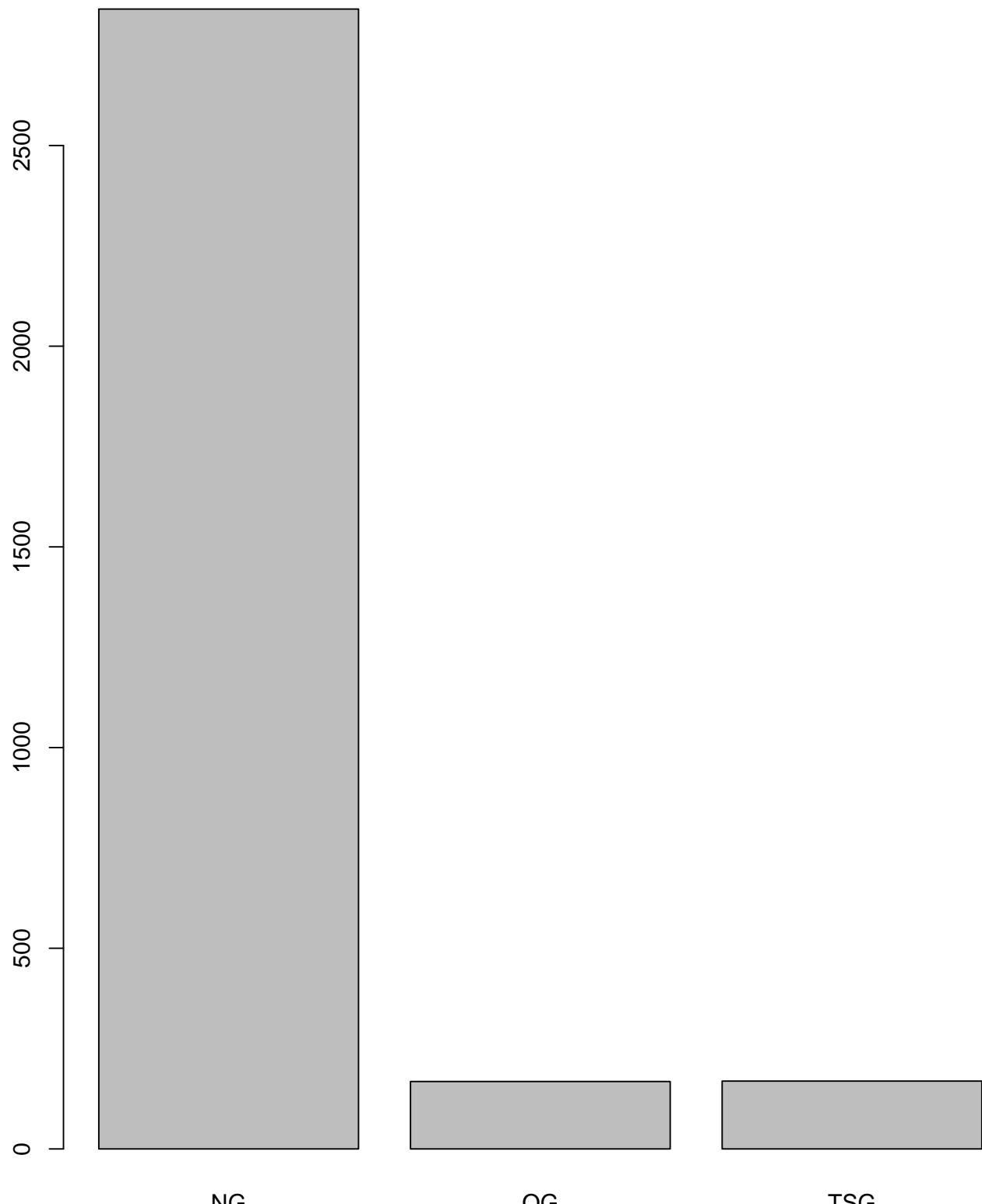
```
barplot(table(training$class))
table(training$class) / nrow(training)
```

```
NG          OG          TSG
0.89392509 0.05288008 0.05319484
```

```
any(is.na(training))
```

```
[1] FALSE
```

```
library(ggplot2)
```

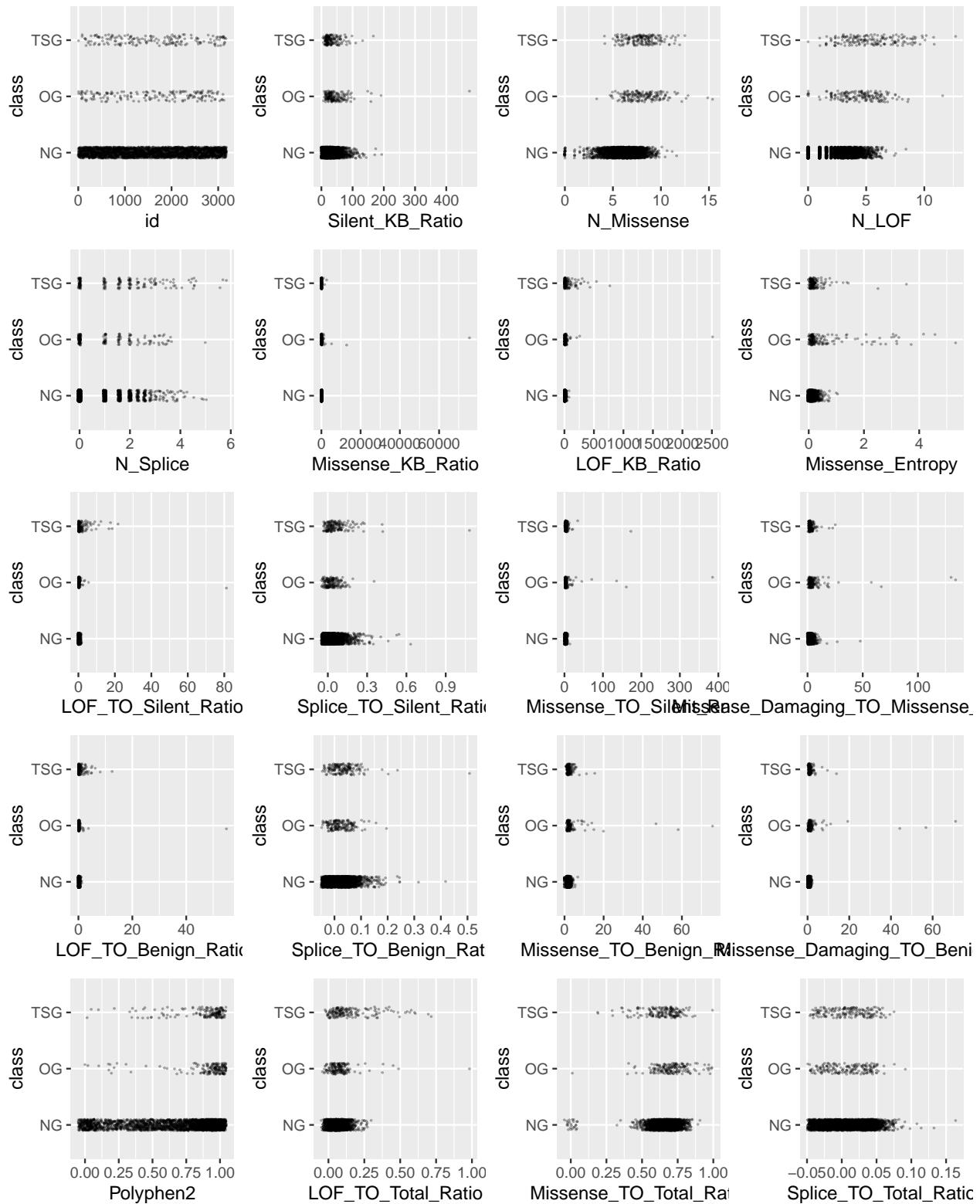


```
scatter <- function(var) {  
  ggplot(training, aes_string(var, "class")) +  
    geom_jitter(width = 0.05, height = 0.1, size = 0.1,  
               colour = rgb(0, 0, 0, alpha = 1 / 3))  
}
```

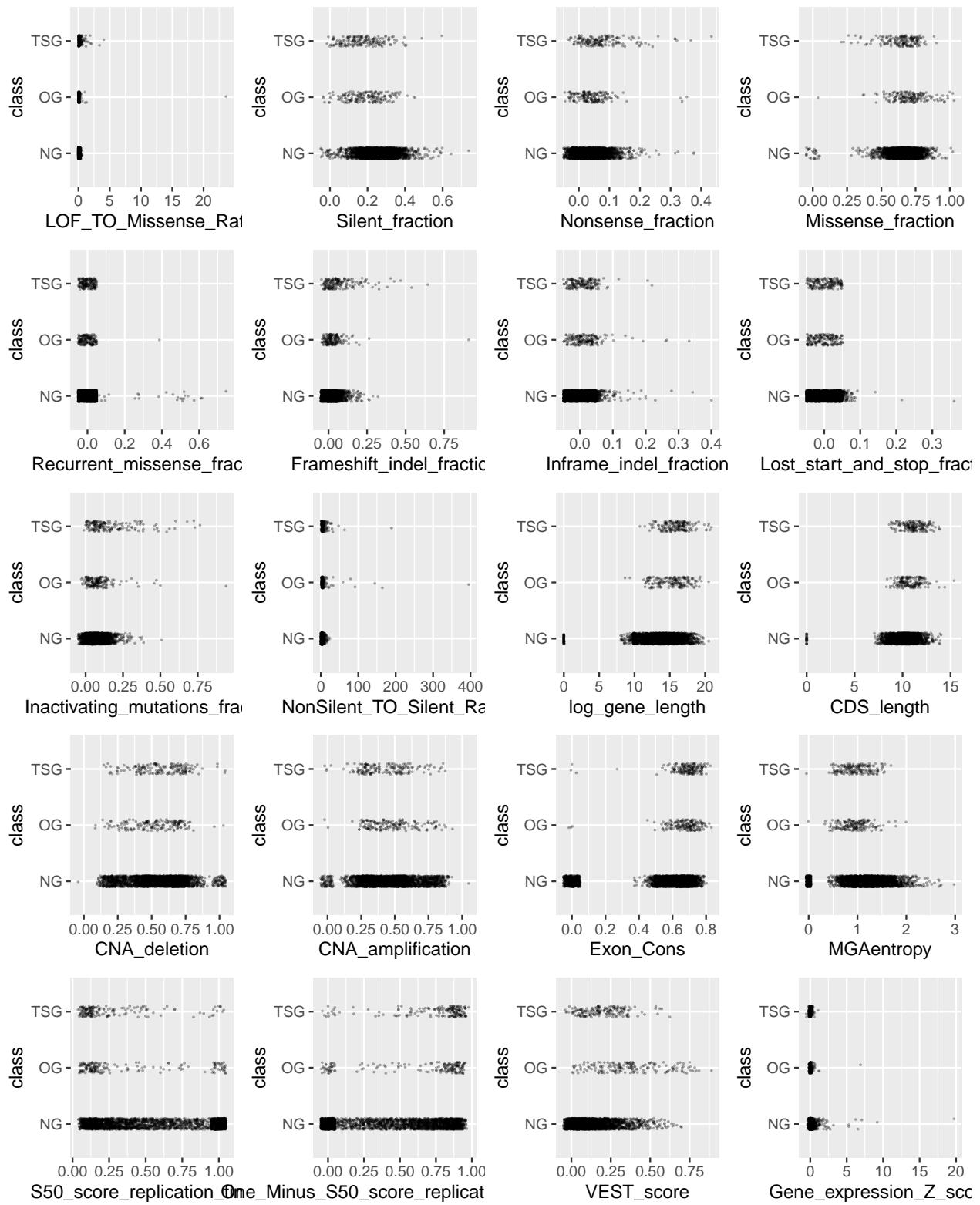
```

scat_plot <- lapply(names(training)[-99], scatter)
library(gridExtra)
grid.arrange(grobs = scat_plot[1:20], ncol = 4)

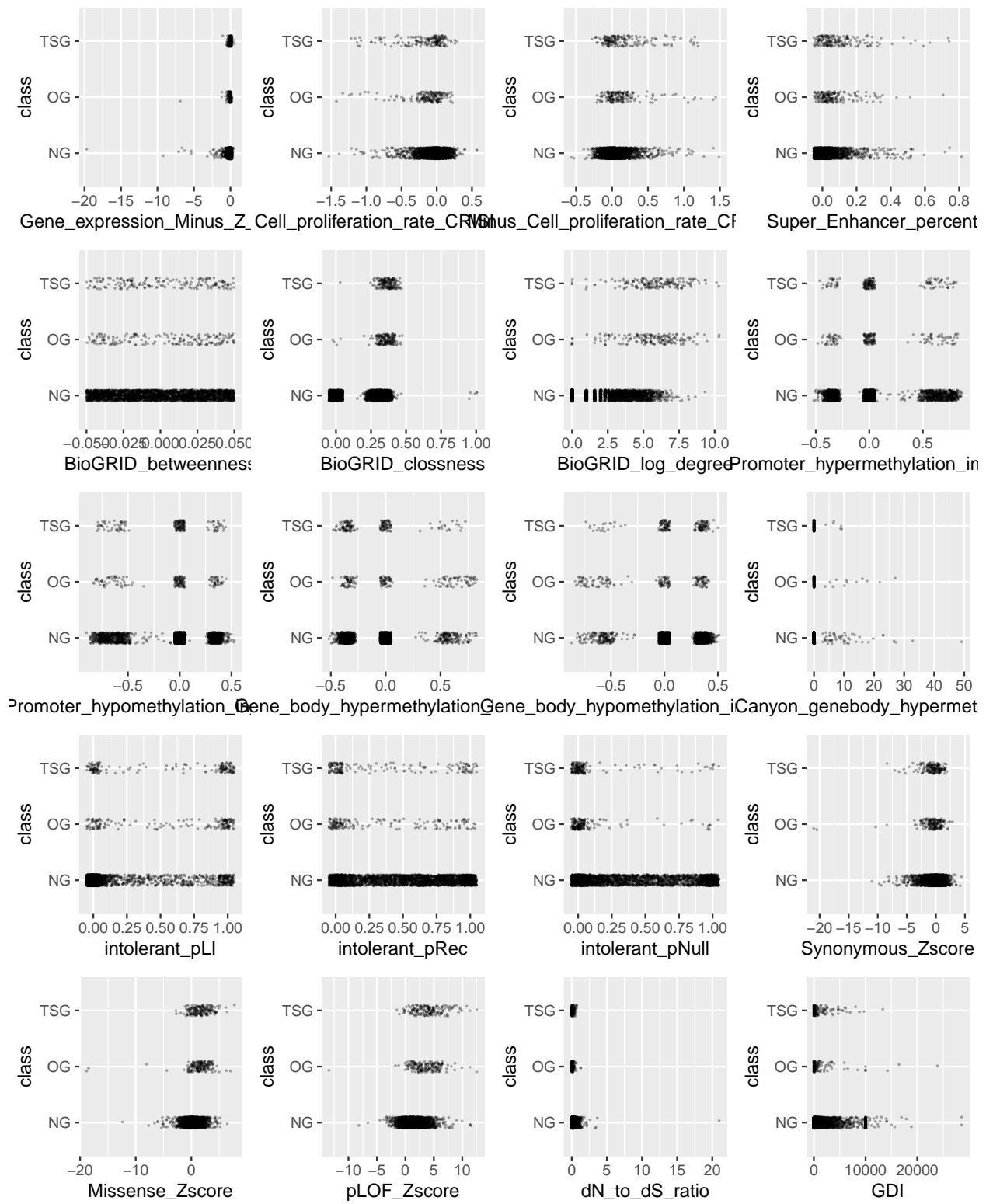
```



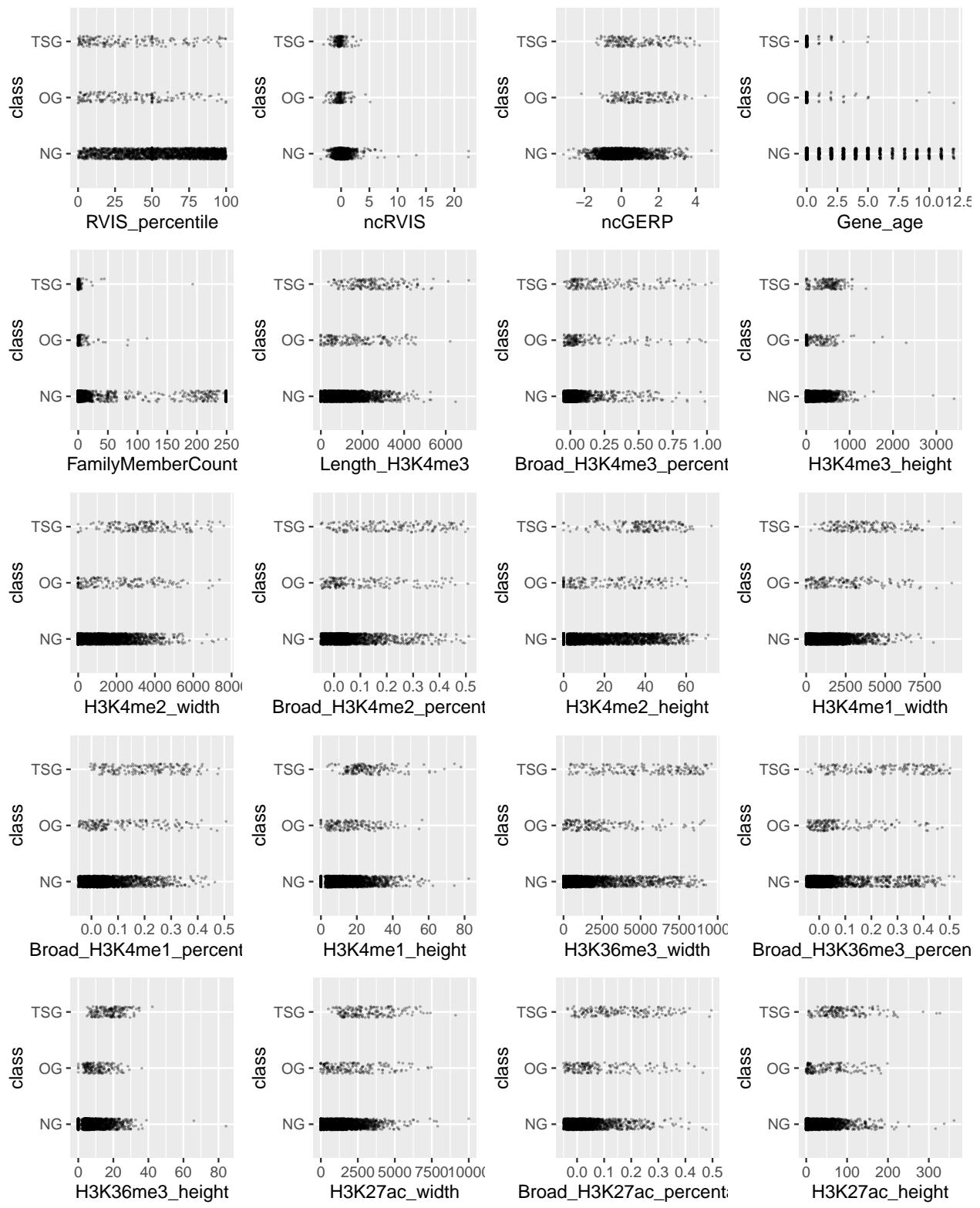
```
grid.arrange(grobs = scat_plot[21:40], ncol = 4)
```



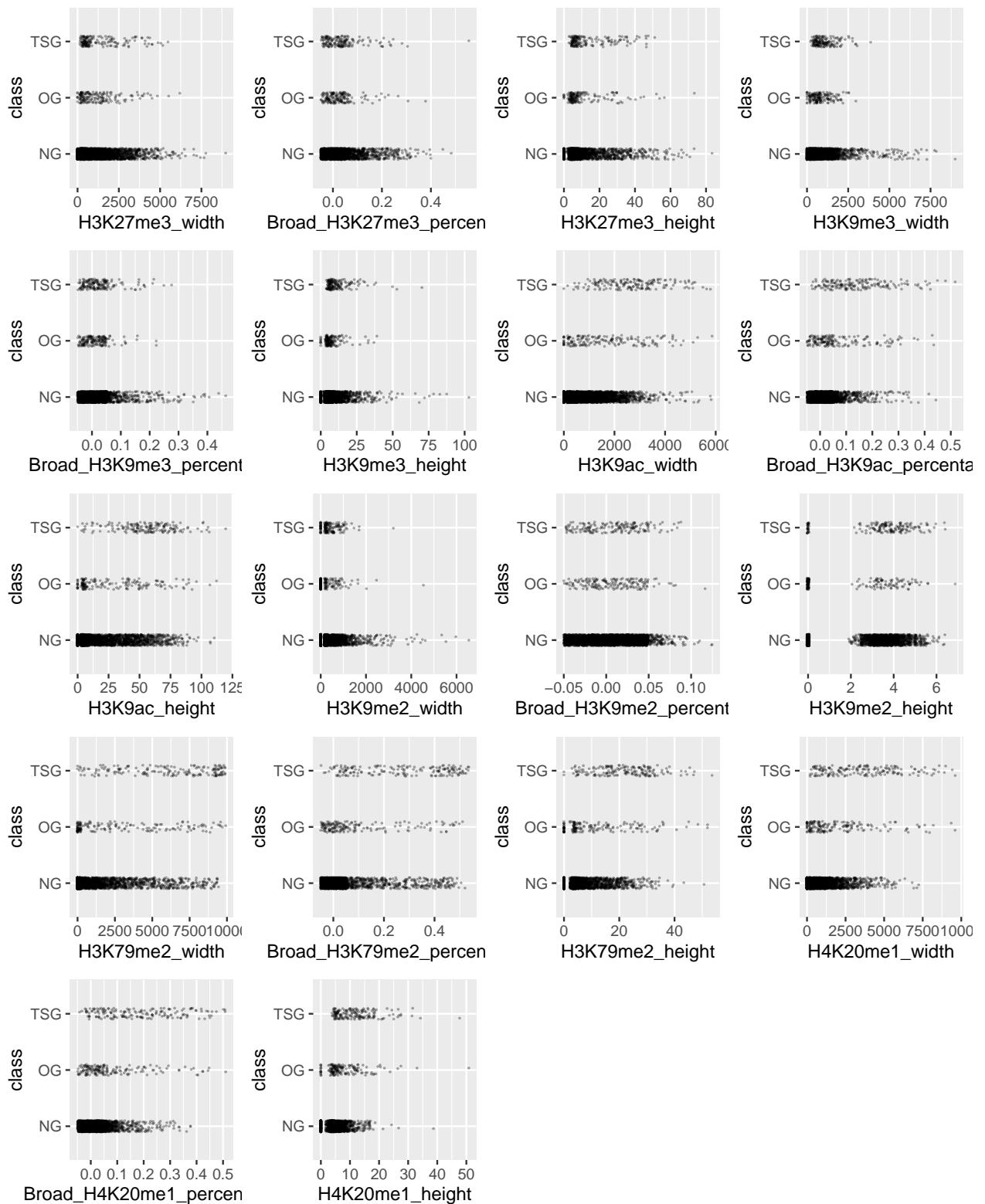
```
grid.arrange(grobs = scat_plot[41:60], ncol = 4)
```



```
grid.arrange(grobs = scat_plot[61:80], ncol = 4)
```



```
grid.arrange(grobs = scat_plot[81:98], ncol = 4)
```



```
sig <- logical(98)
names(sig) <- names(training)[-99]
k <- 1
diffs <- logical(98)
```

```

for (var in names(training)[-99]) {
  model <- aov(training[[var]] ~ factor(training$class))
  sig[k] <- summary(model)[[1]][1, 5]
  diffs[k] <- all(TukeyHSD(model)$`factor(training$class)`[, 4] < 0.05)
  k <- k + 1
}
score <- function (conf_mat) {
  print(sum(diag(conf_mat) * c(1, 20, 20)))
  print(sum(diag(conf_mat) * c(1, 20, 20)) / sum(apply(conf_mat, 2, sum) * c(1, 20, 20)))
}

library(dplyr)

```

Attaching package: 'dplyr'

The following object is masked from 'package:gridExtra':

combine

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

filter, lag

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

intersect, setdiff, setequal, union

```

vars <- training %>% select(Broad_H3K9ac_percentage, N_LOF,
                               N_Splice, LOF_TO_Total_Ratio, VEST_score, Missense_Entropy, BioGRID_clossness)
vars$class <- factor(vars$class)
levels(vars$class) <- c("NG", "OG", "TSG")
library(caret)

```

Loading required package: lattice

```

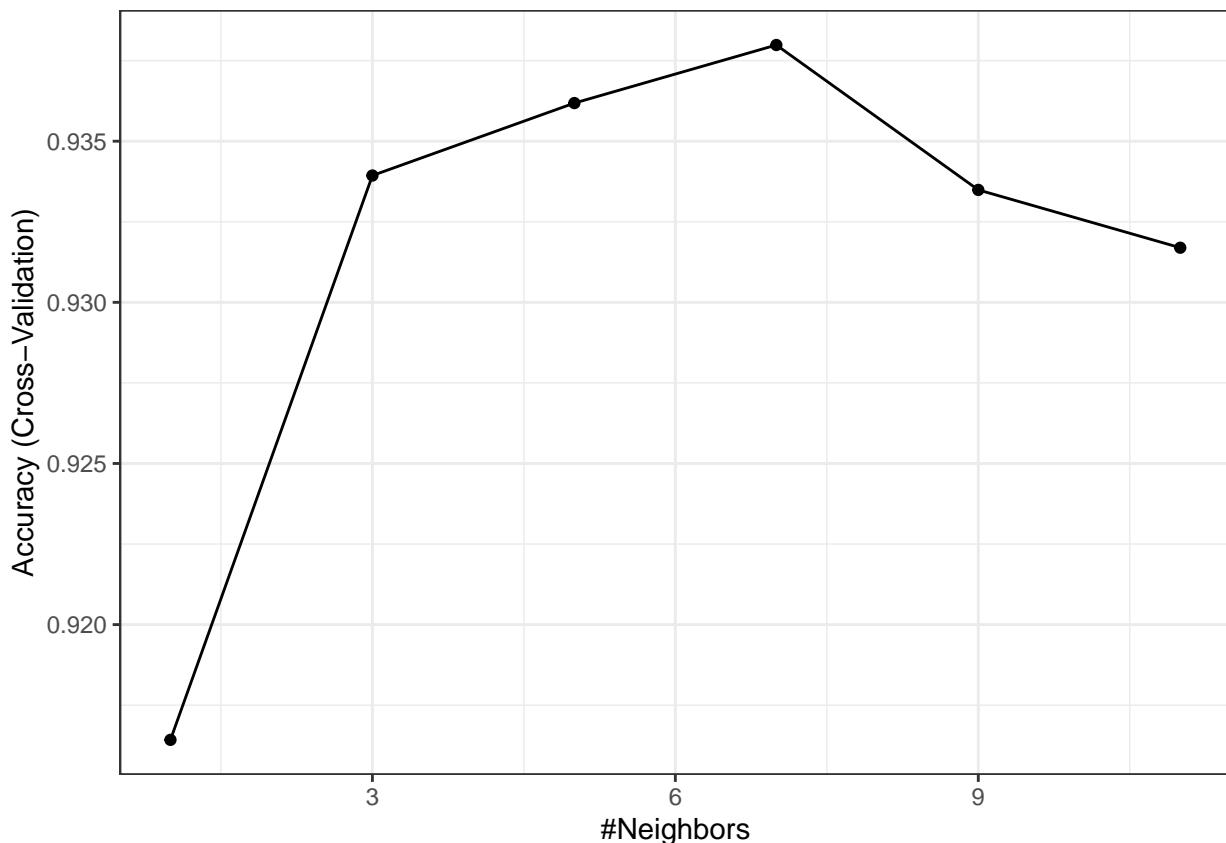
vars_test <- createDataPartition(vars$class, p = 0.7,
                                 list = FALSE)
vars_train <- vars[vars_test, ]
vars_test <- vars[vars_test, ]

```

```

train_cont <- trainControl(method = "cv", number = 10, classProbs = TRUE, savePredictions = TRUE)
knn_ft <- train(class ~ ., data = vars_train, method = "knn", preProc = c("center", "scale"),
                 trControl = train_cont, tuneGrid = expand.grid(k = seq(from = 1, to = 11, by = 2)))
ggplot(knn_ft) + theme_bw()

```



```

for (k in seq(from = 1, to = 11, by = 2)) {
  knn_mod <- table("pred"=predict(knn_ft, newdata = vars_test), "obs" = vars_test$class)
  knn_mod
  score(knn_mod)
}

[1] 4399
[1] 0.6538347
[1] 4402
[1] 0.6542806
[1] 4379
[1] 0.6508621
[1] 4399
[1] 0.6538347
[1] 4439
[1] 0.65978
[1] 4400
[1] 0.6539834

train_cont <- trainControl(method = "cv", number = 10, classProbs = TRUE, savePredictions = TRUE)
knn_ft <- train(class ~ ., data = vars_train, method = "qda", preProc = c("center", "scale"),
                 trControl = train_cont)
qda_mod <- table("pred"=predict(knn_ft, newdata = vars_test), "obs" = vars_test$class)
qda_mod

```

		obs			
		pred	NG	OG	TSG
pred	obs	NG	1901	27	25

```

OG      19    63     7
TSG     68    28    87

score(qda_mod)

[1] 4901
[1] 0.7284483

train_cont <- trainControl(method = "cv", number = 10, classProbs = TRUE, savePredictions = TRUE)
knn_ft <- train(class ~ ., data = vars_train, method = "lda", preProc = c("center", "scale"),
                 trControl = train_cont)
lda_mod <- table("pred"=predict(knn_ft, newdata = vars_test), "obs" = vars_test$class)
lda_mod

  obs
pred   NG   OG  TSG
  NG 1966   36   38
  OG     7   67     9
  TSG    15   15   72

score(lda_mod)

[1] 4746
[1] 0.7054102

library(leaps)
best_subset <- regsubsets(class~., data = training, nbest = 1, nvmax = 10,
                           intercept = TRUE, method = "forward",
                           really.big = TRUE)

Warning in leaps.setup(x, y, wt = wt, nbest = nbest, nvmax = nvmax, force.in =
force.in, : 5 linear dependencies found

Reordering variables and trying again:
sumBS <- summary(best_subset)
sumBS

Subset selection object
Call: regsubsets.formula(class ~ ., data = training, nbest = 1, nvmax = 10,
                         intercept = TRUE, method = "forward", really.big = TRUE)
98 Variables (and intercept)

          Forced in  Forced out
id           FALSE    FALSE
Silent_KB_Ratio  FALSE    FALSE
N_Missense      FALSE    FALSE
N_LOF            FALSE    FALSE
N_Splice          FALSE    FALSE
Missense_KB_Ratio FALSE    FALSE
LOF_KB_Ratio      FALSE    FALSE
Missense_Entropy FALSE    FALSE
LOF_TO_Silent_Ratio FALSE    FALSE
Splice_TO_Silent_Ratio FALSE    FALSE
Missense_TO_Silent_Ratio FALSE    FALSE
Missense_Damaging_TO_Missense_Benign_Ratio FALSE    FALSE
LOF_TO_Benign_Ratio FALSE    FALSE
Splice_TO_Benign_Ratio FALSE    FALSE
Missense_TO_Benign_Ratio FALSE    FALSE

```

Missense_Damaging_T0_Benign_Ratio	FALSE	FALSE
Polyphen2	FALSE	FALSE
LOF_T0_Total_Ratio	FALSE	FALSE
Missense_T0_Total_Ratio	FALSE	FALSE
Splice_T0_Total_Ratio	FALSE	FALSE
LOF_T0_Missense_Ratio	FALSE	FALSE
Silent_fraction	FALSE	FALSE
Nonsense_fraction	FALSE	FALSE
Missense_fraction	FALSE	FALSE
Recurrent_missense_fraction	FALSE	FALSE
Frameshift_indel_fraction	FALSE	FALSE
Inframe_indel_fraction	FALSE	FALSE
Lost_start_and_stop_fraction	FALSE	FALSE
Inactivating_mutations_fraction	FALSE	FALSE
NonSilent_T0_Silent_Ratio	FALSE	FALSE
log_gene_length	FALSE	FALSE
CDS_length	FALSE	FALSE
CNA_deletion	FALSE	FALSE
CNA_amplification	FALSE	FALSE
Exon_Cons	FALSE	FALSE
MGAentropy	FALSE	FALSE
S50_score_replication_timing	FALSE	FALSE
VEST_score	FALSE	FALSE
Gene_expression_Z_score	FALSE	FALSE
Cell_proliferation_rate_CRISPR_KD	FALSE	FALSE
Super_Enhancer_percentage	FALSE	FALSE
BioGRID_betweenness	FALSE	FALSE
BioGRID_closeness	FALSE	FALSE
BioGRID_log_degree	FALSE	FALSE
Promoter_hypermethylation_in_cancer	FALSE	FALSE
Gene_body_hypermethylation_in_cancer	FALSE	FALSE
Canyon_genebody_hypermethylation	FALSE	FALSE
intolerant_pLI	FALSE	FALSE
intolerant_pRec	FALSE	FALSE
intolerant_pNull	FALSE	FALSE
Synonymous_Zscore	FALSE	FALSE
Missense_Zscore	FALSE	FALSE
pLOF_Zscore	FALSE	FALSE
dN_to_dS_ratio	FALSE	FALSE
GDI	FALSE	FALSE
RVIS_percentile	FALSE	FALSE
ncRVIS	FALSE	FALSE
ncGERP	FALSE	FALSE
Gene_age	FALSE	FALSE
FamilyMemberCount	FALSE	FALSE
Length_H3K4me3	FALSE	FALSE
Broad_H3K4me3_percentage	FALSE	FALSE
H3K4me3_height	FALSE	FALSE
H3K4me2_width	FALSE	FALSE
Broad_H3K4me2_percentage	FALSE	FALSE
H3K4me2_height	FALSE	FALSE
H3K4me1_width	FALSE	FALSE
Broad_H3K4me1_percentage	FALSE	FALSE
H3K4me1_height	FALSE	FALSE

H3K36me3_width		FALSE	FALSE			
Broad_H3K36me3_percentage		FALSE	FALSE			
H3K36me3_height		FALSE	FALSE			
H3K27ac_width		FALSE	FALSE			
Broad_H3K27ac_percentage		FALSE	FALSE			
H3K27ac_height		FALSE	FALSE			
H3K27me3_width		FALSE	FALSE			
Broad_H3K27me3_percentage		FALSE	FALSE			
H3K27me3_height		FALSE	FALSE			
H3K9me3_width		FALSE	FALSE			
Broad_H3K9me3_percentage		FALSE	FALSE			
H3K9me3_height		FALSE	FALSE			
H3K9ac_width		FALSE	FALSE			
Broad_H3K9ac_percentage		FALSE	FALSE			
H3K9ac_height		FALSE	FALSE			
H3K9me2_width		FALSE	FALSE			
Broad_H3K9me2_percentage		FALSE	FALSE			
H3K9me2_height		FALSE	FALSE			
H3K79me2_width		FALSE	FALSE			
Broad_H3K79me2_percentage		FALSE	FALSE			
H3K79me2_height		FALSE	FALSE			
H4K20me1_width		FALSE	FALSE			
Broad_H4K20me1_percentage		FALSE	FALSE			
H4K20me1_height		FALSE	FALSE			
One_Minus_S50_score_replication_timing		FALSE	FALSE			
Gene_expression_Minus_Z_score		FALSE	FALSE			
Minus_Cell_proliferation_rate_CRISPR_KD		FALSE	FALSE			
Promoter_hypomethylation_in_cancer		FALSE	FALSE			
Gene_body_hypomethylation_in_cancer		FALSE	FALSE			
1 subsets of each size up to 11						
Selection Algorithm: forward						
	id	Silent_KB_Ratio	N_Missense	N_LOF	N_Splice	Missense_KB_Ratio
1	(1)	" " "	" "	" "	" "	" "
2	(1)	" " "	" "	"*"	" "	" "
3	(1)	" " "	" "	"*"	" "	" "
4	(1)	" " "	" "	"*"	" "	" "
5	(1)	" " "	" "	"*"	" "	" "
6	(1)	" " "	" "	"*"	" "	" "
7	(1)	" " "	" "	"*"	" "	" "
8	(1)	" " "	" "	"*"	" "	" "
9	(1)	" " "	" "	"*"	" "	" "
10	(1)	" " "	" "	"*"	"*"	" "
11	(1)	" " "	" "	"*"	"*"	" "
		LOF_KB_Ratio	Missense_Entropy	LOF_TO_Silent_Ratio		
1	(1)	" "	" "	" "		
2	(1)	" "	" "	" "		
3	(1)	" "	" "	" "		
4	(1)	" "	" "	" "		
5	(1)	" "	" "	" "		
6	(1)	" "	" "	" "		
7	(1)	" "	" "	" "		
8	(1)	" "	"*"	" "		
9	(1)	" "	"*"	" "		
10	(1)	" "	"*"	" "		

```

11  ( 1 ) " "          "*"          " "
      Splice_TO_Silent_Ratio Missense_TO_Silent_Ratio
1  ( 1 ) " "          " "
2  ( 1 ) " "          " "
3  ( 1 ) " "          " "
4  ( 1 ) " "          " "
5  ( 1 ) " "          " "
6  ( 1 ) " "          " "
7  ( 1 ) " "          " "
8  ( 1 ) " "          " "
9  ( 1 ) " "          " "
10 ( 1 ) " "          " "
11 ( 1 ) " "          " "
                  Missense_Damaging_TO_Missense_Benign_Ratio LOF_TO_Benign_Ratio
1  ( 1 ) " "          " "
2  ( 1 ) " "          " "
3  ( 1 ) " "          " "
4  ( 1 ) " "          " "
5  ( 1 ) " "          " "
6  ( 1 ) " "          " "
7  ( 1 ) " "          " "
8  ( 1 ) " "          " "
9  ( 1 ) " "          " "
10 ( 1 ) " "          " "
11 ( 1 ) " "          " "
                  Splice_TO_Benign_Ratio Missense_TO_Benign_Ratio
1  ( 1 ) " "          " "
2  ( 1 ) " "          " "
3  ( 1 ) " "          " "
4  ( 1 ) " "          " "
5  ( 1 ) " "          " "
6  ( 1 ) " "          " "
7  ( 1 ) " "          " "
8  ( 1 ) " "          " "
9  ( 1 ) " "          " "
10 ( 1 ) " "          " "
11 ( 1 ) " "          " "
                  Missense_Damaging_TO_Benign_Ratio Polyphen2 LOF_TO_Total_Ratio
1  ( 1 ) " "          " "          " "
2  ( 1 ) " "          " "          " "
3  ( 1 ) " "          " "          " "
4  ( 1 ) " "          " "          " "
5  ( 1 ) " "          " "          " "
6  ( 1 ) " "          " "          " "
7  ( 1 ) " "          " "          "*"
8  ( 1 ) " "          " "          "*"
9  ( 1 ) " "          " "          "*"
10 ( 1 ) " "          " "          "*"
11 ( 1 ) " "          " "          "*"
                  Missense_TO_Total_Ratio Splice_TO_Total_Ratio LOF_TO_Missense_Ratio
1  ( 1 ) " "          " "          " "
2  ( 1 ) " "          " "          " "
3  ( 1 ) " "          " "          " "
4  ( 1 ) " "          " "          " "

```

```

5  ( 1 ) " "          " "          " "
6  ( 1 ) " "          " "          " "
7  ( 1 ) " "          " "          " "
8  ( 1 ) " "          " "          " "
9  ( 1 ) " "          " "          " "
10 ( 1 ) " "          " "          " "
11 ( 1 ) " "          "*"         " "
                                Silent_fraction Nonsense_fraction Missense_fraction
1  ( 1 ) " "          " "          " "
2  ( 1 ) " "          " "          " "
3  ( 1 ) " "          " "          " "
4  ( 1 ) " "          " "          " "
5  ( 1 ) " "          " "          " "
6  ( 1 ) " "          " "          " "
7  ( 1 ) " "          " "          " "
8  ( 1 ) " "          " "          " "
9  ( 1 ) " "          " "          " "
10 ( 1 ) " "          " "          " "
11 ( 1 ) " "          " "          " "
                                Recurrent_missense_fraction Frameshift_indel_fraction
1  ( 1 ) " "          " "          " "
2  ( 1 ) " "          " "          " "
3  ( 1 ) " "          " "          " "
4  ( 1 ) " "          " "          " "
5  ( 1 ) " "          " "          " "
6  ( 1 ) " "          " "          " "
7  ( 1 ) " "          " "          " "
8  ( 1 ) " "          " "          " "
9  ( 1 ) " "          " "          " "
10 ( 1 ) " "          " "          " "
11 ( 1 ) " "          " "          " "
                                Inframe_indel_fraction Lost_start_and_stop_fraction
1  ( 1 ) " "          " "          " "
2  ( 1 ) " "          " "          " "
3  ( 1 ) " "          " "          " "
4  ( 1 ) " "          " "          " "
5  ( 1 ) " "          " "          " "
6  ( 1 ) " "          " "          " "
7  ( 1 ) " "          " "          " "
8  ( 1 ) " "          " "          " "
9  ( 1 ) " "          " "          " "
10 ( 1 ) " "          " "          " "
11 ( 1 ) " "          " "          " "
                                Inactivating_mutations_fraction NonSilent_TO_Silent_Ratio
1  ( 1 ) " "          " "          " "
2  ( 1 ) " "          " "          " "
3  ( 1 ) " "          " "          " "
4  ( 1 ) " "          " "          " "
5  ( 1 ) " "          " "          " "
6  ( 1 ) " "          " "          " "
7  ( 1 ) " "          " "          " "
8  ( 1 ) " "          " "          " "
9  ( 1 ) " "          " "          " "
10 ( 1 ) " "          " "          " "

```

```

11  ( 1 ) " "
          " "
          log_gene_length CDS_length CNA_deletion CNA_amplification Exon_Cons
1  ( 1 ) " "      " "      " "      " "      " "
2  ( 1 ) " "      " "      " "      " "      " "
3  ( 1 ) " "      " "      " "      " "      " "
4  ( 1 ) " "      " "      " "      " "      " "
5  ( 1 ) " "      " "      " "      " "      " "
6  ( 1 ) " "      " "      " "      " "      " "
7  ( 1 ) " "      " "      " "      " "      " "
8  ( 1 ) " "      " "      " "      " "      " "
9  ( 1 ) " "      " "      " "      " "      " "
10 ( 1 ) " "      " "      " "      " "      " "
11 ( 1 ) " "      " "      " "      " "      " "
          MGAentropy S50_score_replication_timing
1  ( 1 ) " "      " "
2  ( 1 ) " "      " "
3  ( 1 ) " "      " "
4  ( 1 ) " "      " "
5  ( 1 ) " "      " "
6  ( 1 ) " "      " "
7  ( 1 ) " "      " "
8  ( 1 ) " "      " "
9  ( 1 ) " "      " "
10 ( 1 ) " "      " "
11 ( 1 ) " "      " "
          One_Minus_S50_score_replication_timing VEST_score
1  ( 1 ) " "      " "
2  ( 1 ) " "      " "
3  ( 1 ) " "      " "
4  ( 1 ) " "      "*"
5  ( 1 ) " "      "*"
6  ( 1 ) " "      "*"
7  ( 1 ) " "      "*"
8  ( 1 ) " "      "*"
9  ( 1 ) " "      "*"
10 ( 1 ) " "      "*"
11 ( 1 ) " "      "*"
          Gene_expression_Z_score Gene_expression_Minus_Z_score
1  ( 1 ) " "      " "
2  ( 1 ) " "      " "
3  ( 1 ) " "      " "
4  ( 1 ) " "      " "
5  ( 1 ) " "      " "
6  ( 1 ) " "      " "
7  ( 1 ) " "      " "
8  ( 1 ) " "      " "
9  ( 1 ) " "      " "
10 ( 1 ) " "      " "
11 ( 1 ) " "      " "
          Cell_proliferation_rate_CRISPR_KD
1  ( 1 ) " "
2  ( 1 ) " "
3  ( 1 ) " "
4  ( 1 ) " "

```

```

5  ( 1 ) " "
6  ( 1 ) " "
7  ( 1 ) " "
8  ( 1 ) " "
9  ( 1 ) " "
10 ( 1 ) " "
11 ( 1 ) " "
      Minus_Cell_proliferation_rate_CRISPR_KD Super_Enhancer_percentage
1  ( 1 ) " " " "
2  ( 1 ) " " " "
3  ( 1 ) " " " "
4  ( 1 ) " " " "
5  ( 1 ) " " " "
6  ( 1 ) " " " "
7  ( 1 ) " " " "
8  ( 1 ) " " " "
9  ( 1 ) " " " "
10 ( 1 ) " " " "
11 ( 1 ) " " " "
      BioGRID_betweenness BioGRID_closeness BioGRID_log_degree
1  ( 1 ) " " " "
2  ( 1 ) " " " "
3  ( 1 ) " " " *"
4  ( 1 ) " " " *"
5  ( 1 ) " " " *"
6  ( 1 ) " " " *"
7  ( 1 ) " " " *"
8  ( 1 ) " " " *"
9  ( 1 ) " " " *"
10 ( 1 ) " " " *"
11 ( 1 ) " " " *"
      Promoter_hypermethylation_in_cancer
1  ( 1 ) " "
2  ( 1 ) " "
3  ( 1 ) " "
4  ( 1 ) " "
5  ( 1 ) " "
6  ( 1 ) " "
7  ( 1 ) " "
8  ( 1 ) " "
9  ( 1 ) " "
10 ( 1 ) " "
11 ( 1 ) " "
      Promoter_hypomethylation_in_cancer
1  ( 1 ) " "
2  ( 1 ) " "
3  ( 1 ) " "
4  ( 1 ) " "
5  ( 1 ) " "
6  ( 1 ) " "
7  ( 1 ) " "
8  ( 1 ) " "
9  ( 1 ) " "
10 ( 1 ) " "

```

```

11  ( 1 ) " "
          Gene_body_hypermethylation_in_cancer
1  ( 1 ) " "
2  ( 1 ) " "
3  ( 1 ) " "
4  ( 1 ) " "
5  ( 1 ) " "
6  ( 1 ) " "
7  ( 1 ) " "
8  ( 1 ) " "
9  ( 1 ) " "
10 ( 1 ) " "
11 ( 1 ) " "
          Gene_body_hypomethylation_in_cancer Canyon_genebody_hypermethylation
1  ( 1 ) " " " "
2  ( 1 ) " " " "
3  ( 1 ) " " " "
4  ( 1 ) " " " "
5  ( 1 ) " " " "
6  ( 1 ) " " " "
7  ( 1 ) " " " "
8  ( 1 ) " " " "
9  ( 1 ) " " " "
10 ( 1 ) " " " "
11 ( 1 ) " " " "
          intolerant_pLI intolerant_pRec intolerant_pNull Synonymous_Zscore
1  ( 1 ) " " " " " "
2  ( 1 ) " " " " " "
3  ( 1 ) " " " " " "
4  ( 1 ) " " " " " "
5  ( 1 ) " " " " " "
6  ( 1 ) "*" " " " " "
7  ( 1 ) "*" " " " " "
8  ( 1 ) "*" " " " " "
9  ( 1 ) "*" " " " " "
10 ( 1 ) "*" " " " " "
11 ( 1 ) "*" " " " " "
          Missense_Zscore pLOF_Zscore dN_to_dS_ratio GDI RVIS_percentile ncRVIS
1  ( 1 ) " " " " " " " "
2  ( 1 ) " " " " " " " "
3  ( 1 ) " " " " " " " "
4  ( 1 ) " " " " " " " "
5  ( 1 ) " " " " " " " "
6  ( 1 ) " " " " " " " "
7  ( 1 ) " " " " " " " "
8  ( 1 ) " " " " " " " "
9  ( 1 ) " " " " " " " "
10 ( 1 ) " " " " " " " "
11 ( 1 ) " " " " " " " "
          ncGERP Gene_age FamilyMemberCount Length_H3K4me3
1  ( 1 ) " " " " " "
2  ( 1 ) " " " " " "
3  ( 1 ) " " " " " "
4  ( 1 ) " " " " " "

```

```

5  ( 1 ) " " " "
6  ( 1 ) " " " "
7  ( 1 ) " " " "
8  ( 1 ) " " " "
9  ( 1 ) " " " "
10 ( 1 ) " " " "
11 ( 1 ) " " " "
          Broad_H3K4me3_percentage H3K4me3_height H3K4me2_width
1  ( 1 ) " " " "
2  ( 1 ) " " " "
3  ( 1 ) " " " "
4  ( 1 ) " " " "
5  ( 1 ) " " " "
6  ( 1 ) " " " "
7  ( 1 ) " " " "
8  ( 1 ) " " " "
9  ( 1 ) " " " "
10 ( 1 ) " " " "
11 ( 1 ) " " " "
          Broad_H3K4me2_percentage H3K4me2_height H3K4me1_width
1  ( 1 ) " " " "
2  ( 1 ) " " " "
3  ( 1 ) " " " "
4  ( 1 ) " " " "
5  ( 1 ) " " " "
6  ( 1 ) " " " "
7  ( 1 ) " " " "
8  ( 1 ) " " " "
9  ( 1 ) " " " "
10 ( 1 ) " " " "
11 ( 1 ) " " " "
          Broad_H3K4me1_percentage H3K4me1_height H3K36me3_width
1  ( 1 ) " " " "
2  ( 1 ) " " " "
3  ( 1 ) " " " "
4  ( 1 ) " " " "
5  ( 1 ) " " " "
6  ( 1 ) " " " "
7  ( 1 ) " " " "
8  ( 1 ) " " " "
9  ( 1 ) " " " "
10 ( 1 ) " " " "
11 ( 1 ) " " " "
          Broad_H3K36me3_percentage H3K36me3_height H3K27ac_width
1  ( 1 ) " " " "
2  ( 1 ) " " " "
3  ( 1 ) " " " "
4  ( 1 ) " " " "
5  ( 1 ) " " " "
6  ( 1 ) " " " "
7  ( 1 ) " " " "
8  ( 1 ) " " " "
9  ( 1 ) " " " "
10 ( 1 ) " " " "

```

```

11  ( 1 ) " "
          " "           " "
          Broad_H3K27ac_percentage H3K27ac_height H3K27me3_width
1  ( 1 ) " "
2  ( 1 ) " "
3  ( 1 ) " "
4  ( 1 ) " "
5  ( 1 ) " "
6  ( 1 ) " "
7  ( 1 ) " "
8  ( 1 ) " "
9  ( 1 ) " "
10 ( 1 ) " "
11 ( 1 ) " "
          Broad_H3K27me3_percentage H3K27me3_height H3K9me3_width
1  ( 1 ) " "
2  ( 1 ) " "
3  ( 1 ) " "
4  ( 1 ) " "
5  ( 1 ) " "
6  ( 1 ) " "
7  ( 1 ) " "
8  ( 1 ) " "
9  ( 1 ) " "
10 ( 1 ) " "
11 ( 1 ) " "
          Broad_H3K9me3_percentage H3K9me3_height H3K9ac_width
1  ( 1 ) " "
2  ( 1 ) " "
3  ( 1 ) " "
4  ( 1 ) " "
5  ( 1 ) " "
6  ( 1 ) " "
7  ( 1 ) " "
8  ( 1 ) " "
9  ( 1 ) " "
10 ( 1 ) " "
11 ( 1 ) " "
          Broad_H3K9ac_percentage H3K9ac_height H3K9me2_width
1  ( 1 ) " "
2  ( 1 ) " "
3  ( 1 ) " "
4  ( 1 ) " "
5  ( 1 ) "*"
6  ( 1 ) "*"
7  ( 1 ) "*"
8  ( 1 ) "*"
9  ( 1 ) "*"
10 ( 1 ) "*"
11 ( 1 ) "*"
          Broad_H3K9me2_percentage H3K9me2_height H3K79me2_width
1  ( 1 ) " "
2  ( 1 ) " "
3  ( 1 ) " "
4  ( 1 ) " "

```

```

5  ( 1 ) " "
6  ( 1 ) " "
7  ( 1 ) " "
8  ( 1 ) " "
9  ( 1 ) " "
10 ( 1 ) " "
11 ( 1 ) " "
          Broad_H3K79me2_percentage H3K79me2_height H4K20me1_width
1  ( 1 ) " " " "
2  ( 1 ) " " " "
3  ( 1 ) " " " "
4  ( 1 ) " " " "
5  ( 1 ) " " " "
6  ( 1 ) " " " "
7  ( 1 ) " " " "
8  ( 1 ) " " " "
9  ( 1 ) " " " "
10 ( 1 ) " " " "
11 ( 1 ) " " " "
          Broad_H4K20me1_percentage H4K20me1_height
1  ( 1 ) "*" " "
2  ( 1 ) "*" " "
3  ( 1 ) "*" " "
4  ( 1 ) "*" " "
5  ( 1 ) "*" " "
6  ( 1 ) "*" " "
7  ( 1 ) "*" " "
8  ( 1 ) "*" " "
9  ( 1 ) "*" " "
10 ( 1 ) "*" " "
11 ( 1 ) "*" " "

```

```
plot(best_subset)
```

