A statistical approach to colorectal cancer diagnosis based on protein signature

Group 5

Introduction

Colorectal cancer (CRC) is cancer of the last several inches of the colon, the lower part of the large intestine. Colorectal cancer often starts as clumps of cells called polyps. It is the third most common type of cancer in the United States [1]. Cancer cases have decreased with the use of Colonoscopies. The current procedure includes a fecal occult blood test (FOBT) for pre-selection of cases for further colonoscopy evaluation. However, its accuracy is quite low and does not adequately detect subjects with CRC. A non-invasive method for pre-selection of screening for CRC is in need. We have analyzed data collected by Surinova, S. et al. [2] which contains information about protein concentration in CRC and Healthy patients and built a statistical model to predict the possibility of presence of CRC in a subject. This is clearly a classification problem and using Logistic Regression, we were able to identify 6 proteins which correlate more with prevalence of CRC.

Definitions and Abbreviations

1. Training dataset - Dataset used to train the final selected model.
2. Testing dataset - Dataset used to test the final selected model.
3. Model-Fitting dataset - Dataset sampled from the Training dataset used to fit the candidate models.
4. Validation dataset - Dataset sampled from the Training dataset used to select a final model from the candidate models.
5. Explanatory Variables/Features - Normalized log2 values of Protein signatures
6. Class/Labels - Labels determining whether a subject has CRC or not.
7. CRC - Colorectal Cancer VIF - Variance Influence Factors
8. AIC - Akaike Information Criterion
9. SBC/BIC - Schwarz Bayesian Criterion/Bayesian Information Criterion
10. ROC - Receiver Operating Characteristic
11. AUROC - Area under Receiver Operating Characteristic

Methods

Datasets

Two independent datasets were used in this study. The first dataset (henceforth referred as training dataset) was used for training and selection of models and the second(referred as testing dataset) was used for final calculation of model metrics. The training dataset comprised of protein concentration of subjects from a prospective screening study (BLiTz) (Hundt et al, 2009; Brenner et. al, 2010) and a case–control study examining the role of colonoscopy in CRC prevention (DACHS+) (Brenner et. al, 2006, 2007). The

validation dataset included subjects selected at the University Hospital Olomouc [2]. The training dataset consisted of two groups (CRC and non-CRC) of 100 subjects each comprising of logarithmic values of protein signatures. The testing dataset consisted of 202 subjects with CRC and 67 subjects of non-CRC subjects. To make the intensities comparable for the purpose of predictive analysis, the median normalized log2-relative quantifications of the validation cohort were equalized with the median normalized log2-relative quantifications of the training cohort.

Preprocessing

Proteins with more than 25% missing values were removed from both datasets and rest were imputed with minimum value observed in the same dataset, representing the limit of detection of protein signatures [2]. The dataset consisted of large number of highly correlated proteins presenting the problem of multicollinearity among the predictors.

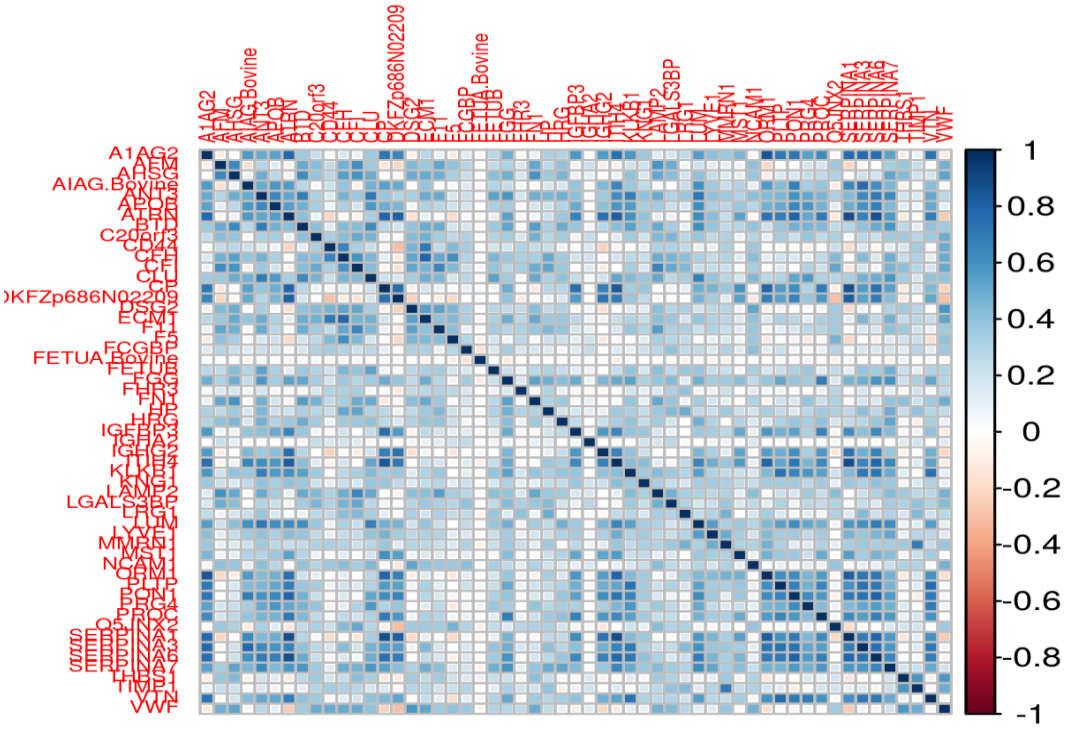


Figure 1 shows the correlation matrix for the training dataset.

Blue represents positive correlation and Red negative. The larger the squares, the larger the correlation. Stepwise Variance Inflation Factors(VIF) was used to eliminate the multicollinearity problem [3]. VIF for an explanatory variable is obtained using the pseudo r-squared value of regression of that variable against all other explanatory variables. A threshold of 5 was used to eliminate a variable.

$$VIF_j = \frac{1}{1 - R_j^2}$$

The training dataset was randomly split in a 80:20 manner into model-fitting and validation datasets.

Selection of Candidate Models

For creation of candidate models, the following four methods were used on the model-fitting dataset: "Stepwise Backward using AIC", "Stepwise Forward using AIC", "Stepwise Both Forward and Backward using AIC" and "Stepwise Both Forward and Backward using SBC Criterion". 2 best models from each method were selected as the Candidate models.

Evaluation of Candidate Models

Homer-Lemeshow Goodness of Fit Test was applied to each model. We consider the model to be unfit if p-value < 0.05. However, none of the candidate models had p-value < 0.05 and hence, none were eliminated [Appendix].
We next plot the deviance residuals of the models with predicted probabilities with Lowess Smooth.

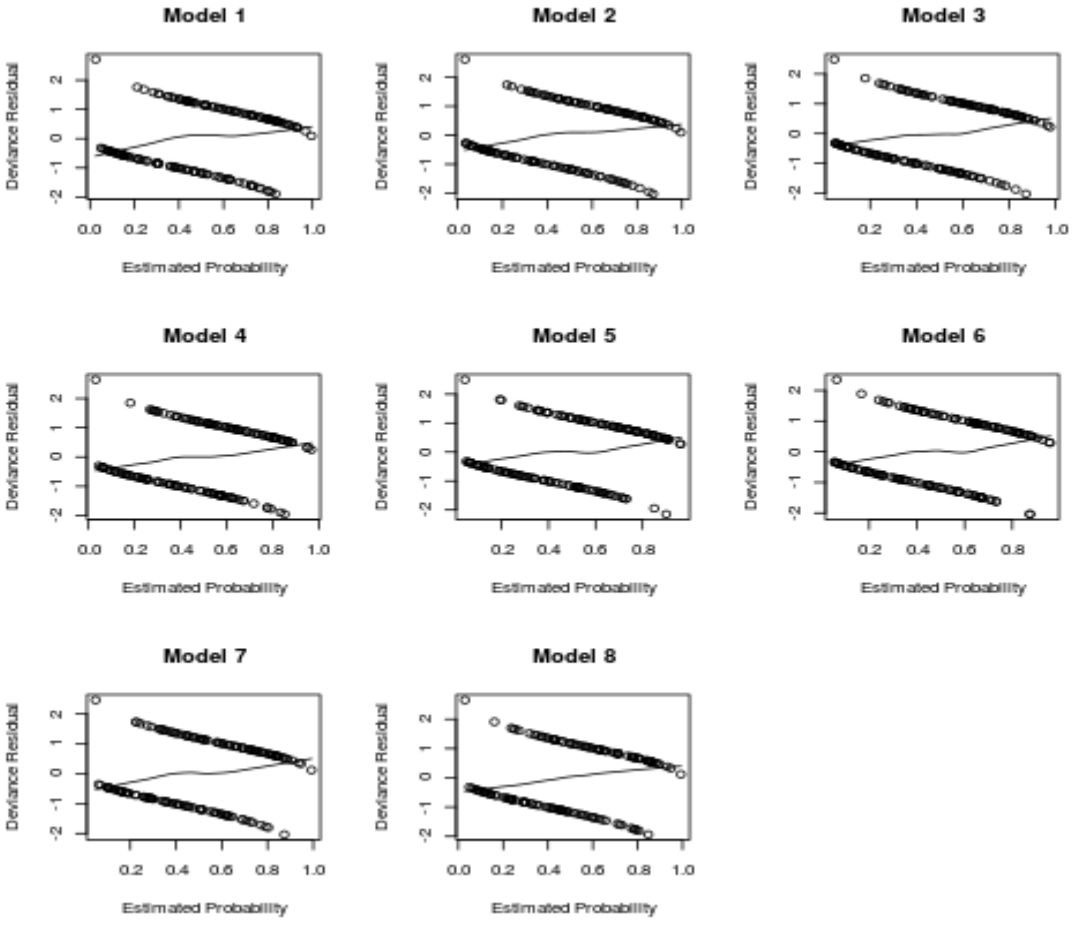


Figure 2 shows the plot of Deviance Residuals Vs Estimated Probability with Lowess Smooth.

If a model is correct, a lowess smooth of the plot of the residuals against the estimated probability $\hat{\pi}_i$ should result approximately in a horizontal line with zero intercept. None of the plots depict any significant departure from this and hence, there is no evidence that any model is inadequate.

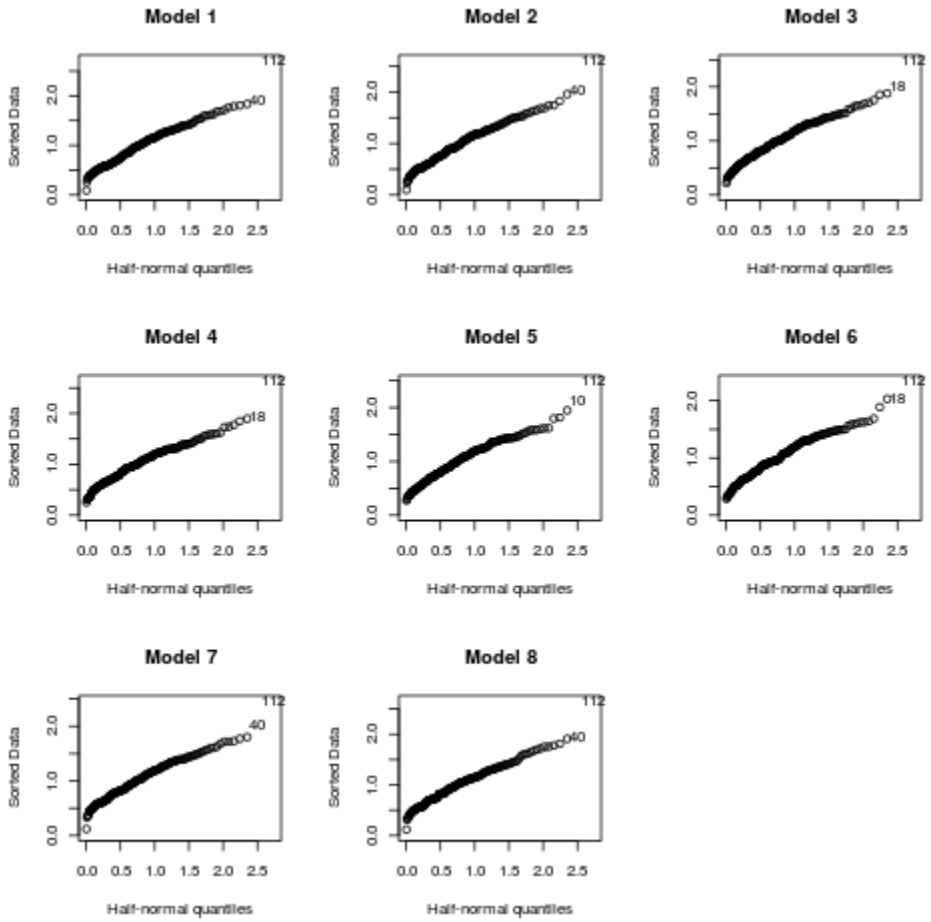


Figure 3 shows the half-normal probability plots for each candidate model.

A Half-normal probability plot helps to highlight outlying deviance residuals even though the residuals are not normally distributed. Outliers appear at the top right of a half-normal probability plot as points separated from the others. From the plots, we observe that observations 18, 40 and 112 might be an outlier. However, more detailed study of these observations is required before we exclude them as outliers. For the purpose of this study, we do not consider them as outliers.

We next consider overdispersion. Sometimes we can get a deviance that is much larger than expected if the model was correct. It can be due to the presence of outliers, sparse data or clustering of data. The approach to deal with overdispersion is to add a dispersion parameter σ^2 . It can be estimated with: $\hat{\sigma}^2 = \frac{\chi^2}{n-p}$ (p = no. of parameters in model).

$$Var\{Y_i\} = \phi n_i \pi_i \{1 - \pi_i\}$$

We consider overdispersion to exist if $\phi >> 1$.
The values of ϕ for the 8 candidate models were 1.17, 1.12, 1.07, 1.13, 1.08, 1.04, 1.06 and 1.14 suggesting dispersion not very different than 1 (no dispersion).

We next consider the predictive ability of the models. We observe how each of the model performs by predicting on the validation set. We generate the ROC plot and calculate the AUROC for each of the models. The Area under the ROC provides an unbiased, and non-parametric measure of the discrimination ability of the model. AOC = 0.5 means that predictions are no better than random guessing. An AUROC value >= 0.80 is considered ideal.

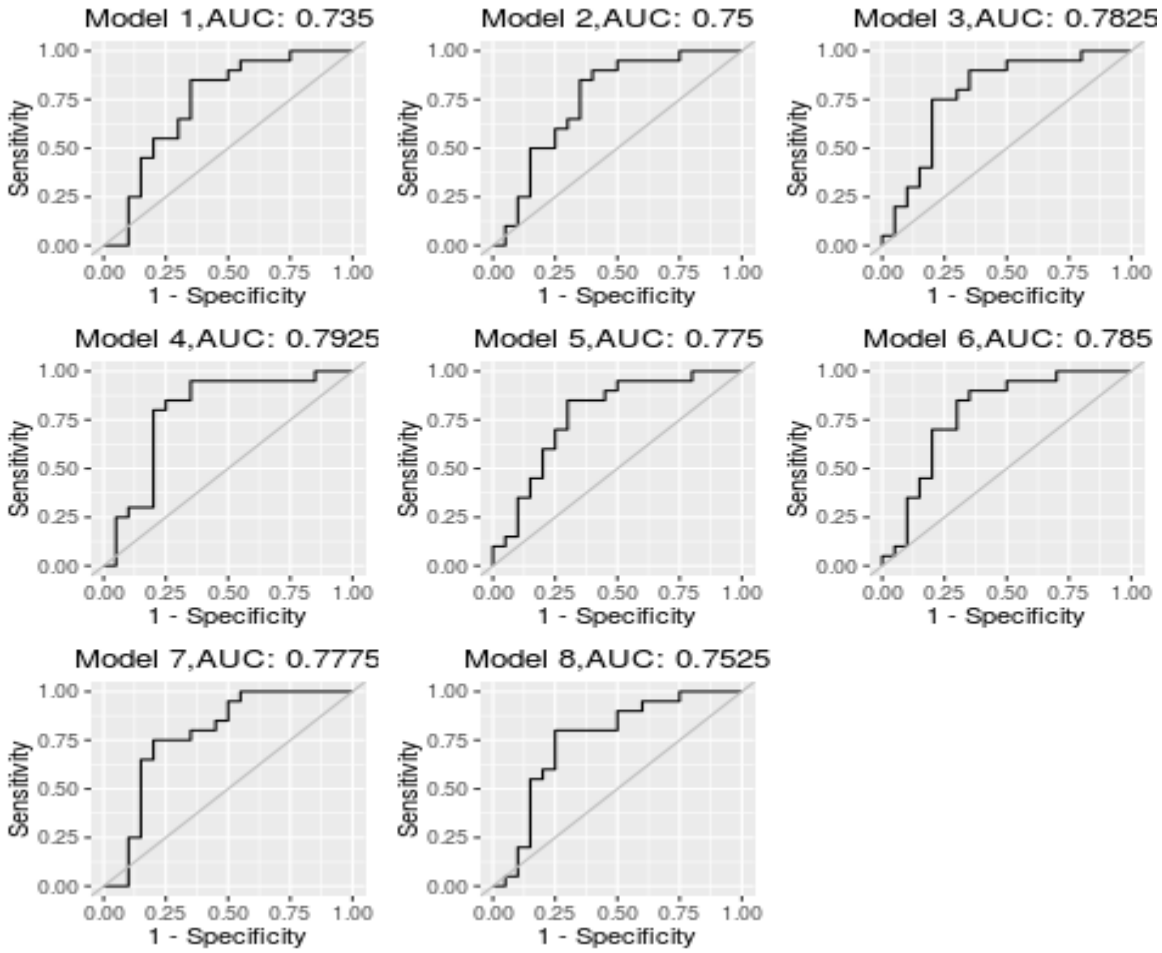


Figure 4 shows the ROC plots for each of the 8 models.

The AUROC is shown in the title of each of the plots. We observe that Model 4 has the maximum AUROC with value 0.7925. We'll consider this as our final model.
Now that we have finalized our model, we still need to determine the best cut-off value for classification. For the purpose of this study, we would like to have a model with high sensitivity as we wish to minimize incorrect label of a CRC subject as non-CRC.

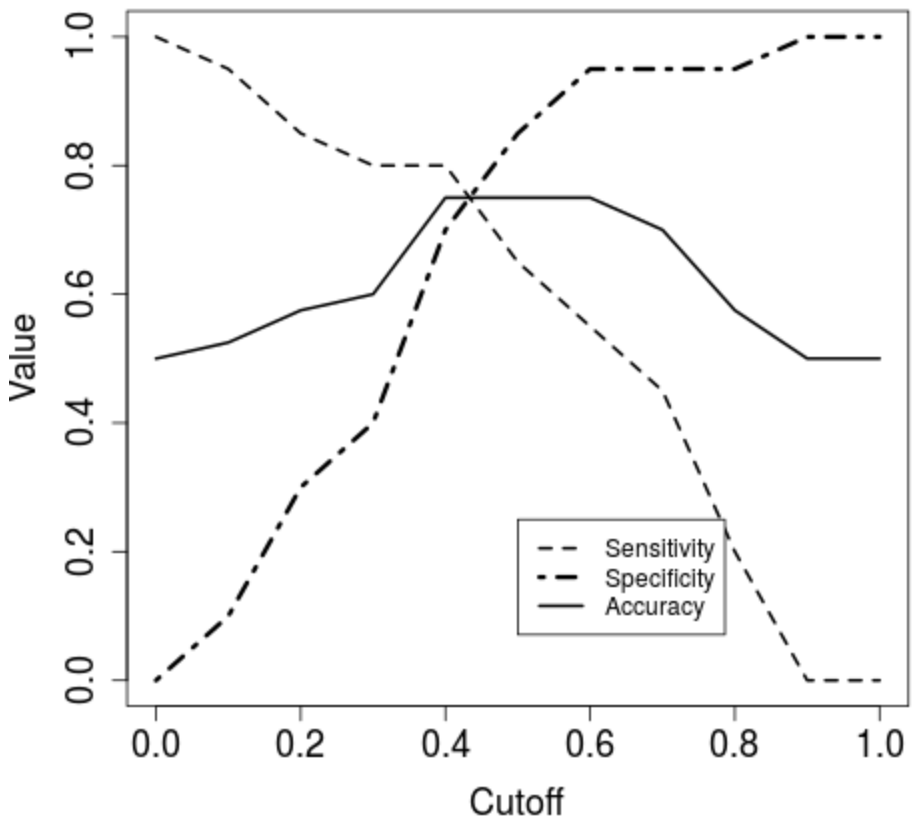


Figure 5 shows a Sensitivity Vs Specificity Vs Accuracy plot.

The sensitivity curve(light-dashed), the specificity curve(darker-dashed) and accuracy curve(solid) all merge at cut-off point 0.45. Since this point also results in a high sensitivity(0.8) for the validation dataset, we use this as our final cut-off point. We now have a final Model and a cut-off point decided. Since the model selection is complete, We merge both the model-fitting and the validation datasets and fit the model with it (while allowing for overdispersion).

Results

Our Final model consists of 6 proteins: TIMP1, LAMP2, HP, LRG1, SERPINA7 and LUM.
Figure 6 shows the predicted probability lines as function of one predictor, while fixing the remaining predictors at their median values.
Prediction of observations from the training dataset results in 70% balanced Accuracy and 0.74 Sensitivity and 0.66 Specificity. Prediction of new observations from the testing dataset for the selected model results in 65.4% balanced Accuracy with 0.995 Sensitivity and 0.31 Specificity.

Discussion

Since we modeled our experiment to have high Sensitivity values, the results are acceptable to us. Also, the balanced accuracy for both training and testing datasets are almost similar (70 and 65.4 respectively). However, we do realize that the specificity and hence, the balanced accuracy has suffered. There will always be a trade-off between Sensitivity and Specificity and due to the critical nature of incorrect labeling of a CRC subject as non-CRC, we have made the choice of choosing higher Sensitivity.
The predictive ability might increase with more data in the training set. Also, usage of techniques like K-Nearest Neighbor for imputation of missing values in the training and testing dataset might yield better results. But since we were limiting our techniques to those learnt in class, we ignored this method. Finally, using an ensemble classification method like Random Forest or Gradient Boosted Trees or even regularization techniques might help.

References

[1] "Colorectal Cancer—Patient Version".Retrieved from https://www.cancer.gov/types/colorectal
[2] "Prediction of colorectal cancer diagnosis based on circulating plasma proteins", Silvia Surinova , Meena Choi, Sha Tao, Peter J Schüffler, Ching-Yun Chang, Timothy Clough, Kamil Vyslouzil, Marta Khoylou, Josef Srovnal, Yansheng Liu, Mariette Matondo, Ruth Hüttenhain, Hendrik Weisser, Joachim M Buhmann, Marián Hajdúch, Hermann Brenner, Olga Vitek & Ruedi Aebersold.
[3] Marcus."Collinearity and stepwise VIF selection".Retrieved from https://beckmw.wordpress.com/2013/02/05/collinearity-and-stepwise-vif-selection/
[4] Applied Linear Statistical Methods, Michael Kutner, Christopher Nachtsheim, John Neter, William Li.

Appendix

Pre-Processing: Cleaning and Formatting Data

Input

```
load(file = "Surinova_training_abun.Rda")
cols <- Surinova_training_abun[,1]
input.train.raw <- t(Surinova_training_abun[,-1])
input.train.raw <- data.frame(input.train.raw)
colnames(input.train.raw) <- cols

load(file = "Surinova_testing_abun.Rda")
cols <- Surinova_testing_abun[,1]
input.test.raw <- t(Surinova_testing_abun[,-1])
input.test.raw <- data.frame(input.test.raw)
colnames(input.test.raw) <- cols

# check to see the column names match in both the input datasets.
which(colnames(input.train.raw) != colnames(input.test.raw))

## integer(0)

# Sicne protein abundances are normalized within each dataset and not
between training and validation set, we cannot merge these both
datasets.

# Formatting Column Names
colnames(input.train.raw)[colnames(input.train.raw) == 'AIAG-Bovine']
<- 'AIAG.Bovine'
colnames(input.train.raw)[colnames(input.train.raw) == 'FETUA-Bovine']
<- 'FETUA.Bovine'

colnames(input.test.raw)[colnames(input.test.raw) == 'AIAG-Bovine'] <-
'AIAG.Bovine'
colnames(input.test.raw)[colnames(input.test.raw) == 'FETUA-Bovine'] <-
'FETUA.Bovine'

# Create Class label: CRC = 1, Healthy = 0
input.train.raw$Class <- gsub("_.*", "\\1",rownames(input.train.raw))
input.train.raw$Class <- as.factor(input.train.raw$Class)
table(input.train.raw$Class)

##
##     CRC Healthy
##     100     100

input.test.raw$Class <- gsub("_.*", "\\1",rownames(input.test.raw))
input.test.raw$Class[input.test.raw$Class == 'Control'] <- 'Healthy'
```

```
input.test.raw$Class <- as.factor(input.test.raw$Class)
table(input.test.raw$Class)

##
##     CRC Healthy
##     202      67

head(input.train.raw)

##              A1AG2      AFM     AHSG AIAG.Bovine     ANT3      AOC3
## CRC_P1A10 14.23816 16.10302 19.95179    15.25354 17.20794 10.033227
## CRC_P1A2  15.02411 16.02071 19.71592    15.15455 17.29790  9.035202
## CRC_P1A4  15.63136 16.14380 19.71085    15.59163 17.59625 10.382938
## CRC_P1A6  15.40136 16.27642 19.70438    15.11819 17.42250  9.504018
## CRC_P1B12 16.00316 16.95821 20.42033    15.58249 17.98820  9.651676
## CRC_P1B2  13.93242 16.52772 19.88985    13.37131 16.32493  9.521379
##               APOB     ATRN      BTD   C20orf3     CADM1    CD163
## CRC_P1A10 15.54477 14.38339 16.28307 10.660954  9.743511 10.94965
## CRC_P1A2  15.13188 13.98172 16.24919 10.702064  9.702175 10.83449
## CRC_P1A4  15.95530 14.63535 16.49916 11.188267 10.373824 11.22385
## CRC_P1A6  15.71493 14.06070 16.27773  9.966157 10.105507 11.17723
## CRC_P1B12 16.24733 14.20360 16.54142 12.574731  9.690320 12.23111
## CRC_P1B2  14.51862 13.70165 15.79196 10.578714  9.190440 12.21289
##                CD44     CDH5      CFH      CFI      CLU       CP
CTSD
## CRC_P1A10  9.943858 8.749720 17.26186 16.52607 19.32642 16.99737
9.275675
## CRC_P1A2  10.425750 9.056710 17.26762 16.86256 19.40444 16.22411
10.865768
## CRC_P1A4  11.026696 9.477187 17.98868 17.20551 19.53746 17.75640
10.029862
## CRC_P1A6  11.103752 9.866905 18.17423 17.52567 19.38313 18.05366
10.203826
## CRC_P1B12 11.166600 8.522353 18.82219 18.47344 20.19405 16.23669
10.941220
## CRC_P1B2   8.038643 9.225166 16.13991 16.94726 18.90905 16.34736
11.858181
##           DKFZp686N02209     DSG2      ECM1      F11       F5
FCGBP
## CRC_P1A10       20.70295 10.22517 13.113391 14.81510 12.23214
12.38189
## CRC_P1A2        21.86027 10.04610 12.569733 14.71162 12.14809
11.68763
## CRC_P1A4        21.84249 11.11479 13.661017 15.56122 12.80789
12.40291
## CRC_P1A6        20.70488 10.95813 13.545199 15.95390 13.18926
13.03992
## CRC_P1B12       21.18504 10.87991 13.371014 15.98843 12.54162
13.23842
## CRC_P1B2        21.77231 10.23335  9.915183 14.69724 11.62907
```

```
12.69986
##            FETUA.Bovine    FETUB       FGA      FGG     FHR3      FN1
## CRC_P1A10     17.04864 13.29704  8.144845 11.34954 11.52154 12.55169
## CRC_P1A2      17.07832 13.50235  9.305395 12.19192 10.73714 13.47176
## CRC_P1A4      17.12503 13.55688  9.432872 11.49952 11.56901 12.36536
## CRC_P1A6      17.06354 13.12742  9.746036 11.87189 13.40319 12.17770
## CRC_P1B12     17.02016 14.95305 10.066428 11.59058 12.56019 14.10296
## CRC_P1B2      17.08612 13.19051        NA 10.68147 10.05570 11.81432
##               GOLM1       HP      HRG    HYOU1    ICAM1    IGFBP3
IGHA2
## CRC_P1A10 8.529159 19.74767 17.47524 7.504040 9.090639 10.918418
18.90573
## CRC_P1A2        NA 20.12543 17.95990 9.764873 9.253566 10.357056
21.04688
## CRC_P1A4        NA 20.78502 17.57534 8.646144 9.369628 11.619532
18.87876
## CRC_P1A6        NA 21.87358 17.38916 8.868061 9.760900 11.425023
20.89932
## CRC_P1B12 9.749398 22.79789 17.95183 8.115907 9.887392 10.175458
21.26316
## CRC_P1B2  9.913008 20.71284 16.76395 8.842497 9.027833  9.527139
21.92661
##              IGHG2    ITIH4    KLKB1     KNG1    LAMP2      LCN2
LGALS3BP
## CRC_P1A10 22.10597 15.64604 14.12489 17.65508 13.68160  9.446541
14.82148
## CRC_P1A2  22.44804 15.36709 14.19172 16.98601 13.48692 10.623488
15.12812
## CRC_P1A4  22.95173 16.40141 14.09215 17.39206 14.12212 10.024225
14.20939
## CRC_P1A6  22.93422 16.18817 13.48183 17.82389 14.41111 10.109405
14.87094
## CRC_P1B12 21.26145 15.87709 15.19997 17.95218 15.17730 11.052435
16.55743
## CRC_P1B2  21.00712 14.94386 13.20866 17.39358 13.94826  9.355121
16.28340
##               LRG1      LUM    LYVE1    MMRN1       MPO     MRC2
MST1
## CRC_P1A10 13.96127 14.99193 11.86759 11.39385  9.729238 10.86842
12.96116
## CRC_P1A2  14.08645 15.41582 11.71390 11.15791 10.202602 10.01434
13.09922
## CRC_P1A4  15.05310 15.06369 12.19818 11.73116  9.782391 10.56038
14.11266
## CRC_P1A6  14.92352 14.60999 11.77828 11.95025 10.761889 10.02103
13.92377
## CRC_P1B12 16.46953 15.93393 12.92101 12.33475 11.727629 11.41932
13.10718
## CRC_P1B2  14.49637 14.47633 12.45132 11.68845  9.834259 10.91711
13.37761
```

```
##               NCAM1     ORM1     PGCP      PIGR     PLTP    PLXDC2
## CRC_P1A10  9.905815 16.74179 8.379739  9.744116 11.85560 10.050261
## CRC_P1A2  10.496499 17.24994 8.406487  9.422932 11.54273  9.974399
## CRC_P1A4  10.521723 18.41524 8.243928  9.306936 12.15338  9.812975
## CRC_P1A6  10.385096 18.09024 9.114961  9.417459 11.19101  9.847176
## CRC_P1B12 10.517894 18.28100 8.778907 11.389195 12.12600 10.414638
## CRC_P1B2   9.854809 17.02723 8.368942  9.466000 11.35824  8.387396
##                PON1     PRG4      PROC     PTPRJ   Q5JNX2 SERPINA1
SERPINA3
## CRC_P1A10 16.87047 11.84463 11.109446  9.758946 19.02813 18.06109
14.21804
## CRC_P1A2  16.55428 12.49583 11.111383  9.934520 19.86254 17.61019
14.83233
## CRC_P1A4  16.34508 11.70096 11.533164 10.328319 19.50177 18.53997
15.30873
## CRC_P1A6  15.46083 11.09706 11.693945  9.741059 19.98466 17.40141
14.25424
## CRC_P1B12 16.46015 14.46691 10.699745 10.568064 20.46745 17.44801
16.23191
## CRC_P1B2  14.47485 11.85023  9.470798  9.922271 18.14204 17.95217
12.60729
##           SERPINA6 SERPINA7    THBS1    TIMP1       TNC      VTN
VWF
## CRC_P1A10 15.79685 13.58185 13.99569 11.57875 10.275261 12.06684
10.66103
## CRC_P1A2  15.81540 13.13844 13.70771 11.96097 10.234434 12.83689
10.77843
## CRC_P1A4  16.40177 13.74233 15.63696 12.15584 10.241626 12.31163
11.18527
## CRC_P1A6  16.31049 13.93689 15.51727 12.37927  9.269195 10.67764
10.88786
## CRC_P1B12 16.48583 14.16713 15.12911 12.16912 10.476927 14.50458
11.25418
## CRC_P1B2  15.29659 13.23403 15.39790 12.52729 10.026299 11.52916
11.03501
##           Class
## CRC_P1A10   CRC
## CRC_P1A2    CRC
## CRC_P1A4    CRC
## CRC_P1A6    CRC
## CRC_P1B12   CRC
## CRC_P1B2    CRC
```

```
str(input.train.raw)
```

```
## 'data.frame':    200 obs. of  73 variables:
##  $ A1AG2         : num  14.2 15 15.6 15.4 16 ...
##  $ AFM           : num  16.1 16 16.1 16.3 17 ...
##  $ AHSG          : num  20 19.7 19.7 19.7 20.4 ...
##  $ AIAG.Bovine   : num  15.3 15.2 15.6 15.1 15.6 ...
```

```
##  $ ANT3           : num  17.2 17.3 17.6 17.4 18 ...
##  $ AOC3           : num  10.03 9.04 10.38 9.5 9.65 ...
##  $ APOB           : num  15.5 15.1 16 15.7 16.2 ...
##  $ ATRN           : num  14.4 14 14.6 14.1 14.2 ...
##  $ BTD            : num  16.3 16.2 16.5 16.3 16.5 ...
##  $ C20orf3        : num  10.66 10.7 11.19 9.97 12.57 ...
##  $ CADM1          : num  9.74 9.7 10.37 10.11 9.69 ...
##  $ CD163          : num  10.9 10.8 11.2 11.2 12.2 ...
##  $ CD44           : num  9.94 10.43 11.03 11.1 11.17 ...
##  $ CDH5           : num  8.75 9.06 9.48 9.87 8.52 ...
##  $ CFH            : num  17.3 17.3 18 18.2 18.8 ...
##  $ CFI            : num  16.5 16.9 17.2 17.5 18.5 ...
##  $ CLU            : num  19.3 19.4 19.5 19.4 20.2 ...
##  $ CP             : num  17 16.2 17.8 18.1 16.2 ...
##  $ CTSD           : num  9.28 10.87 10.03 10.2 10.94 ...
##  $ DKFZp686N02209: num  20.7 21.9 21.8 20.7 21.2 ...
##  $ DSG2           : num  10.2 10 11.1 11 10.9 ...
##  $ ECM1           : num  13.1 12.6 13.7 13.5 13.4 ...
##  $ F11            : num  14.8 14.7 15.6 16 16 ...
##  $ F5             : num  12.2 12.1 12.8 13.2 12.5 ...
##  $ FCGBP          : num  12.4 11.7 12.4 13 13.2 ...
##  $ FETUA.Bovine   : num  17 17.1 17.1 17.1 17 ...
##  $ FETUB          : num  13.3 13.5 13.6 13.1 15 ...
##  $ FGA            : num  8.14 9.31 9.43 9.75 10.07 ...
##  $ FGG            : num  11.3 12.2 11.5 11.9 11.6 ...
##  $ FHR3           : num  11.5 10.7 11.6 13.4 12.6 ...
##  $ FN1            : num  12.6 13.5 12.4 12.2 14.1 ...
##  $ GOLM1          : num  8.53 NA NA NA 9.75 ...
##  $ HP             : num  19.7 20.1 20.8 21.9 22.8 ...
##  $ HRG            : num  17.5 18 17.6 17.4 18 ...
##  $ HYOU1          : num  7.5 9.76 8.65 8.87 8.12 ...
##  $ ICAM1          : num  9.09 9.25 9.37 9.76 9.89 ...
##  $ IGFBP3         : num  10.9 10.4 11.6 11.4 10.2 ...
##  $ IGHA2          : num  18.9 21 18.9 20.9 21.3 ...
##  $ IGHG2          : num  22.1 22.4 23 22.9 21.3 ...
##  $ ITIH4          : num  15.6 15.4 16.4 16.2 15.9 ...
##  $ KLKB1          : num  14.1 14.2 14.1 13.5 15.2 ...
##  $ KNG1           : num  17.7 17 17.4 17.8 18 ...
##  $ LAMP2          : num  13.7 13.5 14.1 14.4 15.2 ...
##  $ LCN2           : num  9.45 10.62 10.02 10.11 11.05 ...
##  $ LGALS3BP       : num  14.8 15.1 14.2 14.9 16.6 ...
##  $ LRG1           : num  14 14.1 15.1 14.9 16.5 ...
##  $ LUM            : num  15 15.4 15.1 14.6 15.9 ...
##  $ LYVE1          : num  11.9 11.7 12.2 11.8 12.9 ...
##  $ MMRN1          : num  11.4 11.2 11.7 12 12.3 ...
##  $ MPO            : num  9.73 10.2 9.78 10.76 11.73 ...
##  $ MRC2           : num  10.9 10 10.6 10 11.4 ...
##  $ MST1           : num  13 13.1 14.1 13.9 13.1 ...
##  $ NCAM1          : num  9.91 10.5 10.52 10.39 10.52 ...
##  $ ORM1           : num  16.7 17.2 18.4 18.1 18.3 ...
```

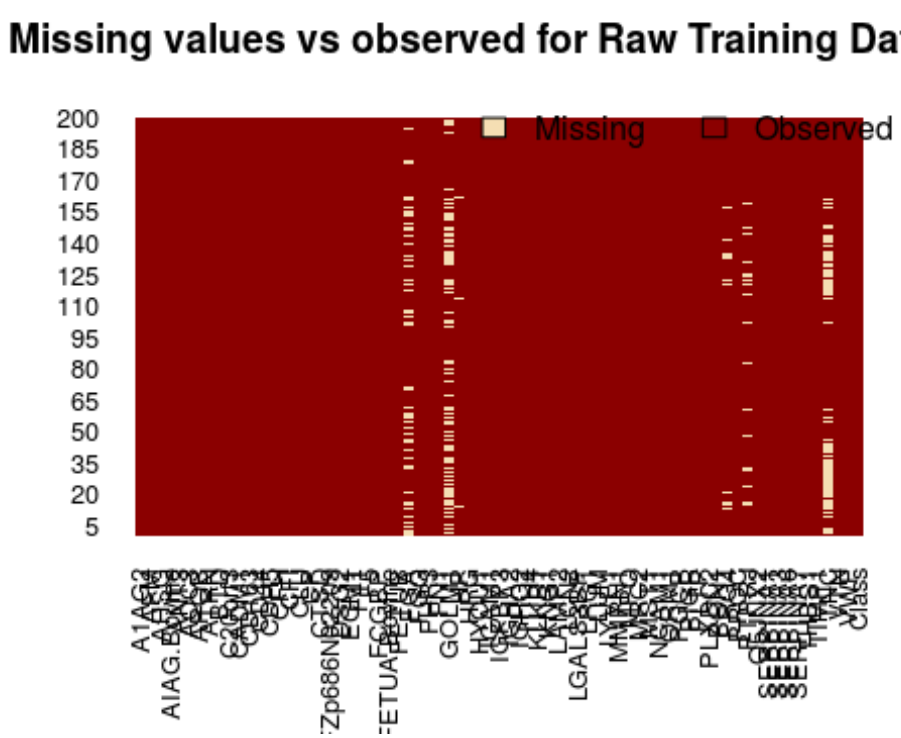
```
##  $ PGCP          : num  8.38 8.41 8.24 9.11 8.78 ...
##  $ PIGR          : num  9.74 9.42 9.31 9.42 11.39 ...
##  $ PLTP          : num  11.9 11.5 12.2 11.2 12.1 ...
##  $ PLXDC2        : num  10.05 9.97 9.81 9.85 10.41 ...
##  $ PON1          : num  16.9 16.6 16.3 15.5 16.5 ...
##  $ PRG4          : num  11.8 12.5 11.7 11.1 14.5 ...
##  $ PROC          : num  11.1 11.1 11.5 11.7 10.7 ...
##  $ PTPRJ         : num  9.76 9.93 10.33 9.74 10.57 ...
##  $ Q5JNX2        : num  19 19.9 19.5 20 20.5 ...
##  $ SERPINA1      : num  18.1 17.6 18.5 17.4 17.4 ...
##  $ SERPINA3      : num  14.2 14.8 15.3 14.3 16.2 ...
##  $ SERPINA6      : num  15.8 15.8 16.4 16.3 16.5 ...
##  $ SERPINA7      : num  13.6 13.1 13.7 13.9 14.2 ...
##  $ THBS1         : num  14 13.7 15.6 15.5 15.1 ...
##  $ TIMP1         : num  11.6 12 12.2 12.4 12.2 ...
##  $ TNC           : num  10.28 10.23 10.24 9.27 10.48 ...
##  $ VTN           : num  12.1 12.8 12.3 10.7 14.5 ...
##  $ VWF           : num  10.7 10.8 11.2 10.9 11.3 ...
##  $ Class         : Factor w/ 2 levels "CRC","Healthy": 1 1 1 1 1 1 1
1 1 1 ...
```

There are 200 subjects and 72 proteins in the training dataset out of which 100 subjects are diagnosed with CRC and 100 are healthy.
There are 269 subjects and 72 proteins in the testing dataset out of which 202 subjects are diagnosed with CRC and 67 are healthy.

Pre-Processing: Missing Values

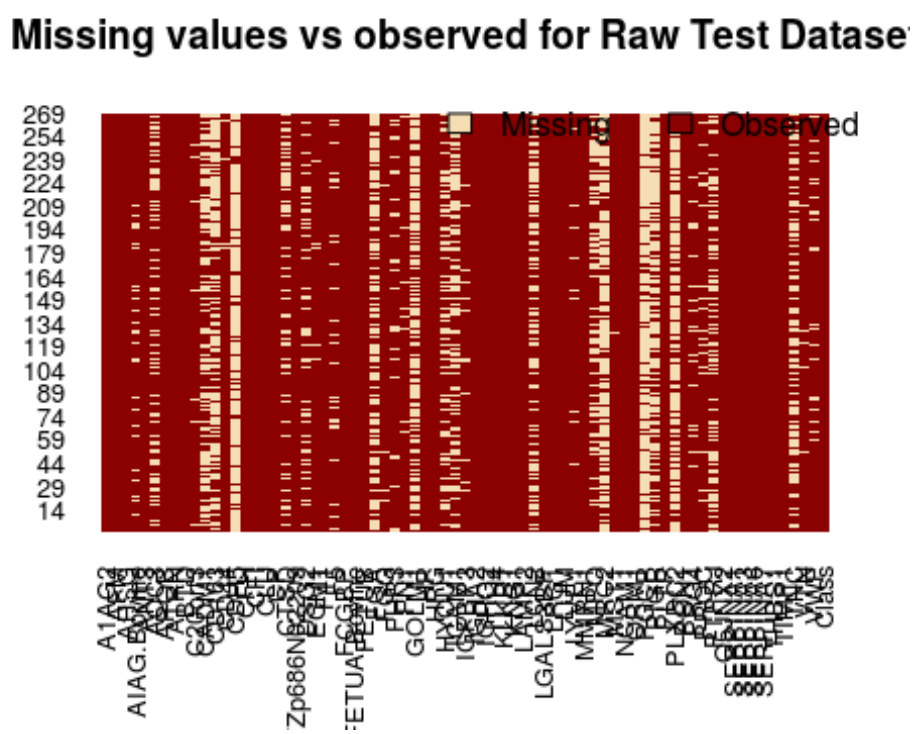
Visualizing Missing Values in Training dataset

```
par(mfrow=c(1,1))
missmap(input.train.raw, main = "Missing values vs observed for Raw
Training Dataset", rank.order=FALSE)
```



Visualizing Missing Values in Testing dataset

```
missmap(input.test.raw, main = "Missing values vs observed for Raw Test Dataset", rank.order=FALSE)
```



Most of the missing values are related to the second dataset("Surinova_testing_abun.Rda")

Dealing with Missing Values Pt. 1: Drop > 25% Missing

Training Dataset

```
# Remove Columns with Missing Values.
missing.percent.train <- sapply(input.train.raw,function(x)
sum(is.na(x)) * 100/length(x))
missing.columns.train <-
sort(missing.percent.train[missing.percent.train != 0])
missing.columns.train

##    HP  PRG4 PTPRJ   FGA   TNC GOLM1
##   1.5   5.0   8.0  23.5  33.5  36.0

length(missing.columns.train)

## [1] 6
```

There are 6 columns with Missing Values in the Training Dataset.

Testing Dataset

```
# Remove Columns with Missing Values.
missing.percent.test <- sapply(input.test.raw, function(x)
sum(is.na(x)) * 100/length(x))
missing.columns.test <- sort(missing.percent.test[missing.percent.test
```

```
!= 0])
missing.columns.test
```

```
##        ATRN        MST1        CD44        ECM1         VTN
FGG
##   0.3717472   0.7434944   1.8587361   2.2304833   2.6022305
2.9739777
##     C20orf3         FN1       LYVE1      IGFBP3        PROC
PRG4
##   3.7174721   4.4609665   5.5762082   7.0631970   8.5501859
9.2936803
##         VWF        FHR3          F5 AIAG.Bovine        DSG2
CTSD
##  11.8959108  12.6394052  13.0111524  13.3828996  23.4200743
28.2527881
##       HYOU1        AOC3         MPO       PTPRJ       ICAM1
LCN2
##  30.4832714  31.5985130  33.8289963  40.5204461  43.8661710
43.8661710
##       CADM1         TNC         FGA        PIGR       CD163
GOLM1
##  44.6096654  46.4684015  55.0185874  56.5055762  59.8513011
61.3382900
##      PLXDC2        MRC2        PGCP        CDH5
##  70.2602230  72.1189591  83.6431227  84.3866171
```

```
length(missing.columns.test)
```

```
## [1] 34
```

There are 34 columns with Missing Values in the Testing Dataset.

As we observe from the above tables, there are large number of columns with more than 25 % missing data.We'll drop these columns from both the datasets.

```
# we'll combine the dropped column list from both training and testing
datasets to keep the remaining columns consistent.
dropped.columns.train <-
names(missing.columns.train[missing.columns.train > 25])
dropped.columns.test <- names(missing.columns.test[missing.columns.test
> 25])
dropped.columns.final <- union(dropped.columns.train,
dropped.columns.test)
dropped.columns.final
```

```
##  [1] "TNC"    "GOLM1"  "CTSD"   "HYOU1"  "AOC3"   "MPO"    "PTPRJ"
##  [8] "ICAM1"  "LCN2"   "CADM1"  "FGA"    "PIGR"   "CD163"  "PLXDC2"
## [15] "MRC2"   "PGCP"   "CDH5"
```

```
input.train.df <- input.train.raw
input.test.df <- input.test.raw
```

```
input.train.df[, dropped.columns.final] <- NULL
input.test.df[, dropped.columns.final] <- NULL
dim(input.train.df)
```

```
## [1] 200  56
```

```
dim(input.test.df)
```

```
## [1] 269  56
```

In total, We dropped 17 columns.

Dealing with Missing Values Pt. 2: Replace with Min

For the rest of the columns with missing values, we'll replace the missing values with the mean.

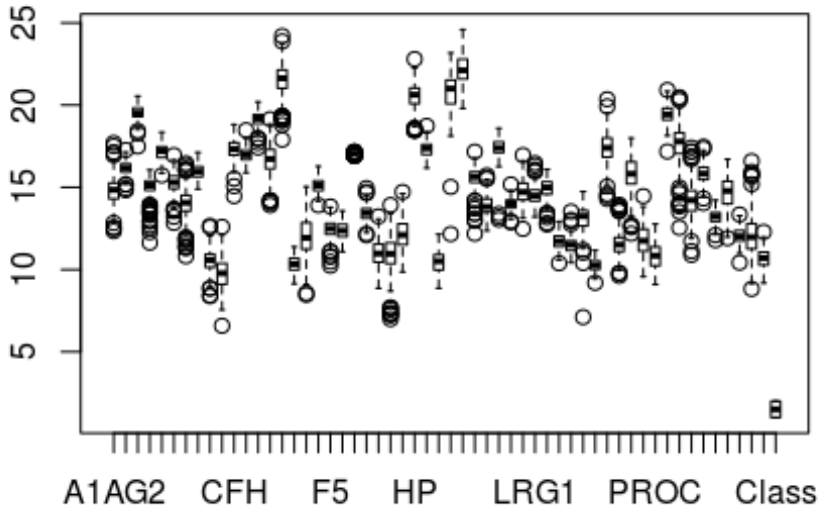
```
# Replace Missing Values with Min
replace.min <- function(x) replace(x, is.na(x), min(x, na.rm=TRUE))
cols <- colnames(subset(input.train.df, select = -Class))
input.train.df[, cols] <- sapply(input.train.df[, cols], replace.min)
input.test.df[, cols] <- sapply(input.test.df[, cols], replace.min)
```

Now, we do not have any missing values in our data.

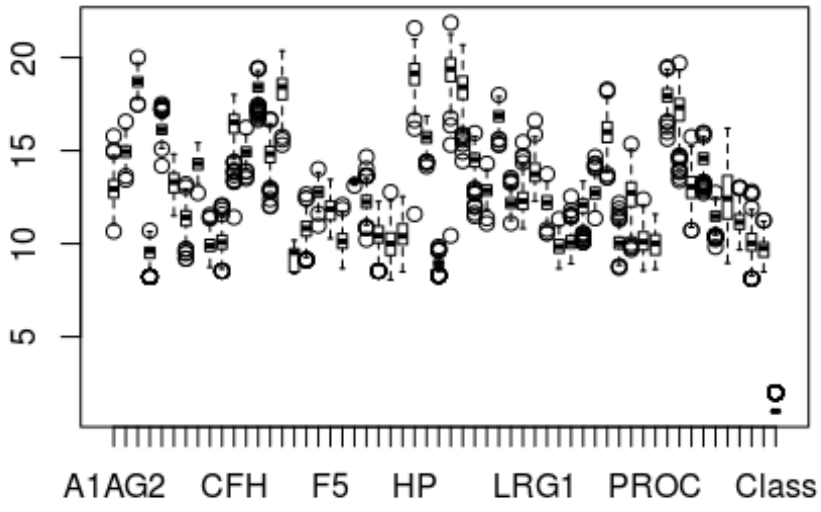
Data Exploration

Boxplot

```
par(mfrow = c(1,1))
boxplot(input.train.df)
```




```
boxplot(input.test.df)
```



The boxplots does show some points as outliers but since they are all logarithmic values,

scale is small. Also, we confirm that the data has indeed been normalized separately for each dataset.

Correlation amongst Predictors.

```
# Correlation between Predictors
correlations <- cor(subset(input.train.df, select=-Class))
png('correlations.png', width=4, height=4, units="in", res=300)
corrplot(correlations, method="square", tl.cex = 0.5)
#dev.off()
#dev.off()

cols <- colnames(subset(input.train.df, select = -Class))
keep.cols <- vif_func(in_frame=input.train.df[, cols],thresh=5,trace=F)

## Loading required package: fmsb

keep.cols

##  [1] "AFM"           "AHSG"          "AIAG.Bovine"   "APOB"
##  [5] "BTD"           "C20orf3"       "CD44"          "CFI"
##  [9] "CLU"           "DKFZp686N02209" "DSG2"         "F11"
## [13] "F5"            "FCGBP"         "FETUA.Bovine"  "FETUB"
## [17] "FHR3"          "FN1"           "HP"            "HRG"
## [21] "IGFBP3"        "IGHA2"         "IGHG2"         "KNG1"
## [25] "LAMP2"         "LGALS3BP"      "LRG1"          "LUM"
## [29] "LYVE1"         "MMRN1"         "MST1"          "NCAM1"
## [33] "PLTP"          "PRG4"          "PROC"          "Q5JNX2"
## [37] "SERPINA7"      "THBS1"         "TIMP1"         "VWF"

input.train.df <- input.train.df[, c(keep.cols, "Class")]
input.test.df <- input.test.df[, c(keep.cols, "Class")]
dim(input.train.df)

## [1] 200  41

dim(input.test.df)

## [1] 269  41
```

Blue represents positive correlation and Red negative. The larger the squares, the larger the correlation.
We observe that some of the predictors are highly correlated with each other.

Preprocessing: Shuffling and Splitting the Datasets

The input datasets consisted of all CRC subject data in the to half and Healthy subject data in the bottom half. The datasets are therefore shuffled to prevent bias in the training and testing cohorts. We'll split the training dataset into 2 subsets: training(80%) and validation sets(20%). We'll not split the testing dataset.

```r
set.seed(seed.id)
training <- input.train.df
inTrain <- createDataPartition(training$Class, p = 0.8, list=FALSE)
train <- training[inTrain, ]
validation <- training[-inTrain, ]

set.seed(seed.id)
testing <- input.test.df[sample(nrow(input.test.df)), ]
```

Model Selection

```r
glm.full <- glm(Class ~ ., data = train, family = binomial)
glm.null <- glm(Class ~ 1, data = train, family = binomial)
```

AIC-based backward selection

```r
model.aic.backward <- step(glm.full, direction = "backward", trace = 1)
```

```
## Start:  AIC=244.75
## Class ~ AFM + AHSG + AIAG.Bovine + APOB + BTD + C20orf3 + CD44 +
##     CFI + CLU + DKFZp686N02209 + DSG2 + F11 + F5 + FCGBP +
FETUA.Bovine +
##     FETUB + FHR3 + FN1 + HP + HRG + IGFBP3 + IGHA2 + IGHG2 +
##     KNG1 + LAMP2 + LGALS3BP + LRG1 + LUM + LYVE1 + MMRN1 + MST1 +
##     NCAM1 + PLTP + PRG4 + PROC + Q5JNX2 + SERPINA7 + THBS1 +
##     TIMP1 + VWF
##
##                  Df Deviance    AIC
## - C20orf3         1   162.75 242.75
## - LYVE1           1   162.75 242.75
## - CFI             1   162.76 242.76
## - PLTP            1   162.76 242.76
## - FN1             1   162.76 242.76
## - Q5JNX2          1   162.77 242.77
## - AFM             1   162.79 242.79
## - CD44            1   162.80 242.80
## - PROC            1   162.80 242.80
## - DSG2            1   162.85 242.85
## - BTD             1   162.86 242.86
## - MST1            1   162.90 242.90
## - AHSG            1   162.95 242.95
## - LRG1            1   162.98 242.98
## - APOB            1   162.98 242.98
## - F11             1   162.99 242.99
## - FETUB           1   163.02 243.02
## - PRG4            1   163.06 243.06
## - HRG             1   163.08 243.08
## - AIAG.Bovine     1   163.12 243.12
## - TIMP1           1   163.16 243.16
## - IGFBP3          1   163.20 243.20
```

```
## - F5              1   163.23 243.23
## - THBS1           1   163.44 243.44
## - VWF             1   163.46 243.46
## - CLU             1   163.48 243.48
## - FCGBP           1   163.65 243.65
## - LGALS3BP        1   163.85 243.85
## - FHR3            1   163.89 243.89
## - IGHG2           1   163.94 243.94
## - KNG1            1   164.13 244.13
## - IGHA2           1   164.42 244.42
## - DKFZp686N02209  1   164.43 244.43
## <none>                162.75 244.75
## - HP              1   164.85 244.85
## - NCAM1           1   164.96 244.96
## - MMRN1           1   164.97 244.97
## - FETUA.Bovine    1   165.47 245.47
## - LUM             1   166.60 246.60
## - SERPINA7        1   171.36 251.36
## - LAMP2           1   174.73 254.73
##
## Step:  AIC=242.75
## Class ~ AFM + AHSG + AIAG.Bovine + APOB + BTD + CD44 + CFI +
##     CLU + DKFZp686N02209 + DSG2 + F11 + F5 + FCGBP + FETUA.Bovine +
##     FETUB + FHR3 + FN1 + HP + HRG + IGFBP3 + IGHA2 + IGHG2 +
##     KNG1 + LAMP2 + LGALS3BP + LRG1 + LUM + LYVE1 + MMRN1 + MST1 +
##     NCAM1 + PLTP + PRG4 + PROC + Q5JNX2 + SERPINA7 + THBS1 +
##     TIMP1 + VWF
##
##                  Df Deviance    AIC
## - LYVE1           1   162.75 240.75
## - CFI             1   162.76 240.76
## - PLTP            1   162.76 240.76
## - FN1             1   162.76 240.76
## - Q5JNX2          1   162.77 240.77
## - AFM             1   162.79 240.79
## - CD44            1   162.80 240.80
## - PROC            1   162.80 240.80
## - DSG2            1   162.85 240.85
## - BTD             1   162.86 240.86
## - MST1            1   162.90 240.90
## - AHSG            1   162.95 240.95
## - LRG1            1   162.98 240.98
## - APOB            1   162.99 240.99
## - F11             1   162.99 240.99
## - FETUB           1   163.02 241.02
## - PRG4            1   163.07 241.07
## - HRG             1   163.08 241.08
## - AIAG.Bovine     1   163.15 241.15
## - TIMP1           1   163.16 241.16
## - IGFBP3          1   163.22 241.22
```

```
## - F5              1   163.23 241.23
## - THBS1           1   163.44 241.44
## - VWF             1   163.46 241.46
## - CLU             1   163.49 241.49
## - FCGBP           1   163.65 241.65
## - LGALS3BP        1   163.88 241.88
## - FHR3            1   163.90 241.90
## - IGHG2           1   163.94 241.94
## - KNG1            1   164.13 242.13
## - DKFZp686N02209  1   164.43 242.43
## - IGHA2           1   164.46 242.46
## <none>                162.75 242.75
## - HP              1   164.87 242.87
## - MMRN1           1   165.00 243.00
## - NCAM1           1   165.00 243.00
## - FETUA.Bovine    1   165.51 243.51
## - LUM             1   166.62 244.62
## - SERPINA7        1   171.45 249.45
## - LAMP2           1   174.73 252.73
##
## Step:  AIC=240.75
## Class ~ AFM + AHSG + AIAG.Bovine + APOB + BTD + CD44 + CFI +
##     CLU + DKFZp686N02209 + DSG2 + F11 + F5 + FCGBP + FETUA.Bovine +
##     FETUB + FHR3 + FN1 + HP + HRG + IGFBP3 + IGHA2 + IGHG2 +
##     KNG1 + LAMP2 + LGALS3BP + LRG1 + LUM + MMRN1 + MST1 + NCAM1 +
##     PLTP + PRG4 + PROC + Q5JNX2 + SERPINA7 + THBS1 + TIMP1 +
##     VWF
##
##                 Df Deviance    AIC
## - CFI            1   162.76 238.76
## - PLTP           1   162.76 238.76
## - FN1            1   162.76 238.76
## - Q5JNX2         1   162.78 238.78
## - AFM            1   162.79 238.79
## - CD44           1   162.80 238.80
## - PROC           1   162.80 238.80
## - DSG2           1   162.85 238.85
## - BTD            1   162.86 238.86
## - MST1           1   162.90 238.90
## - AHSG           1   162.95 238.95
## - LRG1           1   162.99 238.99
## - APOB           1   163.00 239.00
## - F11            1   163.00 239.00
## - FETUB          1   163.03 239.03
## - PRG4           1   163.08 239.08
## - HRG            1   163.08 239.08
## - AIAG.Bovine    1   163.15 239.15
## - TIMP1          1   163.16 239.16
## - IGFBP3         1   163.23 239.23
## - F5             1   163.24 239.24
```

```
## - THBS1           1   163.45 239.45
## - VWF             1   163.46 239.46
## - CLU             1   163.49 239.49
## - FCGBP           1   163.65 239.65
## - FHR3            1   163.90 239.90
## - LGALS3BP        1   163.90 239.90
## - IGHG2           1   163.95 239.95
## - KNG1            1   164.13 240.13
## - DKFZp686N02209  1   164.43 240.43
## - IGHA2           1   164.47 240.47
## <none>                162.75 240.75
## - HP              1   164.94 240.94
## - NCAM1           1   165.01 241.01
## - MMRN1           1   165.08 241.08
## - FETUA.Bovine    1   165.51 241.51
## - LUM             1   167.15 243.15
## - SERPINA7        1   172.02 248.02
## - LAMP2           1   174.74 250.74
##
## Step:  AIC=238.76
## Class ~ AFM + AHSG + AIAG.Bovine + APOB + BTD + CD44 + CLU +
##     DKFZp686N02209 + DSG2 + F11 + F5 + FCGBP + FETUA.Bovine +
##     FETUB + FHR3 + FN1 + HP + HRG + IGFBP3 + IGHA2 + IGHG2 +
##     KNG1 + LAMP2 + LGALS3BP + LRG1 + LUM + MMRN1 + MST1 + NCAM1 +
##     PLTP + PRG4 + PROC + Q5JNX2 + SERPINA7 + THBS1 + TIMP1 +
##     VWF
##
##                  Df Deviance    AIC
## - PLTP            1   162.76 236.76
## - FN1             1   162.77 236.77
## - Q5JNX2          1   162.78 236.78
## - AFM             1   162.79 236.79
## - PROC            1   162.81 236.81
## - CD44            1   162.81 236.81
## - DSG2            1   162.86 236.86
## - BTD             1   162.86 236.86
## - MST1            1   162.91 236.91
## - AHSG            1   162.95 236.95
## - APOB            1   163.01 237.01
## - LRG1            1   163.01 237.01
## - FETUB           1   163.04 237.04
## - F11             1   163.05 237.05
## - PRG4            1   163.08 237.08
## - HRG             1   163.09 237.09
## - TIMP1           1   163.16 237.16
## - AIAG.Bovine     1   163.24 237.24
## - IGFBP3          1   163.27 237.27
## - F5              1   163.28 237.28
## - THBS1           1   163.45 237.45
## - VWF             1   163.46 237.46
```

```
## - CLU              1   163.53 237.53
## - FCGBP            1   163.65 237.65
## - FHR3             1   163.90 237.90
## - IGHG2            1   163.95 237.95
## - LGALS3BP         1   163.97 237.97
## - KNG1             1   164.19 238.19
## - DKFZp686N02209   1   164.47 238.47
## - IGHA2            1   164.47 238.47
## <none>                 162.76 238.76
## - NCAM1            1   165.01 239.01
## - HP               1   165.07 239.07
## - MMRN1            1   165.08 239.08
## - FETUA.Bovine     1   165.72 239.72
## - LUM              1   167.18 241.18
## - SERPINA7         1   173.05 247.05
## - LAMP2            1   174.95 248.95
##
## Step:  AIC=236.76
## Class ~ AFM + AHSG + AIAG.Bovine + APOB + BTD + CD44 + CLU +
##     DKFZp686N02209 + DSG2 + F11 + F5 + FCGBP + FETUA.Bovine +
##     FETUB + FHR3 + FN1 + HP + HRG + IGFBP3 + IGHA2 + IGHG2 +
##     KNG1 + LAMP2 + LGALS3BP + LRG1 + LUM + MMRN1 + MST1 + NCAM1 +
##     PRG4 + PROC + Q5JNX2 + SERPINA7 + THBS1 + TIMP1 + VWF
##
##                  Df Deviance    AIC
## - FN1             1   162.78 234.78
## - Q5JNX2          1   162.79 234.79
## - AFM             1   162.80 234.80
## - PROC            1   162.81 234.81
## - CD44            1   162.81 234.81
## - BTD             1   162.87 234.87
## - DSG2            1   162.87 234.87
## - MST1            1   162.92 234.92
## - AHSG            1   162.96 234.96
## - LRG1            1   163.01 235.01
## - APOB            1   163.02 235.02
## - FETUB           1   163.05 235.05
## - F11             1   163.06 235.06
## - PRG4            1   163.09 235.09
## - HRG             1   163.10 235.10
## - TIMP1           1   163.17 235.17
## - AIAG.Bovine     1   163.25 235.25
## - IGFBP3          1   163.28 235.28
## - F5              1   163.29 235.29
## - THBS1           1   163.45 235.45
## - VWF             1   163.47 235.47
## - CLU             1   163.60 235.60
## - FCGBP           1   163.66 235.66
## - FHR3            1   163.90 235.90
## - IGHG2           1   163.95 235.95
```

```
## - LGALS3BP        1   163.97 235.97
## - KNG1            1   164.23 236.23
## - IGHA2           1   164.47 236.47
## - DKFZp686N02209  1   164.47 236.47
## <none>                162.76 236.76
## - NCAM1           1   165.04 237.04
## - MMRN1           1   165.08 237.08
## - HP              1   165.09 237.09
## - FETUA.Bovine    1   165.74 237.74
## - LUM             1   167.36 239.36
## - SERPINA7        1   173.16 245.16
## - LAMP2           1   175.10 247.10
##
## Step:  AIC=234.78
## Class ~ AFM + AHSG + AIAG.Bovine + APOB + BTD + CD44 + CLU +
##     DKFZp686N02209 + DSG2 + F11 + F5 + FCGBP + FETUA.Bovine +
##     FETUB + FHR3 + HP + HRG + IGFBP3 + IGHA2 + IGHG2 + KNG1 +
##     LAMP2 + LGALS3BP + LRG1 + LUM + MMRN1 + MST1 + NCAM1 + PRG4 +
##     PROC + Q5JNX2 + SERPINA7 + THBS1 + TIMP1 + VWF
##
##                  Df Deviance    AIC
## - AFM             1   162.81 232.81
## - Q5JNX2          1   162.81 232.81
## - CD44            1   162.81 232.81
## - PROC            1   162.82 232.82
## - BTD             1   162.87 232.87
## - DSG2            1   162.90 232.90
## - MST1            1   162.93 232.93
## - AHSG            1   163.00 233.00
## - APOB            1   163.04 233.04
## - LRG1            1   163.06 233.06
## - FETUB           1   163.06 233.06
## - F11             1   163.10 233.10
## - HRG             1   163.11 233.11
## - PRG4            1   163.14 233.14
## - TIMP1           1   163.17 233.17
## - AIAG.Bovine     1   163.25 233.25
## - IGFBP3          1   163.28 233.28
## - F5              1   163.38 233.38
## - THBS1           1   163.47 233.47
## - VWF             1   163.48 233.48
## - CLU             1   163.60 233.60
## - FCGBP           1   163.67 233.67
## - FHR3            1   163.95 233.95
## - LGALS3BP        1   163.97 233.97
## - IGHG2           1   164.03 234.03
## - KNG1            1   164.23 234.23
## - DKFZp686N02209  1   164.47 234.47
## - IGHA2           1   164.48 234.48
## <none>                162.78 234.78
```

```
## - HP              1   165.09 235.09
## - NCAM1           1   165.10 235.10
## - MMRN1           1   165.11 235.11
## - FETUA.Bovine    1   165.74 235.74
## - LUM             1   167.38 237.38
## - SERPINA7        1   173.17 243.17
## - LAMP2           1   175.39 245.39
##
## Step:  AIC=232.81
## Class ~ AHSG + AIAG.Bovine + APOB + BTD + CD44 + CLU +
DKFZp686N02209 +
##     DSG2 + F11 + F5 + FCGBP + FETUA.Bovine + FETUB + FHR3 + HP +
##     HRG + IGFBP3 + IGHA2 + IGHG2 + KNG1 + LAMP2 + LGALS3BP +
##     LRG1 + LUM + MMRN1 + MST1 + NCAM1 + PRG4 + PROC + Q5JNX2 +
##     SERPINA7 + THBS1 + TIMP1 + VWF
##
##                  Df Deviance    AIC
## - Q5JNX2          1   162.83 230.83
## - PROC            1   162.84 230.84
## - CD44            1   162.84 230.84
## - BTD             1   162.89 230.89
## - DSG2            1   162.93 230.93
## - MST1            1   162.96 230.96
## - APOB            1   163.05 231.05
## - FETUB           1   163.08 231.08
## - AHSG            1   163.09 231.09
## - F11             1   163.12 231.12
## - LRG1            1   163.12 231.12
## - HRG             1   163.14 231.14
## - PRG4            1   163.18 231.18
## - TIMP1           1   163.18 231.18
## - AIAG.Bovine     1   163.25 231.25
## - IGFBP3          1   163.34 231.34
## - F5              1   163.38 231.38
## - VWF             1   163.49 231.49
## - THBS1           1   163.51 231.51
## - CLU             1   163.61 231.61
## - FCGBP           1   163.68 231.68
## - LGALS3BP        1   163.97 231.97
## - FHR3            1   163.98 231.98
## - IGHG2           1   164.07 232.07
## - KNG1            1   164.25 232.25
## - IGHA2           1   164.56 232.56
## - DKFZp686N02209  1   164.68 232.68
## <none>                162.81 232.81
## - HP              1   165.13 233.13
## - NCAM1           1   165.15 233.15
## - MMRN1           1   165.28 233.28
## - FETUA.Bovine    1   165.76 233.76
## - LUM             1   167.54 235.54
```

```
## - SERPINA7        1   173.18 241.18
## - LAMP2           1   176.63 244.63
##
## Step:  AIC=230.83
## Class ~ AHSG + AIAG.Bovine + APOB + BTD + CD44 + CLU +
DKFZp686N02209 +
##     DSG2 + F11 + F5 + FCGBP + FETUA.Bovine + FETUB + FHR3 + HP +
##     HRG + IGFBP3 + IGHA2 + IGHG2 + KNG1 + LAMP2 + LGALS3BP +
##     LRG1 + LUM + MMRN1 + MST1 + NCAM1 + PRG4 + PROC + SERPINA7 +
##     THBS1 + TIMP1 + VWF
##
##                  Df Deviance    AIC
## - PROC            1   162.86 228.86
## - CD44            1   162.87 228.87
## - BTD             1   162.91 228.91
## - DSG2            1   162.97 228.97
## - MST1            1   162.98 228.98
## - APOB            1   163.09 229.09
## - AHSG            1   163.11 229.11
## - LRG1            1   163.12 229.12
## - FETUB           1   163.13 229.13
## - F11             1   163.18 229.18
## - TIMP1           1   163.19 229.19
## - HRG             1   163.19 229.19
## - PRG4            1   163.23 229.23
## - AIAG.Bovine     1   163.25 229.25
## - F5              1   163.38 229.38
## - IGFBP3          1   163.38 229.38
## - VWF             1   163.51 229.51
## - THBS1           1   163.51 229.51
## - CLU             1   163.61 229.61
## - FCGBP           1   163.68 229.68
## - FHR3            1   163.99 229.99
## - LGALS3BP        1   164.03 230.03
## - IGHG2           1   164.20 230.20
## - KNG1            1   164.25 230.25
## - IGHA2           1   164.56 230.56
## <none>                162.83 230.83
## - DKFZp686N02209  1   164.97 230.97
## - NCAM1           1   165.15 231.15
## - HP              1   165.16 231.16
## - MMRN1           1   165.50 231.50
## - FETUA.Bovine    1   165.95 231.95
## - LUM             1   167.55 233.55
## - SERPINA7        1   173.48 239.48
## - LAMP2           1   177.04 243.04
##
## Step:  AIC=228.86
## Class ~ AHSG + AIAG.Bovine + APOB + BTD + CD44 + CLU +
DKFZp686N02209 +
```

```
##     DSG2 + F11 + F5 + FCGBP + FETUA.Bovine + FETUB + FHR3 + HP +
##     HRG + IGFBP3 + IGHA2 + IGHG2 + KNG1 + LAMP2 + LGALS3BP +
##     LRG1 + LUM + MMRN1 + MST1 + NCAM1 + PRG4 + SERPINA7 + THBS1 +
##     TIMP1 + VWF
##
##                   Df Deviance    AIC
## - CD44             1   162.91 226.91
## - BTD              1   162.94 226.94
## - DSG2             1   163.00 227.00
## - MST1             1   163.01 227.01
## - AHSG             1   163.13 227.13
## - APOB             1   163.13 227.13
## - LRG1             1   163.16 227.16
## - FETUB            1   163.18 227.18
## - HRG              1   163.20 227.20
## - TIMP1            1   163.21 227.21
## - PRG4             1   163.23 227.23
## - F11              1   163.26 227.26
## - AIAG.Bovine      1   163.27 227.27
## - IGFBP3           1   163.38 227.38
## - F5               1   163.39 227.39
## - VWF              1   163.51 227.51
## - THBS1            1   163.53 227.53
## - CLU              1   163.61 227.61
## - FCGBP            1   163.71 227.71
## - FHR3             1   164.00 228.00
## - LGALS3BP         1   164.09 228.09
## - IGHG2            1   164.20 228.20
## - KNG1             1   164.27 228.27
## - IGHA2            1   164.65 228.65
## <none>                 162.86 228.86
## - NCAM1            1   165.17 229.17
## - DKFZp686N02209   1   165.17 229.17
## - HP               1   165.20 229.20
## - MMRN1            1   165.60 229.60
## - FETUA.Bovine     1   165.99 229.99
## - LUM              1   167.57 231.57
## - SERPINA7         1   173.65 237.65
## - LAMP2            1   177.17 241.17
##
## Step:  AIC=226.91
## Class ~ AHSG + AIAG.Bovine + APOB + BTD + CLU + DKFZp686N02209 +
##     DSG2 + F11 + F5 + FCGBP + FETUA.Bovine + FETUB + FHR3 + HP +
##     HRG + IGFBP3 + IGHA2 + IGHG2 + KNG1 + LAMP2 + LGALS3BP +
##     LRG1 + LUM + MMRN1 + MST1 + NCAM1 + PRG4 + SERPINA7 + THBS1 +
##     TIMP1 + VWF
##
##                   Df Deviance    AIC
## - BTD              1   162.95 224.95
## - DSG2             1   163.02 225.02
```

```
## - MST1             1   163.09 225.09
## - AHSG             1   163.15 225.15
## - LRG1             1   163.17 225.17
## - APOB             1   163.19 225.19
## - HRG              1   163.24 225.24
## - TIMP1            1   163.25 225.25
## - FETUB            1   163.26 225.26
## - AIAG.Bovine      1   163.28 225.28
## - PRG4             1   163.29 225.29
## - F11              1   163.31 225.31
## - IGFBP3           1   163.39 225.39
## - F5               1   163.48 225.48
## - VWF              1   163.52 225.52
## - THBS1            1   163.62 225.62
## - CLU              1   163.65 225.65
## - FCGBP            1   163.74 225.74
## - FHR3             1   164.09 226.09
## - LGALS3BP         1   164.22 226.22
## - IGHG2            1   164.32 226.32
## - KNG1             1   164.41 226.41
## - IGHA2            1   164.65 226.65
## <none>                 162.91 226.91
## - NCAM1            1   165.19 227.19
## - DKFZp686N02209   1   165.21 227.21
## - HP               1   165.28 227.28
## - MMRN1            1   165.95 227.95
## - FETUA.Bovine     1   166.24 228.24
## - LUM              1   167.57 229.57
## - SERPINA7         1   173.81 235.81
## - LAMP2            1   177.27 239.27
##
## Step:  AIC=224.95
## Class ~ AHSG + AIAG.Bovine + APOB + CLU + DKFZp686N02209 + DSG2 +
##     F11 + F5 + FCGBP + FETUA.Bovine + FETUB + FHR3 + HP + HRG +
##     IGFBP3 + IGHA2 + IGHG2 + KNG1 + LAMP2 + LGALS3BP + LRG1 +
##     LUM + MMRN1 + MST1 + NCAM1 + PRG4 + SERPINA7 + THBS1 + TIMP1 +
##     VWF
##
##                   Df Deviance    AIC
## - DSG2             1   163.05 223.05
## - MST1             1   163.15 223.15
## - AHSG             1   163.16 223.16
## - LRG1             1   163.19 223.19
## - TIMP1            1   163.25 223.25
## - APOB             1   163.26 223.26
## - HRG              1   163.28 223.28
## - PRG4             1   163.31 223.31
## - FETUB            1   163.32 223.32
## - AIAG.Bovine      1   163.32 223.32
## - IGFBP3           1   163.39 223.39
```

```
## - F11              1   163.41 223.41
## - F5               1   163.52 223.52
## - VWF              1   163.54 223.54
## - THBS1            1   163.64 223.64
## - FCGBP            1   163.76 223.76
## - CLU              1   163.77 223.77
## - FHR3             1   164.10 224.10
## - IGHG2            1   164.33 224.33
## - LGALS3BP         1   164.34 224.34
## - KNG1             1   164.43 224.43
## - IGHA2            1   164.71 224.71
## <none>                 162.95 224.95
## - DKFZp686N02209   1   165.23 225.23
## - NCAM1            1   165.27 225.27
## - HP               1   165.42 225.42
## - MMRN1            1   166.11 226.11
## - FETUA.Bovine     1   166.38 226.38
## - LUM              1   167.65 227.65
## - SERPINA7         1   173.85 233.85
## - LAMP2            1   177.28 237.28
##
## Step:  AIC=223.05
## Class ~ AHSG + AIAG.Bovine + APOB + CLU + DKFZp686N02209 + F11 +
##     F5 + FCGBP + FETUA.Bovine + FETUB + FHR3 + HP + HRG + IGFBP3 +
##     IGHA2 + IGHG2 + KNG1 + LAMP2 + LGALS3BP + LRG1 + LUM + MMRN1 +
##     MST1 + NCAM1 + PRG4 + SERPINA7 + THBS1 + TIMP1 + VWF
##
##                   Df Deviance    AIC
## - LRG1             1   163.26 221.26
## - AHSG             1   163.27 221.27
## - MST1             1   163.27 221.27
## - APOB             1   163.30 221.30
## - PRG4             1   163.38 221.38
## - TIMP1            1   163.39 221.39
## - F11              1   163.41 221.41
## - FETUB            1   163.41 221.41
## - IGFBP3           1   163.42 221.42
## - AIAG.Bovine      1   163.44 221.44
## - HRG              1   163.47 221.47
## - F5               1   163.56 221.56
## - VWF              1   163.78 221.78
## - FCGBP            1   163.80 221.80
## - THBS1            1   163.82 221.82
## - CLU              1   163.95 221.95
## - FHR3             1   164.21 222.21
## - IGHG2            1   164.36 222.36
## - LGALS3BP         1   164.47 222.47
## - KNG1             1   164.55 222.55
## - IGHA2            1   164.72 222.72
## <none>                 163.05 223.05
```

```
## - DKFZp686N02209  1   165.25 223.25
## - NCAM1           1   165.28 223.28
## - HP              1   165.68 223.68
## - MMRN1           1   166.16 224.16
## - FETUA.Bovine    1   166.40 224.40
## - LUM             1   167.81 225.81
## - SERPINA7        1   173.98 231.98
## - LAMP2           1   177.29 235.29
##
## Step:  AIC=221.26
## Class ~ AHSG + AIAG.Bovine + APOB + CLU + DKFZp686N02209 + F11 +
##     F5 + FCGBP + FETUA.Bovine + FETUB + FHR3 + HP + HRG + IGFBP3 +
##     IGHA2 + IGHG2 + KNG1 + LAMP2 + LGALS3BP + LUM + MMRN1 + MST1 +
##     NCAM1 + PRG4 + SERPINA7 + THBS1 + TIMP1 + VWF
##
##                  Df Deviance    AIC
## - APOB            1   163.48 219.48
## - AHSG            1   163.50 219.50
## - FETUB           1   163.50 219.50
## - MST1            1   163.56 219.56
## - PRG4            1   163.65 219.65
## - AIAG.Bovine     1   163.66 219.66
## - F11             1   163.69 219.69
## - TIMP1           1   163.73 219.73
## - HRG             1   163.76 219.76
## - IGFBP3          1   163.85 219.85
## - F5              1   163.91 219.91
## - FCGBP           1   164.23 220.23
## - CLU             1   164.25 220.25
## - VWF             1   164.26 220.26
## - THBS1           1   164.29 220.29
## - LGALS3BP        1   164.58 220.58
## - IGHG2           1   164.61 220.61
## - KNG1            1   164.79 220.79
## - FHR3            1   164.83 220.83
## - IGHA2           1   165.06 221.06
## <none>                163.26 221.26
## - NCAM1           1   165.34 221.34
## - DKFZp686N02209  1   165.53 221.53
## - HP              1   166.34 222.34
## - FETUA.Bovine    1   166.73 222.73
## - MMRN1           1   167.11 223.11
## - LUM             1   167.87 223.87
## - SERPINA7        1   174.37 230.37
## - LAMP2           1   177.29 233.29
##
## Step:  AIC=219.48
## Class ~ AHSG + AIAG.Bovine + CLU + DKFZp686N02209 + F11 + F5 +
##     FCGBP + FETUA.Bovine + FETUB + FHR3 + HP + HRG + IGFBP3 +
##     IGHA2 + IGHG2 + KNG1 + LAMP2 + LGALS3BP + LUM + MMRN1 + MST1 +
```

```
##     NCAM1 + PRG4 + SERPINA7 + THBS1 + TIMP1 + VWF
##
##                   Df Deviance    AIC
## - FETUB            1   163.72 217.72
## - AHSG             1   163.73 217.73
## - MST1             1   163.79 217.79
## - HRG              1   163.90 217.90
## - F11              1   163.91 217.91
## - F5               1   163.96 217.96
## - TIMP1            1   164.02 218.02
## - AIAG.Bovine      1   164.05 218.05
## - PRG4             1   164.08 218.08
## - IGFBP3           1   164.20 218.20
## - CLU              1   164.39 218.39
## - VWF              1   164.45 218.45
## - FCGBP            1   164.52 218.52
## - THBS1            1   164.59 218.59
## - LGALS3BP         1   164.72 218.72
## - IGHG2            1   164.86 218.86
## - KNG1             1   164.96 218.96
## - FHR3             1   165.16 219.16
## - IGHA2            1   165.19 219.19
## <none>                 163.48 219.48
## - NCAM1            1   165.49 219.49
## - DKFZp686N02209   1   165.63 219.63
## - HP               1   166.56 220.56
## - FETUA.Bovine     1   166.90 220.90
## - MMRN1            1   167.13 221.13
## - LUM              1   168.67 222.67
## - SERPINA7         1   174.87 228.87
## - LAMP2            1   177.58 231.58
##
## Step:  AIC=217.72
## Class ~ AHSG + AIAG.Bovine + CLU + DKFZp686N02209 + F11 + F5 +
##     FCGBP + FETUA.Bovine + FHR3 + HP + HRG + IGFBP3 + IGHA2 +
##     IGHG2 + KNG1 + LAMP2 + LGALS3BP + LUM + MMRN1 + MST1 + NCAM1 +
##     PRG4 + SERPINA7 + THBS1 + TIMP1 + VWF
##
##                   Df Deviance    AIC
## - MST1             1   164.01 216.01
## - AHSG             1   164.02 216.02
## - F11              1   164.10 216.10
## - HRG              1   164.10 216.10
## - F5               1   164.19 216.19
## - TIMP1            1   164.27 216.27
## - PRG4             1   164.40 216.40
## - AIAG.Bovine      1   164.44 216.44
## - IGFBP3           1   164.46 216.46
## - CLU              1   164.50 216.50
## - VWF              1   164.69 216.69
```

```
## - THBS1           1   164.78 216.78
## - FCGBP           1   164.79 216.79
## - LGALS3BP        1   164.97 216.97
## - IGHG2           1   165.00 217.00
## - KNG1            1   165.11 217.11
## - FHR3            1   165.31 217.31
## - IGHA2           1   165.47 217.47
## - NCAM1           1   165.69 217.69
## <none>                163.72 217.72
## - DKFZp686N02209  1   166.12 218.12
## - HP              1   166.83 218.83
## - FETUA.Bovine    1   167.10 219.10
## - MMRN1           1   167.19 219.19
## - LUM             1   168.74 220.74
## - SERPINA7        1   174.94 226.94
## - LAMP2           1   178.45 230.45
##
## Step:  AIC=216.01
## Class ~ AHSG + AIAG.Bovine + CLU + DKFZp686N02209 + F11 + F5 +
##     FCGBP + FETUA.Bovine + FHR3 + HP + HRG + IGFBP3 + IGHA2 +
##     IGHG2 + KNG1 + LAMP2 + LGALS3BP + LUM + MMRN1 + NCAM1 + PRG4 +
##     SERPINA7 + THBS1 + TIMP1 + VWF
##
##                  Df Deviance    AIC
## - AHSG            1   164.26 214.26
## - F11             1   164.33 214.33
## - TIMP1           1   164.44 214.44
## - F5              1   164.48 214.48
## - HRG             1   164.57 214.57
## - PRG4            1   164.65 214.65
## - AIAG.Bovine     1   164.72 214.72
## - CLU             1   164.82 214.82
## - IGFBP3          1   164.88 214.88
## - VWF             1   164.93 214.93
## - FCGBP           1   165.05 215.05
## - THBS1           1   165.10 215.10
## - LGALS3BP        1   165.20 215.20
## - KNG1            1   165.38 215.38
## - IGHG2           1   165.53 215.53
## - IGHA2           1   165.62 215.62
## - FHR3            1   165.89 215.89
## - NCAM1           1   165.89 215.89
## <none>                164.01 216.01
## - DKFZp686N02209  1   166.22 216.22
## - HP              1   167.13 217.13
## - MMRN1           1   167.41 217.41
## - FETUA.Bovine    1   167.48 217.48
## - LUM             1   169.14 219.14
## - SERPINA7        1   175.25 225.25
## - LAMP2           1   178.64 228.64
```

```
##
## Step:  AIC=214.26
## Class ~ AIAG.Bovine + CLU + DKFZp686N02209 + F11 + F5 + FCGBP +
##     FETUA.Bovine + FHR3 + HP + HRG + IGFBP3 + IGHA2 + IGHG2 +
##     KNG1 + LAMP2 + LGALS3BP + LUM + MMRN1 + NCAM1 + PRG4 + SERPINA7
+
##     THBS1 + TIMP1 + VWF
##
##                   Df Deviance    AIC
## - F11              1   164.55 212.55
## - TIMP1            1   164.71 212.71
## - F5               1   164.73 212.73
## - PRG4             1   164.81 212.81
## - AIAG.Bovine      1   164.90 212.90
## - HRG              1   164.97 212.97
## - CLU              1   165.01 213.01
## - IGFBP3           1   165.11 213.11
## - LGALS3BP         1   165.35 213.35
## - FCGBP            1   165.36 213.36
## - VWF              1   165.41 213.41
## - THBS1            1   165.51 213.51
## - KNG1             1   165.62 213.62
## - IGHG2            1   165.87 213.87
## - IGHA2            1   165.98 213.98
## - NCAM1            1   166.09 214.09
## - FHR3             1   166.13 214.13
## <none>                 164.26 214.26
## - DKFZp686N02209   1   166.56 214.56
## - HP               1   167.22 215.22
## - FETUA.Bovine     1   167.69 215.69
## - MMRN1            1   167.97 215.97
## - LUM              1   169.83 217.83
## - SERPINA7         1   175.33 223.33
## - LAMP2            1   179.01 227.01
##
## Step:  AIC=212.55
## Class ~ AIAG.Bovine + CLU + DKFZp686N02209 + F5 + FCGBP +
FETUA.Bovine +
##     FHR3 + HP + HRG + IGFBP3 + IGHA2 + IGHG2 + KNG1 + LAMP2 +
##     LGALS3BP + LUM + MMRN1 + NCAM1 + PRG4 + SERPINA7 + THBS1 +
##     TIMP1 + VWF
##
##                   Df Deviance    AIC
## - TIMP1            1   164.94 210.94
## - PRG4             1   165.06 211.06
## - HRG              1   165.15 211.15
## - F5               1   165.18 211.18
## - IGFBP3           1   165.26 211.26
## - AIAG.Bovine      1   165.35 211.35
## - CLU              1   165.45 211.45
```

```
## - LGALS3BP        1   165.47 211.47
## - FCGBP           1   165.51 211.51
## - THBS1           1   165.61 211.61
## - VWF             1   165.75 211.75
## - KNG1            1   165.84 211.84
## - IGHA2           1   166.16 212.16
## - IGHG2           1   166.24 212.24
## - NCAM1           1   166.54 212.54
## <none>                164.55 212.55
## - DKFZp686N02209  1   166.72 212.72
## - FHR3            1   167.05 213.05
## - FETUA.Bovine    1   167.97 213.97
## - HP              1   168.09 214.09
## - MMRN1           1   168.39 214.39
## - LUM             1   170.14 216.14
## - SERPINA7        1   175.74 221.74
## - LAMP2           1   179.10 225.10
##
## Step:  AIC=210.94
## Class ~ AIAG.Bovine + CLU + DKFZp686N02209 + F5 + FCGBP +
FETUA.Bovine +
##     FHR3 + HP + HRG + IGFBP3 + IGHA2 + IGHG2 + KNG1 + LAMP2 +
##     LGALS3BP + LUM + MMRN1 + NCAM1 + PRG4 + SERPINA7 + THBS1 +
##     VWF
##
##                  Df Deviance    AIC
## - F5              1   165.48 209.48
## - IGFBP3          1   165.55 209.55
## - THBS1           1   165.66 209.66
## - CLU             1   165.70 209.70
## - PRG4            1   165.71 209.71
## - HRG             1   165.72 209.72
## - FCGBP           1   165.90 209.90
## - LGALS3BP        1   165.92 209.92
## - AIAG.Bovine     1   165.93 209.93
## - VWF             1   165.93 209.93
## - KNG1            1   166.28 210.28
## - IGHA2           1   166.37 210.37
## - IGHG2           1   166.70 210.70
## <none>                164.94 210.94
## - NCAM1           1   166.99 210.99
## - DKFZp686N02209  1   167.30 211.30
## - FHR3            1   167.51 211.51
## - FETUA.Bovine    1   168.75 212.75
## - HP              1   169.23 213.23
## - LUM             1   170.30 214.30
## - MMRN1           1   172.70 216.70
## - SERPINA7        1   176.42 220.42
## - LAMP2           1   180.21 224.21
##
```

```
## Step:  AIC=209.48
## Class ~ AIAG.Bovine + CLU + DKFZp686N02209 + FCGBP + FETUA.Bovine +
##     FHR3 + HP + HRG + IGFBP3 + IGHA2 + IGHG2 + KNG1 + LAMP2 +
##     LGALS3BP + LUM + MMRN1 + NCAM1 + PRG4 + SERPINA7 + THBS1 +
##     VWF
##
##                   Df Deviance    AIC
## - IGFBP3           1   165.85 207.85
## - HRG              1   166.02 208.02
## - THBS1            1   166.08 208.08
## - AIAG.Bovine      1   166.24 208.24
## - CLU              1   166.29 208.29
## - VWF              1   166.35 208.35
## - FCGBP            1   166.50 208.50
## - LGALS3BP         1   166.63 208.63
## - IGHA2            1   166.69 208.69
## - KNG1             1   166.81 208.81
## - PRG4             1   166.83 208.83
## <none>                 165.48 209.48
## - IGHG2            1   167.58 209.58
## - DKFZp686N02209   1   167.68 209.68
## - NCAM1            1   167.98 209.98
## - FHR3             1   168.42 210.42
## - FETUA.Bovine     1   169.17 211.17
## - HP               1   169.68 211.68
## - LUM              1   171.48 213.48
## - MMRN1            1   172.93 214.93
## - SERPINA7         1   176.89 218.89
## - LAMP2            1   180.37 222.37
##
## Step:  AIC=207.85
## Class ~ AIAG.Bovine + CLU + DKFZp686N02209 + FCGBP + FETUA.Bovine +
##     FHR3 + HP + HRG + IGHA2 + IGHG2 + KNG1 + LAMP2 + LGALS3BP +
##     LUM + MMRN1 + NCAM1 + PRG4 + SERPINA7 + THBS1 + VWF
##
##                   Df Deviance    AIC
## - THBS1            1   166.45 206.45
## - HRG              1   166.47 206.47
## - VWF              1   166.56 206.56
## - AIAG.Bovine      1   166.66 206.66
## - CLU              1   166.67 206.67
## - FCGBP            1   166.79 206.79
## - IGHA2            1   166.97 206.97
## - LGALS3BP         1   166.98 206.98
## - PRG4             1   167.32 207.32
## - KNG1             1   167.62 207.62
## - DKFZp686N02209   1   167.72 207.72
## <none>                 165.85 207.85
## - NCAM1            1   168.13 208.13
## - IGHG2            1   168.34 208.34
```

```
## - FHR3            1   168.59 208.59
## - FETUA.Bovine    1   169.81 209.81
## - HP              1   169.94 209.94
## - LUM             1   171.51 211.51
## - MMRN1           1   173.24 213.24
## - SERPINA7        1   177.06 217.06
## - LAMP2           1   180.96 220.96
##
## Step:  AIC=206.45
## Class ~ AIAG.Bovine + CLU + DKFZp686N02209 + FCGBP + FETUA.Bovine +
##     FHR3 + HP + HRG + IGHA2 + IGHG2 + KNG1 + LAMP2 + LGALS3BP +
##     LUM + MMRN1 + NCAM1 + PRG4 + SERPINA7 + VWF
##
##                  Df Deviance    AIC
## - HRG             1   166.95 204.95
## - AIAG.Bovine     1   167.07 205.07
## - CLU             1   167.10 205.10
## - LGALS3BP        1   167.52 205.52
## - FCGBP           1   167.61 205.61
## - IGHA2           1   167.70 205.70
## - PRG4            1   167.91 205.91
## - VWF             1   167.96 205.96
## - KNG1            1   168.05 206.05
## - DKFZp686N02209  1   168.07 206.07
## <none>                166.45 206.45
## - NCAM1           1   168.63 206.63
## - IGHG2           1   168.76 206.76
## - FHR3            1   169.50 207.50
## - FETUA.Bovine    1   169.93 207.93
## - HP              1   170.16 208.16
## - LUM             1   171.87 209.87
## - MMRN1           1   173.25 211.25
## - SERPINA7        1   177.73 215.73
## - LAMP2           1   181.77 219.77
##
## Step:  AIC=204.95
## Class ~ AIAG.Bovine + CLU + DKFZp686N02209 + FCGBP + FETUA.Bovine +
##     FHR3 + HP + IGHA2 + IGHG2 + KNG1 + LAMP2 + LGALS3BP + LUM +
##     MMRN1 + NCAM1 + PRG4 + SERPINA7 + VWF
##
##                  Df Deviance    AIC
## - AIAG.Bovine     1   167.38 203.38
## - CLU             1   167.67 203.67
## - LGALS3BP        1   167.81 203.81
## - IGHA2           1   168.09 204.09
## - FCGBP           1   168.32 204.32
## - PRG4            1   168.38 204.38
## - VWF             1   168.42 204.42
## - KNG1            1   168.75 204.75
## - DKFZp686N02209  1   168.80 204.80
```

```
## - NCAM1           1   168.95 204.95
## <none>                166.95 204.95
## - IGHG2           1   169.53 205.53
## - FHR3            1   169.76 205.76
## - FETUA.Bovine    1   170.46 206.46
## - HP              1   170.78 206.78
## - LUM             1   173.86 209.86
## - MMRN1           1   175.43 211.43
## - SERPINA7        1   178.16 214.16
## - LAMP2           1   182.56 218.56
##
## Step:  AIC=203.38
## Class ~ CLU + DKFZp686N02209 + FCGBP + FETUA.Bovine + FHR3 +
##     HP + IGHA2 + IGHG2 + KNG1 + LAMP2 + LGALS3BP + LUM + MMRN1 +
##     NCAM1 + PRG4 + SERPINA7 + VWF
##
##                  Df Deviance    AIC
## - CLU             1   167.84 201.84
## - LGALS3BP        1   168.28 202.28
## - IGHA2           1   168.53 202.53
## - FCGBP           1   168.62 202.62
## - PRG4            1   168.67 202.67
## - VWF             1   168.68 202.68
## - DKFZp686N02209  1   168.98 202.98
## - KNG1            1   169.11 203.11
## <none>                167.38 203.38
## - NCAM1           1   169.40 203.40
## - FHR3            1   169.91 203.91
## - IGHG2           1   170.04 204.04
## - FETUA.Bovine    1   170.81 204.81
## - HP              1   171.16 205.16
## - LUM             1   175.29 209.29
## - MMRN1           1   175.55 209.55
## - SERPINA7        1   178.17 212.17
## - LAMP2           1   182.74 216.74
##
## Step:  AIC=201.84
## Class ~ DKFZp686N02209 + FCGBP + FETUA.Bovine + FHR3 + HP + IGHA2 +
##     IGHG2 + KNG1 + LAMP2 + LGALS3BP + LUM + MMRN1 + NCAM1 + PRG4 +
##     SERPINA7 + VWF
##
##                  Df Deviance    AIC
## - LGALS3BP        1   168.59 200.59
## - PRG4            1   168.75 200.75
## - VWF             1   168.79 200.79
## - FCGBP           1   169.00 201.00
## - IGHA2           1   169.18 201.18
## - KNG1            1   169.25 201.25
## - DKFZp686N02209  1   169.42 201.42
## <none>                167.84 201.84
```

```
## - FHR3            1   170.12 202.12
## - NCAM1           1   170.51 202.51
## - IGHG2           1   170.83 202.83
## - FETUA.Bovine    1   171.21 203.21
## - HP              1   171.56 203.56
## - LUM             1   175.50 207.50
## - MMRN1           1   175.86 207.86
## - SERPINA7        1   179.39 211.39
## - LAMP2           1   182.96 214.96
##
## Step:  AIC=200.59
## Class ~ DKFZp686N02209 + FCGBP + FETUA.Bovine + FHR3 + HP + IGHA2 +
##     IGHG2 + KNG1 + LAMP2 + LUM + MMRN1 + NCAM1 + PRG4 + SERPINA7 +
##     VWF
##
##                  Df Deviance    AIC
## - PRG4            1   169.20 199.20
## - VWF             1   169.29 199.29
## - IGHA2           1   169.52 199.52
## - DKFZp686N02209  1   170.10 200.10
## - KNG1            1   170.11 200.11
## - FCGBP           1   170.44 200.44
## <none>                168.59 200.59
## - NCAM1           1   170.79 200.79
## - IGHG2           1   171.56 201.56
## - FHR3            1   171.57 201.57
## - FETUA.Bovine    1   171.95 201.95
## - HP              1   172.75 202.75
## - LUM             1   176.34 206.34
## - MMRN1           1   177.59 207.59
## - SERPINA7        1   179.98 209.98
## - LAMP2           1   182.97 212.97
##
## Step:  AIC=199.2
## Class ~ DKFZp686N02209 + FCGBP + FETUA.Bovine + FHR3 + HP + IGHA2 +
##     IGHG2 + KNG1 + LAMP2 + LUM + MMRN1 + NCAM1 + SERPINA7 + VWF
##
##                  Df Deviance    AIC
## - VWF             1   169.83 197.83
## - IGHA2           1   170.08 198.08
## - DKFZp686N02209  1   170.49 198.49
## - KNG1            1   170.83 198.83
## <none>                169.20 199.20
## - NCAM1           1   171.28 199.28
## - FCGBP           1   171.36 199.36
## - IGHG2           1   172.06 200.06
## - FETUA.Bovine    1   172.16 200.16
## - FHR3            1   172.25 200.25
## - HP              1   173.15 201.15
## - MMRN1           1   178.13 206.13
```

```
## - LUM             1   178.78 206.78
## - SERPINA7        1   180.12 208.12
## - LAMP2           1   185.62 213.62
##
## Step:  AIC=197.83
## Class ~ DKFZp686N02209 + FCGBP + FETUA.Bovine + FHR3 + HP + IGHA2 +
##     IGHG2 + KNG1 + LAMP2 + LUM + MMRN1 + NCAM1 + SERPINA7
##
##                  Df Deviance    AIC
## - IGHA2           1   170.75 196.75
## - KNG1            1   171.25 197.25
## - DKFZp686N02209  1   171.44 197.44
## - NCAM1           1   171.49 197.49
## - FCGBP           1   171.56 197.56
## <none>                169.83 197.83
## - IGHG2           1   172.32 198.32
## - FHR3            1   172.50 198.50
## - FETUA.Bovine    1   172.51 198.51
## - HP              1   173.86 199.86
## - MMRN1           1   178.13 204.13
## - LUM             1   179.27 205.27
## - SERPINA7        1   180.14 206.14
## - LAMP2           1   187.37 213.37
##
## Step:  AIC=196.75
## Class ~ DKFZp686N02209 + FCGBP + FETUA.Bovine + FHR3 + HP + IGHG2 +
##     KNG1 + LAMP2 + LUM + MMRN1 + NCAM1 + SERPINA7
##
##                  Df Deviance    AIC
## - NCAM1           1   172.03 196.03
## - FCGBP           1   172.03 196.03
## - DKFZp686N02209  1   172.18 196.18
## - KNG1            1   172.75 196.75
## <none>                170.75 196.75
## - IGHG2           1   173.04 197.04
## - FETUA.Bovine    1   173.25 197.25
## - FHR3            1   174.10 198.10
## - HP              1   174.31 198.31
## - MMRN1           1   179.33 203.33
## - LUM             1   179.49 203.49
## - SERPINA7        1   181.11 205.11
## - LAMP2           1   188.15 212.15
##
## Step:  AIC=196.03
## Class ~ DKFZp686N02209 + FCGBP + FETUA.Bovine + FHR3 + HP + IGHG2 +
##     KNG1 + LAMP2 + LUM + MMRN1 + SERPINA7
##
##                  Df Deviance    AIC
## - DKFZp686N02209  1   172.97 194.97
## - KNG1            1   173.88 195.88
```

```
## <none>                172.03 196.03
## - IGHG2           1   174.07 196.07
## - FCGBP           1   174.61 196.61
## - FETUA.Bovine    1   174.72 196.72
## - FHR3            1   175.05 197.05
## - HP              1   175.70 197.70
## - LUM             1   179.92 201.92
## - MMRN1           1   181.53 203.53
## - SERPINA7        1   181.94 203.94
## - LAMP2           1   188.15 210.15
##
## Step:  AIC=194.97
## Class ~ FCGBP + FETUA.Bovine + FHR3 + HP + IGHG2 + KNG1 + LAMP2 +
##     LUM + MMRN1 + SERPINA7
##
##                Df Deviance    AIC
## - IGHG2         1   174.09 194.09
## - KNG1          1   174.52 194.52
## <none>              172.97 194.97
## - FCGBP         1   175.38 195.38
## - FETUA.Bovine  1   175.40 195.40
## - FHR3          1   175.49 195.49
## - HP            1   176.24 196.24
## - LUM           1   180.40 200.40
## - MMRN1         1   182.74 202.74
## - SERPINA7      1   184.85 204.85
## - LAMP2         1   191.38 211.38
##
## Step:  AIC=194.09
## Class ~ FCGBP + FETUA.Bovine + FHR3 + HP + KNG1 + LAMP2 + LUM +
##     MMRN1 + SERPINA7
##
##                Df Deviance    AIC
## <none>              174.09 194.09
## - KNG1          1   176.13 194.13
## - FETUA.Bovine  1   176.27 194.27
## - FCGBP         1   176.56 194.56
## - FHR3          1   176.71 194.71
## - HP            1   177.90 195.90
## - MMRN1         1   183.67 201.67
## - LUM           1   184.52 202.52
## - SERPINA7      1   184.99 202.99
## - LAMP2         1   191.38 209.38
```

```
summary(model.aic.backward)
```

```
##
## Call:
## glm(formula = Class ~ FCGBP + FETUA.Bovine + FHR3 + HP + KNG1 +
##     LAMP2 + LUM + MMRN1 + SERPINA7, family = binomial, data = train)
```

```
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -1.8946  -0.8822  -0.1143   0.9205   2.7140
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  122.1951    83.8466   1.457 0.145016
## FCGBP         -0.5803     0.3745  -1.550 0.121184
## FETUA.Bovine  -7.0967     4.8653  -1.459 0.144665
## FHR3          -0.2556     0.1604  -1.594 0.110989
## HP            -0.5438     0.2848  -1.910 0.056190 .
## KNG1           0.4819     0.3768   1.279 0.200999
## LAMP2          2.5943     0.6818   3.805 0.000142 ***
## LUM            1.3460     0.4363   3.085 0.002037 **
## MMRN1         -1.4466     0.4904  -2.950 0.003180 **
## SERPINA7      -2.1277     0.6782  -3.137 0.001705 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 221.81  on 159  degrees of freedom
## Residual deviance: 174.09  on 150  degrees of freedom
## AIC: 194.09
##
## Number of Fisher Scoring iterations: 4
```

This results in 9 columns in the best model with AIC = 194.09

AIC-based forward selection

```
model.aic.forward <- step(glm.null, direction = "forward", trace = 1,
scope = list(lower=glm.null, upper=glm.full))
```

```
## Start:  AIC=223.81
## Class ~ 1
##
##                  Df Deviance    AIC
## + TIMP1           1   209.97 213.97
## + MMRN1           1   211.66 215.66
## + HP              1   212.07 216.07
## + LRG1            1   213.06 217.06
## + FHR3            1   216.41 220.41
## + SERPINA7        1   217.81 221.81
## + LGALS3BP        1   218.56 222.56
## + CD44            1   218.72 222.72
## + LAMP2           1   218.96 222.96
## + FCGBP           1   219.20 223.20
## <none>                221.81 223.81
## + DSG2            1   219.81 223.81
```

```
## + F5                1   220.07 224.07
## + LYVE1             1   220.11 224.11
## + CFI               1   220.48 224.48
## + HRG               1   220.92 224.92
## + PRG4              1   220.94 224.94
## + NCAM1             1   221.06 225.06
## + VWF               1   221.13 225.13
## + BTD               1   221.22 225.22
## + FETUB             1   221.26 225.26
## + THBS1             1   221.28 225.28
## + PROC              1   221.35 225.35
## + F11               1   221.40 225.40
## + LUM               1   221.41 225.41
## + IGHG2             1   221.46 225.46
## + C20orf3           1   221.47 225.47
## + AIAG.Bovine       1   221.48 225.48
## + AFM               1   221.52 225.52
## + FN1               1   221.54 225.54
## + FETUA.Bovine      1   221.58 225.58
## + APOB              1   221.58 225.58
## + MST1              1   221.62 225.62
## + PLTP              1   221.63 225.63
## + DKFZp686N02209    1   221.69 225.69
## + Q5JNX2            1   221.71 225.71
## + IGHA2             1   221.74 225.74
## + KNG1              1   221.78 225.78
## + CLU               1   221.78 225.78
## + IGFBP3            1   221.80 225.80
## + AHSG              1   221.81 225.81
##
## Step:  AIC=213.97
## Class ~ TIMP1
##
##                   Df Deviance    AIC
## + LAMP2            1   202.29 208.29
## + HP               1   206.23 212.23
## + THBS1            1   206.98 212.98
## + LRG1             1   207.12 213.12
## + FHR3             1   207.18 213.18
## + AFM              1   207.72 213.72
## <none>                 209.97 213.97
## + SERPINA7         1   208.30 214.30
## + FN1              1   208.36 214.36
## + LUM              1   208.38 214.38
## + IGHA2            1   208.43 214.43
## + BTD              1   208.90 214.90
## + MMRN1            1   208.92 214.92
## + VWF              1   209.00 215.00
## + AHSG             1   209.13 215.13
## + PRG4             1   209.17 215.17
```

```
## + KNG1             1   209.19 215.19
## + APOB             1   209.23 215.23
## + CLU              1   209.34 215.34
## + CD44             1   209.38 215.38
## + C20orf3          1   209.59 215.59
## + HRG              1   209.60 215.60
## + F5               1   209.60 215.60
## + AIAG.Bovine      1   209.65 215.65
## + FCGBP            1   209.69 215.69
## + IGHG2            1   209.76 215.76
## + DSG2             1   209.78 215.78
## + MST1             1   209.78 215.78
## + FETUA.Bovine     1   209.83 215.83
## + LGALS3BP         1   209.84 215.84
## + DKFZp686N02209   1   209.84 215.84
## + Q5JNX2           1   209.90 215.90
## + F11              1   209.92 215.92
## + NCAM1            1   209.92 215.92
## + LYVE1            1   209.93 215.93
## + PROC             1   209.94 215.94
## + PLTP             1   209.95 215.95
## + CFI              1   209.95 215.95
## + IGFBP3           1   209.96 215.96
## + FETUB            1   209.96 215.96
##
## Step:  AIC=208.29
## Class ~ TIMP1 + LAMP2
##
##                  Df Deviance    AIC
## + HP              1   194.96 202.96
## + SERPINA7        1   195.46 203.46
## + LRG1            1   196.14 204.14
## + CFI             1   196.86 204.86
## + FHR3            1   198.25 206.25
## + MMRN1           1   199.49 207.49
## + LGALS3BP        1   199.91 207.91
## + F11             1   200.19 208.19
## + DSG2            1   200.21 208.21
## <none>                202.29 208.29
## + CD44            1   200.41 208.41
## + F5              1   200.55 208.55
## + FETUB           1   200.68 208.68
## + FCGBP           1   200.81 208.81
## + THBS1           1   201.34 209.34
## + NCAM1           1   201.48 209.48
## + IGHA2           1   201.51 209.51
## + PROC            1   201.73 209.73
## + Q5JNX2          1   201.74 209.74
## + LYVE1           1   201.80 209.80
## + AIAG.Bovine     1   201.92 209.92
```

```
## + C20orf3          1   201.93 209.93
## + CLU              1   202.00 210.00
## + LUM              1   202.06 210.06
## + IGHG2            1   202.07 210.07
## + PLTP             1   202.08 210.08
## + HRG              1   202.09 210.09
## + AHSG             1   202.10 210.10
## + BTD              1   202.10 210.10
## + IGFBP3           1   202.19 210.19
## + FETUA.Bovine     1   202.19 210.19
## + MST1             1   202.24 210.24
## + DKFZp686N02209   1   202.24 210.24
## + PRG4             1   202.27 210.27
## + AFM              1   202.28 210.28
## + VWF              1   202.28 210.28
## + KNG1             1   202.29 210.29
## + APOB             1   202.29 210.29
## + FN1              1   202.29 210.29
##
## Step:  AIC=202.96
## Class ~ TIMP1 + LAMP2 + HP
##
##                  Df Deviance    AIC
## + LRG1            1   191.82 201.82
## + SERPINA7        1   191.88 201.88
## + CFI             1   192.65 202.65
## + FHR3            1   192.70 202.70
## + MMRN1           1   192.82 202.82
## <none>                194.96 202.96
## + DSG2            1   193.23 203.23
## + FCGBP           1   193.23 203.23
## + F5              1   193.47 203.47
## + CD44            1   193.59 203.59
## + LUM             1   193.75 203.75
## + LGALS3BP        1   193.78 203.78
## + IGHA2           1   194.15 204.15
## + NCAM1           1   194.23 204.23
## + THBS1           1   194.24 204.24
## + IGHG2           1   194.34 204.34
## + FETUB           1   194.52 204.52
## + F11             1   194.61 204.61
## + LYVE1           1   194.64 204.64
## + KNG1            1   194.71 204.71
## + MST1            1   194.71 204.71
## + PRG4            1   194.73 204.73
## + FETUA.Bovine    1   194.74 204.74
## + C20orf3         1   194.75 204.75
## + FN1             1   194.79 204.79
## + VWF             1   194.86 204.86
## + Q5JNX2          1   194.86 204.86
```

```
## + BTD              1   194.87 204.87
## + APOB             1   194.87 204.87
## + AHSG             1   194.92 204.92
## + HRG              1   194.94 204.94
## + PROC             1   194.95 204.95
## + AFM              1   194.95 204.95
## + CLU              1   194.95 204.95
## + IGFBP3           1   194.95 204.95
## + PLTP             1   194.95 204.95
## + AIAG.Bovine      1   194.95 204.95
## + DKFZp686N02209   1   194.96 204.96
##
## Step:  AIC=201.82
## Class ~ TIMP1 + LAMP2 + HP + LRG1
##
##                   Df Deviance    AIC
## + SERPINA7         1   189.22 201.22
## + LUM              1   189.76 201.76
## <none>                 191.82 201.82
## + CFI              1   190.34 202.34
## + FCGBP            1   190.61 202.61
## + LGALS3BP         1   190.64 202.64
## + F5               1   190.65 202.65
## + DSG2             1   190.65 202.65
## + CD44             1   190.69 202.69
## + MMRN1            1   190.70 202.70
## + FHR3             1   190.78 202.78
## + NCAM1            1   190.93 202.93
## + IGHG2            1   191.24 203.24
## + IGHA2            1   191.36 203.36
## + FETUA.Bovine     1   191.46 203.46
## + THBS1            1   191.52 203.52
## + APOB             1   191.52 203.52
## + KNG1             1   191.52 203.52
## + PRG4             1   191.54 203.54
## + F11              1   191.66 203.66
## + AFM              1   191.69 203.69
## + C20orf3          1   191.69 203.69
## + AIAG.Bovine      1   191.69 203.69
## + VWF              1   191.70 203.70
## + PROC             1   191.76 203.76
## + MST1             1   191.76 203.76
## + LYVE1            1   191.77 203.77
## + CLU              1   191.78 203.78
## + IGFBP3           1   191.81 203.81
## + AHSG             1   191.81 203.81
## + Q5JNX2           1   191.81 203.81
## + PLTP             1   191.81 203.81
## + BTD              1   191.82 203.82
## + FETUB            1   191.82 203.82
```

```
## + DKFZp686N02209  1   191.82 203.82
## + FN1             1   191.82 203.82
## + HRG             1   191.82 203.82
##
## Step:  AIC=201.23
## Class ~ TIMP1 + LAMP2 + HP + LRG1 + SERPINA7
##
##                  Df Deviance    AIC
## + LUM             1   180.75 194.75
## + IGHG2           1   186.45 200.45
## <none>                189.22 201.22
## + AIAG.Bovine     1   187.32 201.32
## + APOB            1   187.59 201.59
## + PRG4            1   187.69 201.69
## + FHR3            1   187.84 201.84
## + KNG1            1   187.87 201.87
## + F5              1   187.97 201.97
## + CLU             1   188.08 202.08
## + CD44            1   188.11 202.11
## + CFI             1   188.16 202.16
## + NCAM1           1   188.24 202.24
## + MST1            1   188.24 202.24
## + LGALS3BP        1   188.24 202.24
## + DKFZp686N02209  1   188.31 202.31
## + HRG             1   188.38 202.38
## + FCGBP           1   188.39 202.39
## + BTD             1   188.42 202.42
## + DSG2            1   188.43 202.43
## + FETUA.Bovine    1   188.53 202.53
## + MMRN1           1   188.68 202.68
## + PLTP            1   188.71 202.71
## + LYVE1           1   188.72 202.72
## + AHSG            1   188.75 202.75
## + IGHA2           1   188.80 202.80
## + PROC            1   188.83 202.83
## + IGFBP3          1   188.86 202.86
## + THBS1           1   189.03 203.03
## + FETUB           1   189.04 203.04
## + VWF             1   189.09 203.09
## + FN1             1   189.10 203.10
## + F11             1   189.10 203.10
## + Q5JNX2          1   189.11 203.11
## + AFM             1   189.19 203.19
## + C20orf3         1   189.22 203.22
##
## Step:  AIC=194.75
## Class ~ TIMP1 + LAMP2 + HP + LRG1 + SERPINA7 + LUM
##
##                  Df Deviance    AIC
## + NCAM1           1   178.47 194.47
```

```
## + MMRN1           1   178.72 194.72
## <none>                180.75 194.75
## + FHR3            1   179.01 195.01
## + FCGBP           1   179.15 195.15
## + LGALS3BP        1   179.46 195.46
## + CD44            1   179.81 195.81
## + IGHA2           1   179.98 195.98
## + FETUA.Bovine    1   179.98 195.98
## + DSG2            1   180.10 196.10
## + KNG1            1   180.21 196.21
## + CFI             1   180.23 196.23
## + F5              1   180.28 196.28
## + FETUB           1   180.40 196.40
## + IGHG2           1   180.41 196.41
## + C20orf3         1   180.42 196.42
## + THBS1           1   180.52 196.52
## + AHSG            1   180.61 196.61
## + BTD             1   180.63 196.63
## + VWF             1   180.64 196.64
## + CLU             1   180.65 196.65
## + PRG4            1   180.66 196.66
## + LYVE1           1   180.66 196.66
## + F11             1   180.66 196.66
## + HRG             1   180.66 196.66
## + MST1            1   180.68 196.68
## + Q5JNX2          1   180.69 196.69
## + IGFBP3          1   180.72 196.72
## + AFM             1   180.72 196.72
## + PLTP            1   180.73 196.73
## + AIAG.Bovine     1   180.73 196.73
## + FN1             1   180.74 196.74
## + APOB            1   180.75 196.75
## + PROC            1   180.75 196.75
## + DKFZp686N02209  1   180.75 196.75
##
## Step:  AIC=194.47
## Class ~ TIMP1 + LAMP2 + HP + LRG1 + SERPINA7 + LUM + NCAM1
##
##                  Df Deviance    AIC
## <none>               178.47 194.47
## + MMRN1           1   176.70 194.70
## + FHR3            1   176.86 194.86
## + LGALS3BP        1   177.08 195.08
## + IGHA2           1   177.10 195.10
## + FETUA.Bovine    1   177.71 195.71
## + FETUB           1   177.75 195.75
## + FCGBP           1   177.78 195.78
## + KNG1            1   177.82 195.82
## + AHSG            1   178.07 196.07
## + THBS1           1   178.08 196.08
```

```
## + CD44             1   178.09 196.09
## + CFI              1   178.12 196.12
## + IGHG2            1   178.24 196.24
## + C20orf3          1   178.25 196.25
## + HRG              1   178.32 196.32
## + LYVE1            1   178.36 196.36
## + PLTP             1   178.37 196.37
## + DKFZp686N02209   1   178.37 196.37
## + APOB             1   178.40 196.40
## + Q5JNX2           1   178.40 196.40
## + AIAG.Bovine      1   178.40 196.40
## + DSG2             1   178.41 196.41
## + PRG4             1   178.42 196.42
## + F5               1   178.42 196.42
## + MST1             1   178.42 196.42
## + IGFBP3           1   178.43 196.43
## + PROC             1   178.44 196.44
## + VWF              1   178.46 196.46
## + CLU              1   178.46 196.46
## + BTD              1   178.46 196.46
## + AFM              1   178.47 196.47
## + FN1              1   178.47 196.47
## + F11              1   178.47 196.47
```

```
summary(model.aic.forward)
```

```
##
## Call:
## glm(formula = Class ~ TIMP1 + LAMP2 + HP + LRG1 + SERPINA7 +
##     LUM + NCAM1, family = binomial, data = train)
##
## Deviance Residuals:
##      Min        1Q    Median        3Q       Max
## -2.02726  -0.94320  -0.05077   0.93724   2.48858
##
## Coefficients:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)   7.3902     8.2991   0.890  0.37320
## TIMP1        -0.9190     0.4498  -2.043  0.04102 *
## LAMP2         2.7202     0.6946   3.916 8.99e-05 ***
## HP           -0.3650     0.2814  -1.297  0.19463
## LRG1         -0.7993     0.3743  -2.135  0.03273 *
## SERPINA7     -1.9639     0.6460  -3.040  0.00237 **
## LUM           1.2721     0.4218   3.016  0.00256 **
## NCAM1        -0.8055     0.5401  -1.491  0.13585
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
```

```
##     Null deviance: 221.81  on 159  degrees of freedom
## Residual deviance: 178.47  on 152  degrees of freedom
## AIC: 194.47
##
## Number of Fisher Scoring iterations: 4
```

This results in 7 columns in the best model with AIC = 194.47

AIC-based forward-backward selection

```
model.aic.both <- step(glm.null, direction = "both", trace = 1,
scope=list(lower=glm.null, upper=glm.full))

## Start:  AIC=223.81
## Class ~ 1
##
##                  Df Deviance    AIC
## + TIMP1           1   209.97 213.97
## + MMRN1           1   211.66 215.66
## + HP              1   212.07 216.07
## + LRG1            1   213.06 217.06
## + FHR3            1   216.41 220.41
## + SERPINA7        1   217.81 221.81
## + LGALS3BP        1   218.56 222.56
## + CD44            1   218.72 222.72
## + LAMP2           1   218.96 222.96
## + FCGBP           1   219.20 223.20
## <none>                221.81 223.81
## + DSG2            1   219.81 223.81
## + F5              1   220.07 224.07
## + LYVE1           1   220.11 224.11
## + CFI             1   220.48 224.48
## + HRG             1   220.92 224.92
## + PRG4            1   220.94 224.94
## + NCAM1           1   221.06 225.06
## + VWF             1   221.13 225.13
## + BTD             1   221.22 225.22
## + FETUB           1   221.26 225.26
## + THBS1           1   221.28 225.28
## + PROC            1   221.35 225.35
## + F11             1   221.40 225.40
## + LUM             1   221.41 225.41
## + IGHG2           1   221.46 225.46
## + C20orf3         1   221.47 225.47
## + AIAG.Bovine     1   221.48 225.48
## + AFM             1   221.52 225.52
## + FN1             1   221.54 225.54
## + FETUA.Bovine    1   221.58 225.58
## + APOB            1   221.58 225.58
## + MST1            1   221.62 225.62
## + PLTP            1   221.63 225.63
```

```
## + DKFZp686N02209  1   221.69 225.69
## + Q5JNX2          1   221.71 225.71
## + IGHA2           1   221.74 225.74
## + KNG1            1   221.78 225.78
## + CLU             1   221.78 225.78
## + IGFBP3          1   221.80 225.80
## + AHSG            1   221.81 225.81
##
## Step:  AIC=213.97
## Class ~ TIMP1
##
##                  Df Deviance    AIC
## + LAMP2           1   202.29 208.29
## + HP              1   206.23 212.23
## + THBS1           1   206.98 212.98
## + LRG1            1   207.12 213.12
## + FHR3            1   207.18 213.18
## + AFM             1   207.72 213.72
## <none>                209.97 213.97
## + SERPINA7        1   208.30 214.30
## + FN1             1   208.36 214.36
## + LUM             1   208.38 214.38
## + IGHA2           1   208.43 214.43
## + BTD             1   208.90 214.90
## + MMRN1           1   208.92 214.92
## + VWF             1   209.00 215.00
## + AHSG            1   209.13 215.13
## + PRG4            1   209.17 215.17
## + KNG1            1   209.19 215.19
## + APOB            1   209.23 215.23
## + CLU             1   209.34 215.34
## + CD44            1   209.38 215.38
## + C20orf3         1   209.59 215.59
## + HRG             1   209.60 215.60
## + F5              1   209.60 215.60
## + AIAG.Bovine     1   209.65 215.65
## + FCGBP           1   209.69 215.69
## + IGHG2           1   209.76 215.76
## + DSG2            1   209.78 215.78
## + MST1            1   209.78 215.78
## + FETUA.Bovine    1   209.83 215.83
## + LGALS3BP        1   209.84 215.84
## + DKFZp686N02209  1   209.84 215.84
## + Q5JNX2          1   209.90 215.90
## + F11             1   209.92 215.92
## + NCAM1           1   209.92 215.92
## + LYVE1           1   209.93 215.93
## + PROC            1   209.94 215.94
## + PLTP            1   209.95 215.95
## + CFI             1   209.95 215.95
```

```
## + IGFBP3          1   209.96 215.96
## + FETUB           1   209.96 215.96
## - TIMP1           1   221.81 223.81
##
## Step:  AIC=208.29
## Class ~ TIMP1 + LAMP2
##
##                  Df Deviance    AIC
## + HP              1   194.96 202.96
## + SERPINA7        1   195.46 203.46
## + LRG1            1   196.14 204.14
## + CFI             1   196.86 204.86
## + FHR3            1   198.25 206.25
## + MMRN1           1   199.49 207.49
## + LGALS3BP        1   199.91 207.91
## + F11             1   200.19 208.19
## + DSG2            1   200.21 208.21
## <none>                202.29 208.29
## + CD44            1   200.41 208.41
## + F5              1   200.55 208.55
## + FETUB           1   200.68 208.68
## + FCGBP           1   200.81 208.81
## + THBS1           1   201.34 209.34
## + NCAM1           1   201.48 209.48
## + IGHA2           1   201.51 209.51
## + PROC            1   201.73 209.73
## + Q5JNX2          1   201.74 209.74
## + LYVE1           1   201.80 209.80
## + AIAG.Bovine     1   201.92 209.92
## + C20orf3         1   201.93 209.93
## + CLU             1   202.00 210.00
## + LUM             1   202.06 210.06
## + IGHG2           1   202.07 210.07
## + PLTP            1   202.08 210.08
## + HRG             1   202.09 210.09
## + AHSG            1   202.10 210.10
## + BTD             1   202.10 210.10
## + IGFBP3          1   202.19 210.19
## + FETUA.Bovine    1   202.19 210.19
## + MST1            1   202.24 210.24
## + DKFZp686N02209  1   202.24 210.24
## + PRG4            1   202.27 210.27
## + AFM             1   202.28 210.28
## + VWF             1   202.28 210.28
## + KNG1            1   202.29 210.29
## + APOB            1   202.29 210.29
## + FN1             1   202.29 210.29
## - LAMP2           1   209.97 213.97
## - TIMP1           1   218.96 222.96
##
```

```
## Step:  AIC=202.96
## Class ~ TIMP1 + LAMP2 + HP
##
##                   Df Deviance    AIC
## + LRG1             1   191.82 201.82
## + SERPINA7         1   191.88 201.88
## + CFI              1   192.65 202.65
## + FHR3             1   192.70 202.70
## + MMRN1            1   192.82 202.82
## <none>                 194.96 202.96
## + DSG2             1   193.23 203.23
## + FCGBP            1   193.23 203.23
## + F5               1   193.47 203.47
## + CD44             1   193.59 203.59
## + LUM              1   193.75 203.75
## + LGALS3BP         1   193.78 203.78
## + IGHA2            1   194.15 204.15
## + NCAM1            1   194.23 204.23
## + THBS1            1   194.24 204.24
## + IGHG2            1   194.34 204.34
## + FETUB            1   194.52 204.52
## + F11              1   194.61 204.61
## + LYVE1            1   194.64 204.64
## + KNG1             1   194.71 204.71
## + MST1             1   194.71 204.71
## + PRG4             1   194.73 204.73
## + FETUA.Bovine     1   194.74 204.74
## + C20orf3          1   194.75 204.75
## + FN1              1   194.79 204.79
## + VWF              1   194.86 204.86
## + Q5JNX2           1   194.86 204.86
## + BTD              1   194.87 204.87
## + APOB             1   194.87 204.87
## + AHSG             1   194.92 204.92
## + HRG              1   194.94 204.94
## + PROC             1   194.95 204.95
## + AFM              1   194.95 204.95
## + CLU              1   194.95 204.95
## + IGFBP3           1   194.95 204.95
## + PLTP             1   194.95 204.95
## + AIAG.Bovine      1   194.95 204.95
## + DKFZp686N02209   1   194.96 204.96
## - HP               1   202.29 208.29
## - TIMP1            1   203.53 209.53
## - LAMP2            1   206.23 212.23
##
## Step:  AIC=201.82
## Class ~ TIMP1 + LAMP2 + HP + LRG1
##
##                   Df Deviance    AIC
```

```
## + SERPINA7         1   189.22 201.22
## + LUM              1   189.76 201.76
## <none>                 191.82 201.82
## + CFI              1   190.34 202.34
## + FCGBP            1   190.61 202.61
## + LGALS3BP         1   190.64 202.64
## + F5               1   190.65 202.65
## + DSG2             1   190.65 202.65
## + CD44             1   190.69 202.69
## + MMRN1            1   190.70 202.70
## + FHR3             1   190.78 202.78
## + NCAM1            1   190.93 202.93
## - LRG1             1   194.96 202.96
## + IGHG2            1   191.24 203.24
## + IGHA2            1   191.36 203.36
## + FETUA.Bovine     1   191.46 203.46
## + THBS1            1   191.52 203.52
## + APOB             1   191.52 203.52
## + KNG1             1   191.52 203.52
## + PRG4             1   191.54 203.54
## + F11              1   191.66 203.66
## + AFM              1   191.69 203.69
## + C20orf3          1   191.69 203.69
## + AIAG.Bovine      1   191.69 203.69
## + VWF              1   191.70 203.70
## + PROC             1   191.76 203.76
## + MST1             1   191.76 203.76
## + LYVE1            1   191.77 203.77
## + CLU              1   191.78 203.78
## + IGFBP3           1   191.81 203.81
## + AHSG             1   191.81 203.81
## + Q5JNX2           1   191.81 203.81
## + PLTP             1   191.81 203.81
## + BTD              1   191.82 203.82
## + FETUB            1   191.82 203.82
## + DKFZp686N02209   1   191.82 203.82
## + FN1              1   191.82 203.82
## + HRG              1   191.82 203.82
## - HP               1   196.14 204.14
## - TIMP1            1   197.71 205.71
## - LAMP2            1   205.03 213.03
##
## Step:  AIC=201.23
## Class ~ TIMP1 + LAMP2 + HP + LRG1 + SERPINA7
##
##                  Df Deviance    AIC
## + LUM             1   180.75 194.75
## + IGHG2           1   186.45 200.45
## <none>                189.22 201.22
## - HP              1   191.30 201.30
```

```
## + AIAG.Bovine     1   187.32 201.32
## + APOB            1   187.59 201.59
## + PRG4            1   187.69 201.69
## - SERPINA7        1   191.82 201.82
## + FHR3            1   187.84 201.84
## + KNG1            1   187.87 201.87
## - LRG1            1   191.88 201.88
## + F5              1   187.97 201.97
## + CLU             1   188.08 202.08
## + CD44            1   188.11 202.11
## + CFI             1   188.16 202.16
## + NCAM1           1   188.24 202.24
## + MST1            1   188.24 202.24
## + LGALS3BP        1   188.24 202.24
## + DKFZp686N02209  1   188.31 202.31
## + HRG             1   188.38 202.38
## + FCGBP           1   188.39 202.39
## + BTD             1   188.42 202.42
## + DSG2            1   188.43 202.43
## + FETUA.Bovine    1   188.53 202.53
## + MMRN1           1   188.68 202.68
## + PLTP            1   188.71 202.71
## + LYVE1           1   188.72 202.72
## + AHSG            1   188.75 202.75
## + IGHA2           1   188.80 202.80
## + PROC            1   188.83 202.83
## + IGFBP3          1   188.86 202.86
## + THBS1           1   189.03 203.03
## + FETUB           1   189.04 203.04
## + VWF             1   189.09 203.09
## + FN1             1   189.10 203.10
## + F11             1   189.10 203.10
## + Q5JNX2          1   189.11 203.11
## + AFM             1   189.19 203.19
## + C20orf3         1   189.22 203.22
## - TIMP1           1   195.38 205.38
## - LAMP2           1   204.91 214.91
##
## Step:  AIC=194.75
## Class ~ TIMP1 + LAMP2 + HP + LRG1 + SERPINA7 + LUM
##
##                  Df Deviance    AIC
## + NCAM1           1   178.47 194.47
## - HP              1   182.59 194.59
## + MMRN1           1   178.72 194.72
## <none>                180.75 194.75
## + FHR3            1   179.01 195.01
## + FCGBP           1   179.15 195.15
## + LGALS3BP        1   179.46 195.46
## + CD44            1   179.81 195.81
```

```
## + IGHA2           1   179.98 195.98
## + FETUA.Bovine    1   179.98 195.98
## + DSG2            1   180.10 196.10
## + KNG1            1   180.21 196.21
## + CFI             1   180.23 196.23
## + F5              1   180.28 196.28
## + FETUB           1   180.40 196.40
## + IGHG2           1   180.41 196.41
## + C20orf3         1   180.42 196.42
## + THBS1           1   180.52 196.52
## + AHSG            1   180.61 196.61
## + BTD             1   180.63 196.63
## + VWF             1   180.64 196.64
## + CLU             1   180.65 196.65
## + PRG4            1   180.66 196.66
## + LYVE1           1   180.66 196.66
## + F11             1   180.66 196.66
## + HRG             1   180.66 196.66
## + MST1            1   180.68 196.68
## + Q5JNX2          1   180.69 196.69
## + IGFBP3          1   180.72 196.72
## + AFM             1   180.72 196.72
## + PLTP            1   180.73 196.73
## + AIAG.Bovine     1   180.73 196.73
## + FN1             1   180.74 196.74
## + APOB            1   180.75 196.75
## + PROC            1   180.75 196.75
## + DKFZp686N02209  1   180.75 196.75
## - LRG1            1   185.13 197.13
## - TIMP1           1   187.04 199.04
## - LUM             1   189.22 201.22
## - SERPINA7        1   189.76 201.76
## - LAMP2           1   196.66 208.66
##
## Step:  AIC=194.47
## Class ~ TIMP1 + LAMP2 + HP + LRG1 + SERPINA7 + LUM + NCAM1
##
##                  Df Deviance    AIC
## - HP              1   180.19 194.19
## <none>                178.47 194.47
## + MMRN1           1   176.70 194.70
## - NCAM1           1   180.75 194.75
## + FHR3            1   176.86 194.86
## + LGALS3BP        1   177.08 195.08
## + IGHA2           1   177.10 195.10
## + FETUA.Bovine    1   177.71 195.71
## + FETUB           1   177.75 195.75
## + FCGBP           1   177.78 195.78
## + KNG1            1   177.82 195.82
## + AHSG            1   178.07 196.07
```

```
## + THBS1           1   178.08 196.08
## + CD44            1   178.09 196.09
## + CFI             1   178.12 196.12
## + IGHG2           1   178.24 196.24
## + C20orf3         1   178.25 196.25
## + HRG             1   178.32 196.32
## + LYVE1           1   178.36 196.36
## + PLTP            1   178.37 196.37
## + DKFZp686N02209  1   178.37 196.37
## + APOB            1   178.40 196.40
## + Q5JNX2          1   178.40 196.40
## + AIAG.Bovine     1   178.40 196.40
## + DSG2            1   178.41 196.41
## + PRG4            1   178.42 196.42
## + F5              1   178.42 196.42
## + MST1            1   178.42 196.42
## + IGFBP3          1   178.43 196.43
## + PROC            1   178.44 196.44
## + VWF             1   178.46 196.46
## + CLU             1   178.46 196.46
## + BTD             1   178.46 196.46
## + AFM             1   178.47 196.47
## + FN1             1   178.47 196.47
## + F11             1   178.47 196.47
## - TIMP1           1   182.78 196.78
## - LRG1            1   183.20 197.20
## - LUM             1   188.24 202.24
## - SERPINA7        1   188.50 202.50
## - LAMP2           1   196.66 210.66
##
## Step:  AIC=194.19
## Class ~ TIMP1 + LAMP2 + LRG1 + SERPINA7 + LUM + NCAM1
##
##                  Df Deviance    AIC
## + FHR3            1   178.19 194.19
## <none>                180.19 194.19
## + LGALS3BP        1   178.45 194.45
## + HP              1   178.47 194.47
## - NCAM1           1   182.59 194.59
## + MMRN1           1   178.73 194.73
## + IGHA2           1   178.99 194.99
## + CFI             1   179.41 195.41
## + FETUB           1   179.50 195.50
## + FETUA.Bovine    1   179.50 195.50
## + KNG1            1   179.54 195.54
## + CD44            1   179.74 195.74
## + FCGBP           1   179.76 195.76
## + THBS1           1   179.86 195.86
## + IGHG2           1   179.88 195.88
## + C20orf3         1   179.92 195.92
```

```
## + AHSG            1   179.99 195.99
## + APOB            1   180.08 196.08
## + F11             1   180.09 196.09
## + BTD             1   180.10 196.10
## + PLTP            1   180.10 196.10
## + AIAG.Bovine     1   180.10 196.10
## + PROC            1   180.11 196.11
## + HRG             1   180.13 196.13
## + MST1            1   180.14 196.14
## + FN1             1   180.14 196.14
## + F5              1   180.14 196.14
## + VWF             1   180.15 196.15
## + IGFBP3          1   180.15 196.15
## + DSG2            1   180.16 196.16
## + DKFZp686N02209  1   180.16 196.16
## + CLU             1   180.17 196.17
## + AFM             1   180.18 196.18
## + LYVE1           1   180.18 196.18
## + PRG4            1   180.18 196.18
## + Q5JNX2          1   180.19 196.19
## - TIMP1           1   186.01 198.01
## - LRG1            1   186.75 198.75
## - LUM             1   190.23 202.23
## - SERPINA7        1   193.65 205.65
## - LAMP2           1   197.74 209.74
##
## Step:  AIC=194.19
## Class ~ TIMP1 + LAMP2 + LRG1 + SERPINA7 + LUM + NCAM1 + FHR3
##
##                 Df Deviance    AIC
## <none>               178.19 194.19
## - FHR3            1   180.19 194.19
## - NCAM1           1   180.38 194.38
## + MMRN1           1   176.77 194.77
## + HP              1   176.86 194.86
## + LGALS3BP        1   177.17 195.17
## + IGHA2           1   177.33 195.33
## + KNG1            1   177.40 195.40
## + FETUA.Bovine    1   177.49 195.49
## + FETUB           1   177.50 195.50
## + CFI             1   177.73 195.73
## + AHSG            1   177.77 195.77
## + IGHG2           1   177.81 195.81
## + HRG             1   177.89 195.89
## - LRG1            1   181.89 195.89
## + FCGBP           1   177.89 195.89
## + THBS1           1   177.97 195.97
## + VWF             1   177.98 195.98
## + IGFBP3          1   178.04 196.04
## + C20orf3         1   178.04 196.04
```

```
## + CD44            1   178.08 196.08
## + AIAG.Bovine     1   178.10 196.10
## + DKFZp686N02209  1   178.10 196.10
## + Q5JNX2          1   178.12 196.12
## + FN1             1   178.12 196.12
## + BTD             1   178.12 196.12
## + APOB            1   178.13 196.13
## + PLTP            1   178.15 196.15
## + MST1            1   178.16 196.16
## + LYVE1           1   178.17 196.17
## + PROC            1   178.18 196.18
## + AFM             1   178.18 196.18
## + CLU             1   178.18 196.18
## + DSG2            1   178.18 196.18
## + PRG4            1   178.19 196.19
## + F5              1   178.19 196.19
## + F11             1   178.19 196.19
## - TIMP1           1   183.87 197.87
## - LUM             1   188.58 202.58
## - SERPINA7        1   192.29 206.29
## - LAMP2           1   196.03 210.03

summary(model.aic.both)

##
## Call:
## glm(formula = Class ~ TIMP1 + LAMP2 + LRG1 + SERPINA7 + LUM +
##     NCAM1 + FHR3, family = binomial, data = train)
##
## Deviance Residuals:
##      Min        1Q    Median        3Q       Max
## -2.14878  -0.93004  -0.02417   0.93892   2.49909
##
## Coefficients:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)   6.1312     8.3810   0.732 0.464434
## TIMP1        -1.0291     0.4422  -2.328 0.019938 *
## LAMP2         2.6904     0.6946   3.873 0.000107 ***
## LRG1         -0.7291     0.3851  -1.893 0.058322 .
## SERPINA7     -2.2482     0.6394  -3.516 0.000438 ***
## LUM           1.3098     0.4240   3.089 0.002006 **
## NCAM1        -0.7839     0.5356  -1.463 0.143332
## FHR3         -0.2323     0.1666  -1.395 0.163052
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 221.81  on 159  degrees of freedom
## Residual deviance: 178.19  on 152  degrees of freedom
```

```
## AIC: 194.19
##
## Number of Fisher Scoring iterations: 4
```

This results in 7 columns in the model with AIC = 194.19

BIC-based forward-backward selection

```
model.sbc.both <- step(glm.full, direction = "both", trace = 1,
scope=list(lower=glm.null, upper=glm.full),
                       k=log(nrow(train)))

## Start:  AIC=370.83
## Class ~ AFM + AHSG + AIAG.Bovine + APOB + BTD + C20orf3 + CD44 +
##     CFI + CLU + DKFZp686N02209 + DSG2 + F11 + F5 + FCGBP +
FETUA.Bovine +
##     FETUB + FHR3 + FN1 + HP + HRG + IGFBP3 + IGHA2 + IGHG2 +
##     KNG1 + LAMP2 + LGALS3BP + LRG1 + LUM + LYVE1 + MMRN1 + MST1 +
##     NCAM1 + PLTP + PRG4 + PROC + Q5JNX2 + SERPINA7 + THBS1 +
##     TIMP1 + VWF
##
##                  Df Deviance    AIC
## - C20orf3         1   162.75 365.76
## - LYVE1           1   162.75 365.76
## - CFI             1   162.76 365.76
## - PLTP            1   162.76 365.76
## - FN1             1   162.76 365.77
## - Q5JNX2          1   162.77 365.78
## - AFM             1   162.79 365.80
## - CD44            1   162.80 365.81
## - PROC            1   162.80 365.81
## - DSG2            1   162.85 365.85
## - BTD             1   162.86 365.87
## - MST1            1   162.90 365.91
## - AHSG            1   162.95 365.96
## - LRG1            1   162.98 365.98
## - APOB            1   162.98 365.99
## - F11             1   162.99 366.00
## - FETUB           1   163.02 366.03
## - PRG4            1   163.06 366.07
## - HRG             1   163.08 366.09
## - AIAG.Bovine     1   163.12 366.12
## - TIMP1           1   163.16 366.16
## - IGFBP3          1   163.20 366.21
## - F5              1   163.23 366.23
## - THBS1           1   163.44 366.45
## - VWF             1   163.46 366.46
## - CLU             1   163.48 366.49
## - FCGBP           1   163.65 366.65
## - LGALS3BP        1   163.85 366.86
## - FHR3            1   163.89 366.90
```

```
## - IGHG2           1   163.94 366.95
## - KNG1            1   164.13 367.13
## - IGHA2           1   164.42 367.43
## - DKFZp686N02209  1   164.43 367.44
## - HP              1   164.85 367.86
## - NCAM1           1   164.96 367.97
## - MMRN1           1   164.97 367.97
## - FETUA.Bovine    1   165.47 368.48
## - LUM             1   166.60 369.61
## <none>                162.75 370.83
## - SERPINA7        1   171.36 374.37
## - LAMP2           1   174.73 377.73
##
## Step:  AIC=365.76
## Class ~ AFM + AHSG + AIAG.Bovine + APOB + BTD + CD44 + CFI +
##     CLU + DKFZp686N02209 + DSG2 + F11 + F5 + FCGBP + FETUA.Bovine +
##     FETUB + FHR3 + FN1 + HP + HRG + IGFBP3 + IGHA2 + IGHG2 +
##     KNG1 + LAMP2 + LGALS3BP + LRG1 + LUM + LYVE1 + MMRN1 + MST1 +
##     NCAM1 + PLTP + PRG4 + PROC + Q5JNX2 + SERPINA7 + THBS1 +
##     TIMP1 + VWF
##
##                  Df Deviance    AIC
## - LYVE1           1   162.75 360.68
## - CFI             1   162.76 360.69
## - PLTP            1   162.76 360.69
## - FN1             1   162.76 360.69
## - Q5JNX2          1   162.77 360.71
## - AFM             1   162.79 360.72
## - CD44            1   162.80 360.73
## - PROC            1   162.80 360.73
## - DSG2            1   162.85 360.78
## - BTD             1   162.86 360.79
## - MST1            1   162.90 360.83
## - AHSG            1   162.95 360.88
## - LRG1            1   162.98 360.91
## - APOB            1   162.99 360.92
## - F11             1   162.99 360.92
## - FETUB           1   163.02 360.95
## - PRG4            1   163.07 361.00
## - HRG             1   163.08 361.01
## - AIAG.Bovine     1   163.15 361.08
## - TIMP1           1   163.16 361.09
## - IGFBP3          1   163.22 361.15
## - F5              1   163.23 361.16
## - THBS1           1   163.44 361.37
## - VWF             1   163.46 361.39
## - CLU             1   163.49 361.42
## - FCGBP           1   163.65 361.58
## - LGALS3BP        1   163.88 361.81
## - FHR3            1   163.90 361.83
```

```
## - IGHG2           1   163.94 361.87
## - KNG1            1   164.13 362.06
## - DKFZp686N02209  1   164.43 362.36
## - IGHA2           1   164.46 362.39
## - HP              1   164.87 362.80
## - MMRN1           1   165.00 362.93
## - NCAM1           1   165.00 362.93
## - FETUA.Bovine    1   165.51 363.44
## - LUM             1   166.62 364.55
## <none>                162.75 365.76
## - SERPINA7        1   171.45 369.38
## + C20orf3         1   162.75 370.83
## - LAMP2           1   174.73 372.66
##
## Step:  AIC=360.68
## Class ~ AFM + AHSG + AIAG.Bovine + APOB + BTD + CD44 + CFI +
##     CLU + DKFZp686N02209 + DSG2 + F11 + F5 + FCGBP + FETUA.Bovine +
##     FETUB + FHR3 + FN1 + HP + HRG + IGFBP3 + IGHA2 + IGHG2 +
##     KNG1 + LAMP2 + LGALS3BP + LRG1 + LUM + MMRN1 + MST1 + NCAM1 +
##     PLTP + PRG4 + PROC + Q5JNX2 + SERPINA7 + THBS1 + TIMP1 +
##     VWF
##
##                 Df Deviance    AIC
## - CFI            1   162.76 355.61
## - PLTP           1   162.76 355.61
## - FN1            1   162.76 355.62
## - Q5JNX2         1   162.78 355.63
## - AFM            1   162.79 355.65
## - CD44           1   162.80 355.66
## - PROC           1   162.80 355.66
## - DSG2           1   162.85 355.71
## - BTD            1   162.86 355.72
## - MST1           1   162.90 355.76
## - AHSG           1   162.95 355.81
## - LRG1           1   162.99 355.84
## - APOB           1   163.00 355.85
## - F11            1   163.00 355.86
## - FETUB          1   163.03 355.88
## - PRG4           1   163.08 355.93
## - HRG            1   163.08 355.94
## - AIAG.Bovine    1   163.15 356.01
## - TIMP1          1   163.16 356.02
## - IGFBP3         1   163.23 356.08
## - F5             1   163.24 356.10
## - THBS1          1   163.45 356.30
## - VWF            1   163.46 356.32
## - CLU            1   163.49 356.35
## - FCGBP          1   163.65 356.51
## - FHR3           1   163.90 356.75
## - LGALS3BP       1   163.90 356.76
```

```
## - IGHG2           1   163.95 356.80
## - KNG1            1   164.13 356.98
## - DKFZp686N02209  1   164.43 357.29
## - IGHA2           1   164.47 357.32
## - HP              1   164.94 357.80
## - NCAM1           1   165.01 357.87
## - MMRN1           1   165.08 357.94
## - FETUA.Bovine    1   165.51 358.37
## - LUM             1   167.15 360.01
## <none>                162.75 360.68
## - SERPINA7        1   172.02 364.88
## + LYVE1           1   162.75 365.76
## + C20orf3         1   162.75 365.76
## - LAMP2           1   174.74 367.60
##
## Step:  AIC=355.61
## Class ~ AFM + AHSG + AIAG.Bovine + APOB + BTD + CD44 + CLU +
##     DKFZp686N02209 + DSG2 + F11 + F5 + FCGBP + FETUA.Bovine +
##     FETUB + FHR3 + FN1 + HP + HRG + IGFBP3 + IGHA2 + IGHG2 +
##     KNG1 + LAMP2 + LGALS3BP + LRG1 + LUM + MMRN1 + MST1 + NCAM1 +
##     PLTP + PRG4 + PROC + Q5JNX2 + SERPINA7 + THBS1 + TIMP1 +
##     VWF
##
##                  Df Deviance    AIC
## - PLTP            1   162.76 350.54
## - FN1             1   162.77 350.55
## - Q5JNX2          1   162.78 350.56
## - AFM             1   162.79 350.58
## - PROC            1   162.81 350.59
## - CD44            1   162.81 350.59
## - DSG2            1   162.86 350.64
## - BTD             1   162.86 350.64
## - MST1            1   162.91 350.70
## - AHSG            1   162.95 350.73
## - APOB            1   163.01 350.79
## - LRG1            1   163.01 350.79
## - FETUB           1   163.04 350.82
## - F11             1   163.05 350.83
## - PRG4            1   163.08 350.86
## - HRG             1   163.09 350.87
## - TIMP1           1   163.16 350.94
## - AIAG.Bovine     1   163.24 351.02
## - IGFBP3          1   163.27 351.05
## - F5              1   163.28 351.07
## - THBS1           1   163.45 351.23
## - VWF             1   163.46 351.24
## - CLU             1   163.53 351.31
## - FCGBP           1   163.65 351.44
## - FHR3            1   163.90 351.68
## - IGHG2           1   163.95 351.73
```

```
## - LGALS3BP        1   163.97 351.75
## - KNG1            1   164.19 351.97
## - DKFZp686N02209  1   164.47 352.25
## - IGHA2           1   164.47 352.25
## - NCAM1           1   165.01 352.79
## - HP              1   165.07 352.85
## - MMRN1           1   165.08 352.87
## - FETUA.Bovine    1   165.72 353.50
## - LUM             1   167.18 354.96
## <none>                162.76 355.61
## + CFI             1   162.75 360.68
## + LYVE1           1   162.76 360.69
## + C20orf3         1   162.76 360.69
## - SERPINA7        1   173.05 360.83
## - LAMP2           1   174.95 362.73
##
## Step:  AIC=350.54
## Class ~ AFM + AHSG + AIAG.Bovine + APOB + BTD + CD44 + CLU +
##     DKFZp686N02209 + DSG2 + F11 + F5 + FCGBP + FETUA.Bovine +
##     FETUB + FHR3 + FN1 + HP + HRG + IGFBP3 + IGHA2 + IGHG2 +
##     KNG1 + LAMP2 + LGALS3BP + LRG1 + LUM + MMRN1 + MST1 + NCAM1 +
##     PRG4 + PROC + Q5JNX2 + SERPINA7 + THBS1 + TIMP1 + VWF
##
##                  Df Deviance    AIC
## - FN1             1   162.78 345.48
## - Q5JNX2          1   162.79 345.49
## - AFM             1   162.80 345.50
## - PROC            1   162.81 345.51
## - CD44            1   162.81 345.52
## - BTD             1   162.87 345.57
## - DSG2            1   162.87 345.57
## - MST1            1   162.92 345.62
## - AHSG            1   162.96 345.66
## - LRG1            1   163.01 345.72
## - APOB            1   163.02 345.73
## - FETUB           1   163.05 345.76
## - F11             1   163.06 345.77
## - PRG4            1   163.09 345.80
## - HRG             1   163.10 345.81
## - TIMP1           1   163.17 345.87
## - AIAG.Bovine     1   163.25 345.95
## - IGFBP3          1   163.28 345.98
## - F5              1   163.29 345.99
## - THBS1           1   163.45 346.16
## - VWF             1   163.47 346.17
## - CLU             1   163.60 346.31
## - FCGBP           1   163.66 346.37
## - FHR3            1   163.90 346.61
## - IGHG2           1   163.95 346.66
## - LGALS3BP        1   163.97 346.67
```

```
## - KNG1            1   164.23 346.93
## - IGHA2           1   164.47 347.18
## - DKFZp686N02209  1   164.47 347.18
## - NCAM1           1   165.04 347.75
## - MMRN1           1   165.08 347.79
## - HP              1   165.09 347.79
## - FETUA.Bovine    1   165.74 348.44
## - LUM             1   167.36 350.07
## <none>                162.76 350.54
## + PLTP            1   162.76 355.61
## + CFI             1   162.76 355.61
## + LYVE1           1   162.76 355.62
## + C20orf3         1   162.76 355.62
## - SERPINA7        1   173.16 355.86
## - LAMP2           1   175.10 357.80
##
## Step:  AIC=345.48
## Class ~ AFM + AHSG + AIAG.Bovine + APOB + BTD + CD44 + CLU +
##     DKFZp686N02209 + DSG2 + F11 + F5 + FCGBP + FETUA.Bovine +
##     FETUB + FHR3 + HP + HRG + IGFBP3 + IGHA2 + IGHG2 + KNG1 +
##     LAMP2 + LGALS3BP + LRG1 + LUM + MMRN1 + MST1 + NCAM1 + PRG4 +
##     PROC + Q5JNX2 + SERPINA7 + THBS1 + TIMP1 + VWF
##
##                  Df Deviance    AIC
## - AFM             1   162.81 340.44
## - Q5JNX2          1   162.81 340.44
## - CD44            1   162.81 340.45
## - PROC            1   162.82 340.45
## - BTD             1   162.87 340.50
## - DSG2            1   162.90 340.53
## - MST1            1   162.93 340.56
## - AHSG            1   163.00 340.63
## - APOB            1   163.04 340.67
## - LRG1            1   163.06 340.69
## - FETUB           1   163.06 340.69
## - F11             1   163.10 340.73
## - HRG             1   163.11 340.74
## - PRG4            1   163.14 340.77
## - TIMP1           1   163.17 340.80
## - AIAG.Bovine     1   163.25 340.88
## - IGFBP3          1   163.28 340.91
## - F5              1   163.38 341.01
## - THBS1           1   163.47 341.10
## - VWF             1   163.48 341.11
## - CLU             1   163.60 341.23
## - FCGBP           1   163.67 341.30
## - FHR3            1   163.95 341.58
## - LGALS3BP        1   163.97 341.60
## - IGHG2           1   164.03 341.66
## - KNG1            1   164.23 341.86
```

```
## - DKFZp686N02209  1   164.47 342.11
## - IGHA2           1   164.48 342.11
## - HP              1   165.09 342.72
## - NCAM1           1   165.10 342.73
## - MMRN1           1   165.11 342.75
## - FETUA.Bovine    1   165.74 343.37
## - LUM             1   167.38 345.01
## <none>                162.78 345.48
## + FN1             1   162.76 350.54
## + PLTP            1   162.77 350.55
## + CFI             1   162.77 350.55
## + LYVE1           1   162.78 350.56
## + C20orf3         1   162.78 350.56
## - SERPINA7        1   173.17 350.80
## - LAMP2           1   175.39 353.02
##
## Step:  AIC=340.44
## Class ~ AHSG + AIAG.Bovine + APOB + BTD + CD44 + CLU +
DKFZp686N02209 +
##     DSG2 + F11 + F5 + FCGBP + FETUA.Bovine + FETUB + FHR3 + HP +
##     HRG + IGFBP3 + IGHA2 + IGHG2 + KNG1 + LAMP2 + LGALS3BP +
##     LRG1 + LUM + MMRN1 + MST1 + NCAM1 + PRG4 + PROC + Q5JNX2 +
##     SERPINA7 + THBS1 + TIMP1 + VWF
##
##                  Df Deviance    AIC
## - Q5JNX2          1   162.83 335.39
## - PROC            1   162.84 335.40
## - CD44            1   162.84 335.40
## - BTD             1   162.89 335.44
## - DSG2            1   162.93 335.49
## - MST1            1   162.96 335.52
## - APOB            1   163.05 335.60
## - FETUB           1   163.08 335.64
## - AHSG            1   163.09 335.64
## - F11             1   163.12 335.67
## - LRG1            1   163.12 335.67
## - HRG             1   163.14 335.69
## - PRG4            1   163.18 335.73
## - TIMP1           1   163.18 335.74
## - AIAG.Bovine     1   163.25 335.81
## - IGFBP3          1   163.34 335.90
## - F5              1   163.38 335.93
## - VWF             1   163.49 336.05
## - THBS1           1   163.51 336.06
## - CLU             1   163.61 336.17
## - FCGBP           1   163.68 336.23
## - LGALS3BP        1   163.97 336.53
## - FHR3            1   163.98 336.53
## - IGHG2           1   164.07 336.63
## - KNG1            1   164.25 336.80
```

```
## - IGHA2           1   164.56 337.12
## - DKFZp686N02209  1   164.68 337.24
## - HP              1   165.13 337.69
## - NCAM1           1   165.15 337.71
## - MMRN1           1   165.28 337.84
## - FETUA.Bovine    1   165.76 338.32
## - LUM             1   167.54 340.10
## <none>                162.81 340.44
## + AFM             1   162.78 345.48
## + FN1             1   162.80 345.50
## + PLTP            1   162.81 345.51
## + CFI             1   162.81 345.51
## + LYVE1           1   162.81 345.51
## + C20orf3         1   162.81 345.51
## - SERPINA7        1   173.18 345.74
## - LAMP2           1   176.63 349.19
##
## Step:  AIC=335.39
## Class ~ AHSG + AIAG.Bovine + APOB + BTD + CD44 + CLU +
DKFZp686N02209 +
##     DSG2 + F11 + F5 + FCGBP + FETUA.Bovine + FETUB + FHR3 + HP +
##     HRG + IGFBP3 + IGHA2 + IGHG2 + KNG1 + LAMP2 + LGALS3BP +
##     LRG1 + LUM + MMRN1 + MST1 + NCAM1 + PRG4 + PROC + SERPINA7 +
##     THBS1 + TIMP1 + VWF
##
##                  Df Deviance    AIC
## - PROC            1   162.86 330.34
## - CD44            1   162.87 330.35
## - BTD             1   162.91 330.39
## - DSG2            1   162.97 330.45
## - MST1            1   162.98 330.46
## - APOB            1   163.09 330.57
## - AHSG            1   163.11 330.59
## - LRG1            1   163.12 330.61
## - FETUB           1   163.13 330.61
## - F11             1   163.18 330.66
## - TIMP1           1   163.19 330.67
## - HRG             1   163.19 330.67
## - PRG4            1   163.23 330.71
## - AIAG.Bovine     1   163.25 330.73
## - F5              1   163.38 330.86
## - IGFBP3          1   163.38 330.87
## - VWF             1   163.51 330.99
## - THBS1           1   163.51 330.99
## - CLU             1   163.61 331.09
## - FCGBP           1   163.68 331.16
## - FHR3            1   163.99 331.47
## - LGALS3BP        1   164.03 331.51
## - IGHG2           1   164.20 331.68
## - KNG1            1   164.25 331.74
```

```
## - IGHA2           1   164.56 332.04
## - DKFZp686N02209  1   164.97 332.45
## - NCAM1           1   165.15 332.63
## - HP              1   165.16 332.64
## - MMRN1           1   165.50 332.98
## - FETUA.Bovine    1   165.95 333.43
## - LUM             1   167.55 335.03
## <none>                162.83 335.39
## + Q5JNX2          1   162.81 340.44
## + AFM             1   162.81 340.44
## + FN1             1   162.82 340.45
## + PLTP            1   162.83 340.46
## + LYVE1           1   162.83 340.46
## + CFI             1   162.83 340.46
## + C20orf3         1   162.83 340.47
## - SERPINA7        1   173.48 340.96
## - LAMP2           1   177.04 344.52
##
## Step:  AIC=330.34
## Class ~ AHSG + AIAG.Bovine + APOB + BTD + CD44 + CLU +
DKFZp686N02209 +
##     DSG2 + F11 + F5 + FCGBP + FETUA.Bovine + FETUB + FHR3 + HP +
##     HRG + IGFBP3 + IGHA2 + IGHG2 + KNG1 + LAMP2 + LGALS3BP +
##     LRG1 + LUM + MMRN1 + MST1 + NCAM1 + PRG4 + SERPINA7 + THBS1 +
##     TIMP1 + VWF
##
##                  Df Deviance    AIC
## - CD44            1   162.91 325.31
## - BTD             1   162.94 325.34
## - DSG2            1   163.00 325.40
## - MST1            1   163.01 325.42
## - AHSG            1   163.13 325.53
## - APOB            1   163.13 325.54
## - LRG1            1   163.16 325.56
## - FETUB           1   163.18 325.59
## - HRG             1   163.20 325.60
## - TIMP1           1   163.21 325.62
## - PRG4            1   163.23 325.64
## - F11             1   163.26 325.67
## - AIAG.Bovine     1   163.27 325.68
## - IGFBP3          1   163.38 325.79
## - F5              1   163.39 325.79
## - VWF             1   163.51 325.92
## - THBS1           1   163.53 325.93
## - CLU             1   163.61 326.02
## - FCGBP           1   163.71 326.11
## - FHR3            1   164.00 326.41
## - LGALS3BP        1   164.09 326.50
## - IGHG2           1   164.20 326.60
## - KNG1            1   164.27 326.68
```

```
## - IGHA2           1   164.65 327.05
## - NCAM1           1   165.17 327.57
## - DKFZp686N02209  1   165.17 327.58
## - HP              1   165.20 327.61
## - MMRN1           1   165.60 328.01
## - FETUA.Bovine    1   165.99 328.40
## - LUM             1   167.57 329.98
## <none>                162.86 330.34
## + PROC            1   162.83 335.39
## + Q5JNX2          1   162.84 335.40
## + AFM             1   162.84 335.40
## + FN1             1   162.85 335.41
## + LYVE1           1   162.86 335.42
## + CFI             1   162.86 335.42
## + PLTP            1   162.86 335.42
## + C20orf3         1   162.86 335.42
## - SERPINA7        1   173.65 336.05
## - LAMP2           1   177.17 339.57
##
## Step:  AIC=325.31
## Class ~ AHSG + AIAG.Bovine + APOB + BTD + CLU + DKFZp686N02209 +
##     DSG2 + F11 + F5 + FCGBP + FETUA.Bovine + FETUB + FHR3 + HP +
##     HRG + IGFBP3 + IGHA2 + IGHG2 + KNG1 + LAMP2 + LGALS3BP +
##     LRG1 + LUM + MMRN1 + MST1 + NCAM1 + PRG4 + SERPINA7 + THBS1 +
##     TIMP1 + VWF
##
##                  Df Deviance    AIC
## - BTD             1   162.95 320.28
## - DSG2            1   163.02 320.35
## - MST1            1   163.09 320.42
## - AHSG            1   163.15 320.48
## - LRG1            1   163.17 320.50
## - APOB            1   163.19 320.52
## - HRG             1   163.24 320.57
## - TIMP1           1   163.25 320.58
## - FETUB           1   163.26 320.59
## - AIAG.Bovine     1   163.28 320.61
## - PRG4            1   163.29 320.62
## - F11             1   163.31 320.64
## - IGFBP3          1   163.39 320.72
## - F5              1   163.48 320.81
## - VWF             1   163.52 320.85
## - THBS1           1   163.62 320.95
## - CLU             1   163.65 320.98
## - FCGBP           1   163.74 321.07
## - FHR3            1   164.09 321.42
## - LGALS3BP        1   164.22 321.55
## - IGHG2           1   164.32 321.65
## - KNG1            1   164.41 321.74
## - IGHA2           1   164.65 321.98
```

```
## - NCAM1           1   165.19 322.53
## - DKFZp686N02209  1   165.21 322.53
## - HP              1   165.28 322.61
## - MMRN1           1   165.95 323.28
## - FETUA.Bovine    1   166.24 323.57
## - LUM             1   167.57 324.90
## <none>                162.91 325.31
## + CD44            1   162.86 330.34
## + PROC            1   162.87 330.35
## + Q5JNX2          1   162.88 330.36
## + AFM             1   162.89 330.37
## + CFI             1   162.90 330.38
## + FN1             1   162.91 330.39
## + LYVE1           1   162.91 330.39
## + C20orf3         1   162.91 330.39
## + PLTP            1   162.91 330.39
## - SERPINA7        1   173.81 331.14
## - LAMP2           1   177.27 334.60
##
## Step:  AIC=320.28
## Class ~ AHSG + AIAG.Bovine + APOB + CLU + DKFZp686N02209 + DSG2 +
##     F11 + F5 + FCGBP + FETUA.Bovine + FETUB + FHR3 + HP + HRG +
##     IGFBP3 + IGHA2 + IGHG2 + KNG1 + LAMP2 + LGALS3BP + LRG1 +
##     LUM + MMRN1 + MST1 + NCAM1 + PRG4 + SERPINA7 + THBS1 + TIMP1 +
##     VWF
##
##                  Df Deviance    AIC
## - DSG2            1   163.05 315.31
## - MST1            1   163.15 315.40
## - AHSG            1   163.16 315.41
## - LRG1            1   163.19 315.44
## - TIMP1           1   163.25 315.51
## - APOB            1   163.26 315.51
## - HRG             1   163.28 315.54
## - PRG4            1   163.31 315.56
## - FETUB           1   163.32 315.57
## - AIAG.Bovine     1   163.32 315.58
## - IGFBP3          1   163.39 315.65
## - F11             1   163.41 315.66
## - F5              1   163.52 315.78
## - VWF             1   163.54 315.80
## - THBS1           1   163.64 315.90
## - FCGBP           1   163.76 316.02
## - CLU             1   163.77 316.02
## - FHR3            1   164.10 316.35
## - IGHG2           1   164.33 316.59
## - LGALS3BP        1   164.34 316.59
## - KNG1            1   164.43 316.68
## - IGHA2           1   164.71 316.97
## - DKFZp686N02209  1   165.23 317.49
```

```
## - NCAM1           1   165.27 317.52
## - HP              1   165.42 317.68
## - MMRN1           1   166.11 318.37
## - FETUA.Bovine    1   166.38 318.63
## - LUM             1   167.65 319.90
## <none>                162.95 320.28
## + BTD             1   162.91 325.31
## + PROC            1   162.92 325.33
## + Q5JNX2          1   162.94 325.34
## + CD44            1   162.94 325.34
## + AFM             1   162.94 325.35
## + LYVE1           1   162.95 325.36
## + CFI             1   162.95 325.36
## + PLTP            1   162.95 325.36
## + C20orf3         1   162.95 325.36
## + FN1             1   162.95 325.36
## - SERPINA7        1   173.85 326.11
## - LAMP2           1   177.28 329.53
##
## Step:  AIC=315.31
## Class ~ AHSG + AIAG.Bovine + APOB + CLU + DKFZp686N02209 + F11 +
##     F5 + FCGBP + FETUA.Bovine + FETUB + FHR3 + HP + HRG + IGFBP3 +
##     IGHA2 + IGHG2 + KNG1 + LAMP2 + LGALS3BP + LRG1 + LUM + MMRN1 +
##     MST1 + NCAM1 + PRG4 + SERPINA7 + THBS1 + TIMP1 + VWF
##
##                  Df Deviance    AIC
## - LRG1            1   163.26 310.44
## - AHSG            1   163.27 310.45
## - MST1            1   163.27 310.45
## - APOB            1   163.30 310.48
## - PRG4            1   163.38 310.56
## - TIMP1           1   163.39 310.57
## - F11             1   163.41 310.59
## - FETUB           1   163.41 310.59
## - IGFBP3          1   163.42 310.60
## - AIAG.Bovine     1   163.44 310.63
## - HRG             1   163.47 310.65
## - F5              1   163.56 310.74
## - VWF             1   163.78 310.96
## - FCGBP           1   163.80 310.98
## - THBS1           1   163.82 311.00
## - CLU             1   163.95 311.13
## - FHR3            1   164.21 311.39
## - IGHG2           1   164.36 311.54
## - LGALS3BP        1   164.47 311.65
## - KNG1            1   164.55 311.73
## - IGHA2           1   164.72 311.90
## - DKFZp686N02209  1   165.25 312.43
## - NCAM1           1   165.28 312.45
## - HP              1   165.68 312.86
```

```
## - MMRN1           1   166.16 313.34
## - FETUA.Bovine    1   166.40 313.58
## - LUM             1   167.81 314.99
## <none>                163.05 315.31
## + DSG2            1   162.95 320.28
## + PROC            1   163.02 320.35
## + BTD             1   163.02 320.35
## + Q5JNX2          1   163.03 320.36
## + AFM             1   163.04 320.37
## + CD44            1   163.05 320.38
## + FN1             1   163.05 320.38
## + C20orf3         1   163.05 320.38
## + LYVE1           1   163.05 320.38
## + CFI             1   163.05 320.38
## + PLTP            1   163.05 320.38
## - SERPINA7        1   173.98 321.16
## - LAMP2           1   177.29 324.47
##
## Step:  AIC=310.44
## Class ~ AHSG + AIAG.Bovine + APOB + CLU + DKFZp686N02209 + F11 +
##     F5 + FCGBP + FETUA.Bovine + FETUB + FHR3 + HP + HRG + IGFBP3 +
##     IGHA2 + IGHG2 + KNG1 + LAMP2 + LGALS3BP + LUM + MMRN1 + MST1 +
##     NCAM1 + PRG4 + SERPINA7 + THBS1 + TIMP1 + VWF
##
##                  Df Deviance    AIC
## - APOB            1   163.48 305.59
## - AHSG            1   163.50 305.60
## - FETUB           1   163.50 305.61
## - MST1            1   163.56 305.67
## - PRG4            1   163.65 305.76
## - AIAG.Bovine     1   163.66 305.76
## - F11             1   163.69 305.80
## - TIMP1           1   163.73 305.83
## - HRG             1   163.76 305.87
## - IGFBP3          1   163.85 305.95
## - F5              1   163.91 306.02
## - FCGBP           1   164.23 306.33
## - CLU             1   164.25 306.36
## - VWF             1   164.26 306.37
## - THBS1           1   164.29 306.39
## - LGALS3BP        1   164.58 306.69
## - IGHG2           1   164.61 306.72
## - KNG1            1   164.79 306.90
## - FHR3            1   164.83 306.94
## - IGHA2           1   165.06 307.17
## - NCAM1           1   165.34 307.44
## - DKFZp686N02209  1   165.53 307.64
## - HP              1   166.34 308.44
## - FETUA.Bovine    1   166.73 308.84
## - MMRN1           1   167.11 309.21
```

```
## - LUM              1   167.87 309.97
## <none>                 163.26 310.44
## + LRG1             1   163.05 315.31
## + DSG2             1   163.19 315.44
## + PROC             1   163.23 315.48
## + AFM              1   163.23 315.49
## + FN1              1   163.24 315.49
## + CFI              1   163.25 315.50
## + BTD              1   163.25 315.51
## + LYVE1            1   163.25 315.51
## + C20orf3          1   163.25 315.51
## + Q5JNX2           1   163.25 315.51
## + CD44             1   163.26 315.52
## + PLTP             1   163.26 315.52
## - SERPINA7         1   174.37 316.47
## - LAMP2            1   177.29 319.40
##
## Step:  AIC=305.59
## Class ~ AHSG + AIAG.Bovine + CLU + DKFZp686N02209 + F11 + F5 +
##     FCGBP + FETUA.Bovine + FETUB + FHR3 + HP + HRG + IGFBP3 +
##     IGHA2 + IGHG2 + KNG1 + LAMP2 + LGALS3BP + LUM + MMRN1 + MST1 +
##     NCAM1 + PRG4 + SERPINA7 + THBS1 + TIMP1 + VWF
##
##                   Df Deviance    AIC
## - FETUB            1   163.72 300.75
## - AHSG             1   163.73 300.75
## - MST1             1   163.79 300.82
## - HRG              1   163.90 300.93
## - F11              1   163.91 300.94
## - F5               1   163.96 300.99
## - TIMP1            1   164.02 301.05
## - AIAG.Bovine      1   164.05 301.08
## - PRG4             1   164.08 301.11
## - IGFBP3           1   164.20 301.23
## - CLU              1   164.39 301.42
## - VWF              1   164.45 301.48
## - FCGBP            1   164.52 301.55
## - THBS1            1   164.59 301.62
## - LGALS3BP         1   164.72 301.75
## - IGHG2            1   164.86 301.89
## - KNG1             1   164.96 301.99
## - FHR3             1   165.16 302.19
## - IGHA2            1   165.19 302.22
## - NCAM1            1   165.49 302.52
## - DKFZp686N02209   1   165.63 302.66
## - HP               1   166.56 303.59
## - FETUA.Bovine     1   166.90 303.93
## - MMRN1            1   167.13 304.16
## <none>                 163.48 305.59
## - LUM              1   168.67 305.70
```

```
## + APOB             1   163.26 310.44
## + LRG1             1   163.30 310.48
## + PROC             1   163.44 310.62
## + DSG2             1   163.45 310.63
## + C20orf3          1   163.46 310.64
## + BTD              1   163.46 310.64
## + LYVE1            1   163.46 310.64
## + CFI              1   163.46 310.64
## + FN1              1   163.47 310.64
## + Q5JNX2           1   163.47 310.64
## + AFM              1   163.47 310.65
## + CD44             1   163.48 310.66
## + PLTP             1   163.48 310.66
## - SERPINA7         1   174.87 311.90
## - LAMP2            1   177.58 314.61
##
## Step:  AIC=300.75
## Class ~ AHSG + AIAG.Bovine + CLU + DKFZp686N02209 + F11 + F5 +
##     FCGBP + FETUA.Bovine + FHR3 + HP + HRG + IGFBP3 + IGHA2 +
##     IGHG2 + KNG1 + LAMP2 + LGALS3BP + LUM + MMRN1 + MST1 + NCAM1 +
##     PRG4 + SERPINA7 + THBS1 + TIMP1 + VWF
##
##                   Df Deviance    AIC
## - MST1             1   164.01 295.96
## - AHSG             1   164.02 295.97
## - F11              1   164.10 296.05
## - HRG              1   164.10 296.06
## - F5               1   164.19 296.15
## - TIMP1            1   164.27 296.23
## - PRG4             1   164.40 296.36
## - AIAG.Bovine      1   164.44 296.39
## - IGFBP3           1   164.46 296.42
## - CLU              1   164.50 296.45
## - VWF              1   164.69 296.64
## - THBS1            1   164.78 296.73
## - FCGBP            1   164.79 296.75
## - LGALS3BP         1   164.97 296.92
## - IGHG2            1   165.00 296.96
## - KNG1             1   165.11 297.07
## - FHR3             1   165.31 297.27
## - IGHA2            1   165.47 297.42
## - NCAM1            1   165.69 297.64
## - DKFZp686N02209   1   166.12 298.08
## - HP               1   166.83 298.78
## - FETUA.Bovine     1   167.10 299.05
## - MMRN1            1   167.19 299.14
## - LUM              1   168.74 300.69
## <none>                 163.72 300.75
## + FETUB            1   163.48 305.59
## + APOB             1   163.50 305.61
```

```
## + PROC            1   163.64 305.75
## + LRG1            1   163.65 305.76
## + BTD             1   163.68 305.79
## + Q5JNX2          1   163.69 305.79
## + DSG2            1   163.69 305.80
## + LYVE1           1   163.69 305.80
## + CD44            1   163.70 305.81
## + C20orf3         1   163.71 305.81
## + CFI             1   163.71 305.81
## + PLTP            1   163.72 305.82
## + FN1             1   163.72 305.82
## + AFM             1   163.72 305.83
## - SERPINA7        1   174.94 306.90
## - LAMP2           1   178.45 310.40
##
## Step:  AIC=295.96
## Class ~ AHSG + AIAG.Bovine + CLU + DKFZp686N02209 + F11 + F5 +
##     FCGBP + FETUA.Bovine + FHR3 + HP + HRG + IGFBP3 + IGHA2 +
##     IGHG2 + KNG1 + LAMP2 + LGALS3BP + LUM + MMRN1 + NCAM1 + PRG4 +
##     SERPINA7 + THBS1 + TIMP1 + VWF
##
##                  Df Deviance    AIC
## - AHSG            1   164.26 291.14
## - F11             1   164.33 291.21
## - TIMP1           1   164.44 291.32
## - F5              1   164.48 291.36
## - HRG             1   164.57 291.45
## - PRG4            1   164.65 291.53
## - AIAG.Bovine     1   164.72 291.60
## - CLU             1   164.82 291.70
## - IGFBP3          1   164.88 291.76
## - VWF             1   164.93 291.81
## - FCGBP           1   165.05 291.93
## - THBS1           1   165.10 291.98
## - LGALS3BP        1   165.20 292.08
## - KNG1            1   165.38 292.26
## - IGHG2           1   165.53 292.40
## - IGHA2           1   165.62 292.49
## - FHR3            1   165.89 292.77
## - NCAM1           1   165.89 292.77
## - DKFZp686N02209  1   166.22 293.10
## - HP              1   167.13 294.01
## - MMRN1           1   167.41 294.29
## - FETUA.Bovine    1   167.48 294.36
## <none>                164.01 295.96
## - LUM             1   169.14 296.02
## + MST1            1   163.72 300.75
## + APOB            1   163.78 300.81
## + FETUB           1   163.79 300.82
## + LRG1            1   163.88 300.91
```

```
## + PROC            1   163.91 300.94
## + BTD             1   163.96 300.99
## + DSG2            1   163.96 300.99
## + CFI             1   163.97 301.00
## + LYVE1           1   163.97 301.00
## + CD44            1   163.98 301.01
## + Q5JNX2          1   163.99 301.02
## + C20orf3         1   164.00 301.03
## + AFM             1   164.01 301.04
## + PLTP            1   164.01 301.04
## + FN1             1   164.01 301.04
## - SERPINA7        1   175.25 302.12
## - LAMP2           1   178.64 305.52
##
## Step:  AIC=291.14
## Class ~ AIAG.Bovine + CLU + DKFZp686N02209 + F11 + F5 + FCGBP +
##     FETUA.Bovine + FHR3 + HP + HRG + IGFBP3 + IGHA2 + IGHG2 +
##     KNG1 + LAMP2 + LGALS3BP + LUM + MMRN1 + NCAM1 + PRG4 + SERPINA7
+
##     THBS1 + TIMP1 + VWF
##
##                   Df Deviance    AIC
## - F11              1   164.55 286.36
## - TIMP1            1   164.71 286.52
## - F5               1   164.73 286.53
## - PRG4             1   164.81 286.61
## - AIAG.Bovine      1   164.90 286.70
## - HRG              1   164.97 286.77
## - CLU              1   165.01 286.81
## - IGFBP3           1   165.11 286.92
## - LGALS3BP         1   165.35 287.15
## - FCGBP            1   165.36 287.17
## - VWF              1   165.41 287.22
## - THBS1            1   165.51 287.31
## - KNG1             1   165.62 287.43
## - IGHG2            1   165.87 287.67
## - IGHA2            1   165.98 287.78
## - NCAM1            1   166.09 287.89
## - FHR3             1   166.13 287.93
## - DKFZp686N02209   1   166.56 288.37
## - HP               1   167.22 289.03
## - FETUA.Bovine     1   167.69 289.50
## - MMRN1            1   167.97 289.77
## <none>                 164.26 291.14
## - LUM              1   169.83 291.63
## + FETUB            1   164.00 295.96
## + AHSG             1   164.01 295.96
## + MST1             1   164.02 295.97
## + APOB             1   164.03 295.98
## + LRG1             1   164.14 296.09
```

```
## + PROC            1   164.19 296.15
## + DSG2            1   164.21 296.16
## + LYVE1           1   164.22 296.18
## + CD44            1   164.23 296.19
## + FN1             1   164.24 296.19
## + Q5JNX2          1   164.24 296.19
## + AFM             1   164.24 296.19
## + CFI             1   164.25 296.20
## + BTD             1   164.26 296.21
## + C20orf3         1   164.26 296.22
## + PLTP            1   164.26 296.22
## - SERPINA7        1   175.33 297.13
## - LAMP2           1   179.01 300.82
##
## Step:  AIC=286.36
## Class ~ AIAG.Bovine + CLU + DKFZp686N02209 + F5 + FCGBP +
FETUA.Bovine +
##     FHR3 + HP + HRG + IGFBP3 + IGHA2 + IGHG2 + KNG1 + LAMP2 +
##     LGALS3BP + LUM + MMRN1 + NCAM1 + PRG4 + SERPINA7 + THBS1 +
##     TIMP1 + VWF
##
##                  Df Deviance    AIC
## - TIMP1           1   164.94 281.67
## - PRG4            1   165.06 281.79
## - HRG             1   165.15 281.88
## - F5              1   165.18 281.91
## - IGFBP3          1   165.26 281.99
## - AIAG.Bovine     1   165.35 282.08
## - CLU             1   165.45 282.18
## - LGALS3BP        1   165.47 282.19
## - FCGBP           1   165.51 282.24
## - THBS1           1   165.61 282.33
## - VWF             1   165.75 282.48
## - KNG1            1   165.84 282.57
## - IGHA2           1   166.16 282.89
## - IGHG2           1   166.24 282.97
## - NCAM1           1   166.54 283.27
## - DKFZp686N02209  1   166.72 283.44
## - FHR3            1   167.05 283.78
## - FETUA.Bovine    1   167.97 284.70
## - HP              1   168.09 284.82
## - MMRN1           1   168.39 285.12
## <none>                164.55 286.36
## - LUM             1   170.14 286.87
## + F11             1   164.26 291.14
## + APOB            1   164.32 291.20
## + AHSG            1   164.33 291.21
## + FETUB           1   164.34 291.22
## + MST1            1   164.35 291.23
## + LRG1            1   164.38 291.26
```

```
## + PROC            1   164.43 291.30
## + CFI             1   164.48 291.36
## + LYVE1           1   164.48 291.36
## + Q5JNX2          1   164.50 291.38
## + CD44            1   164.52 291.40
## + FN1             1   164.52 291.40
## + BTD             1   164.54 291.42
## + AFM             1   164.54 291.43
## + C20orf3         1   164.55 291.43
## + PLTP            1   164.55 291.43
## + DSG2            1   164.55 291.43
## - SERPINA7        1   175.74 292.46
## - LAMP2           1   179.10 295.83
##
## Step:  AIC=281.67
## Class ~ AIAG.Bovine + CLU + DKFZp686N02209 + F5 + FCGBP +
FETUA.Bovine +
##     FHR3 + HP + HRG + IGFBP3 + IGHA2 + IGHG2 + KNG1 + LAMP2 +
##     LGALS3BP + LUM + MMRN1 + NCAM1 + PRG4 + SERPINA7 + THBS1 +
##     VWF
##
##                  Df Deviance    AIC
## - F5              1   165.48 277.14
## - IGFBP3          1   165.55 277.20
## - THBS1           1   165.66 277.31
## - CLU             1   165.70 277.35
## - PRG4            1   165.71 277.36
## - HRG             1   165.72 277.38
## - FCGBP           1   165.90 277.55
## - LGALS3BP        1   165.92 277.58
## - AIAG.Bovine     1   165.93 277.58
## - VWF             1   165.93 277.58
## - KNG1            1   166.28 277.93
## - IGHA2           1   166.37 278.02
## - IGHG2           1   166.70 278.35
## - NCAM1           1   166.99 278.65
## - DKFZp686N02209  1   167.30 278.95
## - FHR3            1   167.51 279.17
## - FETUA.Bovine    1   168.75 280.40
## - HP              1   169.23 280.88
## <none>                164.94 281.67
## - LUM             1   170.30 281.96
## - MMRN1           1   172.70 284.35
## + TIMP1           1   164.55 286.36
## + APOB            1   164.65 286.45
## + AHSG            1   164.69 286.49
## + FETUB           1   164.70 286.50
## + LRG1            1   164.70 286.50
## + F11             1   164.71 286.52
## + PROC            1   164.82 286.62
```

```
## + MST1             1   164.83 286.63
## + LYVE1            1   164.85 286.66
## + CFI              1   164.89 286.69
## + FN1              1   164.89 286.70
## + CD44             1   164.90 286.70
## + Q5JNX2           1   164.92 286.72
## + C20orf3          1   164.93 286.73
## + DSG2             1   164.93 286.74
## + BTD              1   164.93 286.74
## + AFM              1   164.93 286.74
## + PLTP             1   164.94 286.74
## - SERPINA7         1   176.42 288.07
## - LAMP2            1   180.21 291.86
##
## Step:  AIC=277.14
## Class ~ AIAG.Bovine + CLU + DKFZp686N02209 + FCGBP + FETUA.Bovine +
##     FHR3 + HP + HRG + IGFBP3 + IGHA2 + IGHG2 + KNG1 + LAMP2 +
##     LGALS3BP + LUM + MMRN1 + NCAM1 + PRG4 + SERPINA7 + THBS1 +
##     VWF
##
##                   Df Deviance    AIC
## - IGFBP3           1   165.85 272.43
## - HRG              1   166.02 272.60
## - THBS1            1   166.08 272.66
## - AIAG.Bovine      1   166.24 272.82
## - CLU              1   166.29 272.87
## - VWF              1   166.35 272.92
## - FCGBP            1   166.50 273.07
## - LGALS3BP         1   166.63 273.21
## - IGHA2            1   166.69 273.27
## - KNG1             1   166.81 273.39
## - PRG4             1   166.83 273.41
## - IGHG2            1   167.58 274.16
## - DKFZp686N02209   1   167.68 274.26
## - NCAM1            1   167.98 274.56
## - FHR3             1   168.42 275.00
## - FETUA.Bovine     1   169.17 275.75
## - HP               1   169.68 276.26
## <none>                 165.48 277.14
## - LUM              1   171.48 278.06
## - MMRN1            1   172.93 279.50
## + F5               1   164.94 281.67
## + LRG1             1   165.09 281.81
## + F11              1   165.12 281.85
## + TIMP1            1   165.18 281.91
## + AHSG             1   165.25 281.98
## + CFI              1   165.25 281.98
## + FETUB            1   165.26 281.99
## + LYVE1            1   165.36 282.09
## + FN1              1   165.37 282.10
```

```
## + MST1            1   165.37 282.10
## + CD44            1   165.40 282.13
## + PROC            1   165.40 282.13
## + APOB            1   165.42 282.15
## + Q5JNX2          1   165.47 282.20
## + DSG2            1   165.48 282.21
## + BTD             1   165.48 282.21
## + AFM             1   165.48 282.21
## + C20orf3         1   165.48 282.21
## + PLTP            1   165.48 282.21
## - SERPINA7        1   176.89 283.47
## - LAMP2           1   180.37 286.95
##
## Step:  AIC=272.42
## Class ~ AIAG.Bovine + CLU + DKFZp686N02209 + FCGBP + FETUA.Bovine +
##     FHR3 + HP + HRG + IGHA2 + IGHG2 + KNG1 + LAMP2 + LGALS3BP +
##     LUM + MMRN1 + NCAM1 + PRG4 + SERPINA7 + THBS1 + VWF
##
##                  Df Deviance    AIC
## - THBS1           1   166.45 267.95
## - HRG             1   166.47 267.98
## - VWF             1   166.56 268.07
## - AIAG.Bovine     1   166.66 268.16
## - CLU             1   166.67 268.17
## - FCGBP           1   166.79 268.29
## - IGHA2           1   166.97 268.47
## - LGALS3BP        1   166.98 268.49
## - PRG4            1   167.32 268.82
## - KNG1            1   167.62 269.13
## - DKFZp686N02209  1   167.72 269.23
## - NCAM1           1   168.13 269.63
## - IGHG2           1   168.34 269.84
## - FHR3            1   168.59 270.09
## - FETUA.Bovine    1   169.81 271.31
## - HP              1   169.94 271.45
## <none>                165.85 272.43
## - LUM             1   171.51 273.02
## - MMRN1           1   173.24 274.74
## + LRG1            1   165.30 276.95
## + IGFBP3          1   165.48 277.14
## + F5              1   165.55 277.20
## + CFI             1   165.57 277.22
## + TIMP1           1   165.59 277.25
## + FETUB           1   165.60 277.25
## + AHSG            1   165.62 277.27
## + F11             1   165.63 277.28
## + MST1            1   165.66 277.32
## + APOB            1   165.69 277.34
## + LYVE1           1   165.76 277.42
## + CD44            1   165.80 277.45
```

```
## + BTD              1   165.81 277.46
## + FN1              1   165.81 277.46
## + DSG2             1   165.82 277.47
## + AFM              1   165.83 277.49
## + C20orf3          1   165.84 277.49
## + PROC             1   165.84 277.49
## + PLTP             1   165.84 277.50
## + Q5JNX2           1   165.85 277.50
## - SERPINA7         1   177.06 278.56
## - LAMP2            1   180.96 282.46
##
## Step:  AIC=267.95
## Class ~ AIAG.Bovine + CLU + DKFZp686N02209 + FCGBP + FETUA.Bovine +
##     FHR3 + HP + HRG + IGHA2 + IGHG2 + KNG1 + LAMP2 + LGALS3BP +
##     LUM + MMRN1 + NCAM1 + PRG4 + SERPINA7 + VWF
##
##                   Df Deviance    AIC
## - HRG              1   166.95 263.38
## - AIAG.Bovine      1   167.07 263.50
## - CLU              1   167.10 263.52
## - LGALS3BP         1   167.52 263.95
## - FCGBP            1   167.61 264.04
## - IGHA2            1   167.70 264.13
## - PRG4             1   167.91 264.34
## - VWF              1   167.96 264.39
## - KNG1             1   168.05 264.48
## - DKFZp686N02209   1   168.07 264.50
## - NCAM1            1   168.63 265.06
## - IGHG2            1   168.76 265.19
## - FHR3             1   169.50 265.93
## - FETUA.Bovine     1   169.93 266.36
## - HP               1   170.16 266.58
## <none>                 166.45 267.95
## - LUM              1   171.87 268.30
## - MMRN1            1   173.25 269.68
## + LRG1             1   165.72 272.30
## + THBS1            1   165.85 272.43
## + IGFBP3           1   166.08 272.66
## + AHSG             1   166.10 272.68
## + MST1             1   166.17 272.75
## + CFI              1   166.23 272.81
## + APOB             1   166.24 272.81
## + F5               1   166.24 272.82
## + FETUB            1   166.25 272.83
## + LYVE1            1   166.36 272.94
## + F11              1   166.36 272.94
## + CD44             1   166.37 272.95
## + AFM              1   166.38 272.96
## + BTD              1   166.39 272.97
## + FN1              1   166.41 272.99
```

```
## + TIMP1           1   166.42 273.00
## + C20orf3         1   166.45 273.02
## + PLTP            1   166.45 273.02
## + DSG2            1   166.45 273.03
## + Q5JNX2          1   166.45 273.03
## + PROC            1   166.45 273.03
## - SERPINA7        1   177.73 274.16
## - LAMP2           1   181.77 278.20
##
## Step:  AIC=263.38
## Class ~ AIAG.Bovine + CLU + DKFZp686N02209 + FCGBP + FETUA.Bovine +
##     FHR3 + HP + IGHA2 + IGHG2 + KNG1 + LAMP2 + LGALS3BP + LUM +
##     MMRN1 + NCAM1 + PRG4 + SERPINA7 + VWF
##
##                   Df Deviance    AIC
## - AIAG.Bovine     1   167.38 258.73
## - CLU             1   167.67 259.02
## - LGALS3BP        1   167.81 259.16
## - IGHA2           1   168.09 259.44
## - FCGBP           1   168.32 259.68
## - PRG4            1   168.38 259.73
## - VWF             1   168.42 259.77
## - KNG1            1   168.75 260.10
## - DKFZp686N02209  1   168.80 260.15
## - NCAM1           1   168.95 260.30
## - IGHG2           1   169.53 260.89
## - FHR3            1   169.76 261.12
## - FETUA.Bovine    1   170.46 261.82
## - HP              1   170.78 262.13
## <none>                166.95 263.38
## - LUM             1   173.86 265.21
## - MMRN1           1   175.43 266.78
## + LRG1            1   166.07 267.57
## + HRG             1   166.45 267.95
## + AHSG            1   166.46 267.96
## + THBS1           1   166.47 267.98
## + IGFBP3          1   166.50 268.00
## + MST1            1   166.51 268.02
## + APOB            1   166.76 268.27
## + FETUB           1   166.77 268.27
## + CFI             1   166.80 268.31
## + AFM             1   166.83 268.33
## + BTD             1   166.84 268.34
## + TIMP1           1   166.86 268.36
## + LYVE1           1   166.86 268.37
## + F5              1   166.88 268.38
## + FN1             1   166.90 268.40
## + CD44            1   166.92 268.42
## + DSG2            1   166.93 268.43
## + PLTP            1   166.93 268.43
```

```
## + F11              1   166.93 268.43
## + PROC             1   166.94 268.44
## + Q5JNX2           1   166.95 268.45
## + C20orf3          1   166.95 268.46
## - SERPINA7         1   178.16 269.51
## - LAMP2            1   182.56 273.91
##
## Step:  AIC=258.73
## Class ~ CLU + DKFZp686N02209 + FCGBP + FETUA.Bovine + FHR3 +
##     HP + IGHA2 + IGHG2 + KNG1 + LAMP2 + LGALS3BP + LUM + MMRN1 +
##     NCAM1 + PRG4 + SERPINA7 + VWF
##
##                   Df Deviance    AIC
## - CLU              1   167.84 254.12
## - LGALS3BP         1   168.28 254.56
## - IGHA2            1   168.53 254.81
## - FCGBP            1   168.62 254.90
## - PRG4             1   168.67 254.94
## - VWF              1   168.68 254.96
## - DKFZp686N02209   1   168.98 255.26
## - KNG1             1   169.11 255.39
## - NCAM1            1   169.40 255.67
## - FHR3             1   169.91 256.19
## - IGHG2            1   170.04 256.32
## - FETUA.Bovine     1   170.81 257.09
## - HP               1   171.16 257.44
## <none>                 167.38 258.73
## - LUM              1   175.29 261.57
## - MMRN1            1   175.55 261.83
## + LRG1             1   166.62 263.05
## + IGFBP3           1   166.90 263.33
## + AIAG.Bovine      1   166.95 263.38
## + APOB             1   167.01 263.44
## + MST1             1   167.02 263.45
## + CFI              1   167.02 263.45
## + THBS1            1   167.03 263.46
## + AHSG             1   167.03 263.46
## + HRG              1   167.07 263.50
## + FETUB            1   167.09 263.52
## + TIMP1            1   167.20 263.63
## + LYVE1            1   167.26 263.69
## + BTD              1   167.29 263.72
## + F11              1   167.31 263.74
## + PLTP             1   167.32 263.75
## + F5               1   167.35 263.77
## + PROC             1   167.36 263.79
## + CD44             1   167.36 263.79
## + AFM              1   167.37 263.80
## + C20orf3          1   167.37 263.80
## + DSG2             1   167.37 263.80
```

```
## + Q5JNX2          1   167.37 263.80
## + FN1             1   167.38 263.81
## - SERPINA7        1   178.17 264.45
## - LAMP2           1   182.74 269.02
##
## Step:  AIC=254.12
## Class ~ DKFZp686N02209 + FCGBP + FETUA.Bovine + FHR3 + HP + IGHA2 +
##     IGHG2 + KNG1 + LAMP2 + LGALS3BP + LUM + MMRN1 + NCAM1 + PRG4 +
##     SERPINA7 + VWF
##
##                  Df Deviance    AIC
## - LGALS3BP        1   168.59 249.79
## - PRG4            1   168.75 249.95
## - VWF             1   168.79 250.00
## - FCGBP           1   169.00 250.20
## - IGHA2           1   169.18 250.38
## - KNG1            1   169.25 250.46
## - DKFZp686N02209  1   169.42 250.62
## - FHR3            1   170.12 251.33
## - NCAM1           1   170.51 251.71
## - IGHG2           1   170.83 252.03
## - FETUA.Bovine    1   171.21 252.41
## - HP              1   171.56 252.76
## <none>                167.84 254.12
## - LUM             1   175.50 256.70
## - MMRN1           1   175.86 257.07
## + LRG1            1   166.87 258.22
## + IGFBP3          1   167.36 258.71
## + CLU             1   167.38 258.73
## + MST1            1   167.41 258.77
## + HRG             1   167.43 258.79
## + CFI             1   167.46 258.81
## + AHSG            1   167.52 258.87
## + THBS1           1   167.58 258.93
## + APOB            1   167.65 259.00
## + AIAG.Bovine     1   167.67 259.02
## + PLTP            1   167.68 259.04
## + TIMP1           1   167.72 259.07
## + LYVE1           1   167.72 259.07
## + F11             1   167.72 259.07
## + FETUB           1   167.72 259.07
## + PROC            1   167.77 259.12
## + F5              1   167.78 259.13
## + BTD             1   167.80 259.15
## + AFM             1   167.81 259.17
## + Q5JNX2          1   167.82 259.17
## + DSG2            1   167.82 259.17
## + CD44            1   167.83 259.18
## + C20orf3         1   167.84 259.19
## + FN1             1   167.84 259.19
```

```
## - SERPINA7        1   179.39 260.59
## - LAMP2           1   182.96 264.17
##
## Step:  AIC=249.79
## Class ~ DKFZp686N02209 + FCGBP + FETUA.Bovine + FHR3 + HP + IGHA2 +
##     IGHG2 + KNG1 + LAMP2 + LUM + MMRN1 + NCAM1 + PRG4 + SERPINA7 +
##     VWF
##
##                  Df Deviance    AIC
## - PRG4            1   169.20 245.33
## - VWF             1   169.29 245.42
## - IGHA2           1   169.52 245.64
## - DKFZp686N02209  1   170.10 246.22
## - KNG1            1   170.11 246.24
## - FCGBP           1   170.44 246.57
## - NCAM1           1   170.79 246.91
## - IGHG2           1   171.56 247.69
## - FHR3            1   171.57 247.70
## - FETUA.Bovine    1   171.95 248.08
## - HP              1   172.75 248.87
## <none>                168.59 249.79
## - LUM             1   176.34 252.47
## - MMRN1           1   177.59 253.71
## + LGALS3BP        1   167.84 254.12
## + LRG1            1   167.86 254.14
## + CFI             1   168.05 254.33
## + IGFBP3          1   168.13 254.41
## + MST1            1   168.27 254.54
## + CLU             1   168.28 254.56
## + THBS1           1   168.34 254.62
## + AIAG.Bovine     1   168.35 254.63
## + HRG             1   168.38 254.66
## + AHSG            1   168.39 254.67
## + FETUB           1   168.42 254.70
## + LYVE1           1   168.42 254.70
## + TIMP1           1   168.44 254.71
## + APOB            1   168.45 254.72
## + F5              1   168.46 254.74
## + PLTP            1   168.51 254.79
## + C20orf3         1   168.53 254.81
## + F11             1   168.53 254.81
## + PROC            1   168.55 254.83
## + DSG2            1   168.56 254.84
## + Q5JNX2          1   168.56 254.84
## + BTD             1   168.57 254.85
## + FN1             1   168.58 254.86
## + AFM             1   168.58 254.86
## + CD44            1   168.59 254.86
## - SERPINA7        1   179.98 256.11
## - LAMP2           1   182.97 259.09
```

```
##
## Step:  AIC=245.33
## Class ~ DKFZp686N02209 + FCGBP + FETUA.Bovine + FHR3 + HP + IGHA2 +
##     IGHG2 + KNG1 + LAMP2 + LUM + MMRN1 + NCAM1 + SERPINA7 + VWF
##
##                   Df Deviance    AIC
## - VWF              1   169.83 240.88
## - IGHA2            1   170.08 241.14
## - DKFZp686N02209   1   170.49 241.54
## - KNG1             1   170.83 241.88
## - NCAM1            1   171.28 242.34
## - FCGBP            1   171.36 242.41
## - IGHG2            1   172.06 243.11
## - FETUA.Bovine     1   172.16 243.22
## - FHR3             1   172.25 243.30
## - HP               1   173.15 244.20
## <none>                 169.20 245.33
## - MMRN1            1   178.13 249.18
## + LRG1             1   168.30 249.51
## + PRG4             1   168.59 249.79
## - LUM              1   178.78 249.84
## + IGFBP3           1   168.65 249.85
## + LGALS3BP         1   168.75 249.95
## + CFI              1   168.76 249.97
## + APOB             1   168.86 250.07
## + LYVE1            1   168.88 250.08
## + F5               1   168.90 250.10
## + TIMP1            1   168.90 250.10
## + THBS1            1   168.91 250.11
## + FETUB            1   168.91 250.12
## + MST1             1   168.94 250.14
## + AIAG.Bovine      1   168.96 250.16
## + HRG              1   169.00 250.20
## + AHSG             1   169.07 250.27
## + PLTP             1   169.07 250.27
## + PROC             1   169.11 250.32
## + BTD              1   169.13 250.34
## + CLU              1   169.14 250.34
## + F11              1   169.17 250.37
## + Q5JNX2           1   169.19 250.39
## + FN1              1   169.19 250.39
## + CD44             1   169.19 250.40
## + AFM              1   169.20 250.40
## + C20orf3          1   169.20 250.40
## + DSG2             1   169.20 250.40
## - SERPINA7         1   180.12 251.18
## - LAMP2            1   185.62 256.67
##
## Step:  AIC=240.88
## Class ~ DKFZp686N02209 + FCGBP + FETUA.Bovine + FHR3 + HP + IGHA2 +
```

```
##     IGHG2 + KNG1 + LAMP2 + LUM + MMRN1 + NCAM1 + SERPINA7
##
##                   Df Deviance    AIC
## - IGHA2            1   170.75 236.73
## - KNG1             1   171.25 237.22
## - DKFZp686N02209   1   171.44 237.42
## - NCAM1            1   171.49 237.47
## - FCGBP            1   171.56 237.54
## - IGHG2            1   172.32 238.30
## - FHR3             1   172.50 238.48
## - FETUA.Bovine     1   172.51 238.48
## - HP               1   173.86 239.84
## <none>                 169.83 240.88
## - MMRN1            1   178.13 244.11
## + LRG1             1   168.68 244.80
## - LUM              1   179.27 245.24
## + THBS1            1   169.13 245.26
## + VWF              1   169.20 245.33
## + PRG4             1   169.29 245.42
## + LYVE1            1   169.44 245.57
## + APOB             1   169.47 245.60
## + AHSG             1   169.47 245.60
## + CFI              1   169.50 245.62
## + IGFBP3           1   169.50 245.63
## + FETUB            1   169.50 245.63
## + LGALS3BP         1   169.53 245.66
## + MST1             1   169.58 245.71
## + AIAG.Bovine      1   169.59 245.72
## + HRG              1   169.63 245.75
## + F5               1   169.64 245.76
## + BTD              1   169.68 245.81
## + PLTP             1   169.70 245.82
## + FN1              1   169.71 245.84
## + PROC             1   169.71 245.84
## + TIMP1            1   169.76 245.88
## + DSG2             1   169.80 245.93
## + F11              1   169.80 245.93
## + C20orf3          1   169.81 245.94
## + Q5JNX2           1   169.81 245.94
## + CD44             1   169.82 245.95
## + AFM              1   169.82 245.95
## + CLU              1   169.83 245.95
## - SERPINA7         1   180.14 246.12
## - LAMP2            1   187.37 253.35
##
## Step:  AIC=236.73
## Class ~ DKFZp686N02209 + FCGBP + FETUA.Bovine + FHR3 + HP + IGHG2 +
##     KNG1 + LAMP2 + LUM + MMRN1 + NCAM1 + SERPINA7
##
##                   Df Deviance    AIC
```

```
## - NCAM1           1   172.03 232.93
## - FCGBP           1   172.03 232.94
## - DKFZp686N02209  1   172.18 233.08
## - KNG1            1   172.75 233.65
## - IGHG2           1   173.04 233.94
## - FETUA.Bovine    1   173.25 234.15
## - FHR3            1   174.10 235.00
## - HP              1   174.31 235.21
## <none>                170.75 236.73
## - MMRN1           1   179.33 240.23
## - LUM             1   179.49 240.39
## + LRG1            1   169.53 240.58
## + IGHA2           1   169.83 240.88
## + THBS1           1   169.97 241.02
## + VWF             1   170.08 241.14
## + AHSG            1   170.27 241.32
## + PRG4            1   170.27 241.32
## + FETUB           1   170.43 241.48
## + APOB            1   170.47 241.52
## + LYVE1           1   170.52 241.57
## + IGFBP3          1   170.53 241.58
## + CFI             1   170.54 241.59
## + HRG             1   170.54 241.60
## + AIAG.Bovine     1   170.56 241.61
## + MST1            1   170.58 241.63
## + FN1             1   170.62 241.68
## + PLTP            1   170.63 241.69
## + BTD             1   170.64 241.69
## + PROC            1   170.66 241.71
## + C20orf3         1   170.66 241.72
## + LGALS3BP        1   170.68 241.73
## + F5              1   170.68 241.73
## + CD44            1   170.69 241.75
## + AFM             1   170.69 241.75
## + Q5JNX2          1   170.72 241.77
## + CLU             1   170.73 241.78
## + F11             1   170.74 241.79
## + TIMP1           1   170.74 241.79
## + DSG2            1   170.75 241.80
## - SERPINA7        1   181.11 242.01
## - LAMP2           1   188.15 249.05
##
## Step:  AIC=232.93
## Class ~ DKFZp686N02209 + FCGBP + FETUA.Bovine + FHR3 + HP + IGHG2 +
##     KNG1 + LAMP2 + LUM + MMRN1 + SERPINA7
##
##                   Df Deviance    AIC
## - DKFZp686N02209  1   172.97 228.80
## - KNG1            1   173.88 229.71
## - IGHG2           1   174.07 229.90
```

```
## - FCGBP           1   174.61 230.44
## - FETUA.Bovine    1   174.72 230.54
## - FHR3            1   175.05 230.88
## - HP              1   175.70 231.53
## <none>                172.03 232.93
## - LUM             1   179.92 235.75
## + NCAM1           1   170.75 236.73
## + LRG1            1   171.06 237.04
## - MMRN1           1   181.53 237.36
## + IGHA2           1   171.49 237.47
## + THBS1           1   171.59 237.57
## + PRG4            1   171.59 237.57
## + F5              1   171.69 237.67
## + AHSG            1   171.75 237.72
## + CFI             1   171.76 237.74
## + LYVE1           1   171.77 237.74
## + VWF             1   171.77 237.75
## + CLU             1   171.78 237.75
## - SERPINA7        1   181.94 237.77
## + PLTP            1   171.81 237.79
## + F11             1   171.85 237.83
## + HRG             1   171.85 237.83
## + Q5JNX2          1   171.87 237.84
## + FETUB           1   171.87 237.85
## + IGFBP3          1   171.89 237.87
## + AIAG.Bovine     1   171.90 237.88
## + PROC            1   171.90 237.88
## + MST1            1   171.91 237.89
## + FN1             1   171.93 237.91
## + APOB            1   171.94 237.91
## + C20orf3         1   171.94 237.91
## + DSG2            1   171.95 237.92
## + TIMP1           1   171.97 237.95
## + LGALS3BP        1   172.00 237.97
## + BTD             1   172.01 237.99
## + AFM             1   172.01 237.99
## + CD44            1   172.02 237.99
## - LAMP2           1   188.15 243.98
##
## Step:  AIC=228.8
## Class ~ FCGBP + FETUA.Bovine + FHR3 + HP + IGHG2 + KNG1 + LAMP2 +
##     LUM + MMRN1 + SERPINA7
##
##                  Df Deviance    AIC
## - IGHG2           1   174.09 224.84
## - KNG1            1   174.52 225.27
## - FCGBP           1   175.38 226.13
## - FETUA.Bovine    1   175.40 226.15
## - FHR3            1   175.49 226.24
## - HP              1   176.24 226.99
```

```
## <none>                172.97 228.80
## - LUM             1   180.40 231.16
## + DKFZp686N02209  1   172.03 232.93
## + LRG1            1   172.03 232.94
## + NCAM1           1   172.18 233.08
## + IGHA2           1   172.47 233.37
## + AHSG            1   172.48 233.38
## + VWF             1   172.49 233.39
## + THBS1           1   172.59 233.49
## - MMRN1           1   182.74 233.50
## + HRG             1   172.64 233.54
## + FETUB           1   172.69 233.59
## + LYVE1           1   172.70 233.60
## + PRG4            1   172.70 233.60
## + F5              1   172.74 233.65
## + FN1             1   172.76 233.67
## + CLU             1   172.79 233.69
## + C20orf3         1   172.82 233.72
## + CFI             1   172.83 233.74
## + AFM             1   172.87 233.77
## + TIMP1           1   172.88 233.78
## + F11             1   172.89 233.79
## + CD44            1   172.90 233.80
## + PLTP            1   172.91 233.81
## + BTD             1   172.92 233.82
## + AIAG.Bovine     1   172.93 233.83
## + LGALS3BP        1   172.93 233.84
## + MST1            1   172.94 233.84
## + PROC            1   172.95 233.86
## + APOB            1   172.96 233.86
## + DSG2            1   172.96 233.86
## + IGFBP3          1   172.96 233.87
## + Q5JNX2          1   172.97 233.87
## - SERPINA7        1   184.85 235.60
## - LAMP2           1   191.38 242.13
##
## Step:  AIC=224.84
## Class ~ FCGBP + FETUA.Bovine + FHR3 + HP + KNG1 + LAMP2 + LUM +
##     MMRN1 + SERPINA7
##
##                  Df Deviance    AIC
## - KNG1            1   176.13 221.80
## - FETUA.Bovine    1   176.27 221.94
## - FCGBP           1   176.56 222.23
## - FHR3            1   176.71 222.38
## - HP              1   177.90 223.57
## <none>                174.09 224.84
## + LRG1            1   172.81 228.63
## + IGHG2           1   172.97 228.80
## + NCAM1           1   173.16 228.99
```

```
## - MMRN1           1   183.67 229.35
## + CFI             1   173.65 229.48
## + IGHA2           1   173.66 229.49
## + CLU             1   173.67 229.49
## + F5              1   173.70 229.52
## + HRG             1   173.71 229.54
## + PRG4            1   173.74 229.57
## + IGFBP3          1   173.76 229.59
## + AHSG            1   173.79 229.61
## + MST1            1   173.79 229.62
## + PROC            1   173.80 229.62
## + LYVE1           1   173.81 229.63
## + PLTP            1   173.84 229.67
## + F11             1   173.87 229.69
## + DSG2            1   173.88 229.71
## + TIMP1           1   173.92 229.75
## + THBS1           1   173.93 229.75
## + FN1             1   173.96 229.78
## + AIAG.Bovine     1   173.97 229.80
## + VWF             1   173.98 229.81
## + APOB            1   174.02 229.84
## + BTD             1   174.02 229.85
## + LGALS3BP        1   174.02 229.85
## + FETUB           1   174.03 229.85
## + AFM             1   174.07 229.89
## + Q5JNX2          1   174.07 229.90
## + DKFZp686N02209  1   174.07 229.90
## + C20orf3         1   174.08 229.90
## + CD44            1   174.08 229.91
## - LUM             1   184.52 230.20
## - SERPINA7        1   184.99 230.67
## - LAMP2           1   191.38 237.06
##
## Step:  AIC=221.8
## Class ~ FCGBP + FETUA.Bovine + FHR3 + HP + LAMP2 + LUM + MMRN1 +
##     SERPINA7
##
##                  Df Deviance    AIC
## - FETUA.Bovine    1   178.02 218.62
## - FCGBP           1   178.21 218.81
## - FHR3            1   178.54 219.14
## - HP              1   179.98 220.59
## <none>                176.13 221.80
## + KNG1            1   174.09 224.84
## + IGHG2           1   174.52 225.27
## - MMRN1           1   184.67 225.27
## + LRG1            1   174.62 225.37
## + IGFBP3          1   175.13 225.88
## + NCAM1           1   175.19 225.94
## + IGHA2           1   175.26 226.01
```

```
## + CFI              1   175.31 226.07
## - SERPINA7         1   185.67 226.27
## + PROC             1   175.54 226.29
## + PRG4             1   175.63 226.38
## + HRG              1   175.69 226.44
## + MST1             1   175.71 226.46
## + PLTP             1   175.71 226.46
## + DSG2             1   175.75 226.51
## + F5               1   175.81 226.56
## + TIMP1            1   175.83 226.58
## + BTD              1   175.90 226.65
## + AIAG.Bovine      1   175.91 226.66
## + AHSG             1   175.92 226.68
## + CLU              1   175.92 226.68
## + APOB             1   175.94 226.70
## + LYVE1            1   175.96 226.71
## + F11              1   176.04 226.79
## + LGALS3BP         1   176.04 226.80
## + CD44             1   176.06 226.81
## + THBS1            1   176.07 226.82
## + Q5JNX2           1   176.07 226.83
## + FETUB            1   176.09 226.85
## + DKFZp686N02209   1   176.10 226.85
## + AFM              1   176.11 226.86
## + FN1              1   176.11 226.86
## + VWF              1   176.12 226.87
## + C20orf3          1   176.12 226.87
## - LUM              1   188.00 228.61
## - LAMP2            1   194.42 235.02
##
## Step:  AIC=218.62
## Class ~ FCGBP + FHR3 + HP + LAMP2 + LUM + MMRN1 + SERPINA7
##
##                   Df Deviance    AIC
## - FCGBP            1   179.68 215.21
## - FHR3             1   180.49 216.02
## - HP               1   181.82 217.35
## <none>                 178.02 218.62
## - MMRN1            1   186.12 221.65
## + FETUA.Bovine     1   176.13 221.80
## + KNG1             1   176.27 221.94
## + LRG1             1   176.50 222.18
## - SERPINA7         1   186.69 222.21
## + IGHG2            1   176.72 222.40
## + CFI              1   176.79 222.47
## + IGFBP3           1   176.87 222.55
## + NCAM1            1   176.89 222.56
## + IGHA2            1   177.31 222.99
## + TIMP1            1   177.39 223.07
## + DSG2             1   177.50 223.18
```

```
## + PROC            1   177.51 223.19
## + HRG             1   177.56 223.24
## + MST1            1   177.58 223.26
## + CLU             1   177.73 223.41
## + F5              1   177.79 223.47
## + PRG4            1   177.80 223.48
## + AIAG.Bovine     1   177.81 223.49
## + PLTP            1   177.83 223.51
## + APOB            1   177.88 223.56
## + AHSG            1   177.88 223.56
## + LGALS3BP        1   177.88 223.56
## + LYVE1           1   177.89 223.56
## + CD44            1   177.91 223.58
## + F11             1   177.93 223.60
## + BTD             1   177.93 223.60
## + C20orf3         1   177.98 223.65
## + DKFZp686N02209  1   177.98 223.66
## + FETUB           1   178.00 223.67
## + FN1             1   178.01 223.69
## + Q5JNX2          1   178.01 223.69
## + AFM             1   178.01 223.69
## + THBS1           1   178.01 223.69
## + VWF             1   178.01 223.69
## - LUM             1   189.59 225.12
## - LAMP2           1   195.13 230.65
##
## Step:  AIC=215.21
## Class ~ FHR3 + HP + LAMP2 + LUM + MMRN1 + SERPINA7
##
##                  Df Deviance    AIC
## - FHR3            1   182.66 213.11
## - HP              1   183.18 213.63
## <none>                179.68 215.21
## + NCAM1           1   177.60 218.20
## + LRG1            1   177.99 218.59
## + FCGBP           1   178.02 218.62
## + FETUA.Bovine    1   178.21 218.81
## + KNG1            1   178.23 218.83
## + IGHG2           1   178.32 218.93
## + CFI             1   178.44 219.04
## - SERPINA7        1   188.62 219.07
## + IGFBP3          1   178.67 219.27
## + TIMP1           1   178.73 219.33
## + DSG2            1   178.74 219.34
## + F5              1   179.06 219.66
## + PROC            1   179.08 219.68
## + HRG             1   179.11 219.71
## + LGALS3BP        1   179.14 219.75
## + PRG4            1   179.22 219.82
## + CD44            1   179.31 219.92
```

```
## + MST1             1   179.33 219.93
## + CLU              1   179.35 219.95
## + IGHA2            1   179.47 220.07
## + VWF              1   179.51 220.11
## + C20orf3          1   179.51 220.11
## + Q5JNX2           1   179.51 220.11
## + APOB             1   179.53 220.13
## + AIAG.Bovine      1   179.54 220.14
## + BTD              1   179.57 220.17
## + PLTP             1   179.61 220.21
## + AHSG             1   179.61 220.21
## + LYVE1            1   179.62 220.22
## + DKFZp686N02209   1   179.63 220.23
## + AFM              1   179.64 220.24
## + FETUB            1   179.64 220.24
## + F11              1   179.66 220.26
## + FN1              1   179.67 220.27
## + THBS1            1   179.68 220.28
## - MMRN1            1   190.04 220.49
## - LUM              1   190.56 221.01
## - LAMP2            1   195.60 226.05
##
## Step:  AIC=213.11
## Class ~ HP + LAMP2 + LUM + MMRN1 + SERPINA7
##
##                  Df Deviance    AIC
## <none>                182.66 213.11
## - HP              1   187.93 213.31
## + LRG1            1   179.58 215.11
## + FHR3            1   179.68 215.21
## - SERPINA7        1   190.60 215.97
## + FCGBP           1   180.49 216.02
## + NCAM1           1   180.54 216.07
## + CFI             1   180.68 216.21
## + DSG2            1   180.99 216.51
## + FETUA.Bovine    1   181.19 216.71
## + IGHG2           1   181.23 216.75
## + F5              1   181.37 216.89
## + KNG1            1   181.43 216.96
## + LGALS3BP        1   181.55 217.07
## + CD44            1   181.57 217.09
## + TIMP1           1   181.75 217.27
## + IGFBP3          1   181.85 217.38
## + PRG4            1   181.91 217.44
## + MST1            1   182.04 217.56
## + Q5JNX2          1   182.14 217.67
## + PROC            1   182.15 217.68
## + VWF             1   182.16 217.69
## + IGHA2           1   182.17 217.70
## + C20orf3         1   182.34 217.86
```

```
## + F11              1   182.37 217.89
## + HRG              1   182.43 217.96
## + DKFZp686N02209   1   182.43 217.96
## + APOB             1   182.46 217.98
## + BTD              1   182.46 217.99
## + CLU              1   182.50 218.03
## + AFM              1   182.55 218.08
## + LYVE1            1   182.58 218.10
## + AIAG.Bovine      1   182.59 218.12
## - LUM              1   192.75 218.13
## + THBS1            1   182.64 218.17
## + PLTP             1   182.64 218.17
## + AHSG             1   182.65 218.18
## + FN1              1   182.66 218.18
## + FETUB            1   182.66 218.18
## - MMRN1            1   194.63 220.01
## - LAMP2            1   197.82 223.19
```

```
summary(model.sbc.both)
```

```
##
## Call:
## glm(formula = Class ~ HP + LAMP2 + LUM + MMRN1 + SERPINA7, family =
binomial,
##     data = train)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -2.0343  -0.9466  -0.1179   0.9503   2.4615
##
## Coefficients:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)   0.9173     7.7240   0.119 0.905464
## HP           -0.6052     0.2700  -2.242 0.024978 *
## LAMP2         2.2758     0.6206   3.667 0.000245 ***
## LUM           1.2678     0.4145   3.058 0.002225 **
## MMRN1        -1.4659     0.4496  -3.260 0.001112 **
## SERPINA7     -1.6917     0.6187  -2.734 0.006250 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 221.81  on 159  degrees of freedom
## Residual deviance: 182.66  on 154  degrees of freedom
## AIC: 194.66
##
## Number of Fisher Scoring iterations: 4
```

As expected, the BIC selection results in a parsimonious model with 5 columns with SBC = 213.11

We select the top 2 models in each of the above selections as our candidate models.

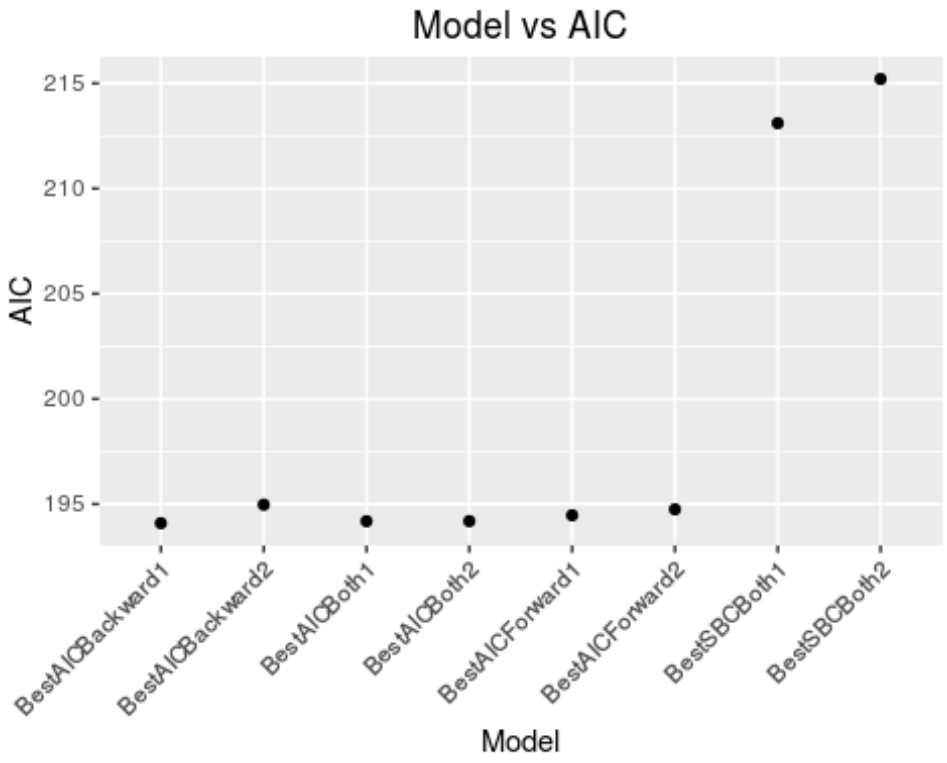
```
formula.BestAICBackward1 <- model.aic.backward$formula
formula.BestAICBackward2 <- as.formula(Class ~ FCGBP + FETUA.Bovine +
FHR3 + HP + IGHG2 + KNG1 + LAMP2 + LUM + MMRN1 + SERPINA7)

formula.BestAICForward1 <- model.aic.forward$formula
formula.BestAICForward2 <- as.formula(Class ~ TIMP1 + LAMP2 + HP + LRG1
+ SERPINA7 + LUM)

formula.BestAICBoth1 <- model.aic.both$formula
formula.BestAICBoth2 <- as.formula(Class ~  TIMP1 + LAMP2 + LRG1 +
SERPINA7 + LUM + NCAM1)

formula.BestSBCBoth1 <- model.sbc.both$formula
formula.BestSBCBoth2 <- as.formula(Class ~ FHR3 + HP + LAMP2 + LUM +
MMRN1 + SERPINA7)

# Plot the Candidate Models' AIC Values
candidateModels.aic <- data.frame(model <- c("BestAICBackward1",
"BestAICBackward2", "BestAICForward1", "BestAICForward2",
"BestAICBoth1", "BestAICBoth2", "BestSBCBoth1", "BestSBCBoth2"), AIC <-
c(model.aic.backward$aic, 194.97, model.aic.forward$aic, 194.75,
model.aic.both$aic, 194.19, 213.11, 215.21))
colnames(candidateModels.aic) <- c("Model", "AIC")
ggplot(candidateModels.aic, aes(x=Model, y=AIC)) + geom_point() +
theme(axis.text.x = element_text(angle = 45, hjust = 1)) +
ggtitle("Model vs AIC")
```



We can ignore the AIC values for BestSBCBoth1 and BestSBCBoth2 in the above plot as it considers the SBC criterion instead.

Candidate Model Fitting

Train Logistic Regression on the candidate Models

```
model.BestAICBackward1 <- glm(formula.BestAICBackward1,
family=binomial(link='logit'), data=train)
summary(model.BestAICBackward1)

##
## Call:
## glm(formula = formula.BestAICBackward1, family = binomial(link =
"logit"),
##     data = train)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -1.8946  -0.8822  -0.1143   0.9205   2.7140
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  122.1951    83.8466   1.457 0.145016
## FCGBP         -0.5803     0.3745  -1.550 0.121184
## FETUA.Bovine  -7.0967     4.8653  -1.459 0.144665
```

```
## FHR3          -0.2556     0.1604  -1.594 0.110989
## HP            -0.5438     0.2848  -1.910 0.056190 .
## KNG1           0.4819     0.3768   1.279 0.200999
## LAMP2          2.5943     0.6818   3.805 0.000142 ***
## LUM            1.3460     0.4363   3.085 0.002037 **
## MMRN1         -1.4466     0.4904  -2.950 0.003180 **
## SERPINA7      -2.1277     0.6782  -3.137 0.001705 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 221.81  on 159  degrees of freedom
## Residual deviance: 174.09  on 150  degrees of freedom
## AIC: 194.09
##
## Number of Fisher Scoring iterations: 4
```

```
model.BestAICBackward2 <- glm(formula.BestAICBackward2,
family=binomial(link='logit'), data=train)
summary(model.BestAICBackward2)
```

```
##
## Call:
## glm(formula = formula.BestAICBackward2, family = binomial(link =
"logit"),
##     data = train)
##
## Deviance Residuals:
##      Min        1Q    Median        3Q       Max
## -2.02513  -0.89869  -0.08684   0.90207   2.61653
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  127.2303    84.9235   1.498 0.134088
## FCGBP         -0.5753     0.3757  -1.531 0.125724
## FETUA.Bovine  -7.6121     4.9520  -1.537 0.124247
## FHR3          -0.2513     0.1604  -1.567 0.117129
## HP            -0.5046     0.2849  -1.771 0.076537 .
## IGHG2          0.2867     0.2720   1.054 0.291882
## KNG1           0.4076     0.3541   1.151 0.249671
## LAMP2          2.7823     0.7174   3.878 0.000105 ***
## LUM            1.2016     0.4560   2.635 0.008419 **
## MMRN1         -1.4901     0.5015  -2.971 0.002966 **
## SERPINA7      -2.2947     0.7078  -3.242 0.001186 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
```

```
##     Null deviance: 221.81  on 159  degrees of freedom
## Residual deviance: 172.97  on 149  degrees of freedom
## AIC: 194.97
##
## Number of Fisher Scoring iterations: 4
```

```
model.BestAICForward1 <- glm(formula.BestAICForward1,
family=binomial(link='logit'), data=train)
summary(model.BestAICForward1)
```

```
##
## Call:
## glm(formula = formula.BestAICForward1, family = binomial(link =
"logit"),
##     data = train)
##
## Deviance Residuals:
##      Min        1Q    Median        3Q       Max
## -2.02726  -0.94320  -0.05077   0.93724   2.48858
##
## Coefficients:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)   7.3902     8.2991   0.890  0.37320
## TIMP1        -0.9190     0.4498  -2.043  0.04102 *
## LAMP2         2.7202     0.6946   3.916 8.99e-05 ***
## HP           -0.3650     0.2814  -1.297  0.19463
## LRG1         -0.7993     0.3743  -2.135  0.03273 *
## SERPINA7     -1.9639     0.6460  -3.040  0.00237 **
## LUM           1.2721     0.4218   3.016  0.00256 **
## NCAM1        -0.8055     0.5401  -1.491  0.13585
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 221.81  on 159  degrees of freedom
## Residual deviance: 178.47  on 152  degrees of freedom
## AIC: 194.47
##
## Number of Fisher Scoring iterations: 4
```

```
model.BestAICForward2 <- glm(formula.BestAICForward2,
family=binomial(link='logit'), data=train)
summary(model.BestAICForward2)
```

```
##
## Call:
## glm(formula = formula.BestAICForward2, family = binomial(link =
"logit"),
##     data = train)
##
```

```
## Deviance Residuals:
##      Min        1Q    Median        3Q       Max
## -1.95503  -0.94785  -0.02557   0.95793   2.64048
##
## Coefficients:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)   5.5052     8.1360   0.677 0.498625
## TIMP1        -1.0630     0.4350  -2.444 0.014529 *
## LAMP2         2.3504     0.6358   3.697 0.000218 ***
## HP           -0.3738     0.2792  -1.339 0.180665
## LRG1         -0.7607     0.3698  -2.057 0.039709 *
## SERPINA7     -1.8355     0.6346  -2.893 0.003821 **
## LUM           1.1676     0.4145   2.817 0.004847 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 221.81  on 159  degrees of freedom
## Residual deviance: 180.75  on 153  degrees of freedom
## AIC: 194.75
##
## Number of Fisher Scoring iterations: 4

model.BestAICBoth1 <- glm(formula.BestAICBoth1,
family=binomial(link='logit'), data=train)
summary(model.BestAICBoth1)

##
## Call:
## glm(formula = formula.BestAICBoth1, family = binomial(link =
"logit"),
##     data = train)
##
## Deviance Residuals:
##      Min        1Q    Median        3Q       Max
## -2.14878  -0.93004  -0.02417   0.93892   2.49909
##
## Coefficients:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)   6.1312     8.3810   0.732 0.464434
## TIMP1        -1.0291     0.4422  -2.328 0.019938 *
## LAMP2         2.6904     0.6946   3.873 0.000107 ***
## LRG1         -0.7291     0.3851  -1.893 0.058322 .
## SERPINA7     -2.2482     0.6394  -3.516 0.000438 ***
## LUM           1.3098     0.4240   3.089 0.002006 **
## NCAM1        -0.7839     0.5356  -1.463 0.143332
## FHR3         -0.2323     0.1666  -1.395 0.163052
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 221.81  on 159  degrees of freedom
## Residual deviance: 178.19  on 152  degrees of freedom
## AIC: 194.19
##
## Number of Fisher Scoring iterations: 4
```

```
model.BestAICBoth2 <- glm(formula.BestAICBoth2,
family=binomial(link='logit'), data=train)
summary(model.BestAICBoth2)
```

```
##
## Call:
## glm(formula = formula.BestAICBoth2, family = binomial(link =
"logit"),
##     data = train)
##
## Deviance Residuals:
##      Min        1Q    Median        3Q       Max
## -2.03763  -0.92744  -0.02168   0.94090   2.34990
##
## Coefficients:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)   6.7812     8.2821   0.819 0.412917
## TIMP1        -1.0332     0.4385  -2.356 0.018462 *
## LAMP2         2.6182     0.6768   3.869 0.000109 ***
## LRG1         -0.9055     0.3639  -2.488 0.012832 *
## SERPINA7     -2.1591     0.6246  -3.457 0.000547 ***
## LUM           1.2807     0.4197   3.051 0.002278 **
## NCAM1        -0.8137     0.5327  -1.527 0.126643
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 221.81  on 159  degrees of freedom
## Residual deviance: 180.19  on 153  degrees of freedom
## AIC: 194.19
##
## Number of Fisher Scoring iterations: 4
```

```
model.BestSBCBoth1 <- glm(formula.BestSBCBoth1,
family=binomial(link='logit'), data=train)
summary(model.BestSBCBoth1)
```

```
##
## Call:
## glm(formula = formula.BestSBCBoth1, family = binomial(link =
"logit"),
```

```
##     data = train)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -2.0343  -0.9466  -0.1179   0.9503   2.4615
##
## Coefficients:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)   0.9173     7.7240   0.119 0.905464
## HP           -0.6052     0.2700  -2.242 0.024978 *
## LAMP2         2.2758     0.6206   3.667 0.000245 ***
## LUM           1.2678     0.4145   3.058 0.002225 **
## MMRN1        -1.4659     0.4496  -3.260 0.001112 **
## SERPINA7     -1.6917     0.6187  -2.734 0.006250 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 221.81  on 159  degrees of freedom
## Residual deviance: 182.66  on 154  degrees of freedom
## AIC: 194.66
##
## Number of Fisher Scoring iterations: 4

model.BestSBCBoth2 <- glm(formula.BestSBCBoth2,
family=binomial(link='logit'), data=train)
summary(model.BestSBCBoth2)

##
## Call:
## glm(formula = formula.BestSBCBoth2, family = binomial(link =
"logit"),
##     data = train)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -1.9395  -0.9310  -0.0993   0.9551   2.6557
##
## Coefficients:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)   0.8137     7.8806   0.103 0.917760
## FHR3         -0.2700     0.1593  -1.695 0.090084 .
## HP           -0.5108     0.2774  -1.842 0.065548 .
## LAMP2         2.3735     0.6380   3.720 0.000199 ***
## LUM           1.3395     0.4262   3.143 0.001671 **
## MMRN1        -1.4132     0.4630  -3.052 0.002271 **
## SERPINA7     -1.8385     0.6389  -2.878 0.004006 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 221.81  on 159  degrees of freedom
## Residual deviance: 179.68  on 153  degrees of freedom
## AIC: 193.68
##
## Number of Fisher Scoring iterations: 4
```

Model Evaluation

Goodness of Fit: Homer Lemeshow Test

We consider the model to be unfit if $p < 0.05$.

```
# Converting Factor label to Numeric, "CRC" -> 1, "Healthy" -> 0
train.numeric <- train
train.numeric$Class <- as.numeric(train$Class)
train.numeric$Class[train.numeric$Class == 1] <- 0
train.numeric$Class[train.numeric$Class == 2] <- 1

res <- hoslem.test(train.numeric$Class, fitted(model.BestAICBackward1),
g=10)
res

##
##  Hosmer and Lemeshow goodness of fit (GOF) test
##
## data:  train.numeric$Class, fitted(model.BestAICBackward1)
## X-squared = 7.1724, df = 8, p-value = 0.5182
```

The p-value is 0.5181539 so there's no evidence the model in incorrect.

```
res <- hoslem.test(train.numeric$Class, fitted(model.BestAICBackward2),
g=10)
res

##
##  Hosmer and Lemeshow goodness of fit (GOF) test
##
## data:  train.numeric$Class, fitted(model.BestAICBackward2)
## X-squared = 7.192, df = 8, p-value = 0.5161
```

The p-value is 0.5160649 so there's no evidence the model in incorrect.

```
res <- hoslem.test(train.numeric$Class, fitted(model.BestAICForward1),
g=10)
res
```

```
##
##  Hosmer and Lemeshow goodness of fit (GOF) test
##
## data:  train.numeric$Class, fitted(model.BestAICForward1)
## X-squared = 3.4505, df = 8, p-value = 0.903
```

The p-value is 0.9029999 so there's no evidence the model in incorrect.

```
res <- hoslem.test(train.numeric$Class, fitted(model.BestAICForward2),
g=10)
res

##
##  Hosmer and Lemeshow goodness of fit (GOF) test
##
## data:  train.numeric$Class, fitted(model.BestAICForward2)
## X-squared = 6.5541, df = 8, p-value = 0.5854
```

The p-value is 0.5854153 so there's no evidence the model in incorrect.

```
res <- hoslem.test(train.numeric$Class, fitted(model.BestAICBoth1),
g=10)
res

##
##  Hosmer and Lemeshow goodness of fit (GOF) test
##
## data:  train.numeric$Class, fitted(model.BestAICBoth1)
## X-squared = 10.933, df = 8, p-value = 0.2055
```

The p-value is 0.2055332 so there's no evidence the model in incorrect.

```
res <- hoslem.test(train.numeric$Class, fitted(model.BestAICBoth2),
g=10)
res

##
##  Hosmer and Lemeshow goodness of fit (GOF) test
##
## data:  train.numeric$Class, fitted(model.BestAICBoth2)
## X-squared = 12.739, df = 8, p-value = 0.1211
```

The p-value is 0.1211393 so there's no evidence the model in incorrect.

```
res <- hoslem.test(train.numeric$Class, fitted(model.BestSBCBoth1),
g=10)
res

##
##  Hosmer and Lemeshow goodness of fit (GOF) test
##
```

```
## data:  train.numeric$Class, fitted(model.BestSBCBoth1)
## X-squared = 4.3334, df = 8, p-value = 0.8259
```

The p-value is 0.8258648 so there's no evidence the model in incorrect.

```
res <- hoslem.test(train.numeric$Class, fitted(model.BestSBCBoth2),
g=10)
res

##
##  Hosmer and Lemeshow goodness of fit (GOF) test
##
## data:  train.numeric$Class, fitted(model.BestSBCBoth2)
## X-squared = 4.7616, df = 8, p-value = 0.7827
```

The p-value is 0.7827256 so there's no evidence the model in incorrect.

From the above tests, we do not eliminate any Model from the candidate models.

Diagnostic Residual Plots

Residuals with Predicted Probabilities with Lowess Smooth

If the model is correct, a lowess smooth of the plot of the residuals against the estimated probability $\hat{\pi}_i$ should result approximately in a horizontal line with zero intercept.

```
png('residuals_lowess.png')
par(mfrow=c(3,3))
scatter.smooth(predict(model.BestAICBackward1, type = "response"),
               residuals(model.BestAICBackward1), xlab = "Estimated
Probability", ylab = "Deviance Residual", main="Model 1")

scatter.smooth(predict(model.BestAICBackward2, type = "response"),
               residuals(model.BestAICBackward2), xlab = "Estimated
Probability", ylab = "Deviance Residual", main="Model 2")

scatter.smooth(predict(model.BestAICForward1, type = "response"),
               residuals(model.BestAICForward1), xlab = "Estimated
Probability", ylab = "Deviance Residual", main="Model 3")

scatter.smooth(predict(model.BestAICForward2, type = "response"),
               residuals(model.BestAICForward2), xlab = "Estimated
Probability", ylab = "Deviance Residual", main="Model 4")

scatter.smooth(predict(model.BestAICBoth1, type = "response"),
               residuals(model.BestAICBoth1), xlab = "Estimated
Probability", ylab = "Deviance Residual", main="Model 5")
```

```
scatter.smooth(predict(model.BestAICBoth2, type = "response"),
               residuals(model.BestAICBoth2), xlab = "Estimated
Probability", ylab = "Deviance Residual", main="Model 6")

scatter.smooth(predict(model.BestSBCBoth1, type = "response"),
               residuals(model.BestSBCBoth1), xlab = "Estimated
Probability", ylab = "Deviance Residual", main="Model 7")

scatter.smooth(predict(model.BestSBCBoth2, type = "response"),
               residuals(model.BestSBCBoth2), xlab = "Estimated
Probability", ylab = "Deviance Residual", main="Model 8")
#dev.off()
```

All the Plots look good.

Half-Normal Probability Plot

A half-normal probability plot helps to highlight outlying deviance residuals even though the residuals are not normally distributed. Outliers will appear at the top right of a half-normal probability plot as points separated from the others.

```
png('half_normal.png')
par(mfrow=c(3,3))
halfnorm(residuals(model.BestAICBackward1), main="Model 1")
halfnorm(residuals(model.BestAICBackward2), main="Model 2")
halfnorm(residuals(model.BestAICForward1), main="Model 3")
halfnorm(residuals(model.BestAICForward2), main="Model 4")
halfnorm(residuals(model.BestAICBoth1), main="Model 5")
halfnorm(residuals(model.BestAICBoth2), main="Model 6")
halfnorm(residuals(model.BestSBCBoth1), main="Model 7")
halfnorm(residuals(model.BestSBCBoth2), main="Model 8")
#dev.off()
```

We observe from the above plots that observation 18, 40 and 112 might be an outlier.

Overdispersion

Sometimes we can get a deviance that is much larger than expected if the model was correct. It can be due to the presence of outliers, sparse data or clustering of data. The approach to deal with overdispersion is to add a dispersion parameter σ^2 . It can be estimated with: $\hat{\sigma}^2 = \frac{\chi^2}{n-p}$ (p = no. of parameters in model).
$Var\{Y_i\} = \phi n_i \pi_i \{1 - \pi_i\}$ We consider overdispersion to exist if ϕ >> 1.

```
par(mfrow=c(1,1))
# No. of observations in the training dataset.
n <- nrow(train)
```

```
# No. of parameters in the model.
p <- length(model.BestAICBackward1$coefficients)
phi <- sum(residuals(model.BestAICBackward1, type = "pearson")^2) / (n
- p)
phi

## [1] 1.16855
```

The dispersion parameter is not very different than one (no dispersion).

```
p <- length(model.BestAICBackward2$coefficients)
phi <- sum(residuals(model.BestAICBackward2, type = "pearson")^2) / (n
- p)
phi

## [1] 1.124246
```

The dispersion parameter is not very different than one (no dispersion).

```
p <- length(model.BestAICForward1$coefficients)
phi <- sum(residuals(model.BestAICForward1, type = "pearson")^2) / (n -
p)
phi

## [1] 1.068189
```

The dispersion parameter is not very different than one (no dispersion).

```
p <- length(model.BestAICForward2$coefficients)
phi <- sum(residuals(model.BestAICForward2, type = "pearson")^2) / (n -
p)
phi

## [1] 1.129269
```

The dispersion parameter is not very different than one (no dispersion).

```
p <- length(model.BestAICBoth1$coefficients)
phi <- sum(residuals(model.BestAICBoth1, type = "pearson")^2) / (n - p)
phi

## [1] 1.078921
```

The dispersion parameter is not very different than one (no dispersion).

```
p <- length(model.BestAICBoth2$coefficients)
phi <- sum(residuals(model.BestAICBoth2, type = "pearson")^2) / (n - p)
phi

## [1] 1.038809
```

The dispersion parameter is not very different than one (no dispersion).

```r
p <- length(model.BestSBCBoth1$coefficients)
phi <- sum(residuals(model.BestSBCBoth1, type = "pearson")^2) / (n - p)
phi

## [1] 1.061533
```

The dispersion parameter is not very different than one (no dispersion).

```r
p <- length(model.BestSBCBoth2$coefficients)
phi <- sum(residuals(model.BestSBCBoth2, type = "pearson")^2) / (n - p)
phi

## [1] 1.142676
```

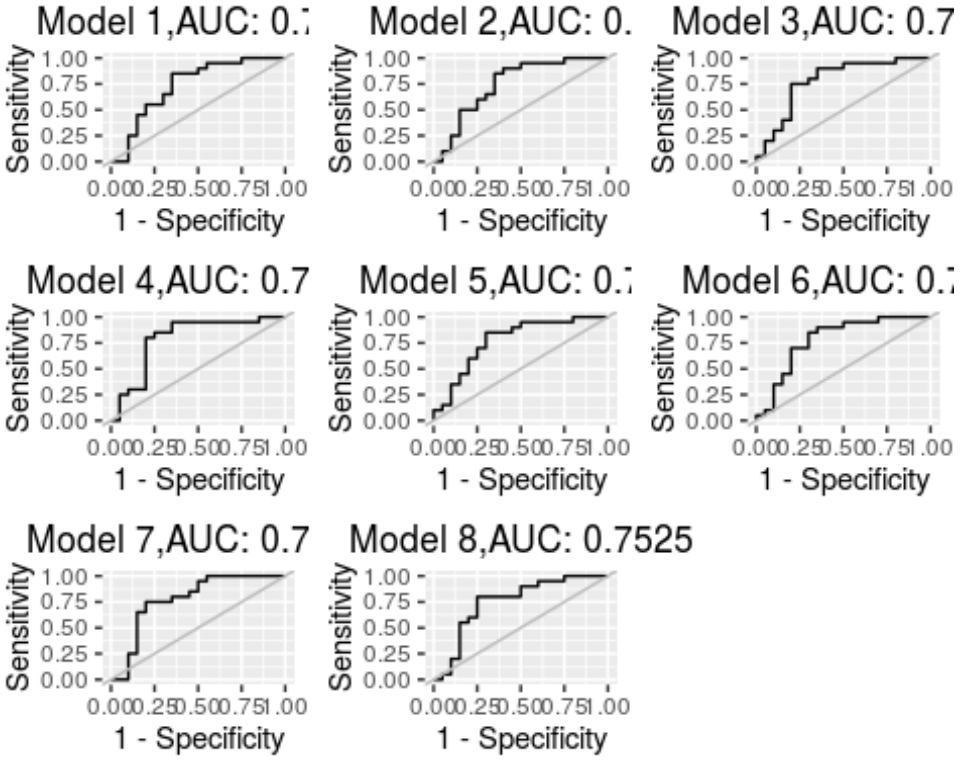
The dispersion parameter is not very different than one (no dispersion).

Predictive Ability of the Model

```r
# Divide training into 10 equal parts, keep one part as validation set
and rest as training.
model.list <- list()
model.list[["BestAICBackward1"]] <- formula.BestAICBackward1
model.list[["BestAICBackward2"]] <- formula.BestAICBackward2
model.list[["BestAICForward1"]] <- formula.BestAICForward1
model.list[["BestAICForward2"]] <- formula.BestAICForward2
model.list[["BestAICBoth1"]] <- formula.BestAICBoth1
model.list[["BestAICBoth2"]] <- formula.BestAICBoth2
model.list[["BestSBCBoth1"]] <- formula.BestSBCBoth1
model.list[["BestSBCBoth2"]] <- formula.BestSBCBoth2

k = 1
roc.mat <- matrix(list(), nrow=length(model.list), ncol=k)
fpr.mat <- matrix(list(), nrow=length(model.list), ncol=k)
tpr.mat <- matrix(list(), nrow=length(model.list), ncol=k)
auc.mat <- matrix(numeric(), nrow=length(model.list), ncol=k)
roc.plots.mat <- matrix(list(), nrow=length(model.list), ncol=k)
#png("ROC_plots.png")
par(mfrow=c(3,3))
for (model.id in 1:length(model.list)){
  cv.id <- 1
  fit <- glm(model.list[[model.id]], data=train,
family=binomial(link='logit'))
  predicted <- predict(fit, newdata=subset(validation, select=-Class))
  prob <- prediction(predicted, validation$Class)
  tprfpr <- performance(prob, "tpr", "fpr")
  tpr <- unlist(slot(tprfpr, "y.values"))
  fpr <- unlist(slot(tprfpr, "x.values"))
  roc <- data.frame(tpr, fpr)
  auc <- performance(prob, measure = "auc")
  auc <- auc@y.values[[1]]
```

```
  roc.mat[[model.id, cv.id]] <- roc
  fpr.mat[[model.id, cv.id]] <- fpr
  tpr.mat[[model.id, cv.id]] <- tpr
  auc.mat[[model.id, cv.id]] <- auc
  roc.plots.mat[[model.id, cv.id]] <- ggplot(roc) + geom_line(aes(x =
fpr, y = tpr)) + geom_abline(intercept = 0, slope = 1, colour = "gray")
+  ylab("Sensitivity") + xlab("1 - Specificity") +
ggtitle(paste0("Model ",model.id, ",AUC: " ,round(auc, 4)))
}
do.call(grid.arrange, roc.plots.mat[, 1])
```




```
#dev.off()
mean.auc <- apply(auc.mat, 1, mean)
names(mean.auc) <- names(model.list)
mean.auc
```

```
## BestAICBackward1 BestAICBackward2  BestAICForward1  BestAICForward2
##           0.7350           0.7500           0.7825           0.7925
##     BestAICBoth1     BestAICBoth2     BestSBCBoth1     BestSBCBoth2
##           0.7750           0.7850           0.7775           0.7525
```

```
#best.model.id <- 7
best.model.id <- which.max(mean.auc)
names(model.list)[best.model.id]
```

```
## [1] "BestAICForward2"
```

The area under the function provides an unbiased, and non-parametric measure of the discrimination ability of the model. AOC = 0.5 means that predictions are no better than random guessing. An AUROC value >= 0.80 is considered ideal.
The Candidate model with best AUROC value is "BestAICForward2" with value 0.7925. We consider this as our final model.

Selecting the best Cutoff

```
par(mfrow=c(1,1))
set.seed(seed.id)
fit <- train(model.list[[best.model.id]], data = train, method = "glm",
family="binomial")
predicted <- predict(fit, newdata=subset(validation, select=-Class),
type="prob", dispersion = 1.129278)[,1]

cutoffs <- seq(0, 1, 0.1)
sens <- c()
spec <- c()
acc <- c()
for (cutoff in cutoffs){
  pred <- ifelse(predicted >= cutoff, "CRC","Healthy")
  cm <- confusionMatrix(pred, validation$Class)
  sens <- c(sens, cm$byClass[["Sensitivity"]])
  spec <- c(spec, cm$byClass[["Specificity"]])
  acc <- c(acc, cm$byClass[["Balanced Accuracy"]])
}

## Warning in confusionMatrix.default(pred, validation$Class): Levels
are not
## in the same order for reference and data. Refactoring data to match.

## Warning in confusionMatrix.default(pred, validation$Class): Levels
are not
## in the same order for reference and data. Refactoring data to match.

## Warning in confusionMatrix.default(pred, validation$Class): Levels
are not
## in the same order for reference and data. Refactoring data to match.

df <- data.frame(cutoffs <- cutoffs, sens <- sens, spec <- spec, acc <-
acc)
names(df) <- c("cutoffs", "sens", "spec", "acc")
df

##    cutoffs sens spec   acc
## 1      0.0 1.00 0.00 0.500
## 2      0.1 0.95 0.10 0.525
## 3      0.2 0.85 0.30 0.575
## 4      0.3 0.80 0.40 0.600
## 5      0.4 0.80 0.70 0.750
```

```
## 6      0.5 0.65 0.85 0.750
## 7      0.6 0.55 0.95 0.750
## 8      0.7 0.45 0.95 0.700
## 9      0.8 0.20 0.95 0.575
## 10     0.9 0.00 1.00 0.500
## 11     1.0 0.00 1.00 0.500

png("sens_spec_acc.png")
plot(round(df$cutoffs, 4),df$sens, xlab="Cutoff",
ylab="Value",cex.lab=1.5,cex.axis=1.5,ylim=c(0,1),type="l",lty=2,lwd=2,
axes=TRUE)
lines(round(df$cutoffs, 4), df$spec,lty=4,lwd=3)
lines(round(df$cutoffs, 4), df$acc,lwd=2, type="l")
legend(0.5,.25,lty=c(2,4,1),lwd=c(2,3,2),c("Sensitivity","Specificity",
"Accuracy"))
#dev.off()

pred <- ifelse(predicted >= 0.45, "CRC","Healthy")
confusionMatrix(pred, validation$Class)

## $positive
## [1] "CRC"
##
## $table
##           Reference
## Prediction CRC Healthy
##    CRC      16       5
##    Healthy   4      15
##
## $overall
##       Accuracy          Kappa  AccuracyLower  AccuracyUpper
AccuracyNull
##   0.7750000000   0.5500000000   0.6154883227   0.8916033610
0.5000000000
## AccuracyPValue  McnemarPValue
##   0.0003397741   1.0000000000
##
## $byClass
##          Sensitivity          Specificity       Pos Pred Value
##            0.8000000            0.7500000            0.7619048
##       Neg Pred Value            Precision               Recall
##            0.7894737            0.7619048            0.8000000
##                   F1           Prevalence       Detection Rate
##            0.7804878            0.5000000            0.4000000
## Detection Prevalence    Balanced Accuracy
##            0.5250000            0.7750000
##
## $mode
## [1] "sens_spec"
##
```

```
## $dots
## list()
##
## attr(,"class")
## [1] "confusionMatrix"
```

Fitting the Best Model

We now fit the Best Model with the full training dataset.

```
set.seed(seed.id)
best.fit <- train(model.list[[best.model.id]], data = training, method
= "glm", family="quasibinomial")
summary(best.fit)

##
## Call:
## NULL
##
## Deviance Residuals:
##      Min        1Q    Median        3Q       Max
## -2.13751  -0.92826   0.00625   0.95266   2.65247
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   5.5676     7.5642   0.736  0.46259
## TIMP1        -1.1197     0.4158  -2.693  0.00771 **
## LAMP2         2.4252     0.6025   4.025 8.18e-05 ***
## HP           -0.2658     0.2683  -0.991  0.32299
## LRG1         -0.9263     0.3614  -2.563  0.01113 *
## SERPINA7     -1.6046     0.5877  -2.730  0.00692 **
## LUM           0.9440     0.3746   2.520  0.01254 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for quasibinomial family taken to be 1.116235)
##
##     Null deviance: 277.26  on 199  degrees of freedom
## Residual deviance: 225.63  on 193  degrees of freedom
## AIC: NA
##
## Number of Fisher Scoring iterations: 4

# Coefficients
exp(coef(best.fit$finalModel))

## (Intercept)       TIMP1       LAMP2          HP        LRG1
SERPINA7
## 261.8171797   0.3263650  11.3041203   0.7665708   0.3960063
0.2009796
```

```
##         LUM
##   2.5702634
```

```
# Confidence Interval of Coefficients
confint(best.fit$finalModel)
```

```
## Waiting for profiling to be done...
```

```
##                   2.5 %     97.5 %
## (Intercept) -9.1325157 20.7092327
## TIMP1       -1.9668537 -0.3259097
## LAMP2        1.2929658  3.6649085
## HP          -0.8031997  0.2565162
## LRG1        -1.6594531 -0.2340107
## SERPINA7    -2.7960502 -0.4790962
## LUM          0.2244853  1.7049338
```

Plotting Predicted Probability Lines

CP

```
training.numeric <- training
training.numeric$Class <- as.numeric(training$Class)
training.numeric$Class[training.numeric$Class == 1] <- 0
training.numeric$Class[training.numeric$Class == 2] <- 1

png("predicted_prob.png")
par(mfrow=c(3,2))
test <- with(training, data.frame(TIMP1=TIMP1, LAMP2=median(LAMP2),
HP=median(HP), LRG1=median(LRG1), SERPINA7= median(SERPINA7),
LUM=median(LUM)))
test$P <- predict(best.fit, newdata=test, type='prob')[, 1]
plot(training$TIMP1, training.numeric$Class, pch=16, xlab="TIMP1",
ylab="Class")
points(training$TIMP1, test$P)

test <- with(training, data.frame(TIMP1=median(TIMP1), LAMP2=LAMP2,
HP=median(HP), LRG1=median(LRG1), SERPINA7= median(SERPINA7),
LUM=median(LUM)))
test$P <- predict(best.fit, newdata=test, type='prob')[, 1]
plot(training$LAMP2, training.numeric$Class, pch=16, xlab="LAMP2",
ylab="Class")
points(training$LAMP2, test$P)

test <- with(training, data.frame(TIMP1=median(TIMP1),
LAMP2=median(LAMP2), HP=HP, LRG1=median(LRG1), SERPINA7=
median(SERPINA7), LUM=median(LUM)))
test$P <- predict(best.fit, newdata=test, type='prob')[, 1]
plot(training$HP, training.numeric$Class, pch=16, xlab="HP",
ylab="Class")
points(training$HP, test$P)
```

```
test <- with(training, data.frame(TIMP1=median(TIMP1),
LAMP2=median(LAMP2), HP=median(HP), LRG1=LRG1, SERPINA7=
median(SERPINA7), LUM=median(LUM)))
test$P <- predict(best.fit, newdata=test, type='prob')[, 1]
plot(training$LRG1, training.numeric$Class, pch=16, xlab="LRG1",
ylab="Class")
points(training$LRG1, test$P)

test <- with(training, data.frame(TIMP1=median(TIMP1),
LAMP2=median(LAMP2), HP=median(HP), LRG1=median(LRG1),
SERPINA7=SERPINA7, LUM=median(LUM)))
test$P <- predict(best.fit, newdata=test, type='prob')[, 1]
plot(training$SERPINA7, training.numeric$Class, pch=16,
xlab="SERPINA7", ylab="Class")
points(training$SERPINA7, test$P)

test <- with(training, data.frame(TIMP1=median(TIMP1),
LAMP2=median(LAMP2), HP=median(HP), LRG1=median(LRG1),
SERPINA7=median(SERPINA7), LUM=LUM))
test$P <- predict(best.fit, newdata=test, type='prob')[, 1]
plot(training$LUM, training.numeric$Class, pch=16, xlab="LUM",
ylab="Class")
points(training$LUM, test$P)
#dev.off()
par(mfrow=c(1,1))
```

Test the Final Model

We now introduce the test dataset for final results.

```
# Confusion Matrix for Training dataset
predicted <- predict(best.fit, newdata=subset(training, select=-Class),
type="prob", dispersion = 1.129278)[,1]
pred <- ifelse(predicted >= 0.45, "CRC", "Healthy")
confusionMatrix(data=pred, training$Class)

## $positive
## [1] "CRC"
##
## $table
##           Reference
## Prediction CRC Healthy
##    CRC      74      34
##    Healthy  26      66
##
## $overall
##       Accuracy          Kappa  AccuracyLower  AccuracyUpper
AccuracyNull
##   7.000000e-01   4.000000e-01   6.313501e-01   7.626104e-01
```

```
5.000000e-01
## AccuracyPValue  McnemarPValue
##   7.535308e-09   3.661566e-01
##
## $byClass
##          Sensitivity          Specificity       Pos Pred Value
##            0.7400000            0.6600000            0.6851852
##       Neg Pred Value            Precision               Recall
##            0.7173913            0.6851852            0.7400000
##                   F1           Prevalence       Detection Rate
##            0.7115385            0.5000000            0.3700000
## Detection Prevalence    Balanced Accuracy
##            0.5400000            0.7000000
##
## $mode
## [1] "sens_spec"
##
## $dots
## list()
##
## attr(,"class")
## [1] "confusionMatrix"

# Confusion Matrix for Testing dataset
predicted <- predict(fit, newdata=subset(testing, select=-Class),
type="prob", dispersion = 1.129278)[,1]
pred <- ifelse(predicted >= 0.45, "CRC","Healthy")
confusionMatrix(data=pred, testing$Class)

## $positive
## [1] "CRC"
##
## $table
##           Reference
## Prediction CRC Healthy
##    CRC     201      46
##    Healthy   1      21
##
## $overall
##       Accuracy          Kappa  AccuracyLower  AccuracyUpper
AccuracyNull
##   8.252788e-01   3.977516e-01   7.745208e-01   8.687063e-01
7.509294e-01
## AccuracyPValue  McnemarPValue
##   2.246670e-03   1.380222e-10
##
## $byClass
##          Sensitivity          Specificity       Pos Pred Value
##            0.9950495            0.3134328            0.8137652
##       Neg Pred Value            Precision               Recall
```

```
##            0.9545455            0.8137652            0.9950495
##                   F1           Prevalence       Detection Rate
##            0.8953229            0.7509294            0.7472119
## Detection Prevalence    Balanced Accuracy
##            0.9182156            0.6542412
##
## $mode
## [1] "sens_spec"
##
## $dots
## list()
##
## attr(,"class")
## [1] "confusionMatrix"
```

The Balanced Accuracy is almost similar for both training and testing datasets.

Statement of Contributions

A.S - Abhijeet Sharma
P.T - Pankaj Tripathi
A.S did selection and evaluation of models and paper write-up.
P.T. did pre-processing and plots.