Prediction of Adar Editing Levelss

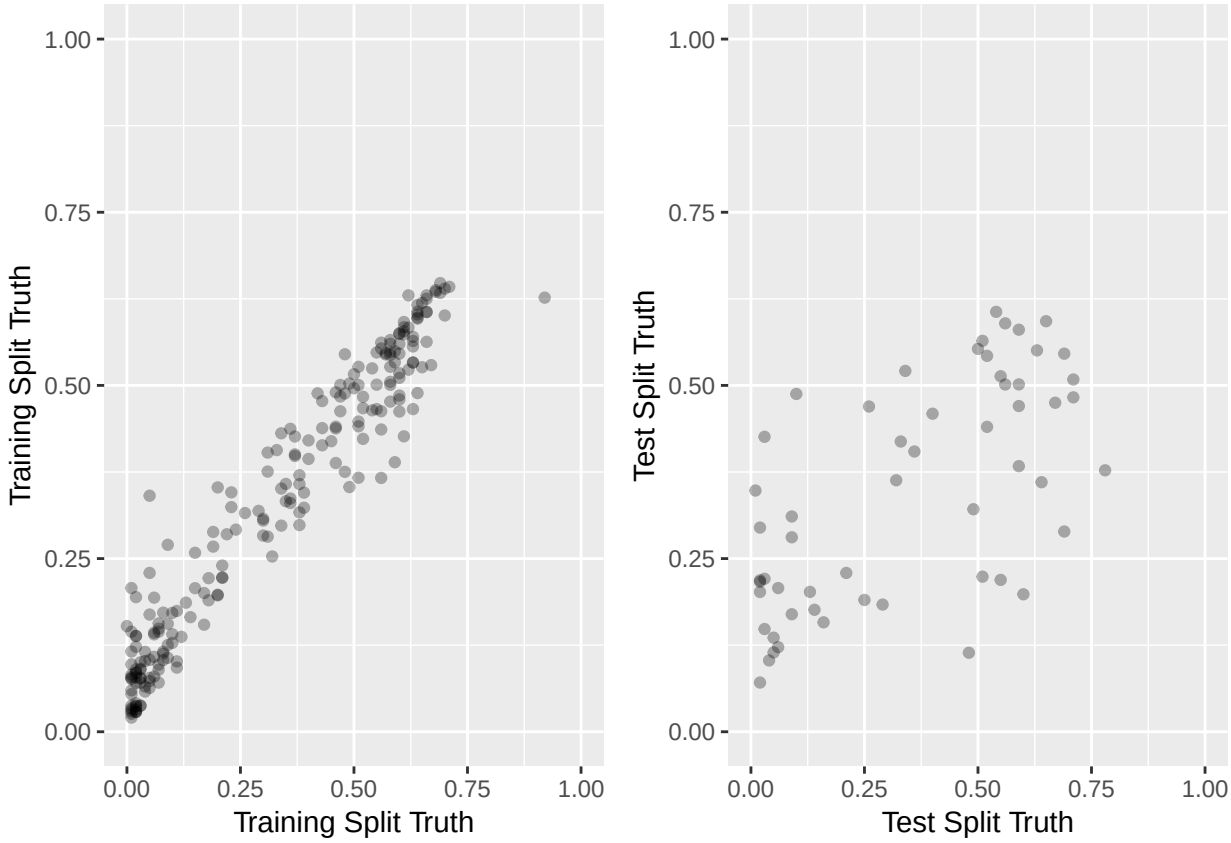
From Frequency of each RNA base in S, H, I, B feature

Note: base positions are 0-indexed

```
##
## Call:
##  randomForest(formula = y ~ ., data = data, importance = TRUE,      ntree = 5000, subset = train_indi
##                Type of random forest: regression
##                      Number of trees: 5000
## No. of variables tried at each split: 108
##
##           Mean of squared residuals: 0.03180788
##                     % Var explained: 47.44
```

Predictions on training and test splits

```
## [1] "MSE on training data:"
## [1] 0.005731044
## [1] "MSE on test data:"
## [1] 0.03698764
```



Feature Importance

```
##              %IncMSE IncNodePurity
## S.base50 53.26333728  1.183432e+00
## S.base46 38.75191607  1.042849e+00
## S.base42 28.34659929  2.913009e-01
## I.base50 27.62327945  2.966094e-01
## H.base42 24.33058511  2.083505e-01
## I.base54 22.31263010  4.463341e-01
## S.base19 21.89065408  2.024418e-01
## B.base29 21.63206576  5.077810e-01
## B.base54 21.22222790  2.144600e-01
## I.base46 19.03041271  9.287520e-02
## S.base47 18.63824365  1.154111e-01
## S.base30 18.60660502  2.021138e-01
## S.base31 18.56140120  2.263280e-01
## I.base31 17.79933325  1.225731e-01
## I.base47 17.22066138  9.087093e-02
## I.base27 16.70269019  1.097791e-01
## H.base40 16.60936745  2.946750e-01
## I.base49 16.38104396  3.407685e-01
## S.base18 15.95093342  7.921829e-02
## S.base35 15.84856237  2.622030e-01
## I.base76 15.30158966  1.367329e-01
## S.base27 14.64121217  9.876258e-02
## I.base19 12.91535369  6.800035e-02
## S.base44 12.79036224  7.678165e-02
## S.base64 12.29118339  1.001292e-01
## I.base67 12.22551868  1.168373e-01
## S.base7  11.75755028  1.958142e-01
## S.base59 11.65106115  5.662151e-02
## I.base28 11.63293210  5.859820e-02
## H.base35 11.61808569  2.216417e-01
## S.base28 11.52601244  5.482869e-02
## I.base32 10.86797165  9.027868e-02
## S.base67 10.85367982  1.304181e-01
## H.base47 10.46382590  8.584349e-02
## H.base41 10.17508279  6.938206e-02
## S.base48 10.13929988  6.132103e-02
## S.base72  9.65602995  9.210900e-02
## I.base53  9.61116599  4.055281e-02
## S.base29  9.45942666  9.533697e-02
## S.base58  9.36145036  4.443575e-02
## S.base79  9.12156291  7.862068e-02
## H.base46  8.78711116  7.293054e-02
## S.base24  8.73371418  3.087423e-02
## B.base64  8.69943579  5.545605e-02
## S.base40  8.56327390  6.463316e-02
## B.base59  8.46977747  2.873426e-02
## S.base70  8.10165844  5.859060e-02
## I.base26  8.02441792  2.980659e-02
## B.base51  7.99683418  2.051722e-02
## B.base52  7.95881409  2.336874e-02
## S.base32  7.93640233  5.530642e-02
```

```
## I.base35  7.86211182  7.103009e-02
## I.base58  7.75394416  2.896064e-02
## I.base30  7.29987007  1.857944e-02
## B.base73  7.27888280  8.219481e-02
## B.base19  7.13179200  1.606509e-02
## S.base54  7.12612075  6.614048e-02
## I.base55  7.06219412  4.043433e-02
## B.base55  6.98229908  2.689643e-02
## S.base56  6.96301530  3.295937e-02
## B.base21  6.71888033  4.047867e-02
## I.base56  6.68429177  3.601035e-02
## I.base60  6.53628807  3.988000e-02
## B.base18  6.52969744  1.376856e-02
## I.base73  6.45337167  4.510084e-02
## I.base70  6.42057102  3.879177e-02
## I.base18  6.24703191  1.247635e-02
## S.base6   6.12217614  3.272493e-02
## I.base21  6.01234999  3.443927e-02
## S.base45  6.00722559  3.899028e-02
## S.base2   5.93416528  4.833588e-02
## S.base15  5.80074469  3.137521e-02
## S.base5   5.69612179  4.978196e-02
## S.base57  5.68773333  3.591206e-02
## S.base55  5.61703938  3.017696e-02
## B.base31  5.53940111  4.416622e-02
## I.base29  5.39431755  4.038818e-02
## S.base53  5.17985915  3.406901e-02
## B.base28  5.17617135  2.835907e-02
## B.base60  4.99600818  5.712833e-02
## B.base44  4.98459217  3.977297e-02
## S.base22  4.92019859  2.234038e-02
## S.base1   4.88137304  3.942798e-02
## H.base38  4.62335592  1.556265e-02
## S.base41  4.49176485  3.608323e-02
## H.base49  4.41026440  1.822405e-02
## B.base8   4.35410973  3.585957e-02
## I.base57  4.35177895  2.466105e-02
## S.base21  4.26742343  3.212976e-02
## B.base25  4.21355367  1.412699e-02
## H.base8   4.12444062  2.341280e-02
## B.base56  4.06262328  1.224809e-02
## I.base48  4.04984185  2.762403e-02
## S.base0   4.04414932  9.835998e-03
## S.base43  4.03210628  3.469901e-02
## B.base26  4.02213736  1.412912e-02
## H.base6   3.94773634  1.858913e-02
## I.base38  3.92652292  5.502061e-03
## S.base69  3.89800907  4.356606e-02
## H.base48  3.86003811  2.002256e-02
## S.base52  3.83675582  3.679682e-02
## S.base26  3.82994545  2.759949e-02
## I.base24  3.81787964  1.060197e-02
## H.base36  3.70683655  1.130272e-02
## H.base33  3.60899266  1.781232e-02
```

```
## H.base5   3.54083004  2.691721e-02
## H.base7   3.53661196  2.680054e-02
## S.base71  3.48467038  2.083597e-02
## I.base39  3.43383730  1.538701e-02
## S.base25  3.42344742  1.569089e-02
## H.base3   3.42318646  8.939501e-03
## I.base15  3.41466932  3.620926e-02
## B.base30  3.22105998  9.109813e-03
## I.base37  3.21376967  5.433132e-03
## I.base59  3.13992286  9.570577e-03
## S.base66  3.11646206  2.227357e-02
## H.base32  3.11138241  1.657355e-02
## B.base35  3.08747295  1.415711e-02
## I.base20  3.05824774  2.286183e-02
## S.base73  3.05817206  5.466208e-02
## S.base51  3.05542805  1.447541e-02
## I.base41  3.02025531  2.191149e-02
## S.base3   2.97856242  2.084081e-02
## B.base9   2.97759717  2.715321e-02
## I.base7   2.96038939  2.091583e-02
## S.base61  2.91070668  4.584177e-02
## I.base69  2.90357915  2.669602e-02
## S.base20  2.89337125  2.204289e-02
## B.base20  2.86779588  6.078273e-03
## B.base22  2.86528688  7.465674e-03
## B.base58  2.84939569  5.777205e-03
## H.base37  2.78991445  1.242957e-02
## S.base37  2.77388377  1.355241e-02
## I.base25  2.76930862  1.219766e-02
## B.base34  2.72105650  6.788100e-03
## I.base16  2.70555557  2.876152e-02
## S.base33  2.67617561  1.766824e-02
## H.base10  2.64365170  2.229714e-02
## B.base33  2.64181331  8.440302e-03
## S.base75  2.63470444  6.240667e-02
## S.base77  2.62053813  3.935042e-02
## H.base50  2.60978461  1.046805e-02
## S.base36  2.50878159  1.461547e-02
## H.base9   2.46051138  1.909569e-02
## I.base40  2.44319990  1.399007e-02
## S.base63  2.44057805  2.775762e-03
## S.base38  2.27973751  1.666236e-02
## S.base9   2.24223466  3.852706e-02
## I.base22  2.15069947  2.192669e-02
## H.base4   2.09460955  8.434678e-03
## B.base57  2.02914909  7.383836e-03
## H.base45  1.99977955  5.735158e-03
## I.base72  1.96291651  4.530692e-02
## I.base13  1.84700607  1.865620e-02
## S.base78  1.83770008  3.128130e-02
## S.base10  1.83350738  3.455813e-02
## S.base16  1.81040226  1.521114e-02
## B.base46  1.76983032  1.339982e-02
## S.base74  1.76581934  5.019377e-02
```

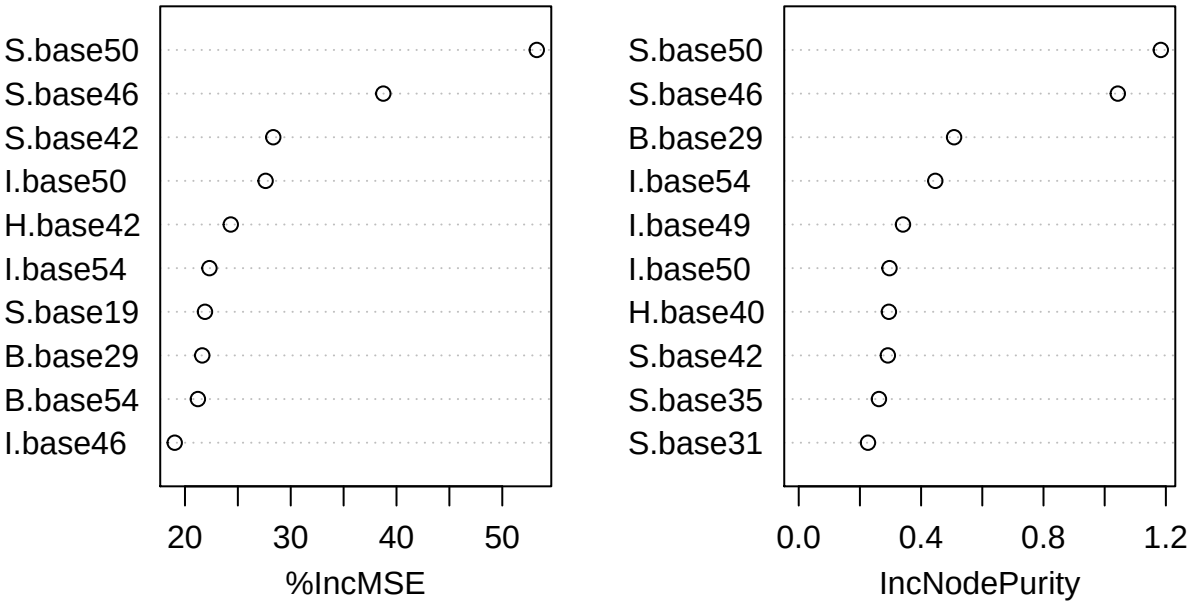
```
## B.base72  1.71874414  8.945205e-04
## S.base49  1.70988403  2.344444e-02
## I.base52  1.67795018  2.719269e-02
## H.base62  1.65512962  1.025909e-03
## S.base76  1.62088583  3.491943e-02
## H.base63  1.60440560  5.039116e-04
## I.base36  1.58806984  3.876790e-02
## H.base27  1.56780867  3.942288e-04
## I.base51  1.50138709  3.838904e-03
## H.base26  1.50033393  4.243257e-04
## I.base71  1.50026496  4.167570e-03
## B.base10  1.50005747  1.456472e-05
## H.base57  1.48894297  2.508400e-03
## I.base14  1.44320694  1.981245e-02
## S.base62  1.41504041  4.175390e-03
## H.base21  1.37780513  9.477982e-04
## H.base29  1.36184405  3.897132e-04
## H.base30  1.32201565  5.815014e-04
## I.base74  1.29592721  3.339557e-02
## B.base24  1.28546999  2.389649e-03
## B.base53  1.25870171  1.938613e-02
## H.base20  1.20522004  5.951724e-04
## B.base65  1.20485001  2.249965e-02
## H.base54  1.17086789  4.922402e-03
## H.base23  1.14022533  6.031967e-04
## S.base13  1.13472293  1.935591e-02
## B.base14  1.12912686  4.713634e-04
## S.base34  1.10986325  3.139191e-02
## H.base22  1.08964223  7.624698e-04
## B.base15  1.06347572  1.372541e-02
## S.base4   1.06310795  1.819180e-02
## B.base12  1.01348836  3.682728e-05
## H.base15  1.00432026  8.373055e-05
## B.base13  0.99982157  1.544566e-05
## S.base12  0.94715053  3.223370e-02
## I.base68  0.93771788  2.092111e-02
## H.base2   0.93323849  2.561832e-04
## S.base14  0.92004026  1.867460e-02
## S.base11  0.89951204  2.733052e-04
## B.base41  0.89284642  6.082983e-03
## H.base28  0.88215219  4.112525e-04
## H.base19  0.86558161  7.988911e-04
## I.base9   0.85360085  6.005670e-04
## I.base66  0.80644276  3.867248e-02
## S.base60  0.77772700  3.928889e-02
## H.base43  0.76050998  1.353466e-02
## H.base18  0.72736902  5.399486e-04
## H.base65  0.71816423  1.012027e-04
## H.base55  0.69762824  3.647889e-03
## B.base61  0.66832852  3.576127e-02
## B.base38  0.63639488  3.940741e-03
## H.base51  0.60960112  9.774135e-03
## H.base24  0.59720591  5.947778e-04
## S.base65  0.56205831  3.164081e-02
```

```
## I.base12  0.50351754  3.272151e-02
## H.base25  0.50182284  6.094826e-04
## B.base48  0.50061703  1.511910e-02
## I.base61  0.32879613  1.013141e-02
## I.base75  0.29084633  2.224766e-02
## H.base16  0.18291572  5.322397e-04
## B.base37  0.13180066  8.510859e-04
## I.base8   0.05698269  3.528160e-02
## H.base0   0.00000000  0.000000e+00
## H.base1   0.00000000  0.000000e+00
## H.base61  0.00000000  8.431446e-05
## H.base66  0.00000000  5.663529e-05
## H.base67  0.00000000  0.000000e+00
## H.base68  0.00000000  0.000000e+00
## H.base69  0.00000000  0.000000e+00
## H.base70  0.00000000  0.000000e+00
## H.base71  0.00000000  0.000000e+00
## H.base72  0.00000000  0.000000e+00
## H.base73  0.00000000  0.000000e+00
## H.base74  0.00000000  0.000000e+00
## H.base75  0.00000000  0.000000e+00
## H.base76  0.00000000  0.000000e+00
## H.base77  0.00000000  0.000000e+00
## H.base78  0.00000000  0.000000e+00
## H.base79  0.00000000  0.000000e+00
## H.base80  0.00000000  0.000000e+00
## I.base0   0.00000000  0.000000e+00
## I.base1   0.00000000  0.000000e+00
## I.base2   0.00000000  0.000000e+00
## I.base3   0.00000000  0.000000e+00
## I.base4   0.00000000  0.000000e+00
## I.base5   0.00000000  0.000000e+00
## I.base6   0.00000000  2.642455e-05
## I.base11  0.00000000  2.336957e-06
## I.base80  0.00000000  0.000000e+00
## B.base0   0.00000000  0.000000e+00
## B.base1   0.00000000  0.000000e+00
## B.base2   0.00000000  0.000000e+00
## B.base3   0.00000000  0.000000e+00
## B.base4   0.00000000  0.000000e+00
## B.base5   0.00000000  0.000000e+00
## B.base6   0.00000000  0.000000e+00
## B.base7   0.00000000  0.000000e+00
## B.base69  0.00000000  8.771613e-05
## B.base70  0.00000000  0.000000e+00
## B.base78  0.00000000  0.000000e+00
## B.base79  0.00000000  0.000000e+00
## B.base80  0.00000000  0.000000e+00
## B.base66 -0.13077023  1.999754e-02
## B.base43 -0.15607629  1.038597e-02
## H.base34 -0.17685348  3.662508e-02
## I.base10 -0.20469031  2.628595e-02
## I.base17 -0.20893736  2.638595e-03
## B.base32 -0.24306047  7.929018e-03
```

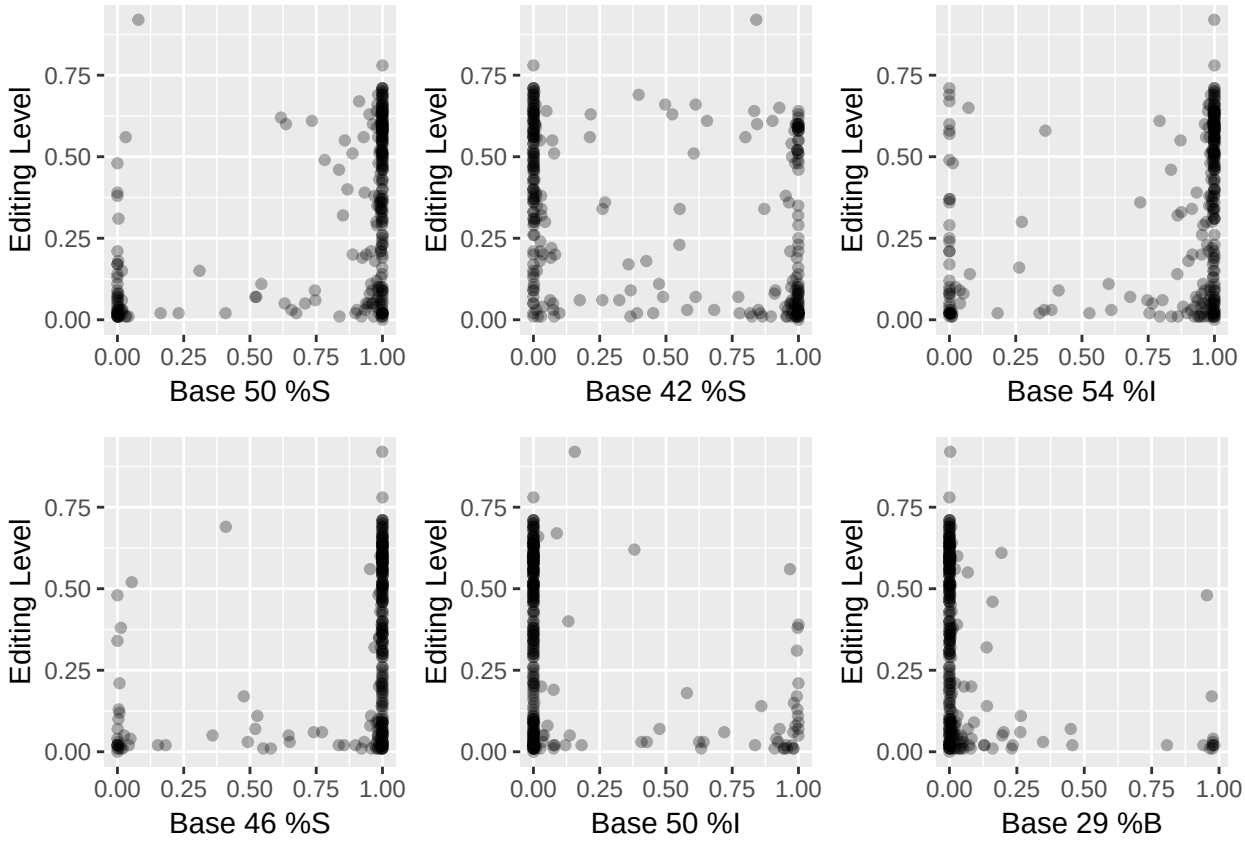
```
## B.base11 -0.26990306  5.620768e-06
## B.base40 -0.30640866  7.737383e-03
## B.base62 -0.35431259  2.136967e-03
## I.base43 -0.37618957  2.247765e-02
## I.base42 -0.50961341  1.120053e-02
## H.base60 -0.51350292  1.543230e-04
## H.base64 -0.63964011  6.552246e-04
## B.base17 -0.65091051  1.630929e-02
## B.base47 -0.66213033  4.402085e-02
## H.base44 -0.66850523  4.347580e-03
## B.base16 -0.82247008  1.191800e-02
## H.base58 -0.83403495  1.253832e-03
## H.base56 -0.90362320  3.461822e-03
## S.base80 -0.90787746  3.686001e-03
## S.base23 -0.92659329  5.819703e-03
## H.base11 -0.93797389  1.989404e-04
## H.base59 -0.95899372  4.680785e-05
## I.base34 -0.98295543  2.480908e-03
## H.base14 -0.99527885  1.421730e-04
## H.base12 -1.00010002  9.573214e-05
## B.base23 -1.05174217  3.032130e-03
## B.base67 -1.05831221  1.688045e-02
## H.base13 -1.07206960  1.698474e-04
## I.base77 -1.13570517  9.820119e-04
## I.base63 -1.14766494  1.139222e-04
## S.base17 -1.21037350  1.567897e-02
## B.base76 -1.26328928  6.309726e-04
## H.base52 -1.46869075  1.079435e-02
## S.base8  -1.48245127  2.489720e-02
## B.base49 -1.50608544  2.297283e-02
## I.base45 -1.61978049  2.146156e-03
## I.base64 -1.62116963  1.262792e-05
## B.base71 -1.84835530  4.119168e-04
## B.base63 -1.87373457  5.692906e-04
## B.base74 -1.90595790  8.795405e-03
## H.base17 -1.91128945  4.487961e-04
## B.base68 -1.93680814  2.195652e-02
## I.base62 -1.95056012  3.738734e-04
## I.base23 -1.95509918  4.669756e-03
## S.base39 -2.08204493  4.778334e-02
## I.base44 -2.08800089  1.839381e-02
## H.base39 -2.15523300  6.350739e-02
## B.base42 -2.31591122  5.313305e-03
## H.base31 -2.36319480  4.017020e-03
## B.base39 -2.47566032  8.897581e-03
## B.base27 -2.48788142  2.570217e-02
## I.base33 -2.58122707  8.349855e-03
## H.base53 -2.58417318  1.603727e-03
## S.base68 -2.66436288  3.697672e-02
## I.base78 -2.76112731  9.744641e-04
## I.base79 -3.10075598  5.813690e-04
## B.base77 -3.33908810  2.031158e-03
## B.base50 -3.59662436  3.152117e-02
## B.base36 -3.95254618  3.921988e-03
```

```
## B.base75 -4.54000231  1.803891e-02
## I.base65 -4.91471253  5.747144e-03
## B.base45 -5.12103507  2.231553e-02
```

Importance of position along the RNA



Feature value vs Editing Level



Editing Level Prediction on Inferred Structure from Both Position-Specific and Structure-Specific Features

```
##   editing_level mp1 mp2 mref1 mref2 malt1 malt2   mtype1 mtype2 adist1
## 1          0.58  44  NA     G  <NA>     A  <NA> mismatch   <NA>      6
## 2          0.46  45  NA     G  <NA>     A  <NA> mismatch   <NA>      5
## 3          0.49  46  NA     G  <NA>     A  <NA> mismatch   <NA>      4
## 4          0.48  47  NA     G  <NA>     A  <NA> mismatch   <NA>      3
## 5          0.59  48  NA     C  <NA>     A  <NA> mismatch   <NA>      2
## 6          0.39  52  NA     G  <NA>     A  <NA> mismatch   <NA>     -2
##   adist2 editing_feature mfeat1 mfeat1_prev mfeat1_next
## 1     NA               I      S           H           I
## 2     NA               I      I           S           S
## 3     NA               B      S           H           B
## 4     NA               I      S           H           I
## 5     NA               I      S           H           I
## 6     NA               I      S           I           I
##   mfeat1_same_as_edit mfeat2 mfeat2_prev mfeat2_next mfeat2_same_as_edit
## 1                   0   <NA>        <NA>        <NA>                  NA
## 2                   0   <NA>        <NA>        <NA>                  NA
## 3                   0   <NA>        <NA>        <NA>                  NA
## 4                   0   <NA>        <NA>        <NA>                  NA
```

```
## 5                   0   <NA>        <NA>        <NA>                  NA
## 6                   0   <NA>        <NA>        <NA>                  NA
##   stem_length hairpin_length feat_3prime_e feat_3prime_e_length
## 1           7             5             I                    3
## 2           2             8             I                    3
## 3           6            10             B                    2
## 4           6             4             I                    1
## 5           7             8             I                    1
## 6           6             4             I                    2
##   feat_3prime_e_length_5prime feat_3prime_e_cp1 feat_3prime_e_cp2
## 1                           2              G:U              U:A
## 2                           3              G:C              U:A
## 3                           0              C:G              A:U
## 4                           1              A:U              U:A
## 5                           1              A:U              U:A
## 6                           1              A:U              U:A
##   feat_3prime_e_distal feat_3prime_e_length_distal
## 1                    I                           3
## 2                    I                           3
## 3                    B                           2
## 4                    I                           1
## 5                    I                           1
## 6                    I                           2
##   feat_3prime_e_length_5prime_distal feat_3prime_e_cp1_distal
## 1                                  2                     G:U
## 2                                  3                     G:C
## 3                                  0                     C:G
## 4                                  1                     A:U
## 5                                  1                     A:U
## 6                                  1                     A:U
##   feat_3prime_e_cp2_distal
## 1                      U:A
## 2                      U:A
## 3                      A:U
## 4                      U:A
## 5                      U:A
## 6                      U:A
```

```
##      |      Out-of-bag   |
## Tree |      MSE  %Var(y) |
##  300 |  0.01682    27.51 |
##      |      Out-of-bag   |
## Tree |      MSE  %Var(y) |
##  300 |  0.01988    32.51 |
##      |      Out-of-bag   |
## Tree |      MSE  %Var(y) |
##  300 |  0.01923    31.44 |
##      |      Out-of-bag   |
## Tree |      MSE  %Var(y) |
##  300 |  0.01973    32.26 |
##      |      Out-of-bag   |
## Tree |      MSE  %Var(y) |
##  300 |  0.02031    33.21 |
```

```
##
```

```
## Call:
##  randomForest(formula = y ~ ., data = data, importance = TRUE,      ntree = 10000, subset = train_inc
##                Type of random forest: regression
##                      Number of trees: 10000
## No. of variables tried at each split: 10
##
##           Mean of squared residuals: 0.02037068
##                     % Var explained: 67.12
```

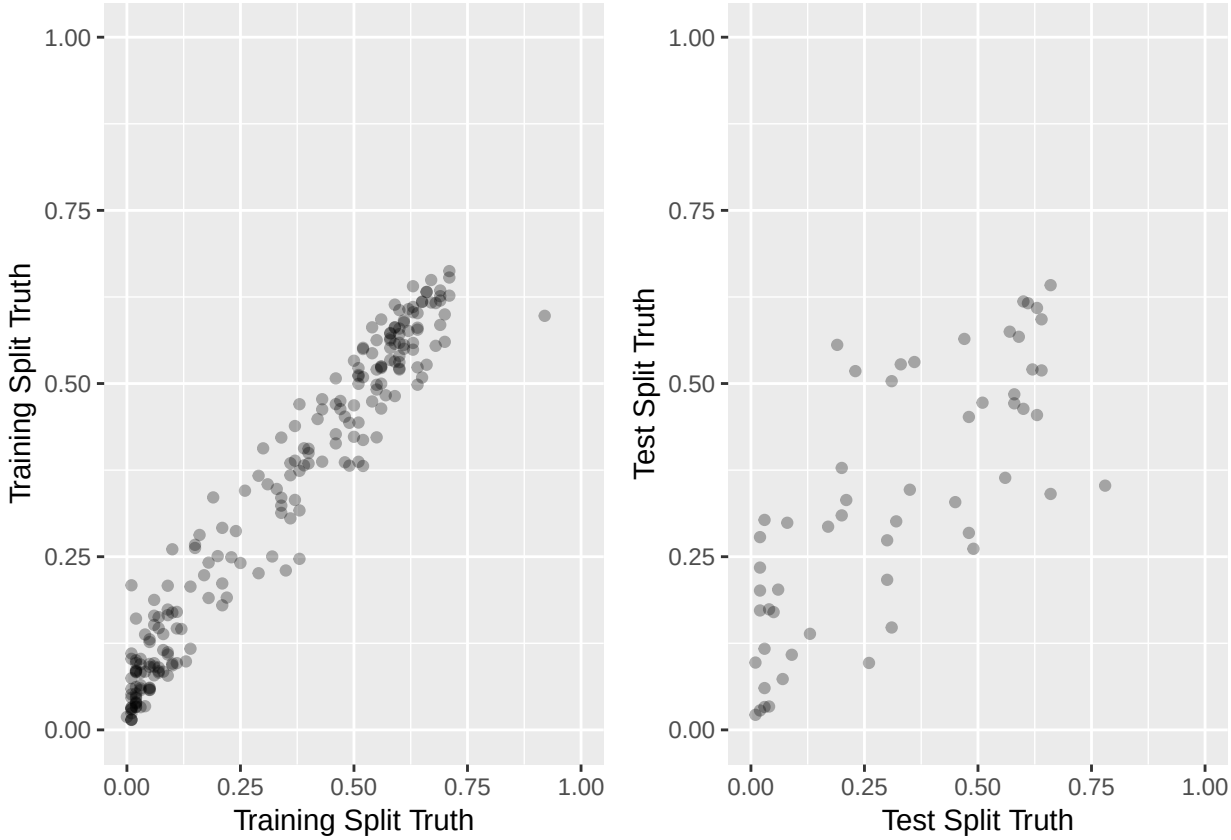
Predictions on training and test splits

```
## [1] "MSE on training data:"
```

```
## [1] 0.00436655
```

```
## [1] "MSE on test data:"
```

```
## [1] 0.02421978
```

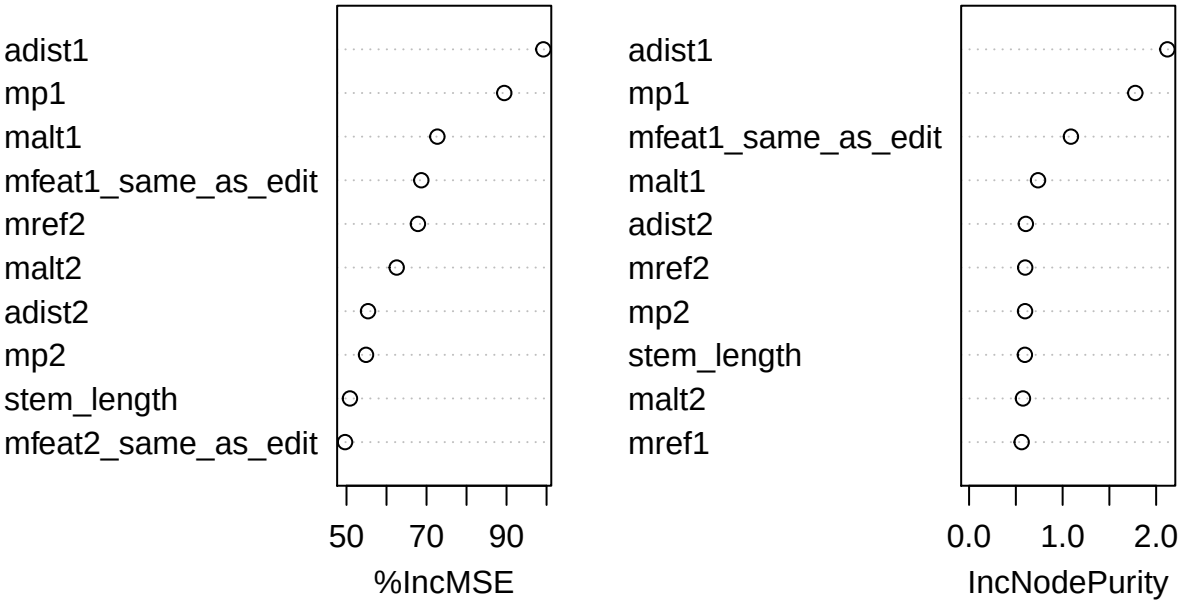


##Feature Importance

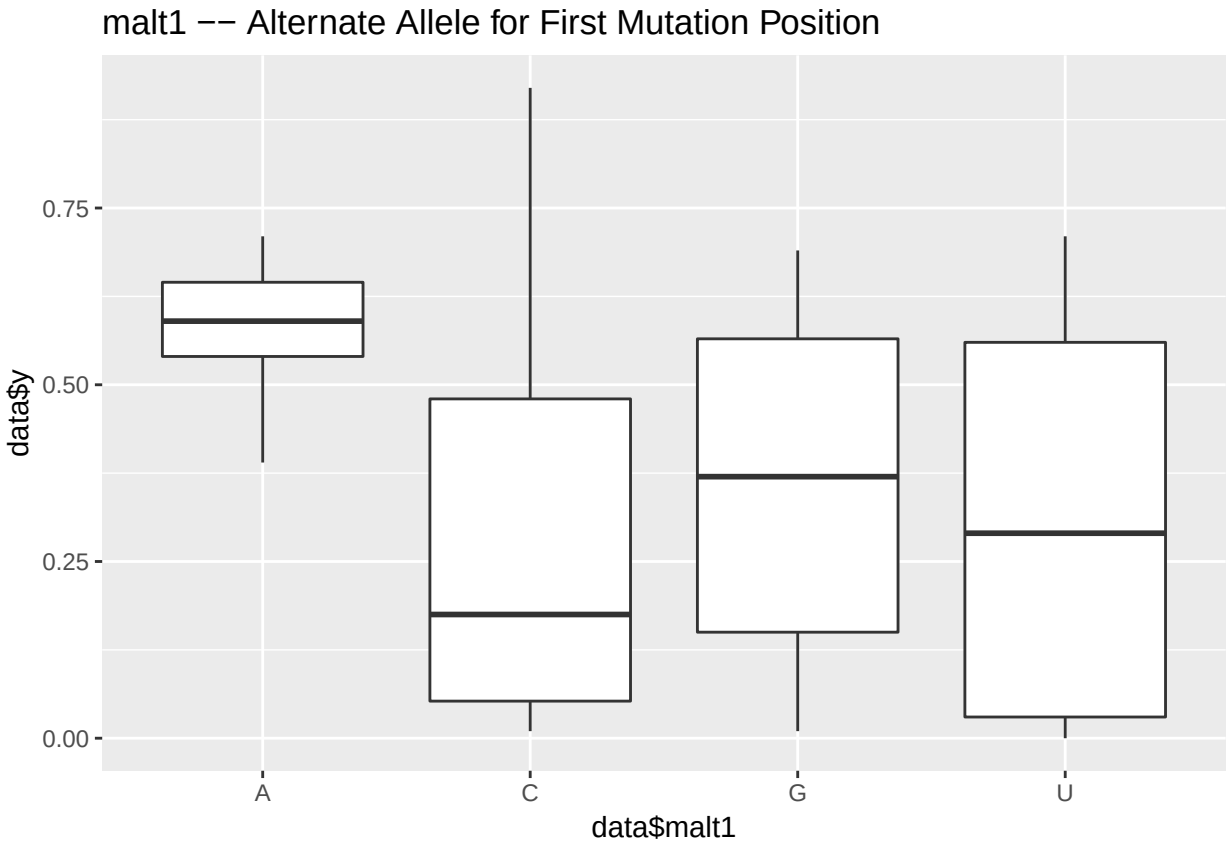
```
##                                %IncMSE IncNodePurity
## adist1                       99.200586    2.11985169
## mp1                          89.429439    1.77693792
## malt1                        72.710320    0.73712756
## mfeat1_same_as_edit          68.696536    1.08869140
## mref2                        67.849012    0.59974446
## malt2                        62.540736    0.57581795
## adist2                       55.385916    0.60670394
```

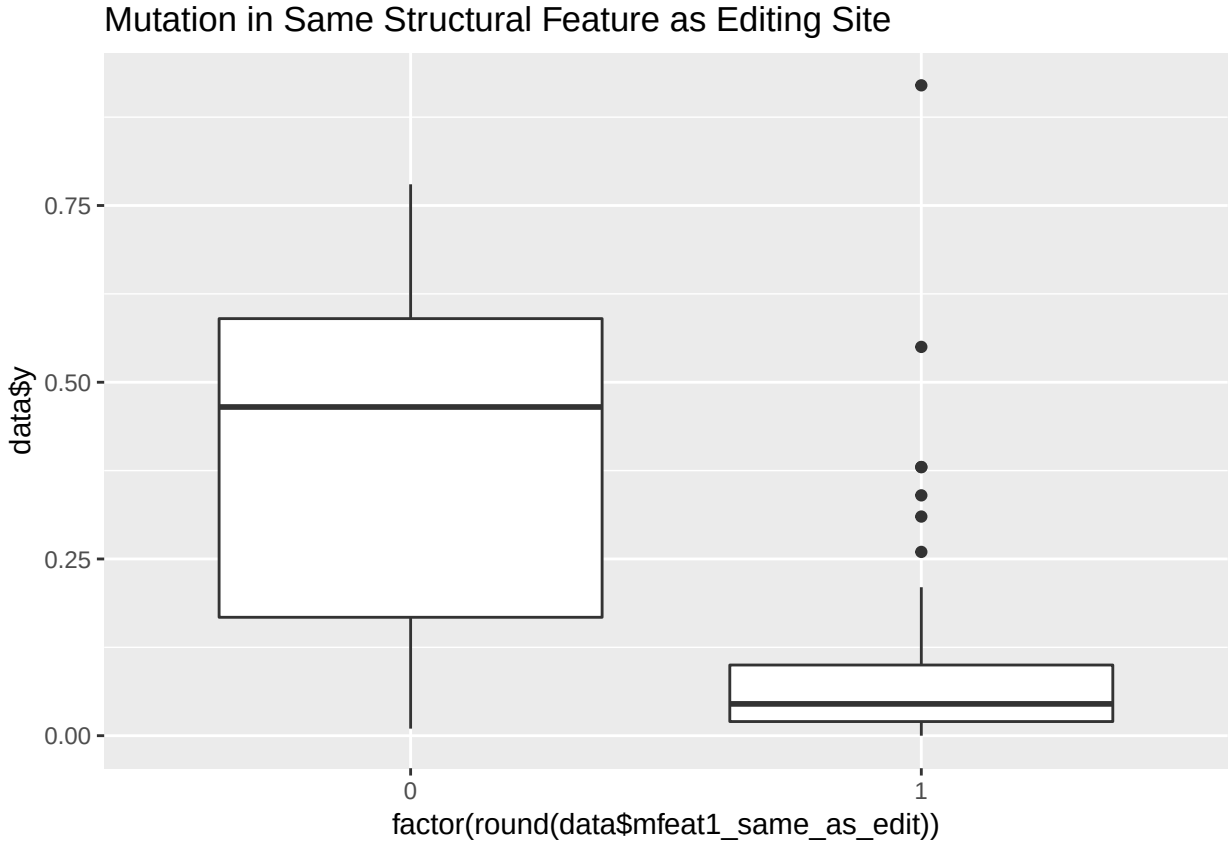
```
## mp2                                  54.878442    0.59853918
## stem_length                          50.855093    0.59654475
## mfeat2_same_as_edit                  49.638572    0.44006956
## mfeat2                               46.812395    0.36633356
## mref1                                41.647376    0.56213007
## mfeat2_prev                          41.624083    0.29913956
## mfeat2_next                          38.531322    0.40230169
## feat_3prime_e_length_distal          38.088409    0.27248928
## feat_3prime_e_length                 37.897661    0.26958926
## mfeat1                               33.637215    0.22726464
## mfeat1_prev                          31.975373    0.26181755
## hairpin_length                       27.412423    0.16509034
## feat_3prime_e_cp2_distal             21.969674    0.22811640
## mtype1                               21.917468    0.06780997
## feat_3prime_e_cp2                    21.706516    0.22891921
## feat_3prime_e_length_5prime_distal   18.893305    0.08671540
## mfeat1_next                          18.137265    0.12142573
## feat_3prime_e_length_5prime          17.479674    0.08086709
## feat_3prime_e_cp1                     9.087820    0.08124108
## feat_3prime_e_cp1_distal              8.901085    0.08143912
## editing_feature                       4.843355    0.08058079
## feat_3prime_e_distal                  1.295613    0.04825592
## feat_3prime_e                         1.294021    0.04654927
## mtype2                                0.000000    0.00000000
```

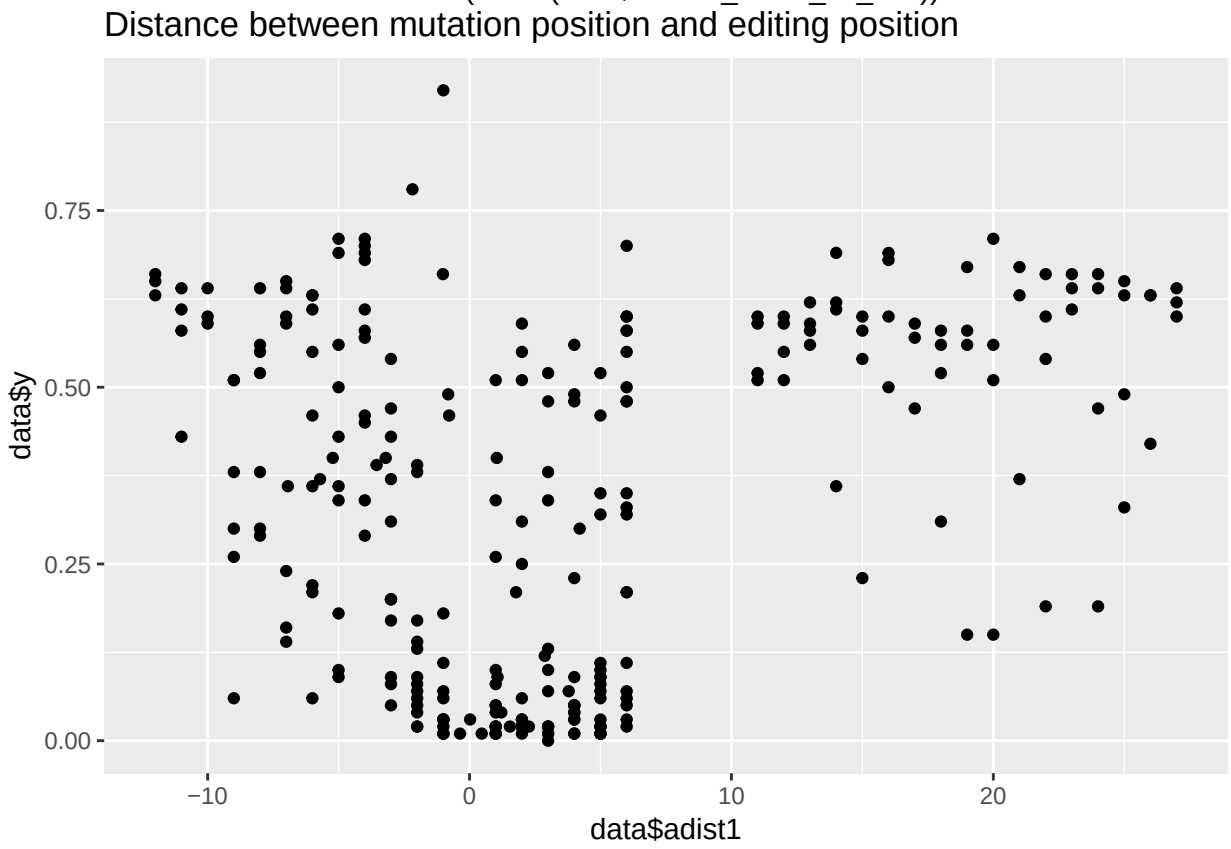
Top 10 Most Important Features



Feature values vs Editing Level




Mutation in Same Structural Feature as Editing Site
0.75
0.50
0.25
0.00
data$y
0
1
factor(round(data$mfeat1_same_as_edit))


Distance between mutation position and editing position
0.75
0.50
0.25
0.00
data$y
−10
0
10
20
data$adist1

Predicting Editing Level from Structure-Specific Features Only (i.e. No Mutation Information Used in Model)

```
##   editing_level editing_feature mfeat1 mfeat1_prev mfeat1_next
## 1          0.58              I      S           H           I
## 2          0.46              I      I           S           S
## 3          0.49              B      S           H           B
## 4          0.48              I      S           H           I
## 5          0.59              I      S           H           I
## 6          0.39              I      S           I           I
##   mfeat1_same_as_edit mfeat2 mfeat2_prev mfeat2_next mfeat2_same_as_edit
## 1                   0   <NA>        <NA>        <NA>                  NA
## 2                   0   <NA>        <NA>        <NA>                  NA
## 3                   0   <NA>        <NA>        <NA>                  NA
## 4                   0   <NA>        <NA>        <NA>                  NA
## 5                   0   <NA>        <NA>        <NA>                  NA
## 6                   0   <NA>        <NA>        <NA>                  NA
##   stem_length hairpin_length feat_3prime_e feat_3prime_e_length
## 1           7              5             I                    3
## 2           2              8             I                    3
## 3           6             10             B                    2
## 4           6              4             I                    1
## 5           7              8             I                    1
## 6           6              4             I                    2
##   feat_3prime_e_cp1 feat_3prime_e_cp2
## 1               G:U               U:A
## 2               G:C               U:A
## 3               C:G               A:U
## 4               A:U               U:A
## 5               A:U               U:A
## 6               A:U               U:A
```

```
##      |      Out-of-bag   |
## Tree |      MSE  %Var(y) |
##  300 |  0.03356    54.88 |
##      |      Out-of-bag   |
## Tree |      MSE  %Var(y) |
##  300 |  0.02852    46.63 |
##      |      Out-of-bag   |
## Tree |      MSE  %Var(y) |
##  300 |  0.02883    47.14 |
##      |      Out-of-bag   |
## Tree |      MSE  %Var(y) |
##  300 |   0.0287    46.93 |
##      |      Out-of-bag   |
## Tree |      MSE  %Var(y) |
##  300 |   0.0287    46.92 |
```

```
##
## Call:
##  randomForest(formula = y ~ ., data = data, importance = TRUE,      ntree = 10000, subset = train_in
##                Type of random forest: regression
##                      Number of trees: 10000
## No. of variables tried at each split: 5
```

```
##
##           Mean of squared residuals: 0.0307335
##                     % Var explained: 50.41
```

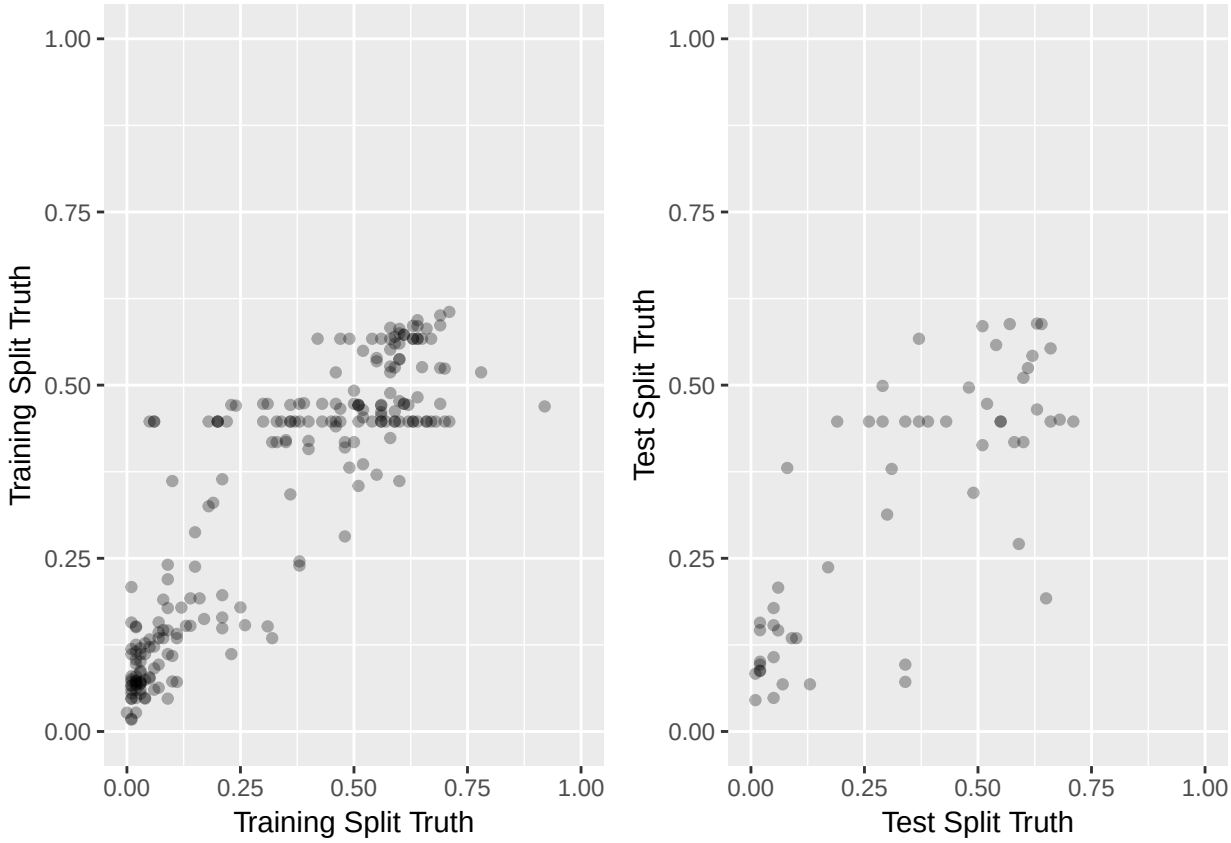
Predictions on training and test splits

```
## [1] "MSE on training data:"
```

```
## [1] 0.01459266
```

```
## [1] "MSE on test data:"
```

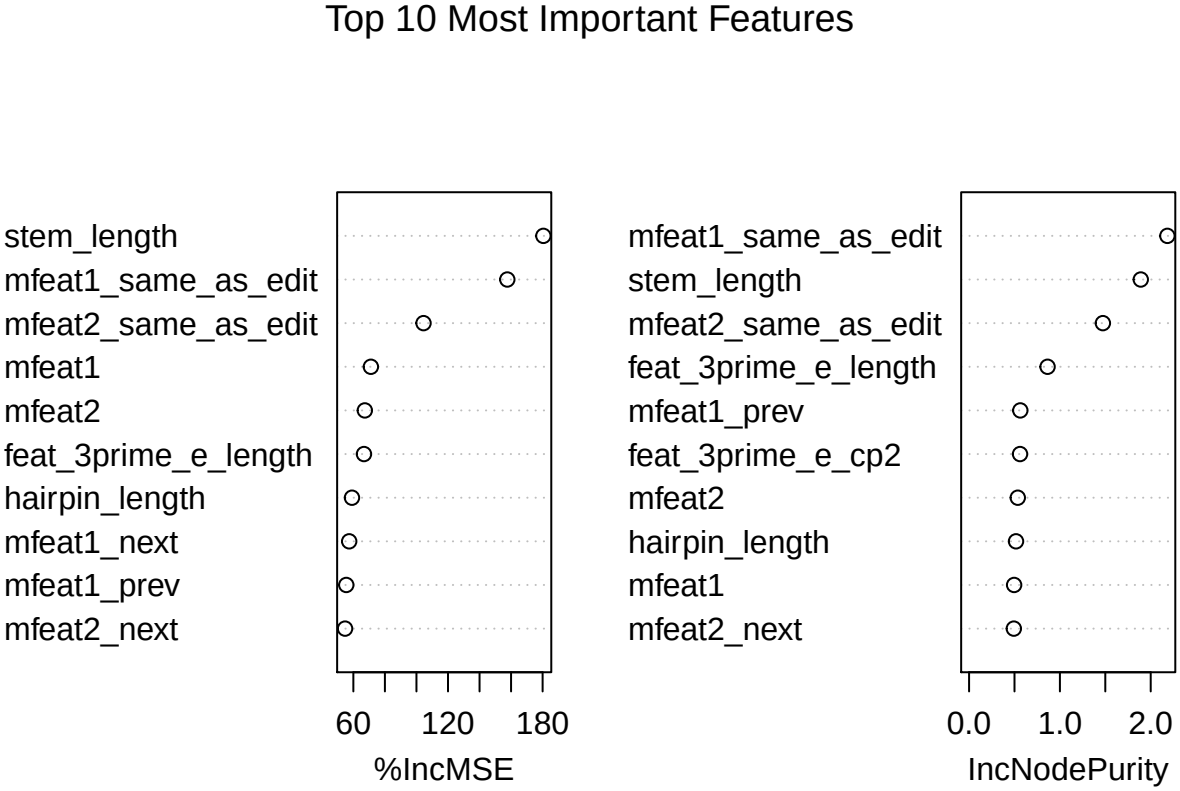
```
## [1] 0.0223757
```



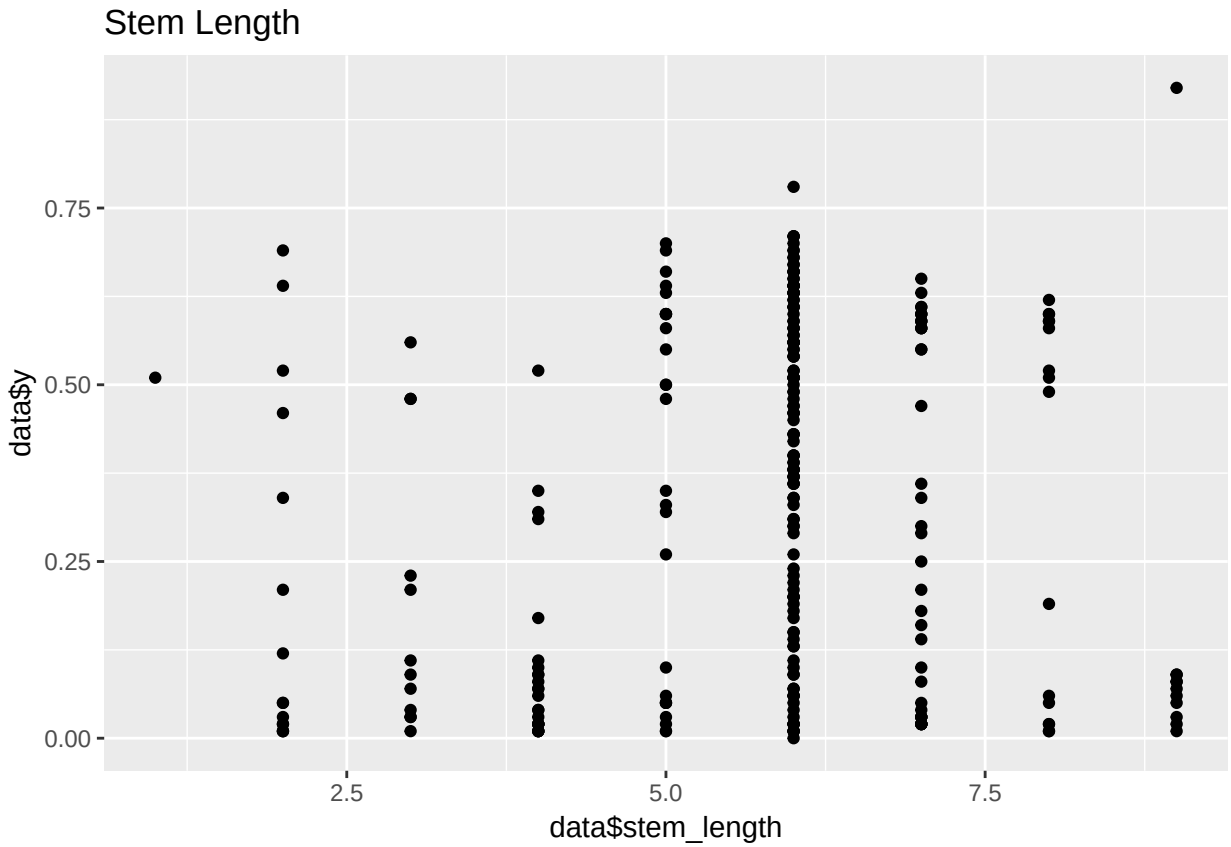
##Feature Importance

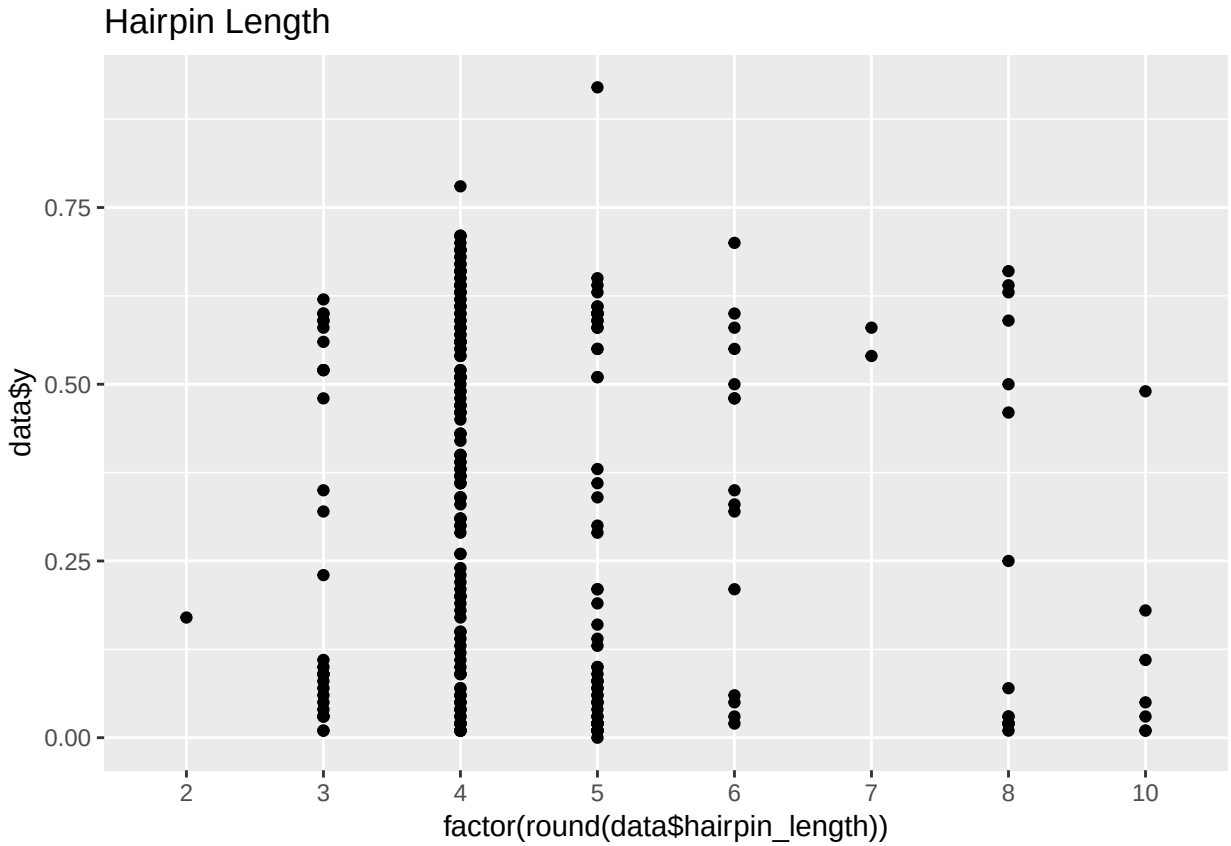
```
##                         %IncMSE IncNodePurity
## stem_length           180.46072     1.8905742
## mfeat1_same_as_edit   157.60555     2.1833565
## mfeat2_same_as_edit   104.47843     1.4731725
## mfeat1                 71.10283     0.4960650
## mfeat2                 67.27690     0.5360758
## feat_3prime_e_length   66.73800     0.8642990
## hairpin_length         59.11610     0.5166704
## mfeat1_next            57.37012     0.2787318
## mfeat1_prev            55.51387     0.5640146
## mfeat2_next            54.74268     0.4927063
## mfeat2_prev            44.91924     0.3618502
## feat_3prime_e_cp2      44.33424     0.5610486
```

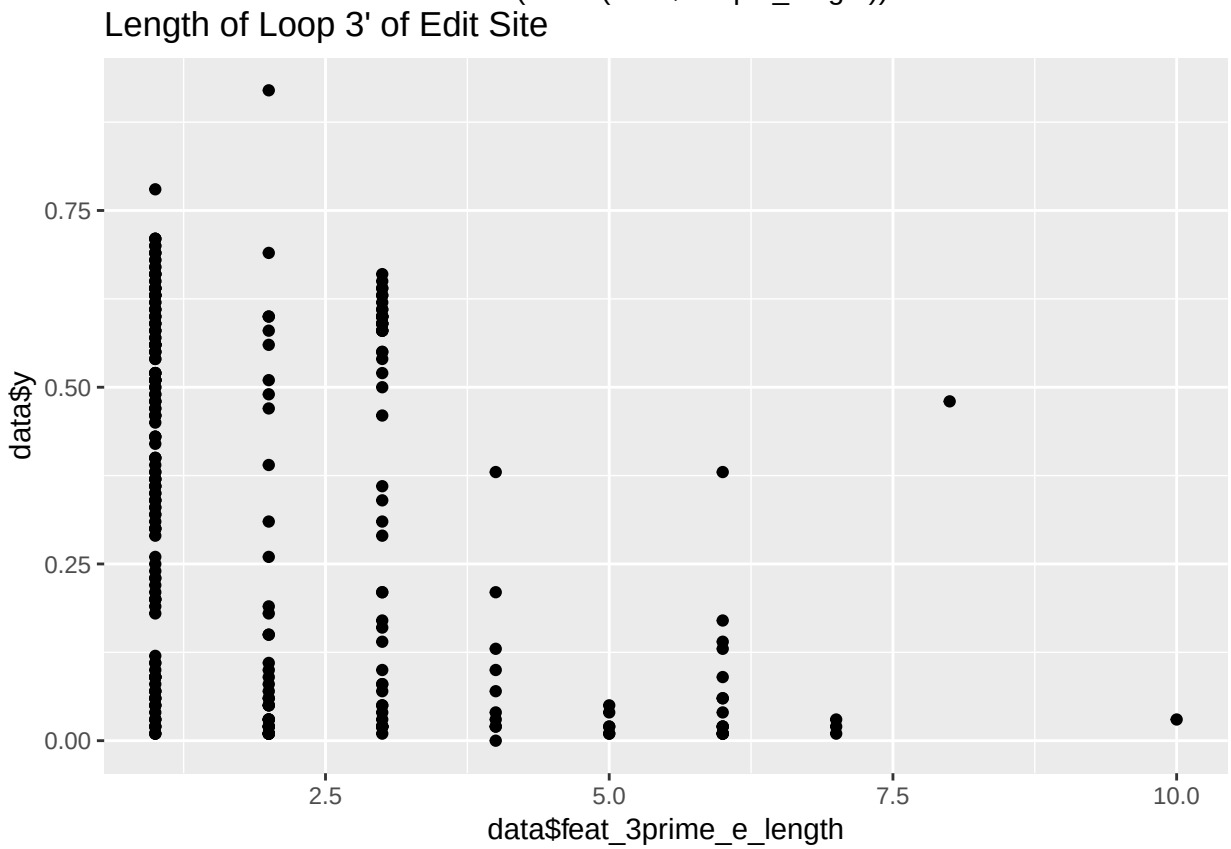
```
## feat_3prime_e_cp1     30.89480     0.2648874
## editing_feature       20.08312     0.2612761
## feat_3prime_e         19.86301     0.1672702
```

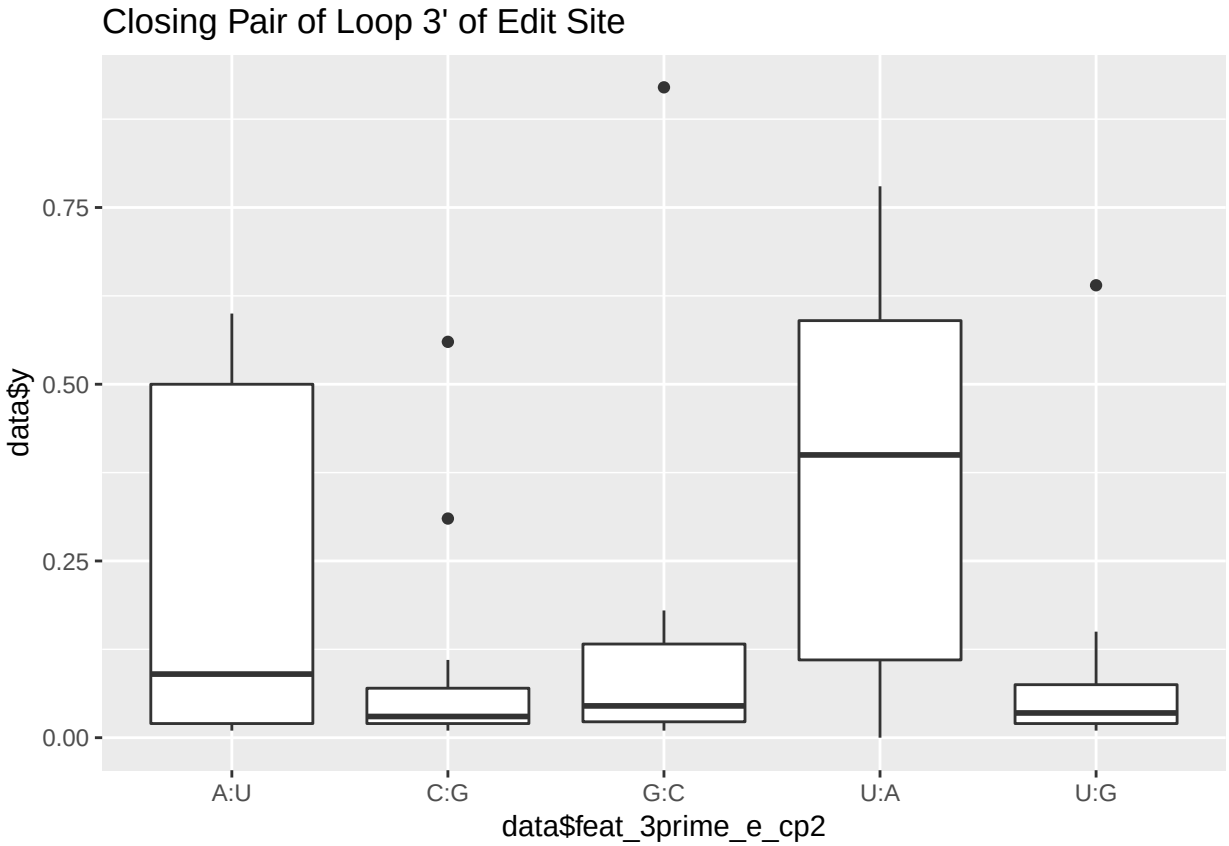


Feature values vs Editing Level




Hairpin Length
data$y
0.75
0.50
0.25
0.00
2
3
4
5
6
7
8
10
factor(round(data$hairpin_length))


Length of Loop 3' of Edit Site
data$y
0.75
0.50
0.25
0.00
2.5
5.0
7.5
10.0
data$feat_3prime_e_length


Closing Pair of Loop 3' of Edit Site
data$y
0.75
0.50
0.25
0.00
A:U
C:G
G:C
U:A
U:G
data$feat_3prime_e_cp2