Untitled

Yuwen

November 20, 2016

Study the effect of promoter, coding, and cg content on mutation rate.

```
load("../analysis/161119_mutation_rate_calibration_new.Rdata")
```

258 control

```
summary(control_model_258)
```

```
##
## Call:
## glm(formula = mut_count ~ coding + promoter + cg + offset(log(2 *
##     mutrate * sample_size)), family = poisson, data = coverage)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -0.0411  -0.0199  -0.0181  -0.0165   5.9030
##
## Coefficients:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)  0.38730    0.09163   4.227 2.37e-05 ***
## coding       0.12557    0.07136   1.760   0.0785 .
## promoter    -0.18504    0.13035  -1.420   0.1557
## cg          -2.34934    0.20142 -11.664  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for poisson family taken to be 1)
##
##     Null deviance: 23023  on 7765273  degrees of freedom
## Residual deviance: 22872  on 7765270  degrees of freedom
## AIC: 25514
##
## Number of Fisher Scoring iterations: 11
```

manuscript cases

```
summary(ASD_model_manuscript)
```

```
##
## Call:
```

```
## glm(formula = mut_count ~ coding + promoter + cg + offset(log(2 *
##     mutrate * sample_size)), family = poisson, data = coverage)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -0.0549  -0.0276  -0.0242  -0.0229   6.8426
##
## Coefficients:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)  0.09386    0.06832   1.374   0.1695
## coding       0.11462    0.04995   2.295   0.0218 *
## promoter    -0.20172    0.08776  -2.299   0.0215 *
## cg          -0.65158    0.14089  -4.625 3.75e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for poisson family taken to be 1)
##
##     Null deviance: 41553  on 7765273  degrees of freedom
## Residual deviance: 41517  on 7765270  degrees of freedom
## AIC: 46702
##
## Number of Fisher Scoring iterations: 10
```

manusript control

```
summary(control_model_693)
```

```
##
## Call:
## glm(formula = mut_count ~ coding + promoter + cg + offset(log(2 *
##     mutrate * sample_size)), family = poisson, data = coverage)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -0.0726  -0.0341  -0.0306  -0.0279   5.3451
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -0.286638   0.054706  -5.240 1.61e-07 ***
## coding       0.308890   0.038324   8.060 7.63e-16 ***
## promoter    -0.005361   0.066404  -0.081    0.936
## cg          -0.708357   0.112633  -6.289 3.19e-10 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for poisson family taken to be 1)
##
##     Null deviance: 60809  on 7765273  degrees of freedom
## Residual deviance: 60717  on 7765270  degrees of freedom
## AIC: 68769
```

```
##
## Number of Fisher Scoring iterations: 9
```

scherer cases, partly overlapped with manuscript cases

```
summary(ASD_model_Scherer)
```

```
##
## Call:
## glm(formula = mut_count ~ coding + promoter + cg + offset(log(2 *
##     mutrate * sample_size)), family = poisson, data = coverage)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -0.0411  -0.0197  -0.0176  -0.0163   5.8729
##
## Coefficients:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept) -0.05012    0.09477  -0.529   0.5969
## coding       0.27751    0.06672   4.160 3.19e-05 ***
## promoter    -0.29610    0.13081  -2.264   0.0236 *
## cg          -0.86703    0.19660  -4.410 1.03e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for poisson family taken to be 1)
##
##     Null deviance: 23087  on 7765273  degrees of freedom
## Residual deviance: 23043  on 7765270  degrees of freedom
## AIC: 25709
##
## Number of Fisher Scoring iterations: 11
```