

Burden

Only count synonymous mutations identified by Annovar that are in the coding regions used in model training for cases. The rate parameter for synonymous mutations for cases thus should be 1. Notice that in model training, promoter regions have a higher priority than coding regions. LoF here is stop gain, frameshift insertion and frameshift deletion.

```
getwd()
```

```
## [1] "/media/yuwen/F/ASD/analysis"
```

```
knitr::opts_chunk$set(warning=FALSE, message=FALSE)
```

read in data with predicted mutation count and observed mutation count

load in functions and datasets that make the functions work

ASD manuscript data

```
mutation = new.env()
load("../data/debug_region_list_073116_data_matrixtransformed_for_old_code.Rdata", envir = mutation)
coding_data_count = get_coding_mut_number(mutation$mutation, mutation$ref_alt_allele, "../other_annotation.Rdata")
compare_pre_with_obs(ASD_model_manuscript_mutrate$coding, coding_data_count$coding_mut_for_gene, geneset = mutation$geneset,
                     mutation_scaling = list(synonymous = 1, nonsynonymous = 2, LoF = 0.1))
```

##	geneset	mut_type	predict	observed
## 1	stringent_ASD	synonymous	10.1995010727898	3
## 2	stringent_ASD	nonsynonymous	20.3990021455796	6
## 3	stringent_ASD	LoF	1.01995010727898	7
## 4	relaxed_ASD	synonymous	65.4341850095092	13
## 5	relaxed_ASD	nonsynonymous	130.868370019018	57
## 6	relaxed_ASD	LoF	6.54341850095092	14
## 7	nonASD_genes	synonymous	97.8309800461363	22
## 8	nonASD_genes	nonsynonymous	195.661960092273	72
## 9	nonASD_genes	LoF	9.78309800461363	3
## 10	constraint_union	synonymous	75.8908770590831	22
## 11	constraint_union	nonsynonymous	151.781754118166	79
## 12	constraint_union	LoF	7.58908770590831	9
##	pvalue			
## 1	0.997649127533647			
## 2	0.999947039006398			
## 3	9.39559341669494e-05			
## 4	0.999999999999999			
## 5	0.9999999999999876			
## 6	0.00749625562806148			
## 7	1			
## 8	1			
## 9	0.996693026484367			
## 10	0.999999999999991			
## 11	0.999999999970451			
## 12	0.350301275135162			

control_model_693_mutrate

##	geneset	mut_type	predict	observed
## 1	stringent_ASD	synonymous	3.42165058765449	7
## 2	stringent_ASD	nonsynonymous	8.55412646913622	10
## 3	stringent_ASD	LoF	0.171082529382724	0
## 4	relaxed_ASD	synonymous	21.9195738522145	27
## 5	relaxed_ASD	nonsynonymous	54.7989346305362	78
## 6	relaxed_ASD	LoF	1.09597869261072	3
## 7	nonASD_genes	synonymous	32.610312443929	33
## 8	nonASD_genes	nonsynonymous	81.5257811098225	126
## 9	nonASD_genes	LoF	1.63051562219645	3
## 10	constraint_union	synonymous	25.2355243262198	34
## 11	constraint_union	nonsynonymous	63.0888108155496	89
## 12	constraint_union	LoF	1.26177621631099	6
##	pvalue			
## 1	0.059416339330749			
## 2	0.354066446486951			
## 3	1			
## 4	0.163162942968121			
## 5	0.00183037746177939			
## 6	0.0987752139385415			
## 7	0.496018603647879			
## 8	3.00626664394074e-06			
## 9	0.224556792655147			
## 10	0.0550801916976613			
## 11	0.00120700169039726			
## 12	0.00192541643232518			

control_258_model,

##	geneset	mut_type	predict	observed
## 1	stringent_ASD	synonymous	0.878807779203229	0
## 2	stringent_ASD	nonsynonymous	2.19701944800807	1
## 3	stringent_ASD	LoF	0.0439403889601615	1
## 4	relaxed_ASD	synonymous	5.40963536115468	6
## 5	relaxed_ASD	nonsynonymous	13.5240884028867	15
## 6	relaxed_ASD	LoF	0.270481768057734	2
## 7	nonASD_genes	synonymous	8.1877345941075	13
## 8	nonASD_genes	nonsynonymous	20.4693364852688	30
## 9	nonASD_genes	LoF	0.409386729705375	0
## 10	constraint_union	synonymous	6.47706953792414	13
## 11	constraint_union	nonsynonymous	16.1926738448104	28
## 12	constraint_union	LoF	0.323853476896207	3
##	pvalue			
## 1	1			
## 2	0.888866094403488			
## 3	0.0429889958061805			
## 4	0.455532493080279			
## 5	0.379259847103119			
## 6	0.0306074052397577			
## 7	0.0732593526945869			
## 8	0.0283403932458275			

```
## 9 1
## 10 0.0156212225315957
## 11 0.00479555350584289
## 12 0.00444916017083179
```

Scherer case data

##	geneset	mut_type	predict	observed
## 1	stringent_AS	synonymous	1.20281416399903	1
## 2	stringent_AS	nonsynonymous	3.00703540999757	4
## 3	stringent_AS	LoF	0.0601407081999514	1
## 4	relaxed_AS	synonymous	7.69472324062565	8
## 5	relaxed_AS	nonsynonymous	19.2368081015641	28
## 6	relaxed_AS	LoF	0.384736162031283	4
## 7	nonASD_genes	synonymous	11.4885480910204	12
## 8	nonASD_genes	nonsynonymous	28.7213702275511	47
## 9	nonASD_genes	LoF	0.574427404551022	2
## 10	constraint_union	synonymous	8.91336472966085	11
## 11	constraint_union	nonsynonymous	22.2834118241521	43
## 12	constraint_union	LoF	0.445668236483042	3

##	pvalue
## 1	0.699652206457023
## 2	0.354344332861948
## 3	0.0583679710855846
## 4	0.503679021301917
## 5	0.0355702109421059
## 6	0.000672428212878325
## 7	0.478849462643799
## 8	0.00106965636799344
## 9	0.113554604566569
## 10	0.283789332721496
## 11	6.34167613326605e-05
## 12	0.0106017563587607