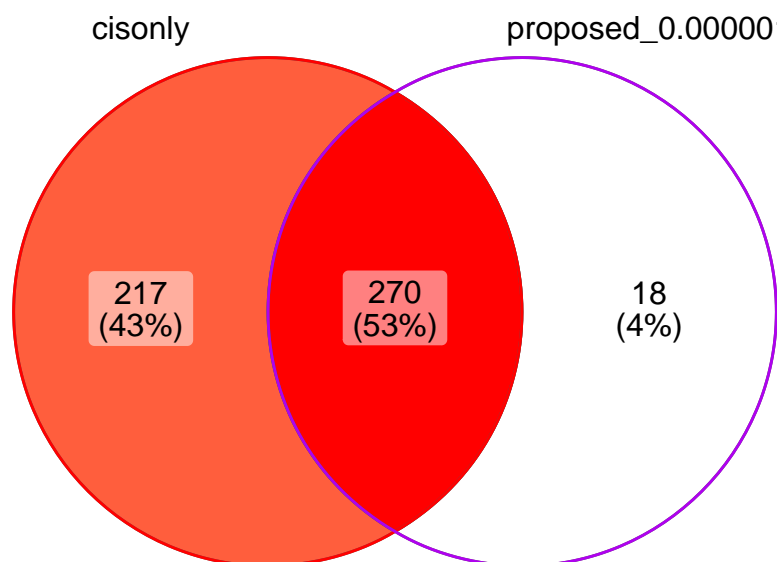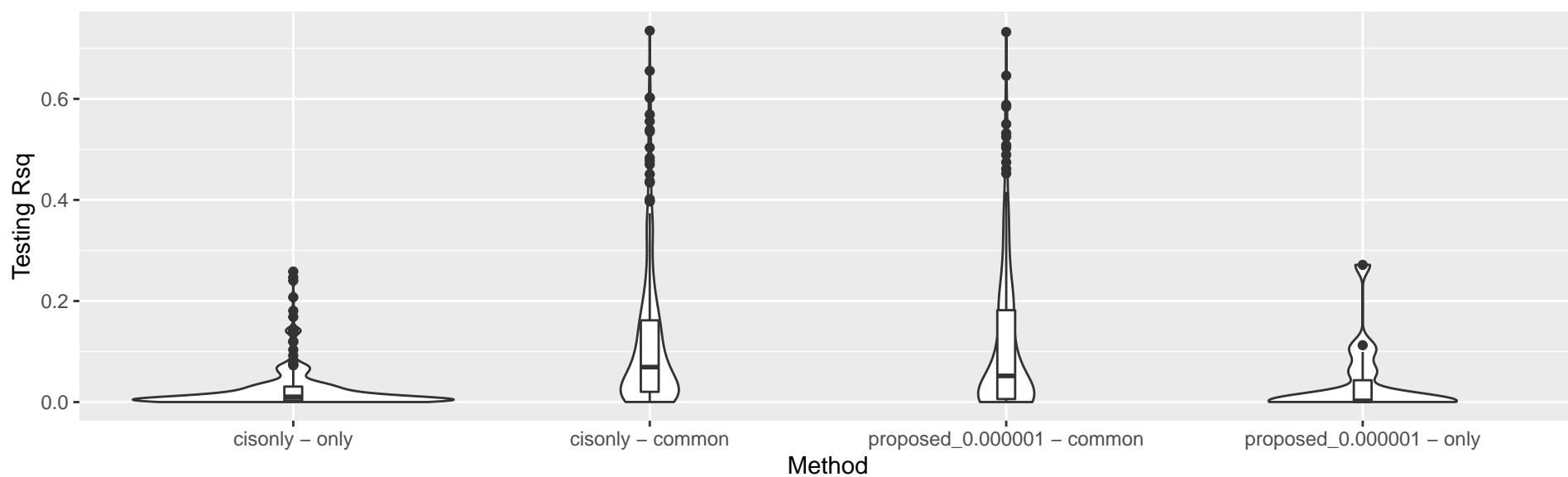## Number of successfully trained protein prediction model

Testing Rsq Summary statistics

	cisonly – only	cisonly – common	proposed_0.000001 – common	proposed_0.000001 – only
Min.	0.0001098	0.0002089	0.0001039	0.0002338
1st Qu.	0.003128	0.02029	0.006033	0.001421
Median	0.01047	0.06923	0.05186	0.002414
Mean	0.0272	0.1193	0.1116	0.03512
3rd Qu.	0.03077	0.1618	0.1819	0.04326
Max.	0.2583	0.7347	0.7323	0.2715

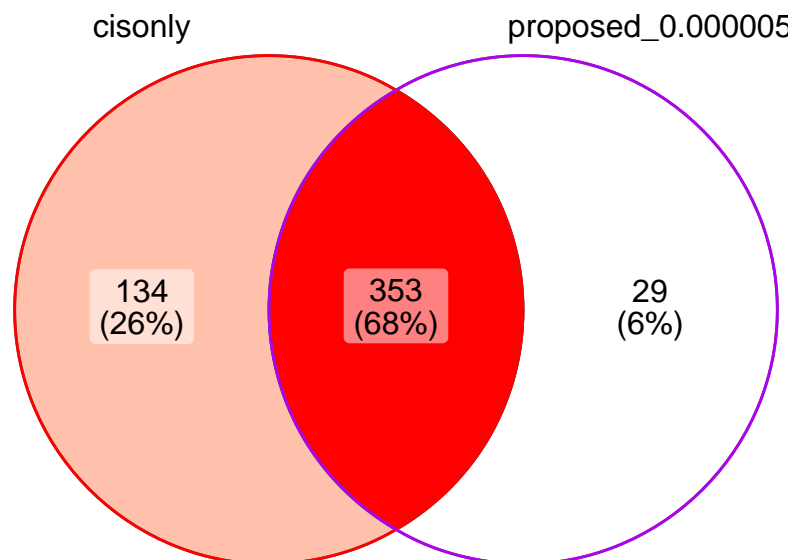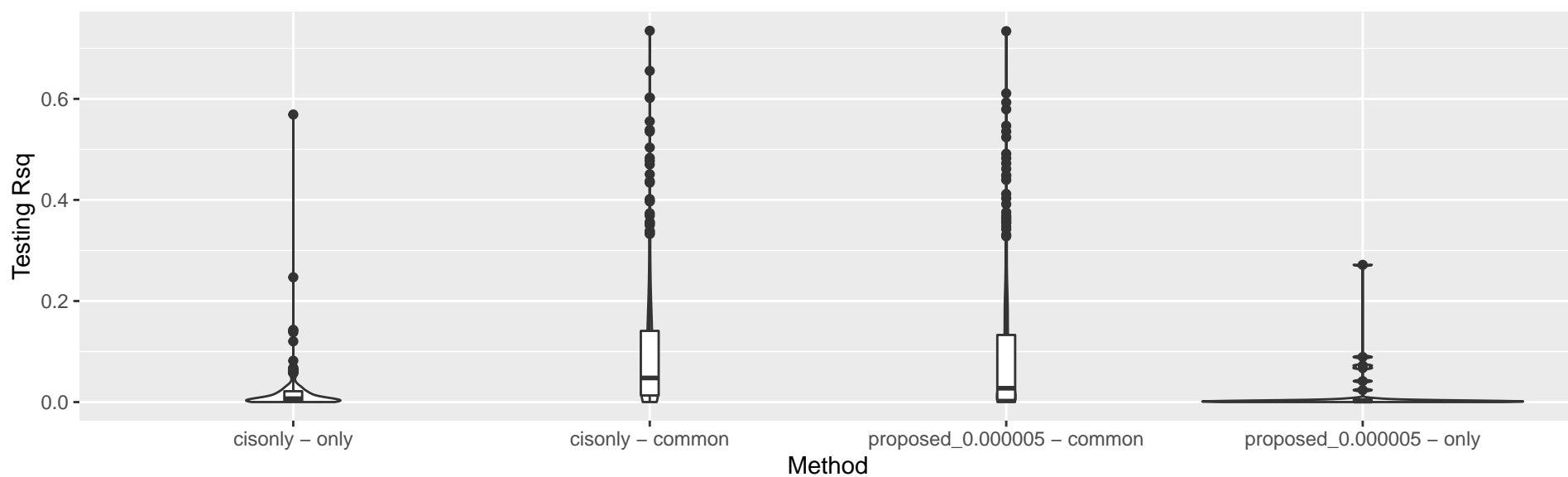## Testing Rsq Violin Plots

Number of successfully trained protein prediction model

Testing Rsq Summary statistics

	cisonly – only	cisonly – common	proposed_0.000005 – common	proposed_0.000005 – only
Min.	0.0001133	0.0001098	0.0001364	0.000119
1st Qu.	0.001532	0.01314	0.004207	0.0004343
Median	0.006785	0.04775	0.02745	0.001917
Mean	0.0233	0.09914	0.0921	0.02088
3rd Qu.	0.02155	0.1409	0.1329	0.005433
Max.	0.5691	0.7347	0.734	0.2716

Testing Rsq Violin Plots
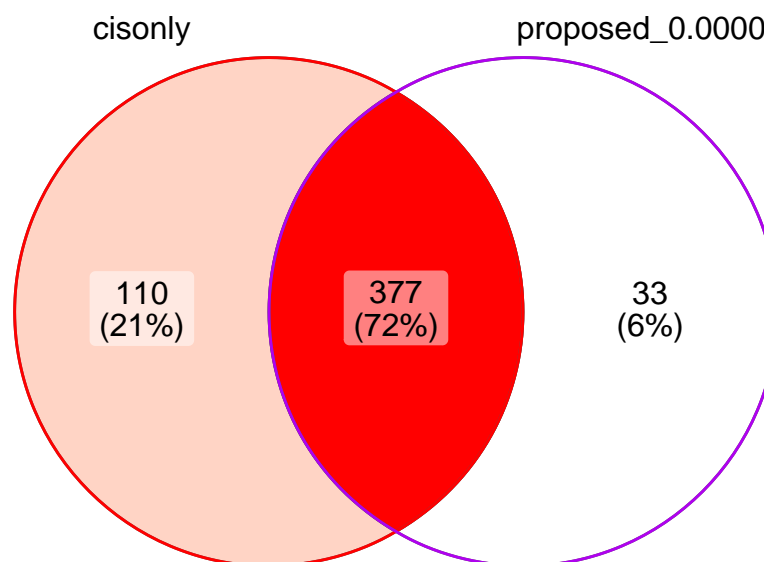
Number of successfully trained protein prediction model

cisonly   proposed_0.00001

| | 110
(21%) | 377
(72%) | 33
(6%) |

Testing Rsq Summary statistics

	cisonly – only	cisonly – common	proposed_0.00001 – common	proposed_0.00001 – only
Min.	0.0001133	0.0001098	0.0001198	0.0001104
1st Qu.	0.001807	0.01119	0.005346	0.0004372
Median	0.007024	0.04075	0.0285	0.002055
Mean	0.01717	0.0961	0.08882	0.01729
3rd Qu.	0.02105	0.1328	0.1259	0.009186
Max.	0.1428	0.7347	0.728	0.2849

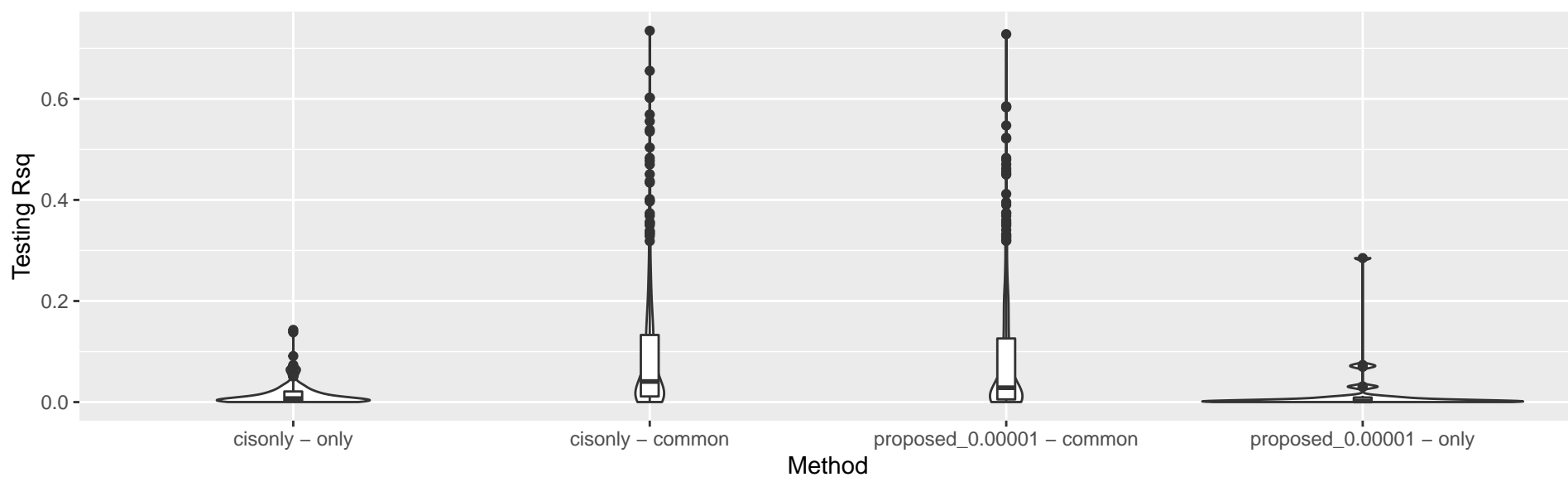## Testing Rsq Violin Plots
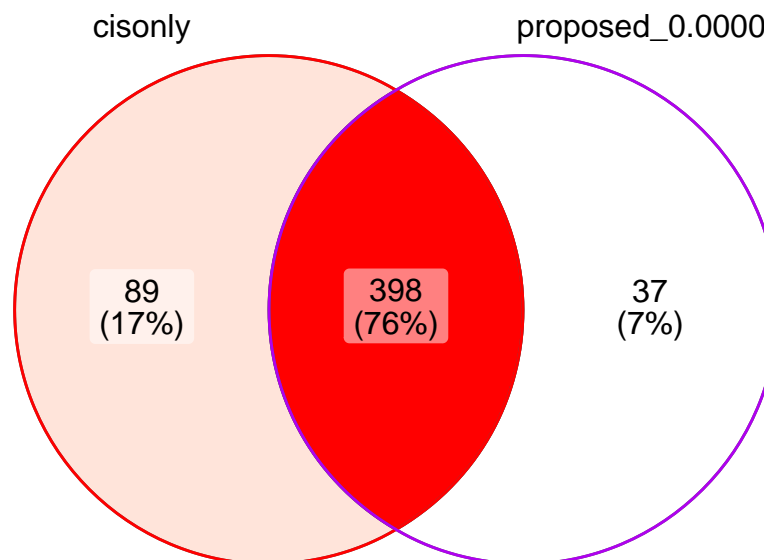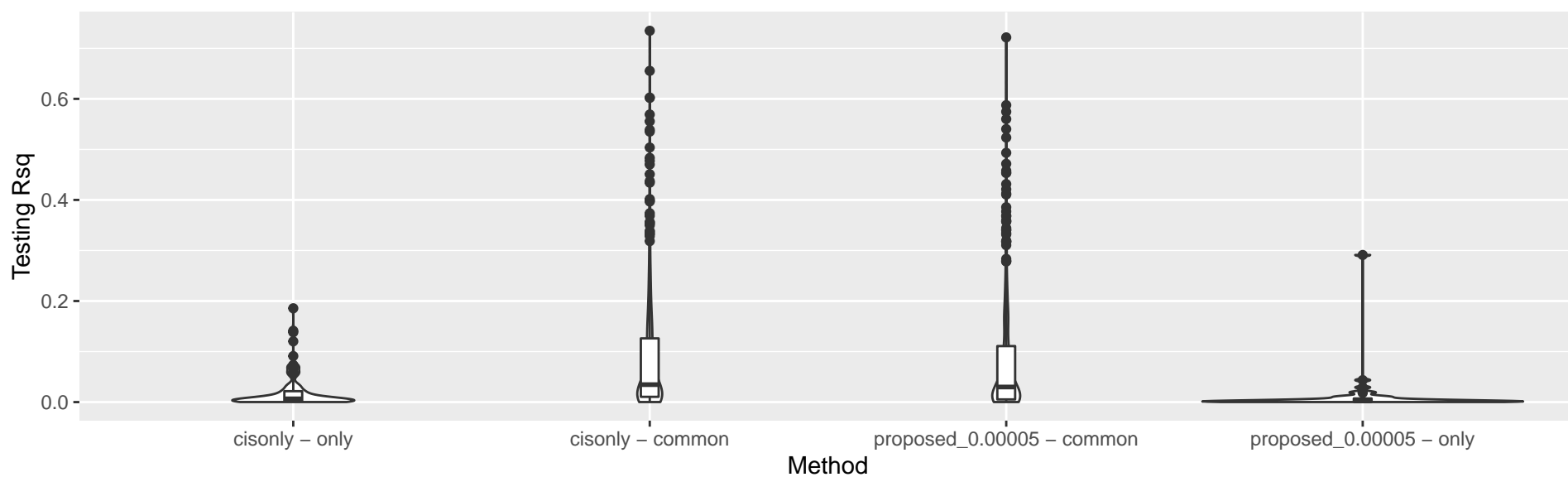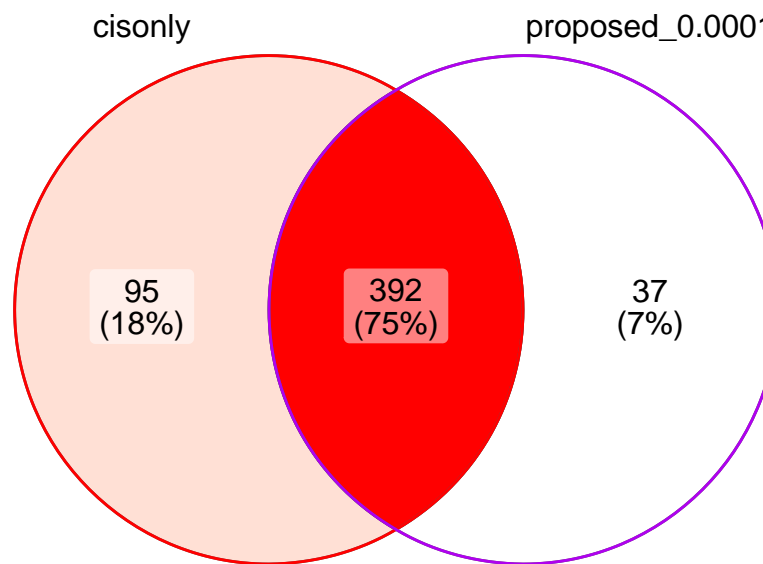
Number of successfully trained protein prediction model

Testing Rsq Summary statistics

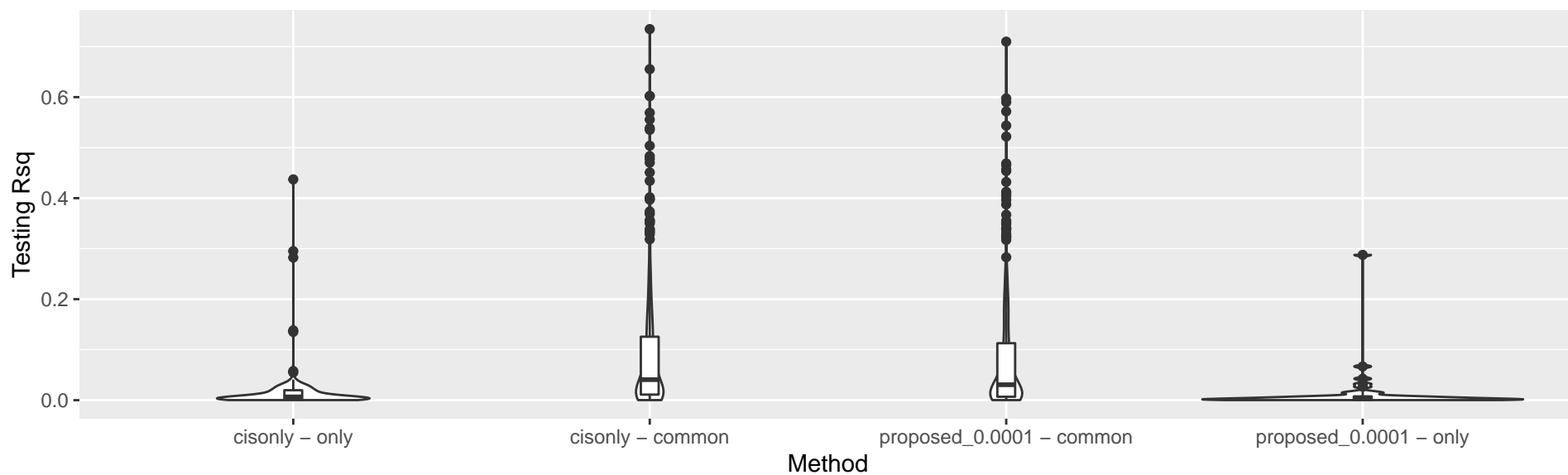|	cisonly – only	cisonly – common	proposed_0.00005 – common	proposed_0.00005 – only
Min.	0.0001133	0.0001098	0.0001114	0.0001268
1st Qu.	0.00164	0.01043	0.005325	0.0007787
Median	0.006449	0.03458	0.02988	0.002216
Mean	0.02139	0.09099	0.08479	0.01361
3rd Qu.	0.02161	0.1261	0.1108	0.007208
Max.	0.1856	0.7347	0.7216	0.2911

Testing Rsq Violin Plots

Number of successfully trained protein prediction model

Testing Rsq Summary statistics

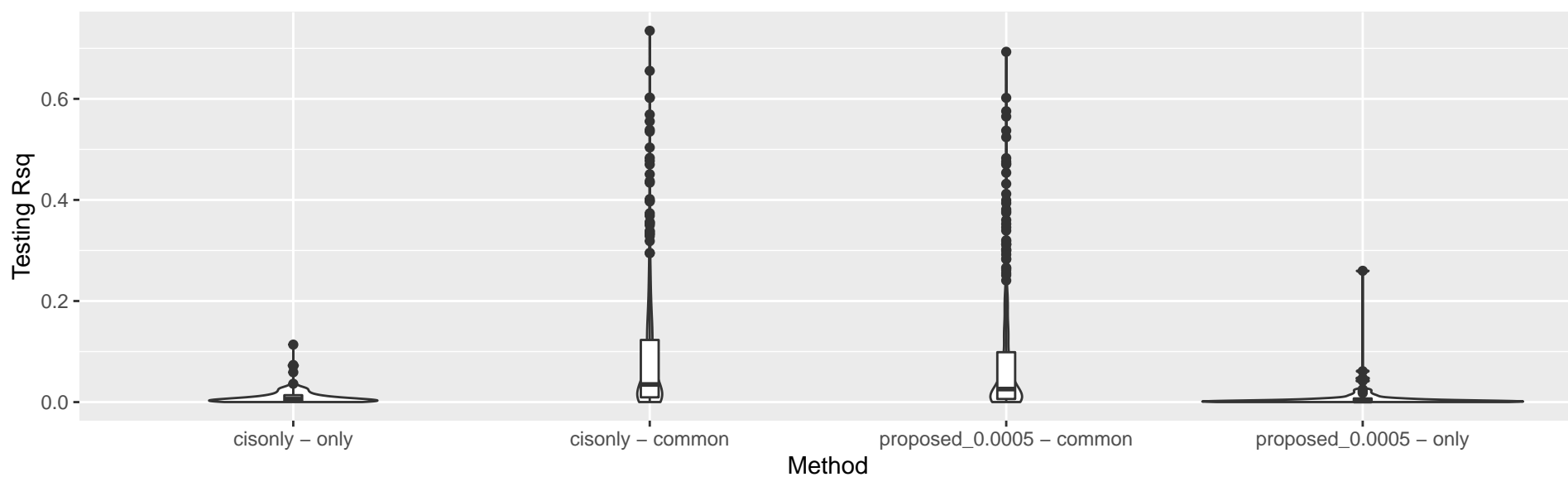|	cisonly – only	cisonly – common	proposed_0.0001 – common	proposed_0.0001 – only
Min.	0.0001098	0.0001133	0.0001102	0.0001034
1st Qu.	0.0017	0.01101	0.006716	0.0008864
Median	0.006449	0.04059	0.03051	0.003572
Mean	0.02362	0.09152	0.08315	0.01577
3rd Qu.	0.01968	0.1256	0.1128	0.007445
Max.	0.4372	0.7347	0.7099	0.2876

Testing Rsq Violin Plots
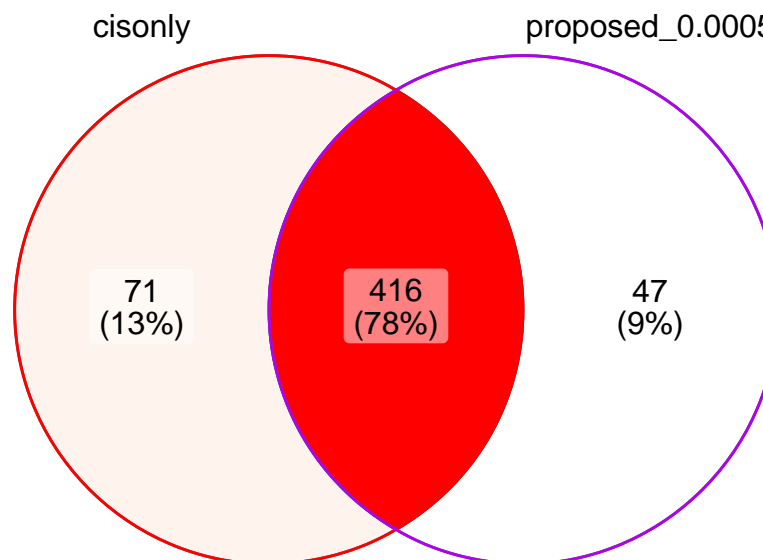
Number of successfully trained protein prediction model

Testing Rsq Summary statistics

	cisonly – only	cisonly – common	proposed_0.0005 – common	proposed_0.0005 – only
Min.	0.0001098	0.0001201	0.0001282	0.0001419
1st Qu.	0.001586	0.009441	0.00611	0.0005112
Median	0.005929	0.0349	0.02583	0.001874
Mean	0.01243	0.08951	0.07897	0.01268
3rd Qu.	0.01368	0.1229	0.09872	0.007233
Max.	0.1137	0.7347	0.6931	0.2598

Testing Rsq Violin Plots
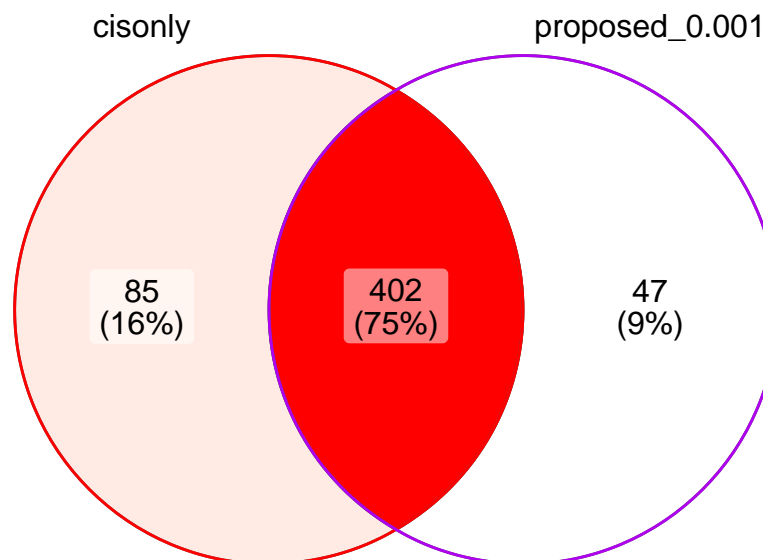
Number of successfully trained protein prediction model

cisonly   proposed_0.001

| | 85 (16%) | 402 (75%) | 47 (9%) |

Testing Rsq Summary statistics

|	cisonly – only	cisonly – common	proposed_0.001 – common	proposed_0.001 – only
Min.	0.0001133	0.0001098	0.0001092	0.0001213
1st Qu.	0.002331	0.01034	0.006256	0.0006594
Median	0.006729	0.0349	0.02636	0.00221
Mean	0.03394	0.08764	0.0766	0.01367
3rd Qu.	0.01696	0.12	0.09489	0.007792
Max.	0.5387	0.7347	0.7087	0.2638

Testing Rsq Violin Plots
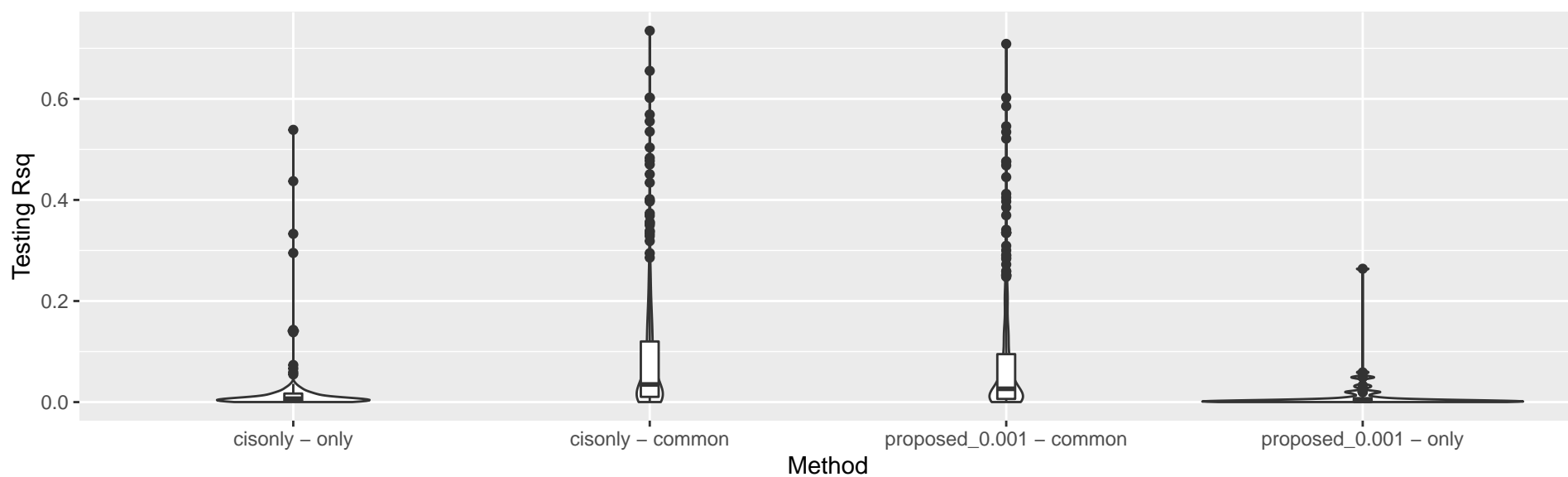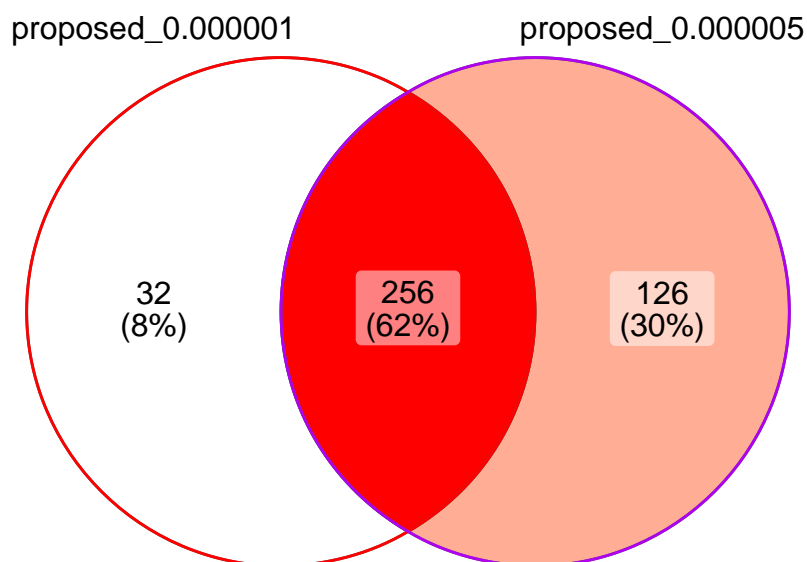
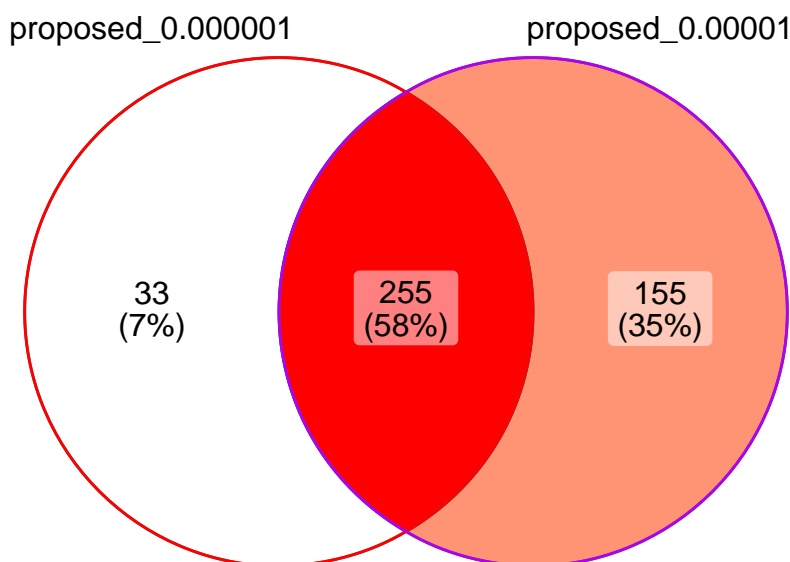# Number of successfully trained protein prediction model

Testing Rsq Summary statistics

	proposed_0.000001 – only	proposed_0.000001 – common	proposed_0.000005 – common	proposed_0.000005 – onl
Min.	0.0002242	0.0001039	0.0001364	0.000119
1st Qu.	0.000744	0.008426	0.01308	0.001048
Median	0.001668	0.06163	0.06665	0.003248
Mean	0.004018	0.1197	0.1225	0.01384
3rd Qu.	0.005393	0.1882	0.1851	0.01056
Max.	0.02734	0.7323	0.734	0.257

Testing Rsq Violin Plots
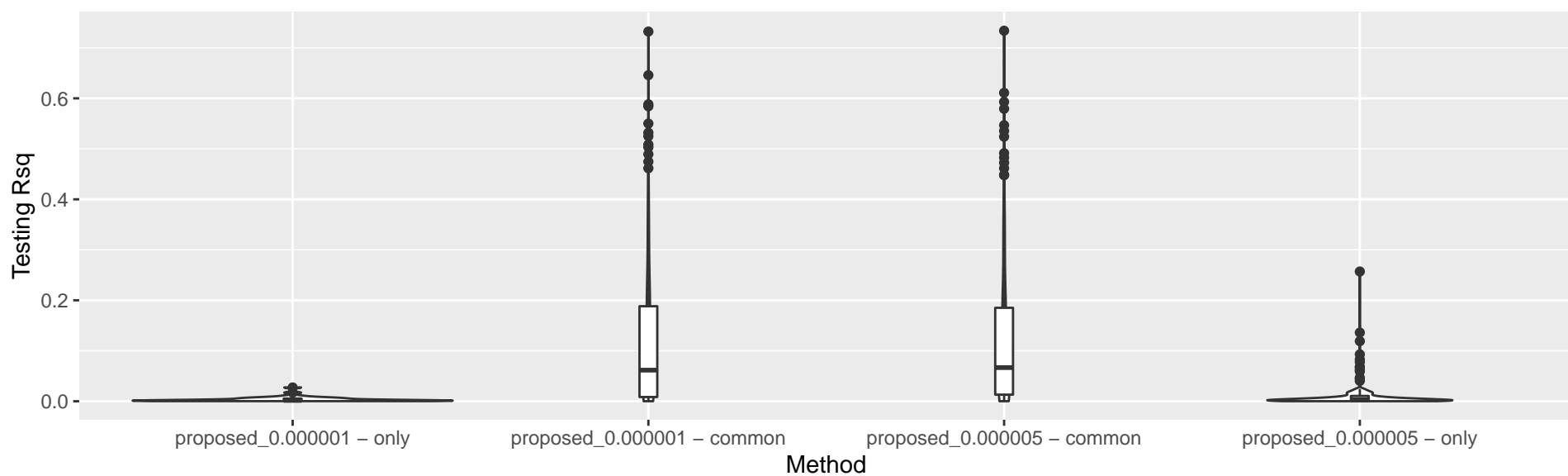
Number of successfully trained protein prediction model

proposed_0.000001   proposed_0.00001

	proposed_0.000001 – only	proposed_0.000001 – common	proposed_0.00001 – common	proposed_0.00001 – only
Min.	0.0002049	0.0001039	0.000119	0.0001104
1st Qu.	0.0008221	0.007971	0.01343	0.001425
Median	0.002122	0.0618	0.06977	0.005671
Mean	0.005839	0.1199	0.1237	0.01626
3rd Qu.	0.006071	0.1885	0.1885	0.01727
Max.	0.04127	0.7323	0.728	0.2592

Testing Rsq Summary statistics
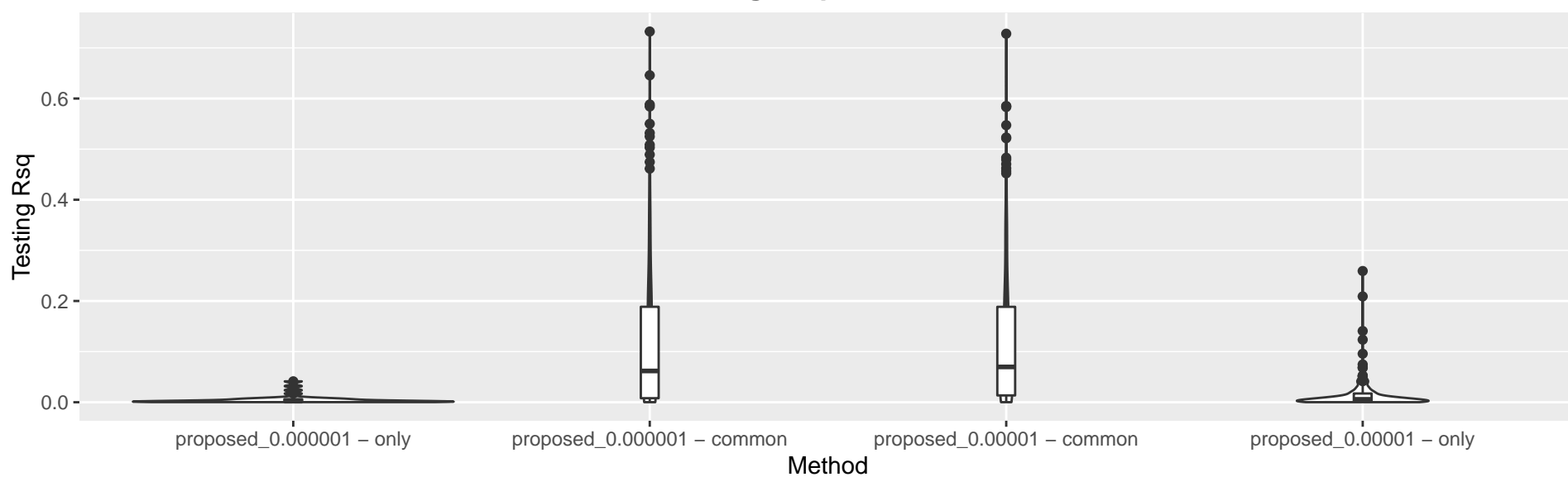
Testing Rsq Violin Plots

Number of successfully trained protein prediction model

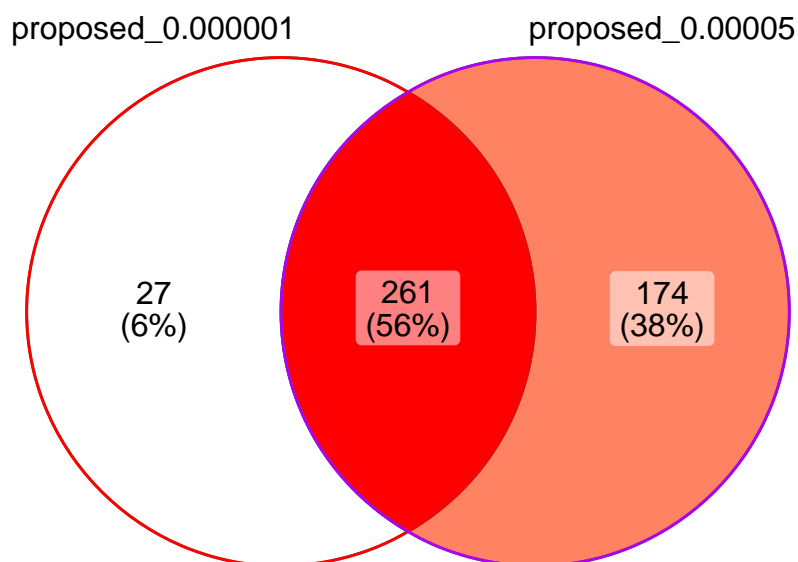proposed_0.000001   proposed_0.00005

27
(6%)

261
(56%)

174
(38%)
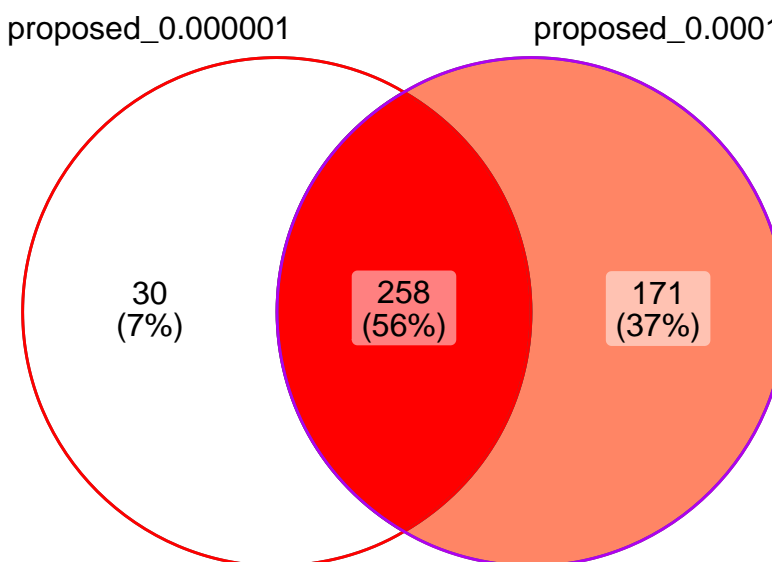
Testing Rsq Summary statistics

	proposed_0.000001 – only	proposed_0.000001 – common	proposed_0.00005 – common	proposed_0.00005 – only
Min.	0.0001228	0.0001039	0.0001255	0.0001114
1st Qu.	0.0008077	0.007089	0.0189	0.001614
Median	0.002313	0.05521	0.06354	0.00543
Mean	0.0184	0.116	0.119	0.01841
3rd Qu.	0.01228	0.1869	0.1727	0.0202
Max.	0.1819	0.7323	0.7216	0.239

Testing Rsq Violin Plots
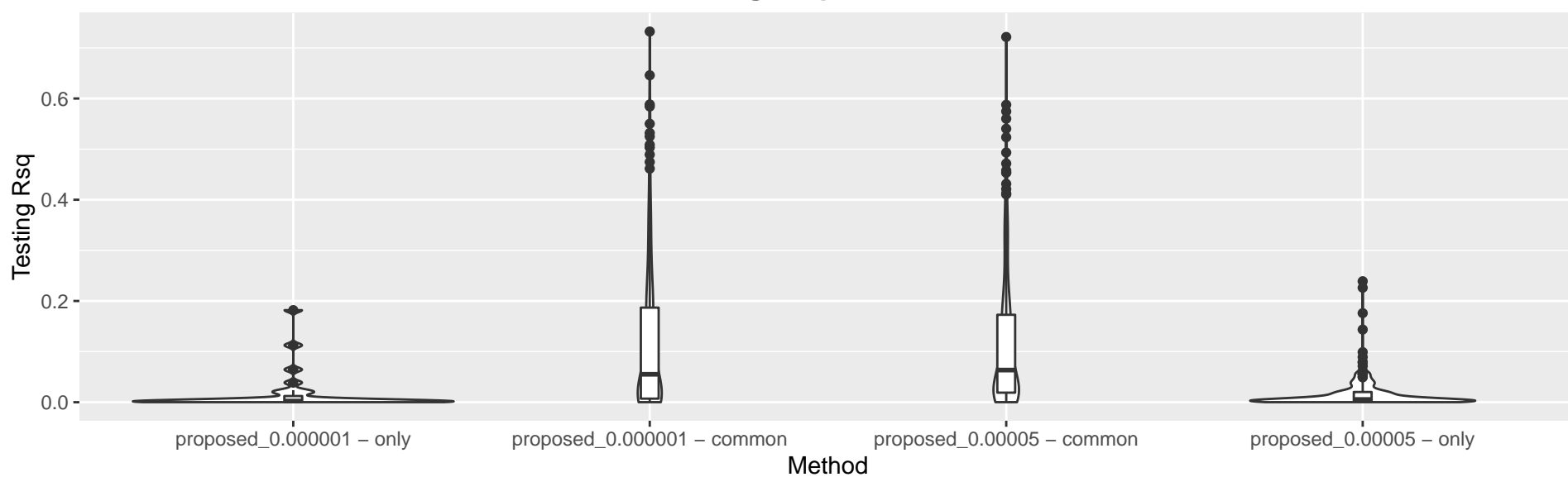
Number of successfully trained protein prediction model

proposed_0.000001   proposed_0.0001

30
(7%)
258
(56%)
171
(37%)

Testing Rsq Summary statistics

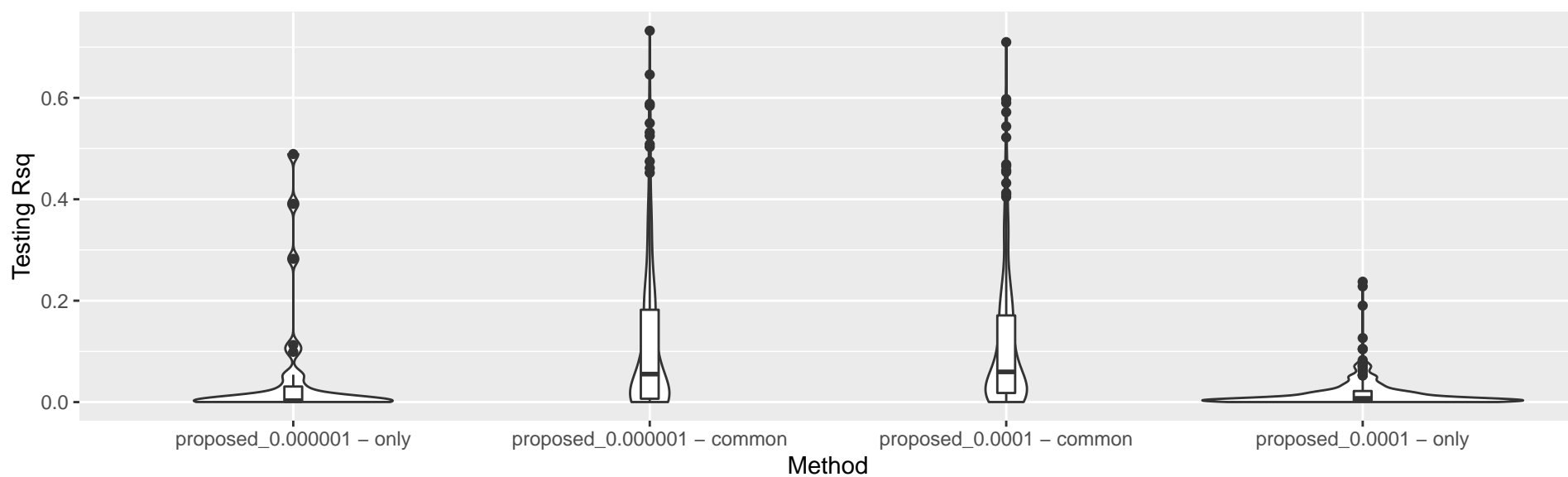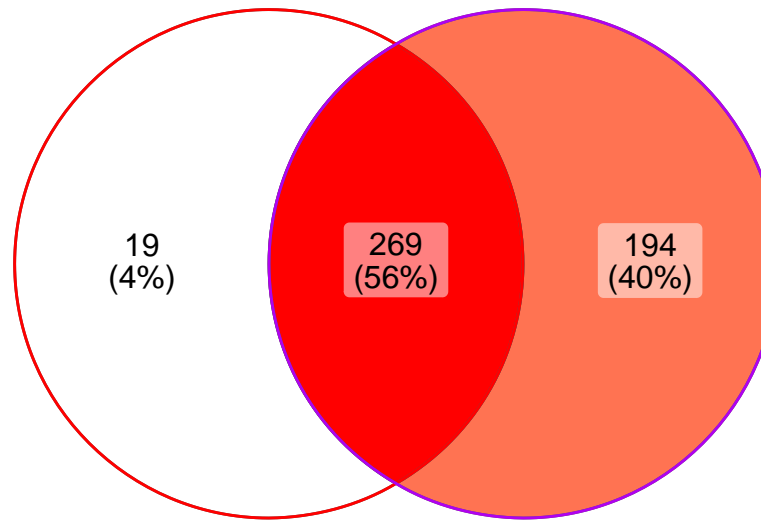|	proposed_0.000001 – only	proposed_0.000001 – common	proposed_0.0001 – common	proposed_0.0001 – only
Min.	0.0002049	0.0001039	0.0001102	0.0001034
1st Qu.	0.001341	0.006613	0.01795	0.002417
Median	0.002864	0.05519	0.05958	0.007357
Mean	0.05348	0.113	0.1152	0.02018
3rd Qu.	0.03072	0.1821	0.1709	0.02183
Max.	0.4892	0.7323	0.7099	0.2374

Testing Rsq Violin Plots

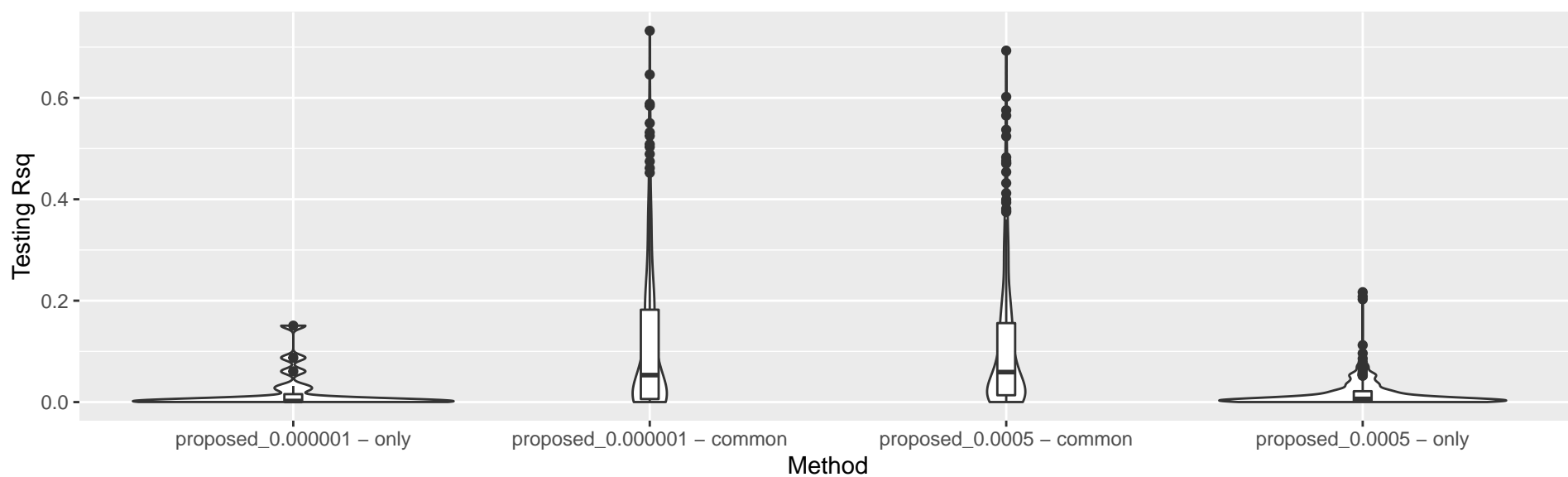# Number of successfully trained protein prediction model

Testing Rsq Summary statistics

	proposed_0.000001 – only	proposed_0.000001 – common	proposed_0.0005 – common	proposed_0.0005 – only
Min.	0.0002049	0.0001039	0.0001282	0.0001309
1st Qu.	0.0008401	0.006071	0.01353	0.001652
Median	0.002122	0.05326	0.05915	0.006263
Mean	0.02014	0.113	0.111	0.01852
3rd Qu.	0.01571	0.1822	0.1558	0.02151
Max.	0.1507	0.7323	0.6931	0.2168

Testing Rsq Violin Plots

Number of successfully trained protein prediction model

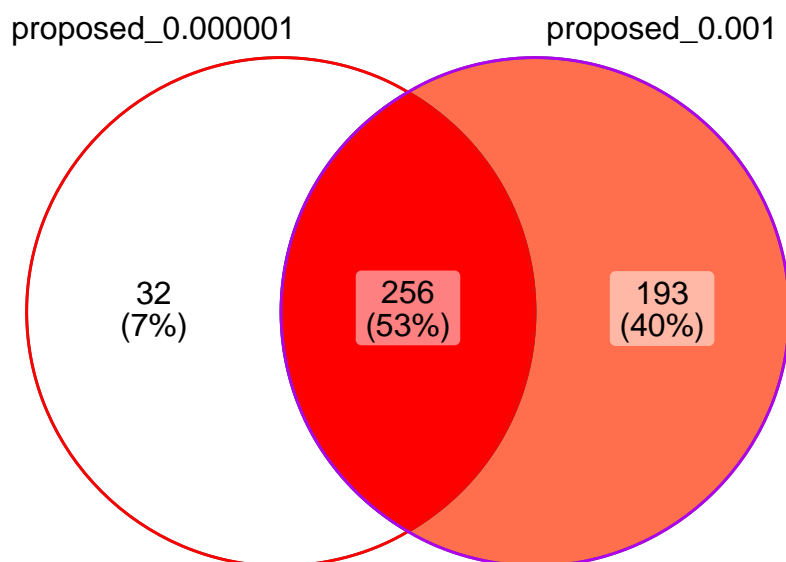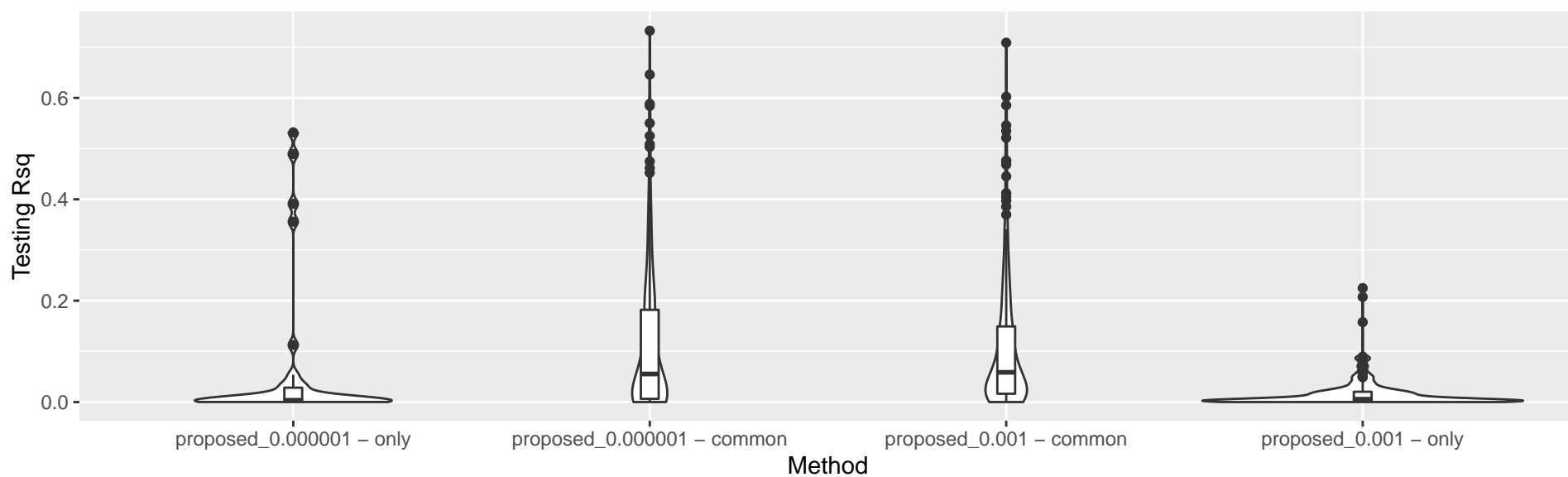## Testing Rsq Summary statistics

	proposed_0.000001 – only	proposed_0.000001 – common	proposed_0.001 – common	proposed_0.001 – only
Min.	0.0002049	0.0001039	0.0001142	0.0001092
1st Qu.	0.001112	0.006423	0.01643	0.001745
Median	0.003642	0.05545	0.05869	0.005684
Mean	0.06595	0.1119	0.1096	0.01753
3rd Qu.	0.02847	0.182	0.1492	0.02042
Max.	0.532	0.7323	0.7087	0.2249

Testing Rsq Violin Plots

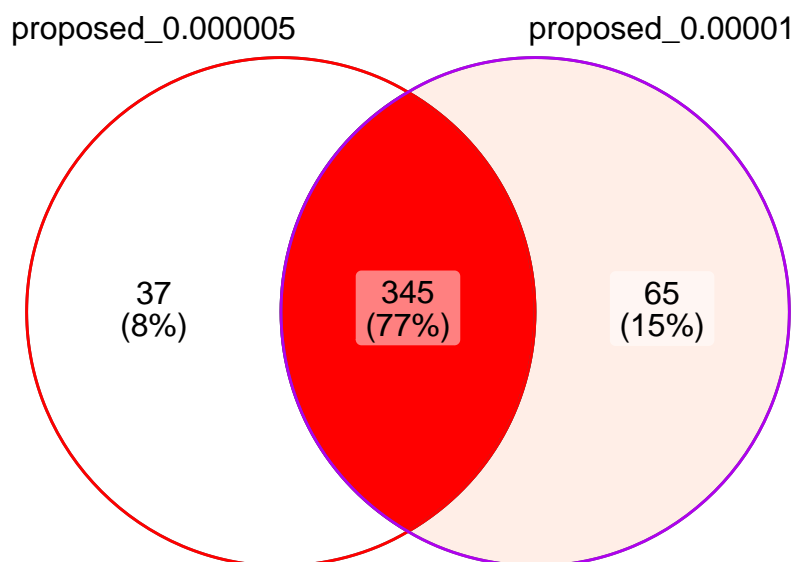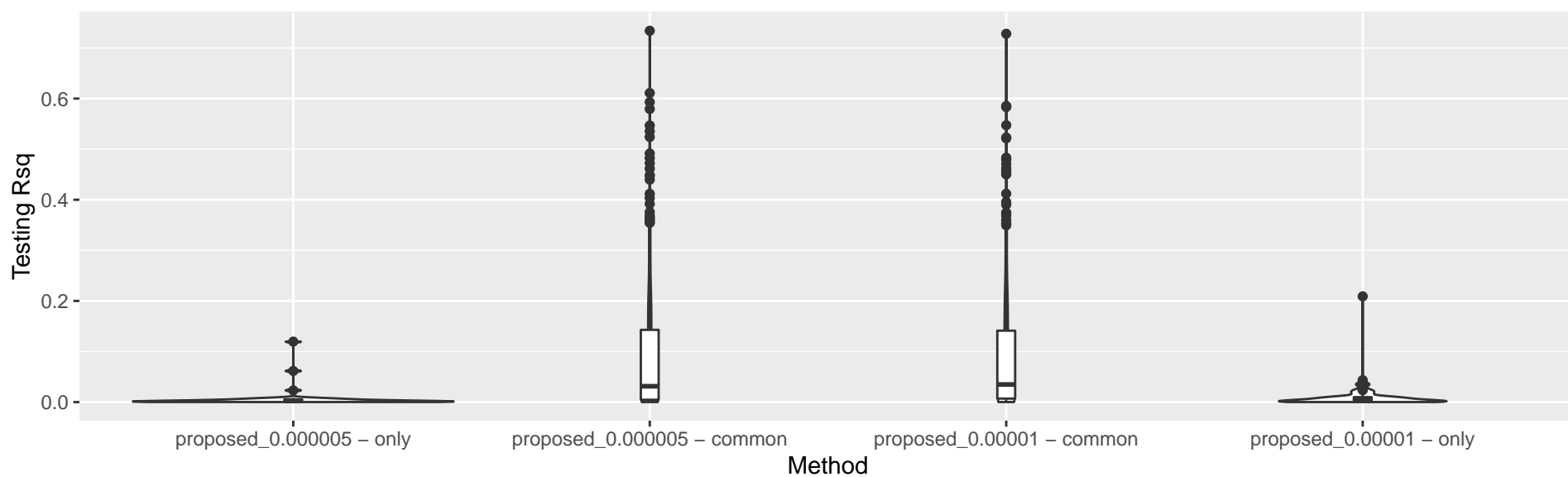## Number of successfully trained protein prediction model

Testing Rsq Summary statistics

	proposed_0.000005 – only	proposed_0.000005 – common	proposed_0.00001 – common	proposed_0.00001 – only
Min.	0.0001572	0.000119	0.0001104	0.0001198
1st Qu.	0.0003817	0.004628	0.006931	0.0008353
Median	0.002206	0.03133	0.03483	0.004578
Mean	0.007805	0.09515	0.09657	0.01137
3rd Qu.	0.005405	0.1428	0.1412	0.009823
Max.	0.1195	0.734	0.728	0.209

Testing Rsq Violin Plots

Number of successfully trained protein prediction model

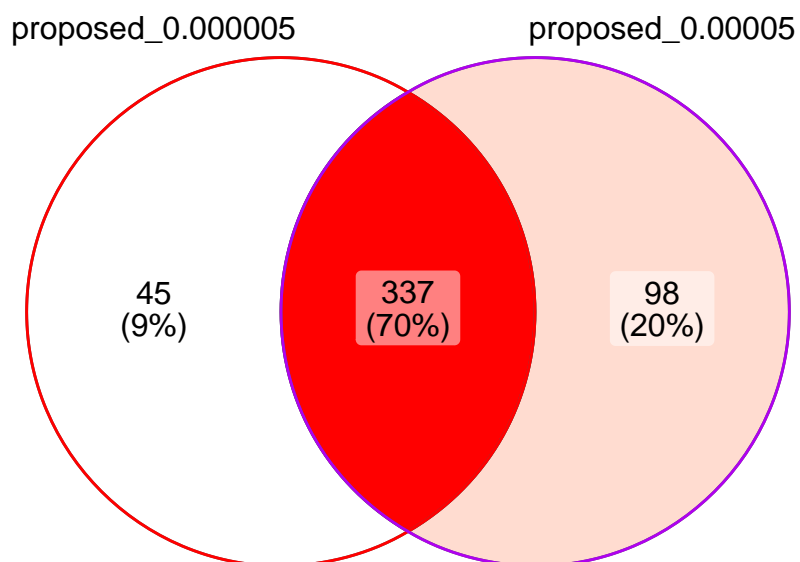## Testing Rsq Summary statistics

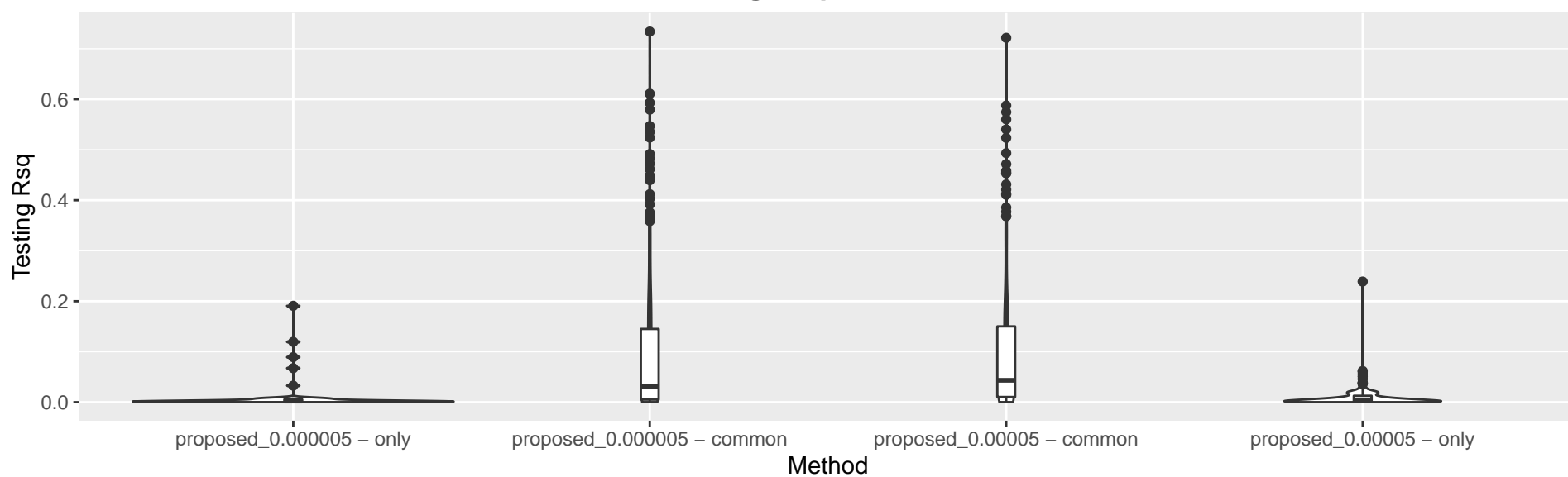|	proposed_0.000005 – only	proposed_0.000005 – common	proposed_0.00005 – common	proposed_0.00005 – only
Min.	0.0001652	0.000119	0.0001186	0.0001114
1st Qu.	0.0005125	0.005181	0.01038	0.001001
Median	0.002237	0.03137	0.04343	0.003963
Mean	0.01341	0.09648	0.0981	0.01215
3rd Qu.	0.005597	0.1451	0.1502	0.01273
Max.	0.1909	0.734	0.7216	0.239

Testing Rsq Violin Plots

Number of successfully trained protein prediction model

Testing Rsq Summary statistics

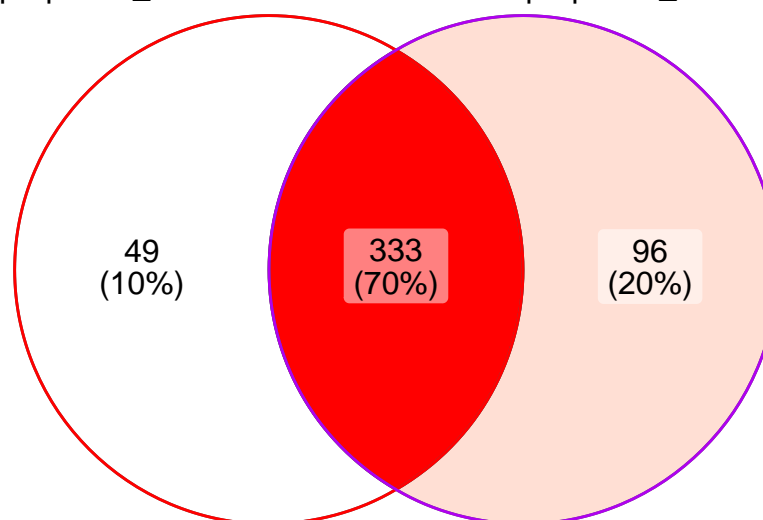|	proposed_0.000005 – only	proposed_0.000005 – common	proposed_0.0001 – common	proposed_0.0001 – only
Min.	0.0001958	0.000119	0.0001034	0.0001204
1st Qu.	0.001088	0.004801	0.01196	0.001602
Median	0.002779	0.03133	0.04105	0.004299
Mean	0.033	0.09459	0.09576	0.01343
3rd Qu.	0.007278	0.1428	0.1343	0.01471
Max.	0.4614	0.734	0.7099	0.2374

Testing Rsq Violin Plots
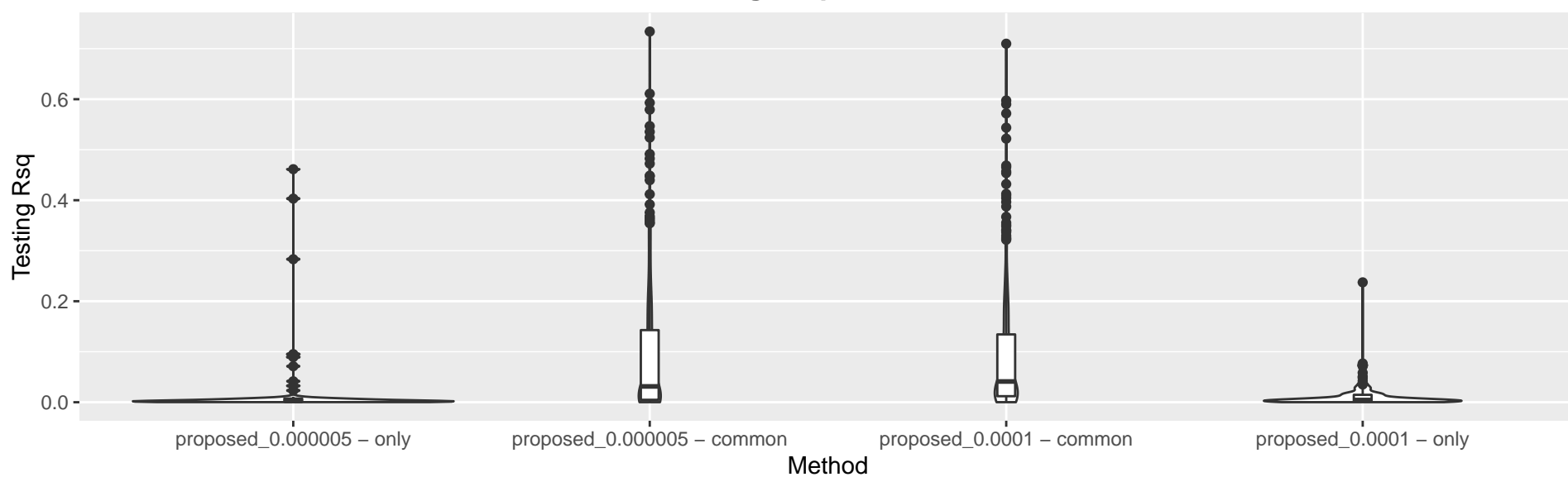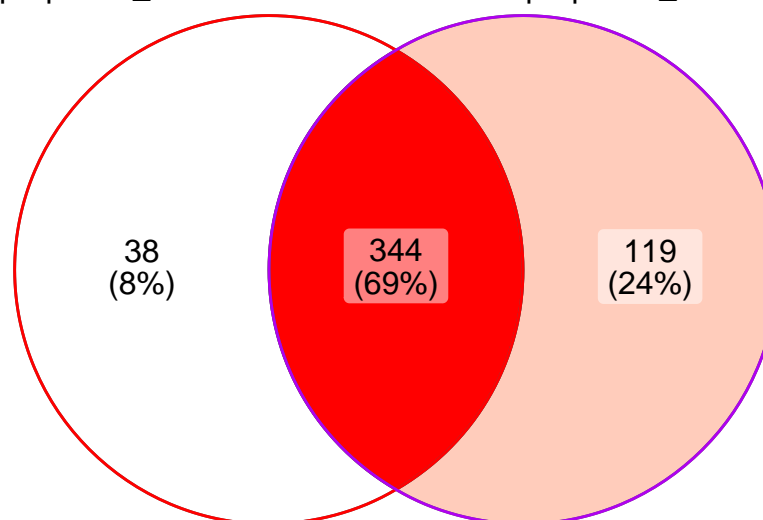
Number of successfully trained protein prediction model

Testing Rsq Summary statistics

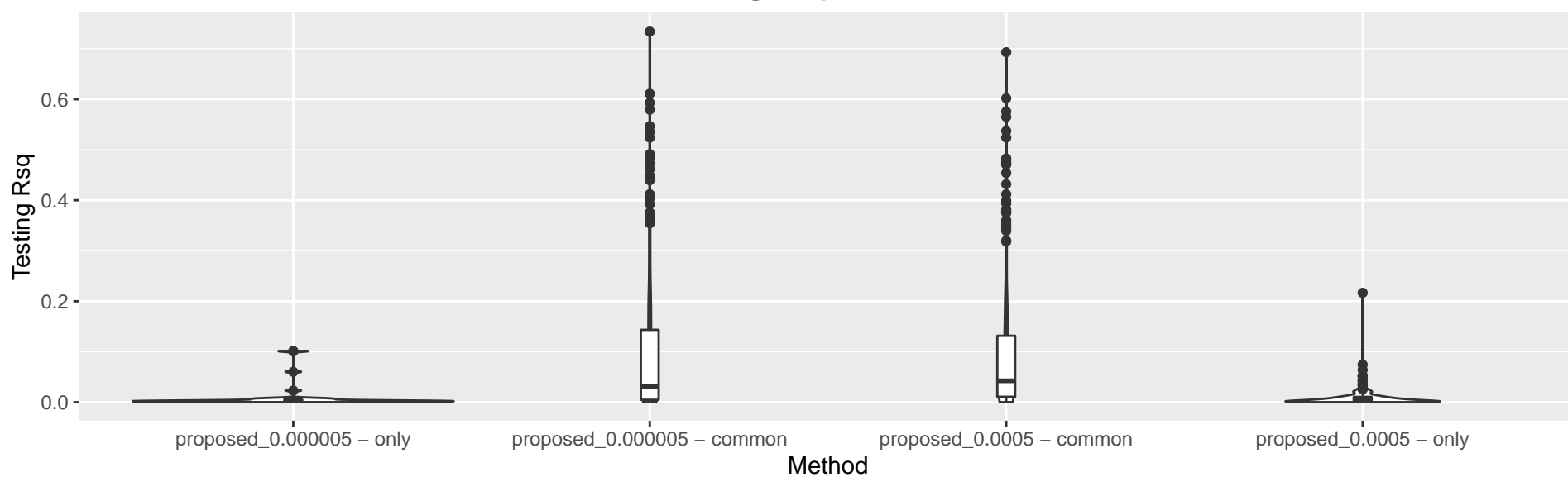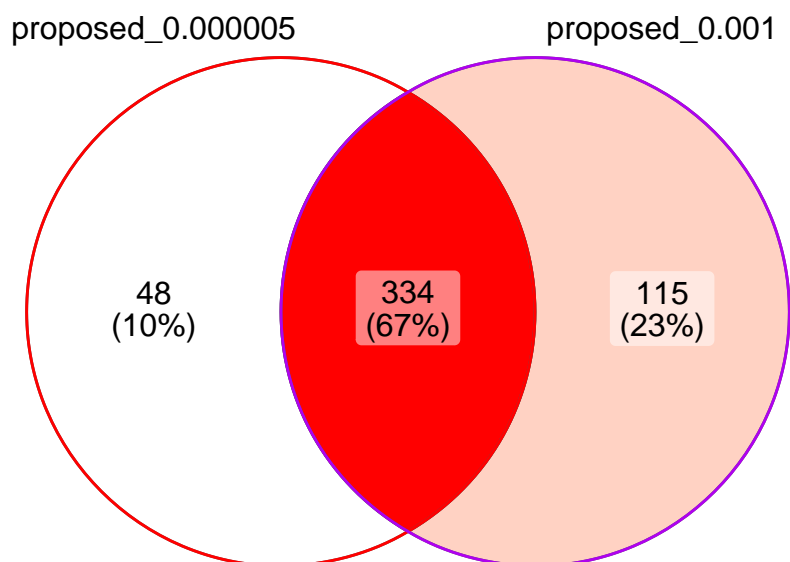|	proposed_0.000005 – only	proposed_0.000005 – common	proposed_0.0005 – common	proposed_0.0005 – only
Min.	0.0001958	0.000119	0.0001282	0.0001309
1st Qu.	0.001739	0.004758	0.01079	0.001206
Median	0.002685	0.03108	0.04249	0.004706
Mean	0.01013	0.09515	0.09345	0.01093
3rd Qu.	0.005717	0.1434	0.1316	0.01065
Max.	0.1016	0.734	0.6931	0.2168

Number of successfully trained protein prediction model

Testing Rsq Summary statistics

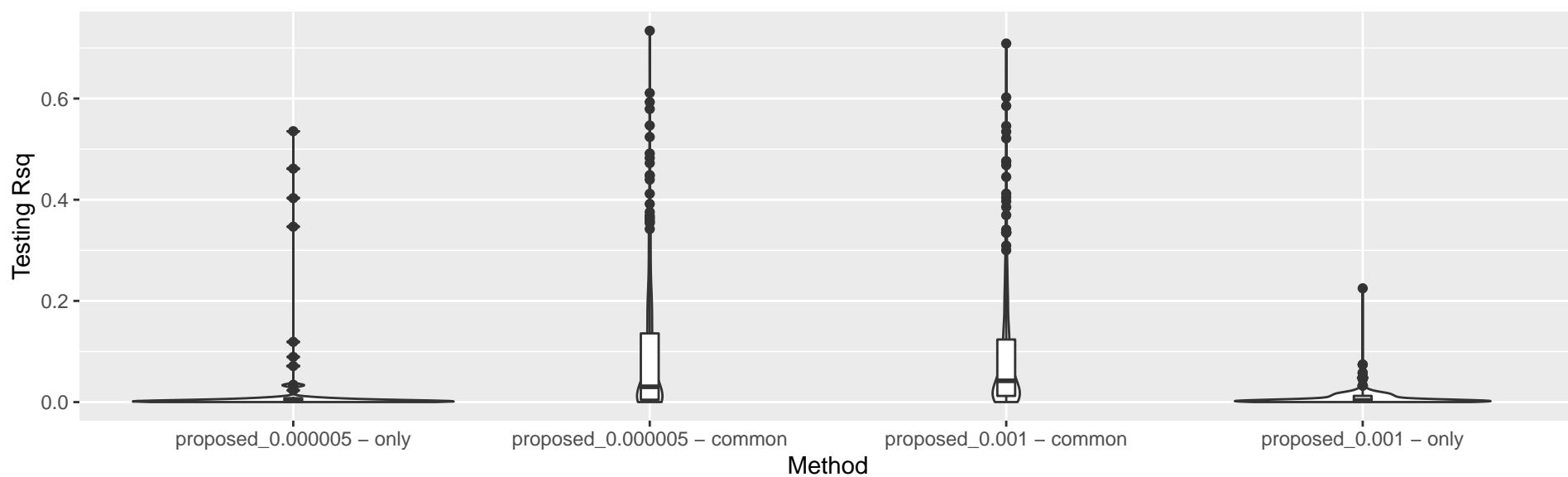|	proposed_0.000005 – only	proposed_0.000005 – common	proposed_0.001 – common	proposed_0.001 – only
Min.	0.000119	0.0001364	0.0001142	0.0001092
1st Qu.	0.0005785	0.004671	0.01215	0.001236
Median	0.002573	0.0303	0.04209	0.003256
Mean	0.04617	0.09251	0.09017	0.01147
3rd Qu.	0.007805	0.1358	0.1237	0.01233
Max.	0.5355	0.734	0.7087	0.2249

Testing Rsq Violin Plots

Number of successfully trained protein prediction model
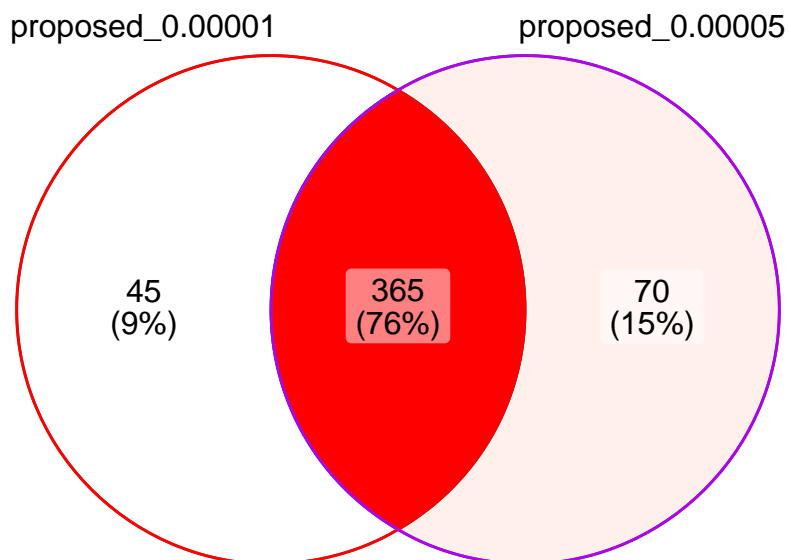
Testing Rsq Summary statistics

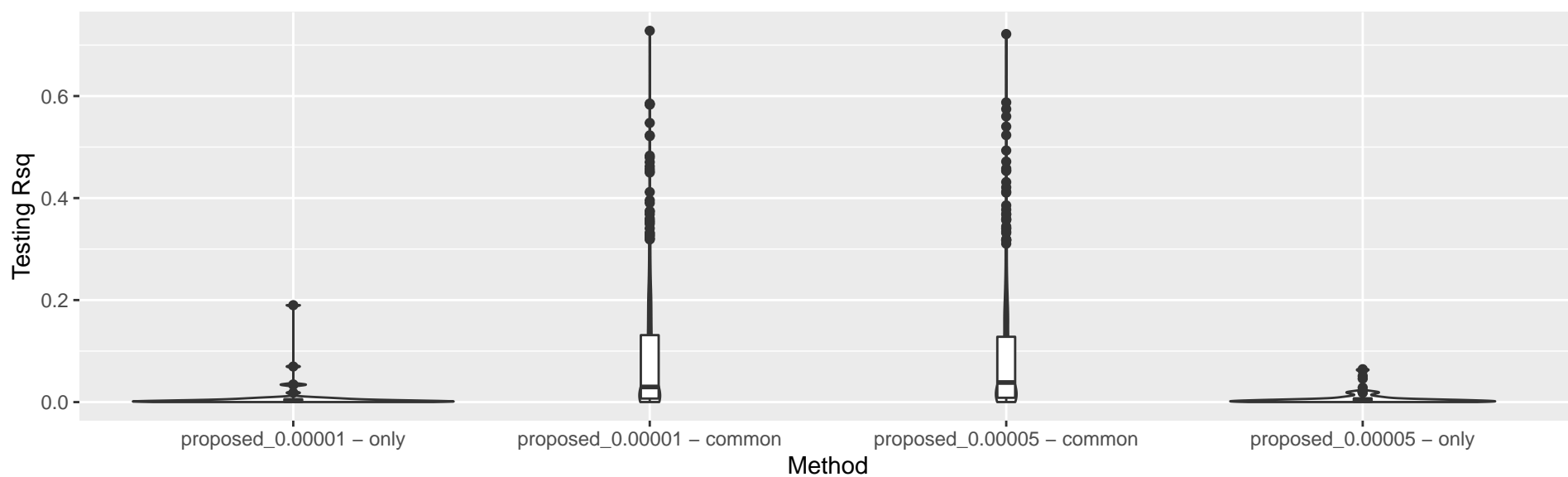|	proposed_0.00001 – only	proposed_0.00001 – common	proposed_0.00005 – common	proposed_0.00005 – only
Min.	0.0001366	0.0001104	0.0001255	0.0001114
1st Qu.	0.0004459	0.006986	0.008484	0.0005487
Median	0.002505	0.02971	0.03839	0.003495
Mean	0.01018	0.09205	0.09235	0.007769
3rd Qu.	0.005822	0.1313	0.1281	0.007533
Max.	0.19	0.728	0.7216	0.0645

Testing Rsq Violin Plots

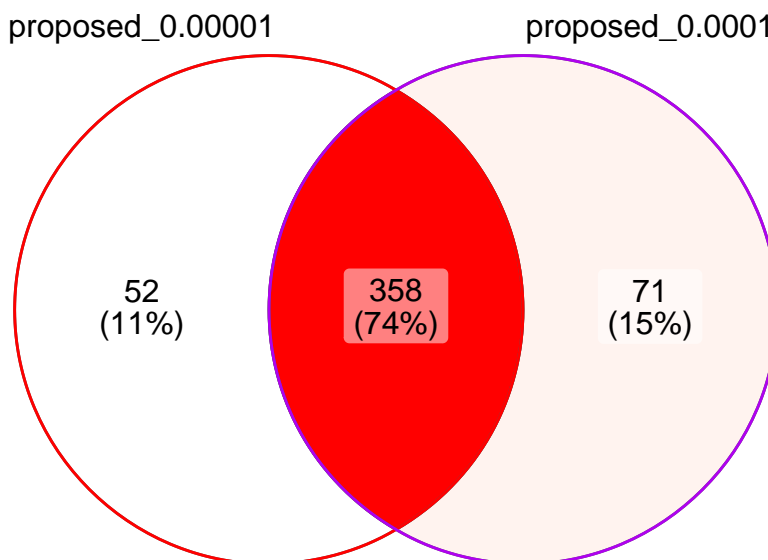## Number of successfully trained protein prediction model

Testing Rsq Summary statistics

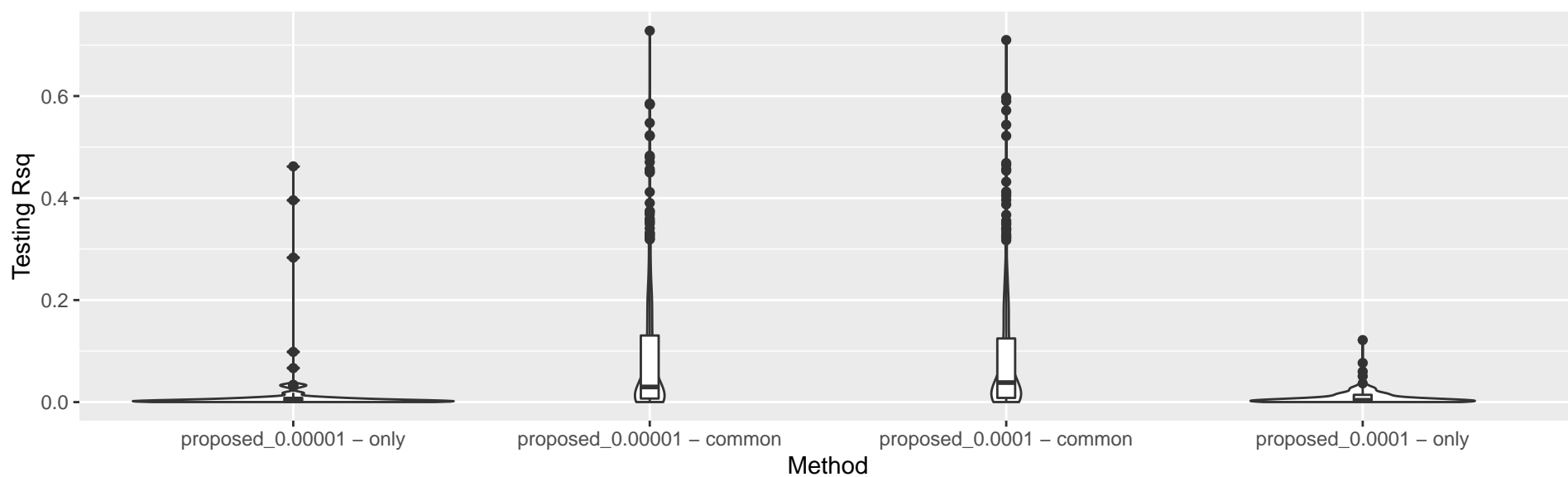|	proposed_0.00001 – only	proposed_0.00001 – common	proposed_0.0001 – common	proposed_0.0001 – only
Min.	0.0001202	0.0001104	0.0001034	0.0001204
1st Qu.	0.0004827	0.006945	0.008198	0.0009872
Median	0.003372	0.02968	0.03825	0.00355
Mean	0.03005	0.09077	0.09048	0.01107
3rd Qu.	0.008401	0.1306	0.1248	0.01426
Max.	0.4621	0.728	0.7099	0.1216

Testing Rsq Violin Plots

Number of successfully trained protein prediction model

Testing Rsq Summary statistics

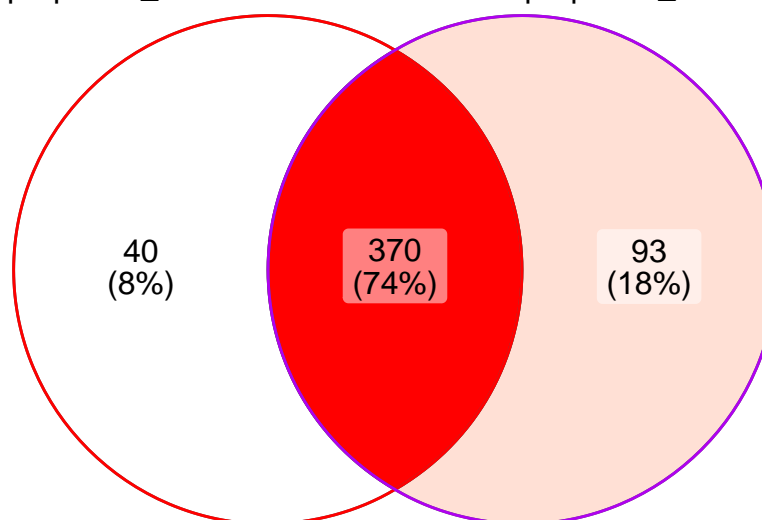|	proposed_0.00001 – only	proposed_0.00001 – common	proposed_0.0005 – common	proposed_0.0005 – only
Min.	0.0001198	0.0001104	0.0001526	0.0001282
1st Qu.	0.0005409	0.006627	0.00863	0.001162
Median	0.001866	0.02955	0.03574	0.003307
Mean	0.009029	0.09107	0.08773	0.01061
3rd Qu.	0.005684	0.1306	0.1239	0.01138
Max.	0.1031	0.728	0.6931	0.1167

Testing Rsq Violin Plots

Number of successfully trained protein prediction model

Testing Rsq Summary statistics

	proposed_0.00001 – only	proposed_0.00001 – common	proposed_0.001 – common	proposed_0.001 – only
Min.	0.0001104	0.0001198	0.0001092	0.0001142
1st Qu.	0.0007206	0.005883	0.009266	0.001203
Median	0.003604	0.02887	0.03492	0.003353
Mean	0.04346	0.08819	0.0841	0.01054
3rd Qu.	0.008432	0.1262	0.1139	0.01244
Max.	0.5215	0.728	0.7087	0.1207

Testing Rsq Violin Plots
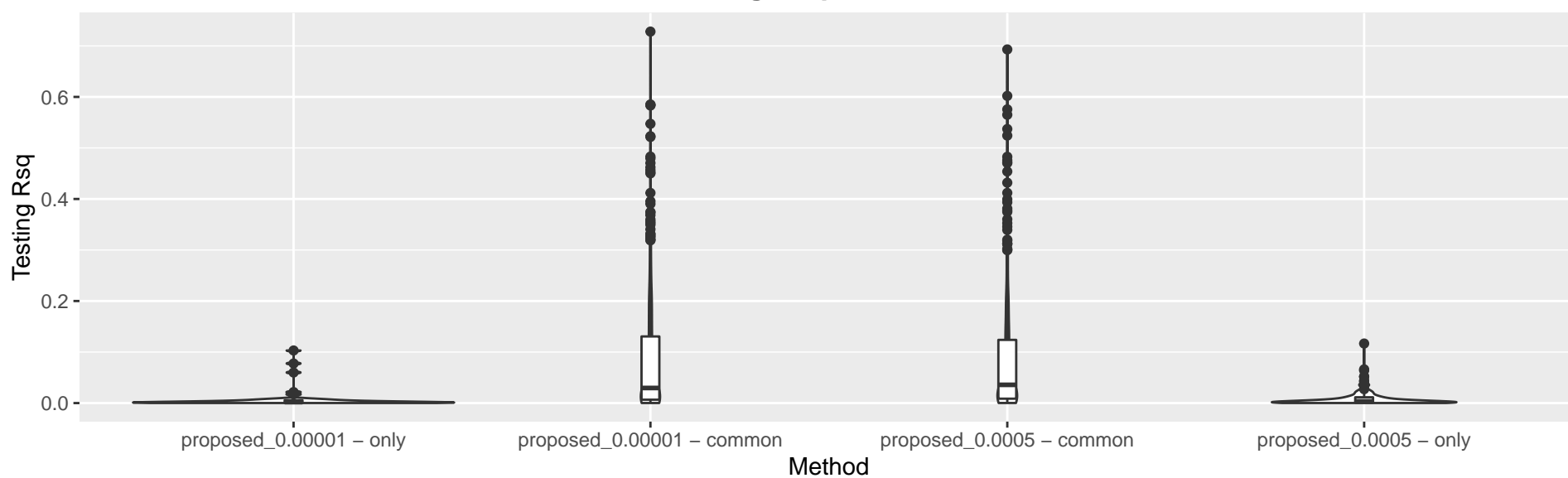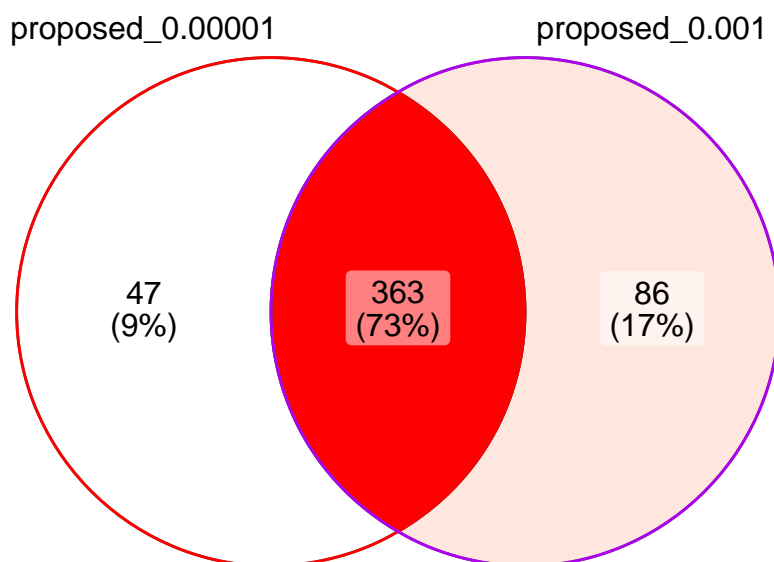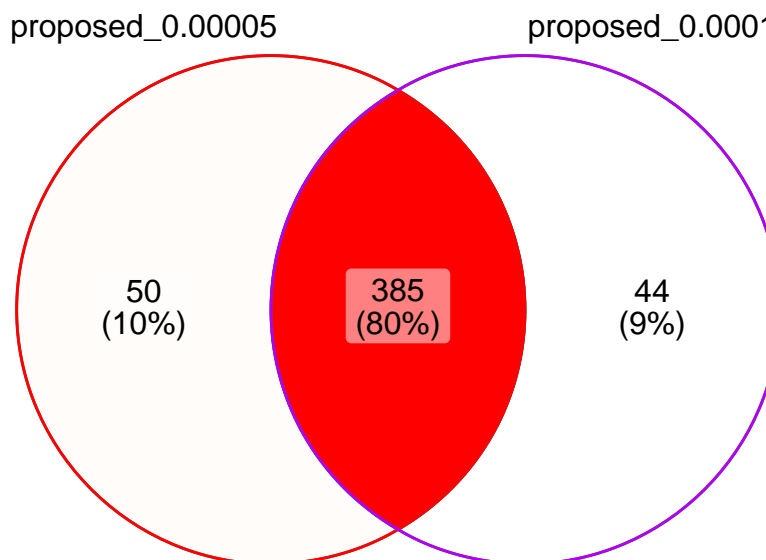
Number of successfully trained protein prediction model

Testing Rsq Summary statistics

	proposed_0.00005 – only	proposed_0.00005 – common	proposed_0.0001 – common	proposed_0.0001 – only
Min.	0.0001114	0.0001186	0.0001034	0.0001204
1st Qu.	0.000845	0.006115	0.00703	0.001114
Median	0.002883	0.02993	0.03138	0.003114
Mean	0.03083	0.08496	0.08418	0.01746
3rd Qu.	0.009823	0.1118	0.1138	0.01464
Max.	0.4314	0.7216	0.7099	0.1948

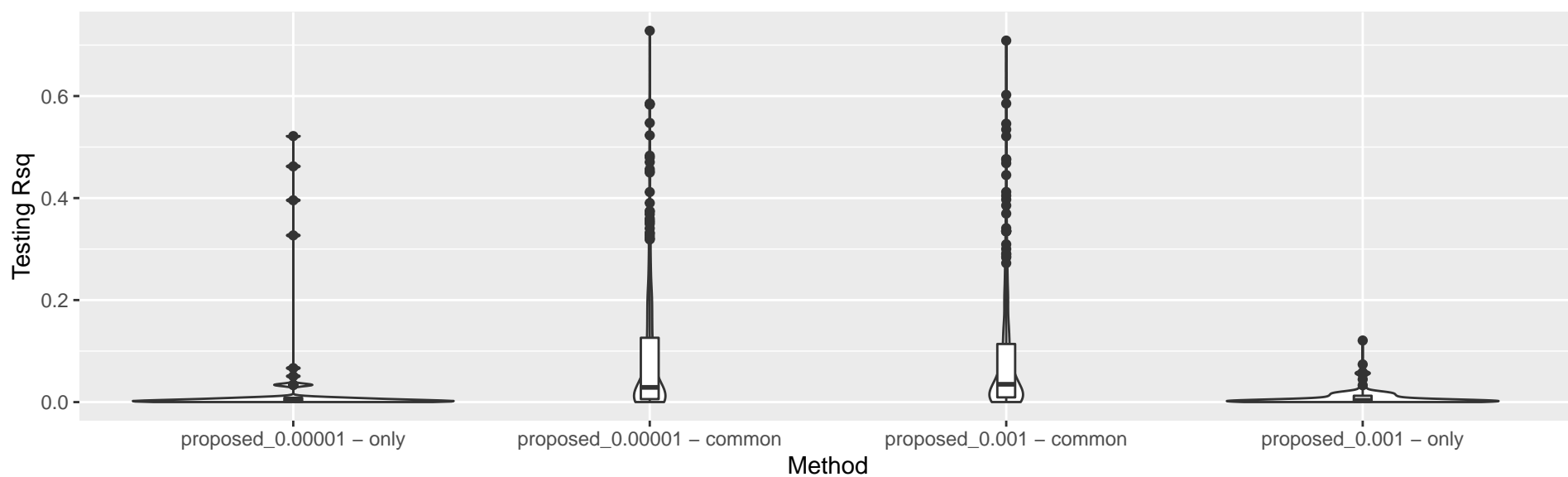## Testing Rsq Violin Plots

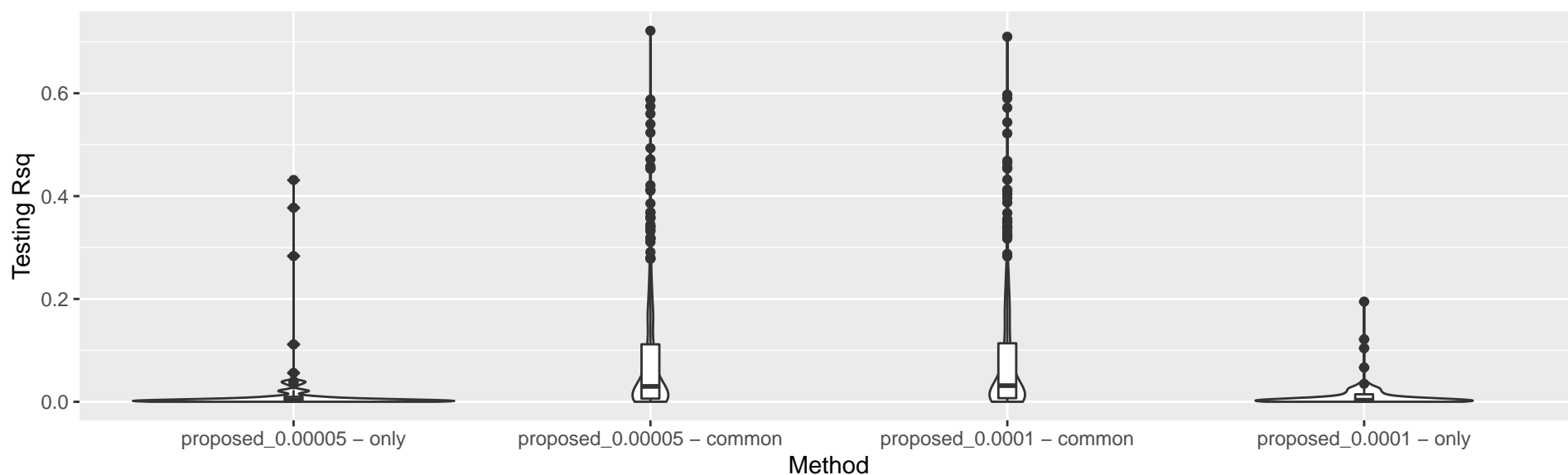# Number of successfully trained protein prediction model

Testing Rsq Summary statistics

	proposed_0.00005 – only	proposed_0.00005 – common	proposed_0.0005 – common	proposed_0.0005 – only
Min.	0.0001114	0.0001186	0.0001282	0.0001419
1st Qu.	0.0007337	0.005873	0.007825	0.001013
Median	0.003386	0.02993	0.0308	0.001949
Mean	0.007971	0.08551	0.08216	0.01259
3rd Qu.	0.005131	0.1118	0.112	0.00802
Max.	0.08233	0.7216	0.6931	0.1818

Testing Rsq Violin Plots

Number of successfully trained protein prediction model

proposed_0.00005   proposed_0.001
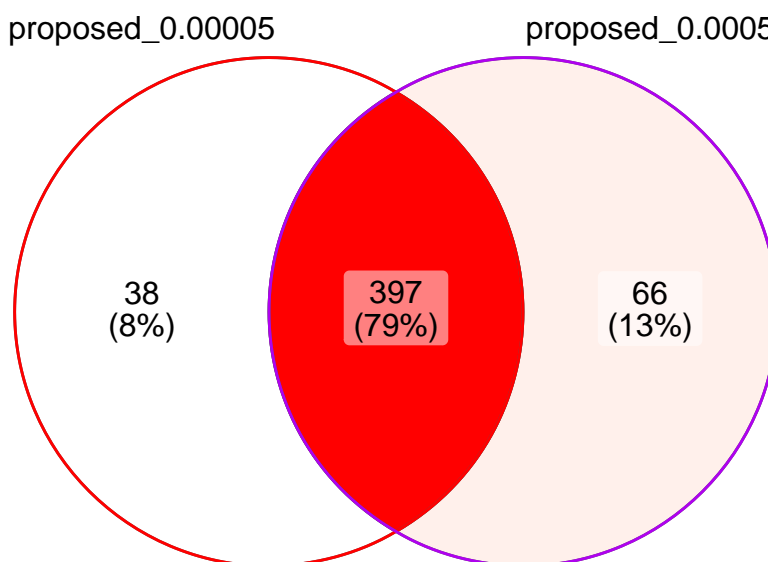

52
(10%)   383
(76%)   66
(13%)

Testing Rsq Summary statistics

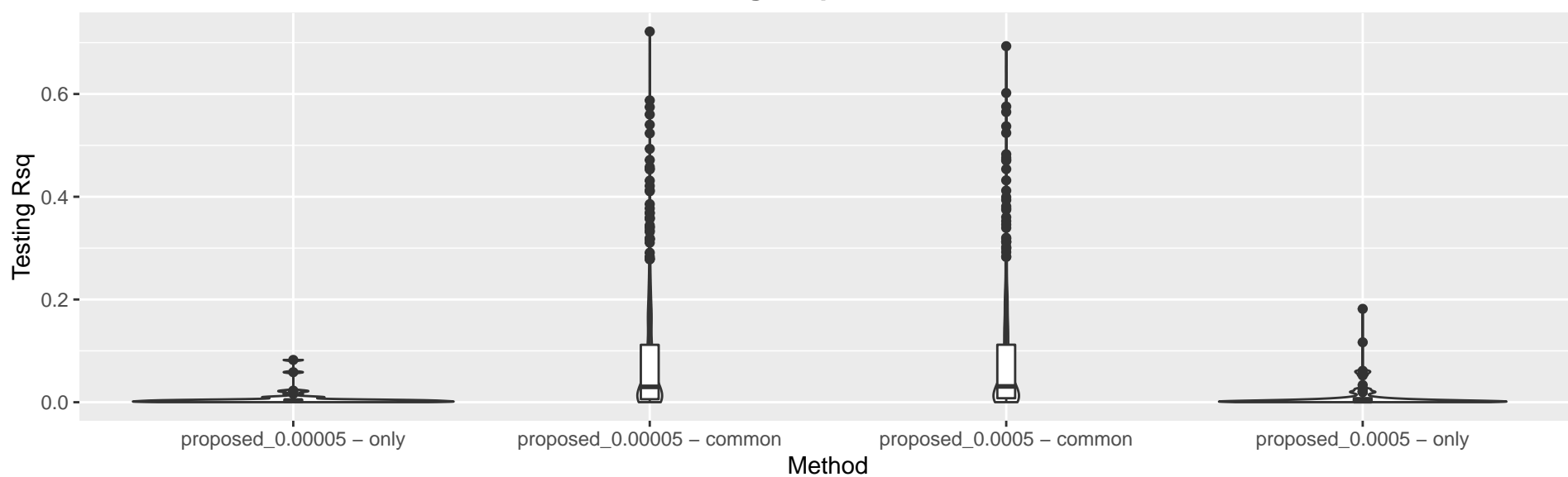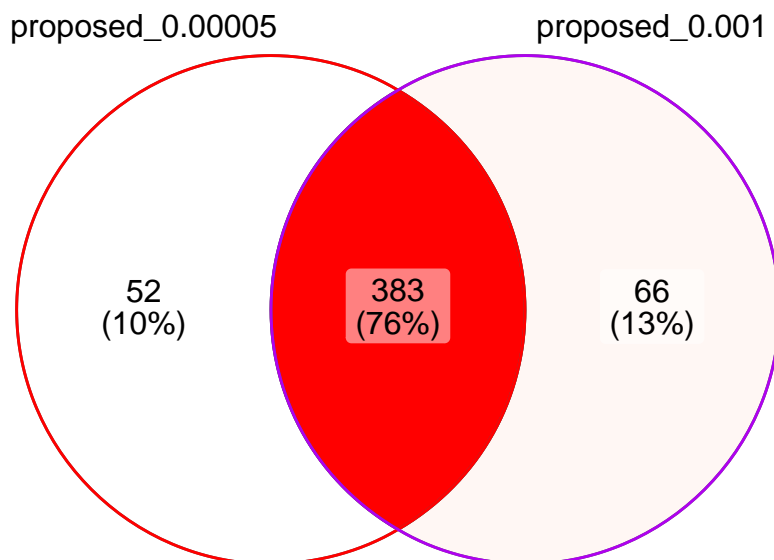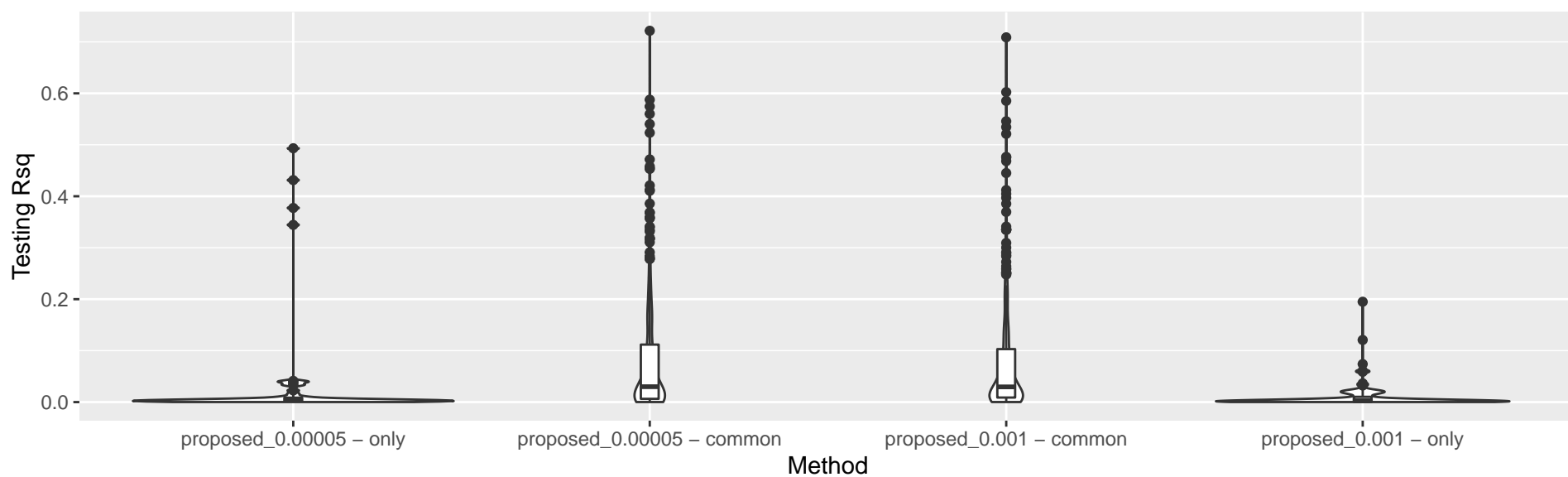|	proposed_0.00005 – only	proposed_0.00005 – common	proposed_0.001 – common	proposed_0.001 – only
Min.	0.0001114	0.0001186	0.0001092	0.0001142
1st Qu.	0.00158	0.006344	0.008825	0.0007038
Median	0.003873	0.02993	0.02959	0.002642
Mean	0.03842	0.08421	0.07982	0.01307
3rd Qu.	0.008577	0.1117	0.103	0.01025
Max.	0.4933	0.7216	0.7087	0.195

Testing Rsq Violin Plots

Number of successfully trained protein prediction model

Testing Rsq Summary statistics

	proposed_0.0001 – only	proposed_0.0001 – common	proposed_0.0005 – common	proposed_0.0005 – only
Min.	0.0001034	0.0001102	0.0001282	0.0001309
1st Qu.	0.001301	0.007063	0.007893	0.0008891
Median	0.00323	0.03078	0.02956	0.001929
Mean	0.007635	0.08333	0.08062	0.02357
3rd Qu.	0.00608	0.1131	0.1023	0.008131
Max.	0.07504	0.7099	0.6931	0.3813

Testing Rsq Violin Plots
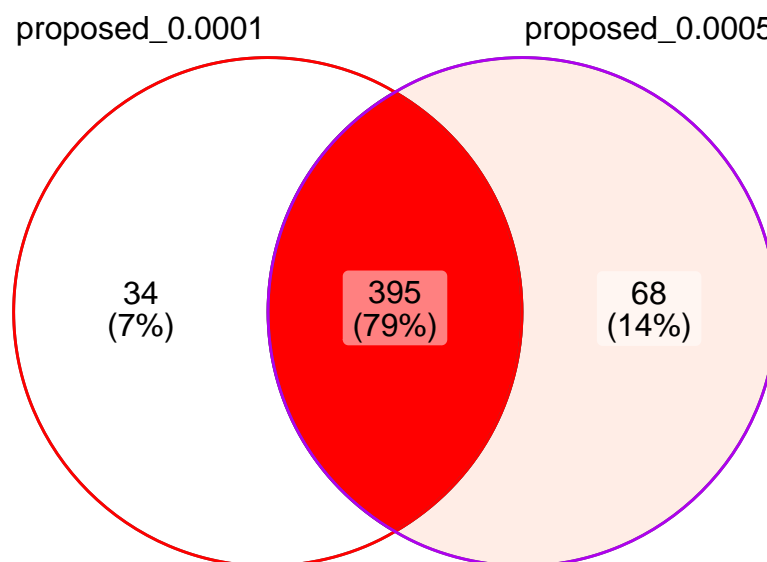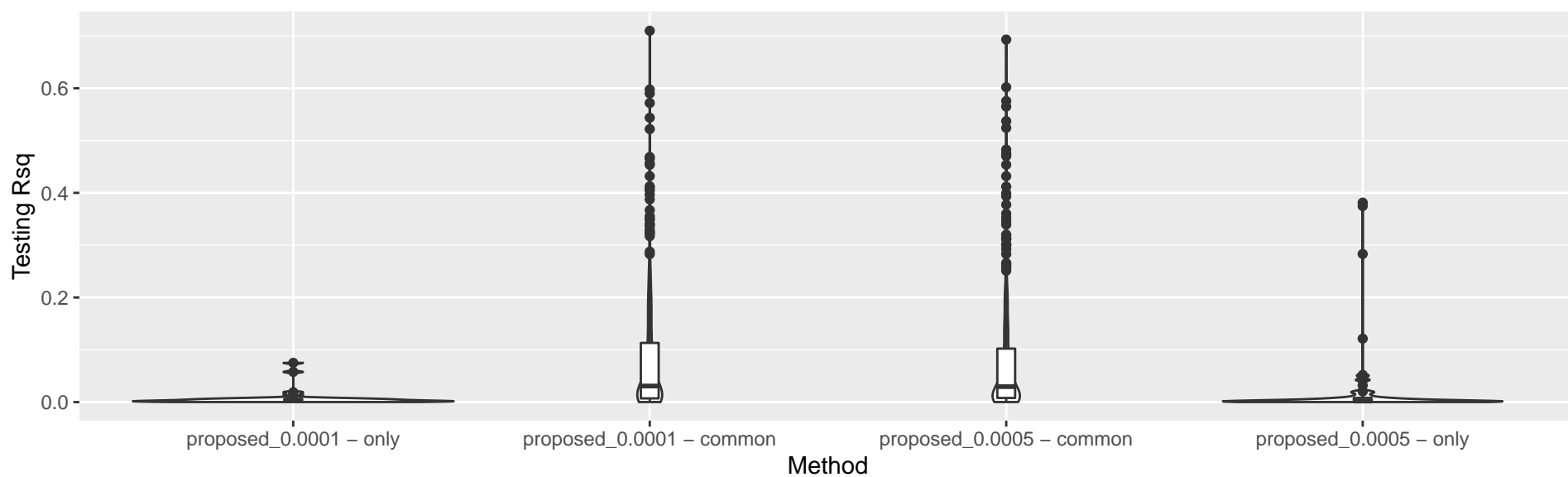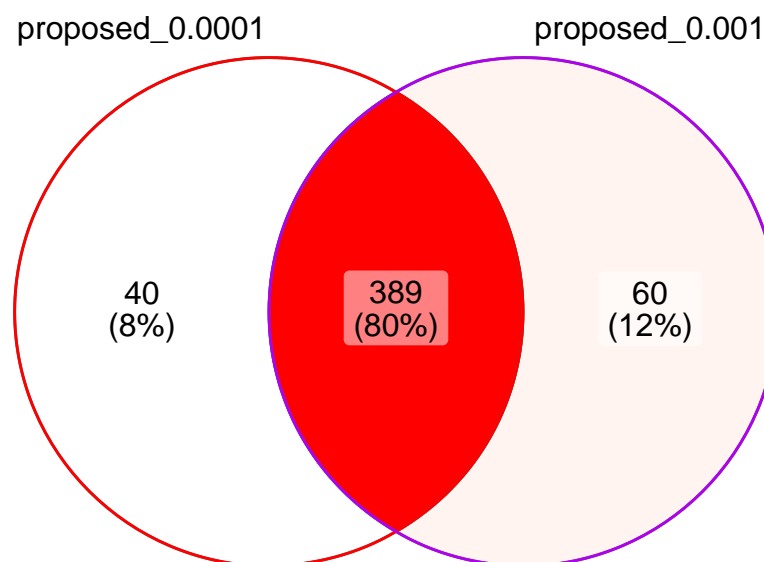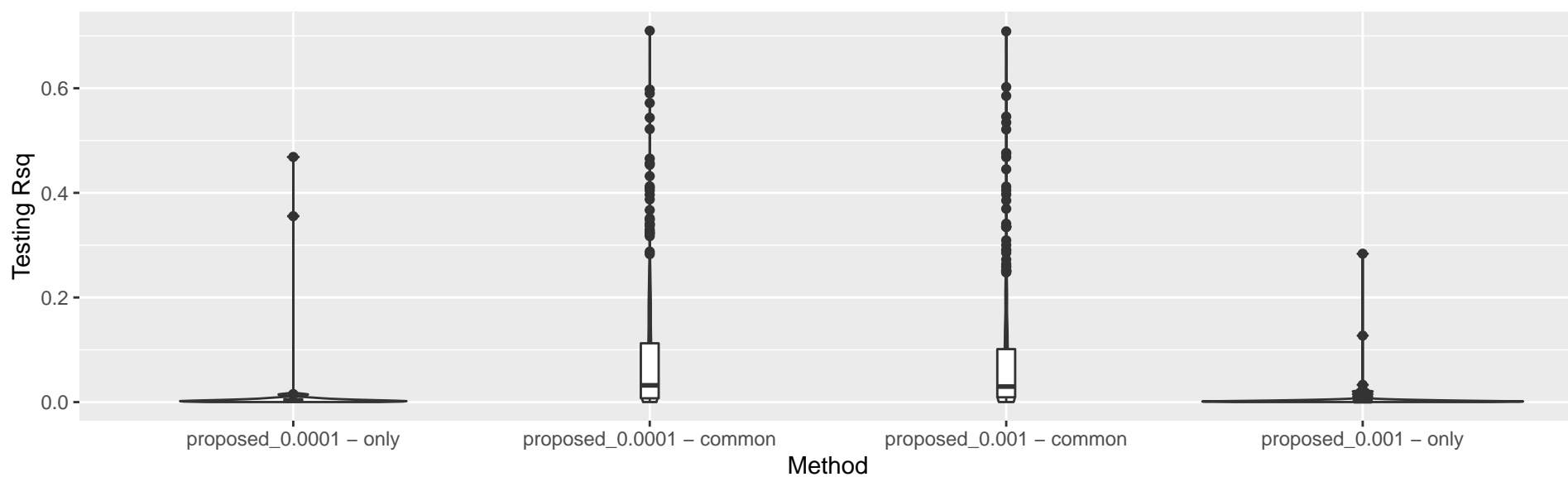
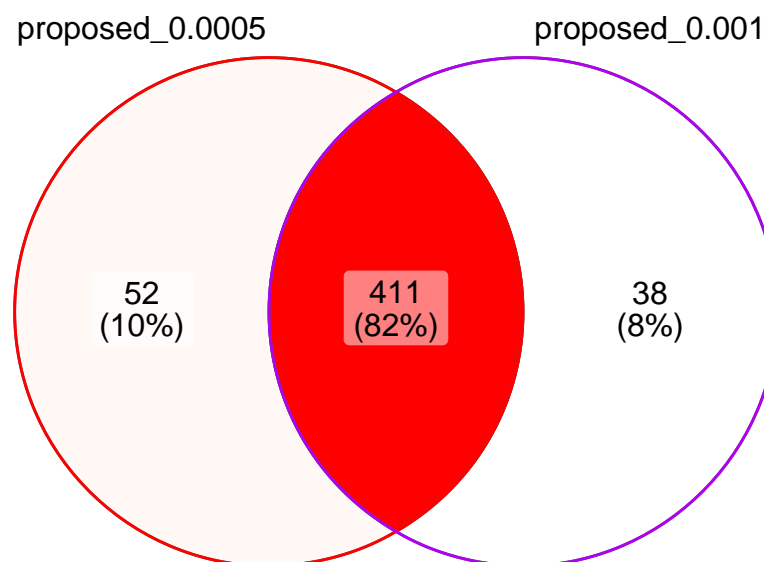# Number of successfully trained protein prediction model

Testing Rsq Summary statistics

	proposed_0.0001 – only	proposed_0.0001 – common	proposed_0.001 – common	proposed_0.001 – only
Min.	0.0001034	0.0001102	0.0001142	0.0001092
1st Qu.	0.001212	0.007237	0.009196	0.0007552
Median	0.002774	0.03201	0.0297	0.001579
Mean	0.02415	0.0828	0.07913	0.01085
3rd Qu.	0.005729	0.1124	0.1013	0.00425
Max.	0.4686	0.7099	0.7087	0.2837

Testing Rsq Violin Plots

Number of successfully trained protein prediction model

Testing Rsq Summary statistics

	proposed_0.0005 – only	proposed_0.0005 – common	proposed_0.001 – common	proposed_0.001 – only
Min.	0.0001327	0.0001282	0.0001511	0.0001092
1st Qu.	0.0005853	0.00643	0.006117	0.0005843
Median	0.002675	0.02581	0.02591	0.001476
Mean	0.03537	0.07691	0.07571	0.008379
3rd Qu.	0.006337	0.09624	0.09271	0.004092
Max.	0.4827	0.6931	0.7087	0.08553

Testing Rsq Violin Plots