week9_work

Eric Weine

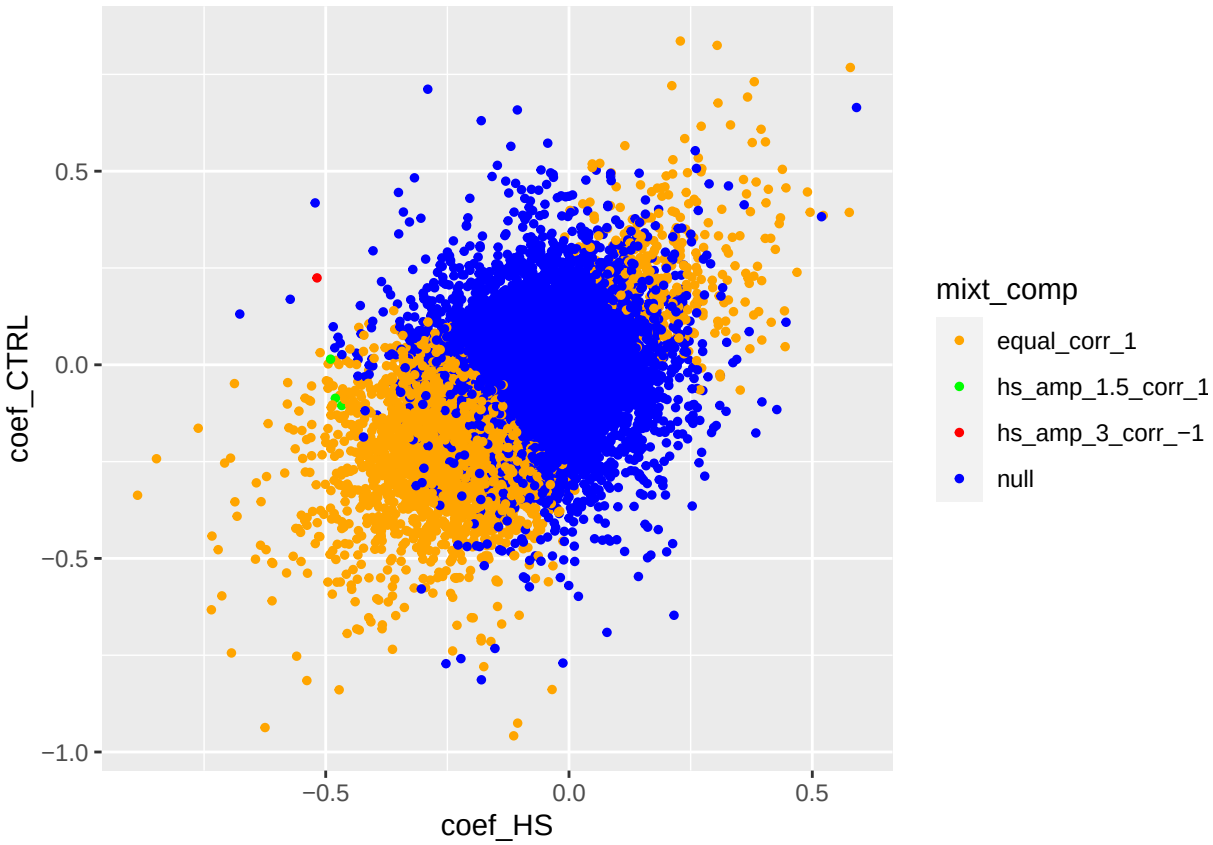
4/28/2022

Pallares Classification

We will use the sparse mash model to classify signals this week. Below is a comparison of the classifications for a sample of 12030 SNPs (1 from each LD block). The rows are Pallares classifications and the columns are MASH classifications.

	equal_corr_1	hs_amp_1.5_corr_1	hs_amp_3_corr_-1	null	ctrl_spec	hs_spec
null	3611	0	0	8313	0	0
shared	79	0	1	0	0	0
hs_spec	23	3	0	0	0	0
ctrl_spec	0	0	0	0	0	0

Below is a plot of the estimated regression coefficients of Pallares colored by mash classification. Note that the columns in the table above that sum to zero are not shown in the plot legend. I felt that this made the plot easier to understand but we can certainly change this.

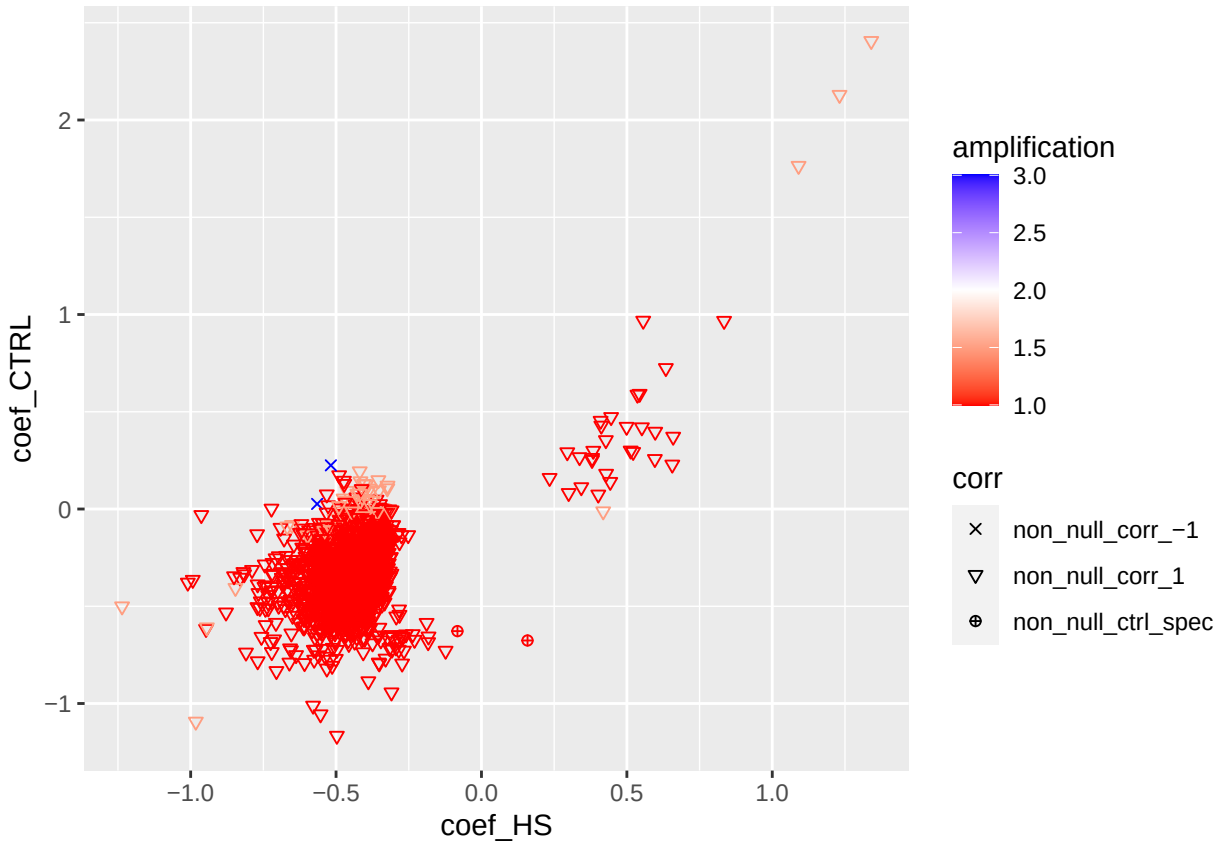


Note that changing the shape and varying the color doesn't look good because the points are too bunched up, even when I downsample to 10% of the data

It is also instructive to look only at the SNPs that Pallares classifies as significant.

	equal_corr_1	hs_amp_1.5_corr_1	hs_amp_3_corr_-1	null	ctrl_spec	hs_spec
shared	1538	7	1	0	0	0
hs_spec	641	59	1	0	0	0
ctrl_spec	1	0	0	0	2	0

Again, below is a plot of the regression coefficients colored by mash classification.


coef_CTRL
coef_HS
2
1
0
−1
−1.0
−0.5
0.0
0.5
1.0
amplification
3.0
2.5
2.0
1.5
1.0
corr
non_null_corr_−1
non_null_corr_1
non_null_ctrl_spec